Hyphenating chromatographic, electrophoretic and on-line immobilized enzymatic strategies for improved oligonucleotide analysis

Thesis submitted to the Faculty of Science in fulfilment of the requirements for the degree of Doctor in Science (Chemistry)

PIOTR WIKTOR ALVAREZ POREBSKI

Promotor
Prof. Dr. Frederic Lynen

2016
Table of Contents

TABLE OF CONTENTS........................................................................................................................................ 1

I GENERAL INTRODUCTION AND SCOPE ........................................................................................................ 1
  I.1 INTRODUCTION ........................................................................................................................................... 1
  I.2 SCOPE ......................................................................................................................................................... 3
  I.3 REFERENCES .............................................................................................................................................. 4

II PRINCIPLES OF THE SEPARATION TECHNIQUES USED IN OLIGONUCLEOTIDE ANALYSIS ................. 7
  SUMMARY ....................................................................................................................................................... 7
 II.1 PRINCIPLES OF HIGH PERFORMANCE LIQUID CHROMATOGRAPHY (HPLC) .............................................. 8
      II.1.1 Ion exchange chromatography ........................................................................................................... 10
      II.1.2 Ion pair chromatography .................................................................................................................. 11
      II.1.3 Hydrophilic interaction liquid chromatography ................................................................................ 12
 II.2 FUNDAMENTAL CHROMATOGRAPHIC PARAMETERS .................................................................................. 13
      II.2.1 Retention factor ................................................................................................................................ 15
      II.2.2 Selectivity factor ................................................................................................................................ 16
      II.2.3 Efficiency .......................................................................................................................................... 16
      II.2.4 Resolution ......................................................................................................................................... 16
      II.2.5 Van Deemter equation ....................................................................................................................... 17
      II.2.6 Peak capacity ...................................................................................................................................... 19
 II.3 ULTRAHIGH PERFORMANCE LIQUID CHROMATOGRAPHY .................................................................. 22
 II.4 PRINCIPLES OF ELECTRODRIVEN SEPARATION TECHNIQUES USED FOR OLIGONUCLEOTIDE ANALYSIS ....... 24
      II.4.1 Basic principles of capillary electrophoresis ....................................................................................... 26
          II.4.1.1 Electrophoresis ........................................................................................................................... 26
          II.4.1.2 Electroosmosis ........................................................................................................................... 27
          II.4.1.3 Apparent electrophoretic mobility and migration time .............................................................. 28
      II.4.2 Zone dispersion in capillary electrophoresis ...................................................................................... 28
      II.4.3 Resolution in CE ............................................................................................................................... 29
      II.4.4 Modes of operation in capillary electrophoresis ............................................................................... 29
Table of contents

II.4.4.1 Capillary zone electrophoresis (CZE) ................................................................. 29
II.4.4.2 Capillary gel electrophoresis (CGE) ................................................................. 30
   II.4.4.2.1 Entangled polymer solutions ................................................................. 30
      II.4.4.2.1.1 Entanglement threshold ............................................................... 31
      II.4.4.2.1.2 Mesh size ....................................................................................... 31
      II.4.4.2.1.3 Mechanisms of electrophoretic migration in entangled polymer solutions ................................................................. 33
II.5 Fundamental aspects of multidimensional separation techniques ........................................ 34
   II.5.1 Orthogonality, sampling frequency and peak capacity ........................................ 37
      II.5.1.1 Orthogonality ............................................................................................. 37
      II.5.1.2 Sampling frequency .................................................................................. 38
      II.5.1.3 Peak capacity in 2D separations ............................................................... 38
II.6 References ..................................................................................................................... 40

III Oligonucleotides: Characteristics and contemporary analytical strategies ........... 43

Summary ............................................................................................................................. 43

III.1 Introduction ................................................................................................................. 44
III.2 Relevant mechanisms of action of antisense oligonucleotides ................................ 49
III.3 Relevance of nuclease resistance of oligonucleotides ............................................. 51
III.4 Toxicity aspects of antisense oligonucleotides ....................................................... 54
III.5 Analytical separation methods for the analysis of oligonucleotides ....................... 55
   III.5.1 Electrophoretic methods for the analysis of oligonucleotides ......................... 56
   III.5.2 Liquid chromatography methods for the analysis of oligonucleotides ............... 59
      III.5.2.1 Ion-exchange chromatography for the analysis of oligonucleotides .......... 59
      III.5.2.2 Ion-pair chromatography for the analysis of oligonucleotides ................. 61
      III.5.2.3 Hydrophilic interaction liquid chromatography for the analysis of Oligonucleotides ................................................................. 63
      III.5.2.4 Mixed mode, affinity and bidimensional separation methods for the analysis of oligonucleotides ................................................................. 65
III.6 References ..................................................................................................................... 68

IV Combining liquid chromatography with multiplexed capillary gel electrophoresis for offline comprehensive analysis of complex oligonucleotide samples ................................................................. 75

Summary ............................................................................................................................. 75

IV.1 Introduction ................................................................................................................. 76
Table of contents

IV.2 EXPERIMENTAL .................................................................................................................. 79
  IV.2.1 Chemicals .................................................................................................................. 79
  IV.2.2 Instrumentation ......................................................................................................... 79
  IV.2.3 Chromatographic conditions ..................................................................................... 80
  IV.2.4 Electrophoretic conditions ........................................................................................ 80
    IV.2.4.1 Single capillary system ....................................................................................... 80
    IV.2.4.2 Multiplexed capillary system .............................................................................. 81

IV.3 RESULTS AND DISCUSSION ............................................................................................ 82
  IV.3.1 Offline comprehensive ON analysis by IEC x CGE ................................................... 82
  IV.3.2 Offline comprehensive ON analysis by IPC x CGE ................................................... 86
  IV.3.3 Peak capacity of IEC x CGE and IPC x CGE approaches .......................................... 89
  IV.3.4 Figures of merit of the developed methodology ....................................................... 91

IV.4 CONCLUSIONS ................................................................................................................ 93

IV.5 REFERENCES .................................................................................................................. 93

V POSSIBILITIES AND LIMITATIONS OF SPECTRAL DECONVOLUTION IN PHOTODIODE ARRAY DETECTION SEPARATIONS OF OLIGONUCLEOTIDES ................................................................................. 95
  SUMMARY ............................................................................................................................. 95
  V.1 INTRODUCTION .............................................................................................................. 100
  V.2 EXPERIMENTAL .............................................................................................................. 100
    V.2.1 Chemicals .................................................................................................................. 100
    V.2.2 Instrumentation and conditions ............................................................................... 100
    V.2.3 Data treatment ......................................................................................................... 101

V.3 RESULTS AND DISCUSSION ............................................................................................ 101

V.4 CONCLUSIONS ................................................................................................................ 106

V.5 REFERENCES .................................................................................................................. 107

VI OVERVIEW OF ENZYME IMMOBILIZATION APPROACHES FOR ON-LINE COUPLING TO HPLC ......................................................................................................................... 109
  SUMMARY ............................................................................................................................. 109
  VI.1 INTRODUCTION ............................................................................................................ 110
  VI.2 ENZYME IMMOBILIZATION ........................................................................................ 110
    VI.2.1 Covalent immobilization ........................................................................................ 111
    VI.2.2 Physical adsorption ............................................................................................... 114
    VI.2.3 Enzyme entrapment .............................................................................................. 115

III
## Table of contents

VI.3 IMMOBILIZATION SUPPORTS .................................................................................................................. 116
   VI.3.1 Inorganic materials ............................................................................................................................. 116
   VI.3.2 Organic materials ............................................................................................................................... 117
VI.4 KINETICS OF IMMOBILIZED ENZYMES ................................................................................................. 118
VI.5 REFERENCES ............................................................................................................................................. 122

VII HYPHENATION OF A DEOXYRIBONUCLEASE I IMMOBILIZED ENZYME REACTOR WITH LIQUID CHROMATOGRAPHY FOR THE ON-LINE STABILITY EVALUATION OF OLIGONUCLEOTIDES ....... 125
   SUMMARY ..................................................................................................................................................... 125
VII.1 INTRODUCTION ....................................................................................................................................... 126
VII.2 EXPERIMENTAL ....................................................................................................................................... 129
   VII.2.1 Chemicals .......................................................................................................................................... 129
   VII.2.2 Instrumentation .................................................................................................................................. 129
   VII.2.3 Preparation of epoxy-functionalized silica ......................................................................................... 130
   VII.2.4 Immobilization of DNase I ................................................................................................................ 130
   VII.2.5 Preparation of the IMERs .................................................................................................................. 130
   VII.2.6 Kinetic determinations ....................................................................................................................... 131
       VII.2.6.1 Activity of free DNase I .............................................................................................................. 131
       VII.2.6.2 Activity of immobilized DNase I ................................................................................................. 131
   VII.2.7 On-line IMER-LC setup .................................................................................................................... 132
VII.3 RESULTS AND DISCUSSION .................................................................................................................... 134
   VII.3.1 Immobilization of DNase I ................................................................................................................ 134
   VII.3.2 Activity of free and immobilized DNase .......................................................................................... 135
   VII.3.3 On-line IMER-LC ............................................................................................................................... 138
       VII.3.3.1 On-line IMER-IPC ...................................................................................................................... 139
       VII.3.3.2 On-line IMER-IEC ...................................................................................................................... 141
   VII.3.4 Stability evaluation of antisense ONs ............................................................................................... 143
VII.4 CONCLUSION ........................................................................................................................................... 147
VII.5 REFERENCES ........................................................................................................................................... 147

VIII OFFLINE COMPREHENSIVE LIQUID CHROMATOGRAPHY IN COMBINATION WITH A DEOXYRIBONUCLEASE I IMMOBILIZED ENZYMATIC REACTOR FOR SELECTIVE SCREENING OF OLIGONUCLEOTIDE MIXTURES ........................................................................................................................ 151
   SUMMARY ..................................................................................................................................................... 151
# Table of contents

<table>
<thead>
<tr>
<th>Section</th>
</tr>
</thead>
<tbody>
<tr>
<td>VIII.1 INTRODUCTION</td>
</tr>
<tr>
<td>VIII.2 EXPERIMENTAL</td>
</tr>
<tr>
<td>VIII.2.1 Chemicals</td>
</tr>
<tr>
<td>VIII.2.2 Instrumentation</td>
</tr>
<tr>
<td>VIII.2.3 Preparation of the IMER</td>
</tr>
<tr>
<td>VIII.2.4 Comprehensive IPC-IMER-IPC setup</td>
</tr>
<tr>
<td>VIII.2.4.1 First dimension, IPC-IMER</td>
</tr>
<tr>
<td>VIII.2.4.2 Second dimension, IPC</td>
</tr>
<tr>
<td>VIII.3 RESULTS AND DISCUSSION</td>
</tr>
<tr>
<td>VIII.4 CONCLUSION</td>
</tr>
<tr>
<td>VIII.5 REFERENCES</td>
</tr>
<tr>
<td>IX GENERAL CONCLUSION</td>
</tr>
<tr>
<td>SUMMARY</td>
</tr>
<tr>
<td>SAMENVATTING</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
</tr>
<tr>
<td>CURRICULUM VITAE</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
</tr>
</tbody>
</table>
Chapter I

General introduction and scope

1.1 Introduction

In last few decades, a gradual increase in the usage of biomolecules has occurred in the therapeutic field. In this sector which has for most of the 20th century almost exclusively been dominated by small molecules (MW < 1000 Da), slowly the larger biomolecules (MW > 1000 Da) have permeated through and are now increasingly allowing pharmaceutical applications [1-3]. Contrary to the needs of pharmaceutical drug development, most of the earlier biochemical and biotechnological research has been aimed at the identification of large biological structures while less emphasis was set on the development of highly accurate quantitative methods [4]. Moreover, in the cases where quantitative analyses were performed, the tolerated levels of accuracy have typically been higher than the ones imposed on the analysis of pharmaceutically active small molecules. Since the incursion of biomolecules in the therapeutic field, those levels of accuracy needed to be narrowed and concomitantly the tools for characterization of the numerous impurities accompanying the active biomolecules needed to be adapted [5]. Additionally also novel analytical technologies have increasingly been required to allow accurate quantitative analysis of samples depicting complexities exceeding the separation capabilities of common separation techniques as can be the case in antisense therapeutics.
Since the elucidation of the human genome sequence, a rise in the antisense technology has been observed not only in the biochemical research field, but also in the commercialization of oligonucleotides with therapeutic activity. Under the antisense technology approach, a single stranded DNA or RNA fragment is used to inhibit the formation of proteins. Those DNA or RNA based single stranded short fragments, better known as oligonucleotides (ONs), have been increasingly implemented as therapeutic agents among other biochemical applications [2,6].

The increased sample complexity observed in oligonucleotide mixtures is a consequence of their synthetic nature and predisposition to generate larger amounts of degradation products. Those impurities are closely related and include mainly sequence nucleotide adducts (X+1, X+2, X+n…) or deletions (X-1, X-2, X-n…) which are challenging to separate from the active ingredients.

The current analytical techniques employed for the analysis of oligonucleotides such as ion-pair chromatography, hydrophilic liquid chromatography, ion exchange chromatography and capillary electrophoresis, provided somewhat limited separation capabilities when dealing with complex mixtures of these molecules. A number of approaches available to increase the resolving power especially in liquid chromatography have been developed. Those technologies relied mainly on the column technology, involving the utilization of porous stationary phases combined with elevated temperatures or sub-2 micron based particles requiring ultra-high pressure (U)HPLC approaches [7-9]. These approaches allowed somewhat faster analysis but unfortunately did not allow to overcome much of the sample complexity of mixtures of oligonucleotides. Additionally, for high throughput analyses, these techniques still proved to be limited in terms of analysis time and cost of instrumentation.

An improved and more straightforward approach for increasing the resolving power is to perform multidimensional separations. Under this approach, the sample is subjected to two different separation mechanisms [10]. With the combination of liquid chromatography and capillary electrophoresis, a highly orthogonal approach can be established, as properties of a pressure driven chromatographic technique are then combined with the orthogonal principles of electrodriven separations [11]. The much differing selectivity of each methodology significantly reduces the potential of co-elution when dealing with mixtures of molecules with high degree of similarity such as mixtures of oligonucleotides.

Furthermore, in the development of a successful therapeutic agent, the in vivo stability of ONs is as important as their actual biological activity [2]. A variety of chemical modifications have been
introduced in the last decades to enhance the stability of oligonucleotides whereby a suitable combination is strived for, achieving both maximal biological activity and in vivo stability [2,12,13]. The number of combinations, and as a consequence also the number of impurities, can drastically increase as the length of the nucleotide sequence does so. The current tests for testing the stability of ONs comprise their batch incubations with enzymes or tissue homogenates followed by their instrumental analysis. Relative stability assessments for various types of ONs can be performed more precisely by the usage of immobilized enzyme reactors (IMERs) coupled in-line with a separation technique while also allowing for higher reproducibility, reduced costs, longer setup life times and higher throughput if developed and implemented in a robust way [14]. The combination of IMERs with liquid chromatography offers the possibility of analyzing the cleaved fragments, which provides an additional valuable set of information. Additionally, by combining the power of multidimensional separation techniques with the specificity and gain in structural information that IMERs might offer, powerful platforms for the analysis and development of ONs can be created.

1.2 Scope

The scope of this thesis comprises the development of improved analytical techniques for the analysis of the composition and the stability of natural and synthetic mixtures of oligonucleotides (ONs). The developed platforms allow achieving high separation capabilities for complex mixtures of oligonucleotides while also allowing selective characterization of natural and modified oligonucleotides.

In chapter II an introduction of the principles of the separation techniques used in the developed platforms is provided. In chapter III the basic aspects of oligonucleotides and the tools for their characterization, relevant in the framework of this thesis are discussed.

The first research component of this work focuses on increasing the separation capabilities for complex mixtures of ONs. Therefore, chapter IV describes the development and evaluation of off-line comprehensive multidimensional platforms, combining liquid chromatography with capillary electrophoresis. In chapter V, the possibilities of an in-house developed in silico tool for the deconvolution of ON signals when using photodiode array detection in combination with separation
techniques, were explored as an additional separation feature to assist in the various available platforms for ON analysis.

In the second research focus of this work, the hyphenation of liquid chromatography with a DNase I immobilized enzyme reactor (IMER) was explored as an alternative high throughput tool for oligonucleotide stability testing. Chapter VI briefly provides the reader with an overview of the most relevant aspects in the manufacturing and characterization of IMERs for subsequent hyphenation with LC. In chapter VII the development of the corresponding online IMER-LC platform is described, while illustrating the discriminative power of the platform for the stability assessment of various types of single and double stranded oligonucleotides. In chapter VIII, the platform is expanded whereby the reactor stability and activity is demonstrated when combining in a post-column way with ion pair chromatography. This allows improved stability determination of ONs as the degradation products can subsequently be reanalyzed in a subsequent chromatographic separation.

This thesis introduces a number of novel approaches for ON analysis, establishing a basis for the future development of more complex analytical platforms possibly involving multidimensional separations with a broader sets of IMERs containing multiple enzymes, not necessarily limited to the study of only therapeutic oligonucleotides.

I.3 References


Chapter II

Principles of the separation techniques used in oligonucleotide analysis

Summary

In this chapter the principles of the separation techniques typically used for the analysis of oligonucleotides will be described. The main chromatographic approaches are discussed first followed by an introduction into the electrodriven solutions which are also broadly used for ON analysis. The discussion covers both the underlying principles and a number of practical aspects of the techniques which are relevant to this work. At the end of the chapter the concept and potential of multidimensional methodologies for ON analysis are discussed.
High performance liquid chromatography (HPLC) is the most widely used separation technique today [1,2]. The broad variety of stationary phases which can be combined with diverse combinations of mobile phases and modifiers can provide suitable conditions for the separation of hydrophilic, hydrophobic, ionic or neutral molecules over several orders of magnitude in the molecular weight. Furthermore, the high robustness, reproducibility and the generally satisfactory figures of merit of the technique, all contribute to its broad implementation in routine and research laboratories, for both analytical and industrial scale separations [1,3].

Basically, an HPLC system consists of a solvent reservoir, a high pressure pump, an injection valve, a chromatographic column and a detection system (Figure II-1). The mobile phase is constantly percolated through the column, allowing the components of an injected mixture to interact with the stationary phase while they are eluted from the column for a further detection. The principle of separation in HPLC relies on the relative affinity of the analytes toward a stationary phase and the mobile phase. After the separation, the solutes are passed through a detector for monitoring the analysis [3].

The choice of the chromatographic column determines the mechanism of separation together with the choice of the mobile phase. The most commonly used operation modes in HPLC are summarized in Table II-1.

**Figure II-1.** Schematic representation of the typical components of an HPLC system.
Table II-1. Schematic overview of the main separation in High performance liquid chromatography (HPLC). The modes of relevance to this work are highlighted.

<table>
<thead>
<tr>
<th>Chromatographic mode</th>
<th>Description</th>
<th>Chromatographic mode</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal Phase chromatography (NPC)</td>
<td>This is the oldest LC mode. The column is of a polar nature while the mobile phase uses less-polar organic solvents. Used mainly for water insoluble samples.</td>
<td>Ion exchange chromatography (IEC)</td>
<td>The stationary phase contains ionizable groups that interact with ions of the opposite charge; the mobile phase generally consists of a buffered water salt solution. Is particularly employed for the analysis of ionizable samples. ONs present an anionic character, therefore can be separated by this technique.</td>
</tr>
<tr>
<td>Reversed phase chromatography (RPC)</td>
<td>The column is non polar while the mobile phase consists of a mixture of water and a miscible organic solvent. It is particularly suitable for water soluble samples.</td>
<td>Ion pair chromatography (IPC)</td>
<td>Similar conditions to RPC are employed with the slight difference of the addition of an ion paring reagent to the mobile phase which interacts with ions of opposite charge. Is applied for ionizable compound that are weakly retained in RPC. ONs present an anionic character, therefore can be separated by this technique.</td>
</tr>
<tr>
<td>Non-aqueous reversed phase chromatography (NARP)</td>
<td>The column is nonpolar and the mobile phase consists of a mixture of organic solvents. It is employed for the analysis of water insoluble and very hydrophobic samples.</td>
<td>Hydrophilic interaction liquid chromatography (HILIC)</td>
<td>The stationary phase is polar while the mobile phase is a mixture of water and an organic mobile phase. Is mainly used for samples that are very polar and poorly retained in RPC. ONs present an anionic character, therefore can be separated by this technique.</td>
</tr>
<tr>
<td>Size exclusion chromatography (SEC)</td>
<td>The separation is based on the molecular weight. An inert column used and the mobile phase can be either of polar and nonpolar nature. It is mainly used for large biomolecules and polymers.</td>
<td>Other modes</td>
<td>Modified stationary phases are employed to increase the selectivity towards specific groups of molecules; <em>e.g.</em> chiral chromatography, immunoaffinity chromatography, etc.</td>
</tr>
</tbody>
</table>

As emphasis in the present work is set on the analysis of polar water soluble molecules such as oligonucleotides, the detailed description of the chromatographic modes is limited to IEC, IPC and HILIC. Examples of the use of these separation modes for the separation of oligonucleotides are described in chapter III.
Principles of the separation techniques used in oligonucleotide analysis

II.1.1 Ion exchange chromatography

Ion exchange chromatography (IEC) is a very effective separation mode which is inherently limited to the separation of charged solutes. The main application ranges of this technique include the analysis of carbohydrates, carboxylic acids, inorganic anions and biomolecules including proteins, peptides, amino acids and oligonucleotides [4-7]. The stationary phases employed in IEC must contain ionized or ionizable groups. The charge that those groups exhibit determines their role as cation or as anion exchanger. In cation exchange chromatography, the stationary phase contains negatively charged groups (e.g., $\text{SO}_3^-$) while in anion exchange chromatography the stationary phase contains positively ionized groups (e.g., $\text{N(CH}_3)_3^+$). Moreover, the stationary phases are further classified as strong or weak ion exchangers depending on the capacity of their functional groups to remain ionized in a broad (2 ≤ pH ≤ 13) or narrow mobile phase pH range, respectively. The retention mechanism is based on the competition between the counter ions present in the mobile phase and the sample ions for interaction with ions of opposite charge on the stationary phase. For instance, the anion exchange mechanism can be represented by the retention of an anionic solute $A^-$ with $Y^-$ as the counter ion present in solution:

$$A^- + S^+Y^- \rightleftharpoons Y^- + S^+A^-$$

Where $S^+$ represents a cationic group present on the stationary phase, which can either interact with the solute $A^-$ or with the mobile phase counter ion $Y^-$ through coulombic attractions.

As the common pH stability range of silica based columns is limited to pH 8 [3], most of the commercial IEC columns are based on a polymeric support. Typically, the mobile phases in IEC consist of water, a buffer and a counter ion to adjust the sample retention. Due to the nature of certain polymeric supports, the addition of organic solvents such as methanol or acetonitrile is also carried out in order to control the selectivity and the solvent strength. Primarily, the concentration of the counter ion is adjusted to control the retention and solvent strength. Moreover, in the presence of weak ion exchangers, or if the analyte ionizability is dependent on the pH of the mobile phase, the pH control is critical for optimizing the separation and assuring the reproducibility of an IEC methodology. As a consequence, parameters such as the buffer type and concentration, the temperature and the percentage of organic solvent, need to be strictly evaluated during the development of IEC separation methodologies [8].
II.1.2 Ion pair chromatography

Ion pair chromatography (IPC) is a variant of reversed phase chromatography (RPC). The difference can be found in the addition of an ion pair forming reagent to the mobile phase. This mobile phase additive interacts with ionized analytes of opposite charge to form a hydrophobic ion-pair that can be retained on an apolar stationary phase as in a reversed phase chromatography (e.g. C18). As in the latter HPLC mode, the separating mobile phase consists of a mixture of water and a miscible organic solvent such as acetonitrile. This separation mode is mainly employed for the separation of ionic water soluble compounds that are poorly or not retained in RPC [3,9]. Assuming a negatively charged analyte $\text{A}^-$ and an ion pair reagent $\text{R}^+$ (e.g. tetraalkylammonium ion), two models are used to explain solute retention and elution order. In one model it is assumed that the ion-pair is formed in solution, resulting in an ion-pair [$\text{A}^-\text{R}^+$] that is retained by the stationary phase:

$$\text{Eq. II-2} \hspace{1cm} [\text{A}^+\text{R}^+]_{\text{mobile phase}} \rightleftharpoons [\text{A}^+\text{R}^+]_{\text{stationary phase}}$$

Under this model, the retention is dependent on three factors: (1) the concentration of the ion-pair reagent and its affinity to form an ion-pair with the analyte; (2) the proportion of the analyte that ionizes in solution, whereby an adequate selection of the mobile phase pH value is necessary in accordance to the pKa value of the analytes; (3) the retention of the ion pair [$\text{A}^+\text{R}^+$], in other words, the affinity that the ion-pair depicts towards the stationary phase. Hence, the use of a more hydrophobic IPC reagent leads to a greater retention of the ion-pair.

The second model assumes that the IPC reagent attaches first to the stationary phase, and consequently, the separation is performed by an ion exchange mechanism:

$$\text{Eq. II-3} \hspace{1cm} \text{A}^-_{\text{mobile phase}} + \text{R}^+\text{X}^-_{\text{stationary phase}} \rightleftharpoons \text{A}^-\text{R}^+_{\text{stationary phase}} + \text{X}^-_{\text{mobile phase}}$$

In practice it can be observed that both of those mechanisms are present during the separation, and depending on conditions either of them can predominate.

Choosing the right pH of the mobile phase is a critical parameter in IPC as its effect can alter the ionization of both, the IPC reagent and the analyte. An adequate value should be chosen to assure the ionization of the IPC reagent and the analytes in order to allow the interaction between them.
Principles of the separation techniques used in oligonucleotide analysis

On the other hand, the selection of the IPC reagent and its concentration also need to be adjusted. When increasing the concentration of the IPC reagent in the mobile phase, the stationary phase reaches a point where it becomes saturated as can be observed in the Langmuir type adsorption isotherm depicted in Figure II-2A, where \([R^+]_m\) and \([R^+]_s\) represent the concentration the IPC reagent in the mobile and stationary phases, respectively. It can be visualized that the differences in the hydrophobicity (TBA\(^+\) > TEA\(^+\)) of the IPC reagent have also a significant effect over the concentration at which the stationary phase becomes saturated. In practice, it is preferable to maintain a concentration of the IPC reagent below the saturation levels of the stationary phase [3]. If the latter is exceeded competition effects are observed between the IPC reagent and the solutes (Figure II-2B).

![Figure II-2. A) Langmuir type adsorption isotherm of two IPC reagents \(R^+\) (TBA: tributylamine. TEA: triethylamine). \([R^+]_m\) and \([R^+]_s\) represent the concentration of the IPC reagent in the mobile and stationary phases, respectively; B) Effect of the concentration of the IPC reagent on the retention factor of the solute; the retention factor of the analytes tends to decrease at too elevated concentrations of the IPC reagent as a consequence of the competition with the counter ions.]

II.1.3 Hydrophilic interaction liquid chromatography

Hydrophilic interaction liquid chromatography (HILIC) is considered to be a variant of normal phase chromatography, in which an aqueous-organic mobile phase is used with polar stationary phases [10]. More polar and/or ionized molecules tend to be more strongly retained in HILIC, meaning that many compounds that show a poor retention in RPC can be therefore retained by HILIC [3,11], making both HPLC modes highly complementary [12]. Next to the simplest, native silica based, HILIC columns other stationary phases including diol-bonded, amide-bonded, polypeptide-bonded, cation- and anion-exchange and zwitterionic phases are also widely employed [13,14]. Water serves as a strong solvent in the mobile phase commonly
combined with acetonitrile which is used as weak solvent in this separation technique. Moreover, the usage of other water soluble solvents such as methanol or ethanol for *e.g.* is also possible; however, their use generally reduces the overall retention of analytes [13]. The addition of formic acid, ammonium acetate or ammonium formate to the mobile has been widely implemented for controlling the separation of ionizable compounds. Furthermore, as the latter buffers are volatile, they allow for an easy coupling to mass spectrometry [13,15].

The retention mechanism is based on the partitioning of the compounds between a polar water layer that is enriched on the surface of the column packing and the less polar mobile phase (Figure II-3). The partitioning process is partially corroborated through the observation that HILIC separations require at least 2-3% of water in the mobile phase, while with water percentages exceeding 40% the HILIC mechanism is practically lost [3,13,16]. Although partitioning is the predominant mechanism a certain degree of absorption and even ion exchange processes are also occurring with many of the HILIC compatible analytes [17].

![Figure II-3. Water layer formation in HILIC responsible for the partition mechanism of the analyte between two liquid phases.](image)

A brief overview of the basic chromatographic parameters allowing better characterization of the oligonucleotide analyses performed in this work is provided in the next section.

### II.2 Fundamental chromatographic parameters

A chromatogram is a representation of the chromatographic process (Figure II-4), consisting of a concentration profile of analytes elution from the column as a function of time. The retention $t_R$ is
Principles of the separation techniques used in oligonucleotide analysis

the time that it takes for a retained analyte to migrate through the column. For Gaussian peak shapes as the ones displayed in Figure II-4 this is the time corresponding to the elapsed time from the injection till the apex of the detected peak. The void time \((t_0)\) on the other hand, is the time that it takes for a non-retained analyte to elute from the column.

In HPLC two general separation settings are available. One, denominated as isocratic separation, is characterized by the fact that the mobile phase composition is held constant through the separation process. The other mode is a gradient separation, in which one or more parameters are continuously varied through the separation. In the latter approach, typically the mobile phase composition is varied from a low to high eluotropic strength.

![Simulated chromatogram illustrating the separation between two peaks with a resolution of 1.5.](image)

The void time \((t_0)\) retention times of peak A \((t_{RA})\) and peak B \((t_{RB})\), the height of peak A \((h)\) and its width at the base \((W_a)\) and width at half height \((W_h)\) are denoted. \(\sigma\): Standard deviation of the peak width.
II.2.1 Retention factor

As already described mentioned in section II.1 the principle of the chromatographic separation is the distribution of the analytes between the stationary and mobile phases. This is described by a distribution constant $K_D$:

$$K_D = \frac{C_s}{C_M} = \frac{m_s}{m_M} \frac{V_M}{V_s} = k' \beta$$

where $C_s$ and $C_M$ represent the concentration of the analyte in the stationary and mobile phase, respectively, $m_s$ and $m_M$ are respectively the quantity of the analyte in the mobile and stationary phase, and $V_M$ and $V_s$ represent the volume of the volume of the stationary phase and the volume of the mobile phase, both in the column, respectively. The ratio between the volumes is also denominated the phase ratio ($\beta$) (Eq. II-1), while the ratio between the quantities of the analyte in the stationary and mobile phase is denominated as the retention factor ($k'$) (Eq. II-6):

$$\beta = \frac{V_M}{V_s}$$

$$k' = \frac{m_s}{m_M}$$

Since the number of molecules in each phase is proportional to the time spend by the solutes in each phase, the retention factor (also referred to as the capacity factor in older literature) can be expressed as:

$$k' = \frac{t_R - t_0}{t_0}$$

$k'$ (Eq. II-7) can easily be calculated from the chromatogram and is independent of the column dimensions.
Principles of the separation techniques used in oligonucleotide analysis

II.2.2 Selectivity factor
For the successful separation of two components in a mixture, the solutes must have different retention times as can be observed in Figure II-4. The parameter that describes this difference is defined as the selectivity factor ($\alpha$). Both the selectivity and the separation factor are related through Eq. II-8.

$$\alpha = \frac{k'_j}{k'_i}, \quad k'_j > k'_i$$

II.2.3 Efficiency
During a chromatographic process, the solute band is subjected to broadening due to the distribution of the solutes along the longitudinal axis of the column during the elution process. This is manifested as a peak width for every chromatographic signal (Figure II-4). The parameter used to quantify this process is denominated as the efficiency ($N$):

$$N = \left(\frac{t_R}{\sigma}\right)^2 = 16 \left(\frac{t_R}{W_b}\right)^2 = 5.54 \left(\frac{t_R}{W_h}\right)^2$$

$N$ also known as the column plate number, depends on the column dimensions, mobile phase properties, temperature and linear velocity of the mobile phase. In order to theoretically evaluate those conditions, the efficiency can be expressed by the plate height ($H$):

$$H = \left(\frac{L}{N}\right)$$

where $L$ is the length of the column. $H$ measures the band broadening that takes place during the elution process and it is dependent on the linear velocity of the mobile phase.

II.2.4 Resolution
The degree of chromatographic separation between two compounds can graphically be described in terms of resolution ($R_s$):

\[ \text{Resol} \]
\[
R_s = 2 \frac{t_{R_j} - t_{R_i}}{W_{bi} + W_{bj}}
\]

Where \( t_r \) and \( W_b \) represent the retention time and the peak width for the compounds \( i \) and \( j \), respectively. An improved separation is consequence of greater differences in retention times and/or of narrower peak shapes. The resolution is thus a measure of the overlap of two adjacent peaks. A base line separation of two adjacent peaks is considered when a resolution of at least 1.5 is obtained (Figure II-4) [3]. From equations Eq. II-7, Eq. II-8, Eq. II-9 and Eq. II-11, Eq. II-12 can be derived:

\[
R_s = \left[ \frac{\sqrt{N}}{4} \right] \left[ \frac{\alpha - 1}{\alpha} \right] \left[ \frac{k'}{k'+1} \right]
\]

It can be observed from Eq. II-12 that the resolution can be controlled in three different ways, and among them, the improvement of the selectivity has the largest influence. Nevertheless, satisfactory chromatography is only obtained if also the other two parameters are optimized. This implies ensuring sufficient retention on the columns (\( k > 3 \)) and maximization of the plate number by using efficient columns operated at the correct linear velocity as explained in the next section.

### 2.5 Van Deemter equation

The measured column efficiency and therefore the plate height are not constant and depict a dependency with the flow rate of the mobile phase and with the characteristics of the stationary phase and of the sample (Eq. II-10). Various models have been designed to explain this relationship of which the most broadly used one, due to its conceptual simplicity, is the Van Deemter model (Eq. II-13) [3,18,19]:

\[
H = A + \frac{B}{u} + C \cdot u
\]

Where \( u \) represents the linear velocity of the mobile phase. The latter is related to the velocity of a solute that is not retained by the column and therefore is measured via void time and the column length as described by:
Principles of the separation techniques used in oligonucleotide analysis

Eq. II-14 \[ u = \frac{L}{t_0} \]

The coefficients A, B and C from Eq. II-13 are constant for a particular molecule, column and set of experimental conditions, and represent the eddy diffusion, longitudinal mass transfer and the mobile phase plus the stationary phase mass transfer, respectively.

Eq. II-15 \[ A = 2\lambda d_p \]

Eq. II-16 \[ B = 2\gamma D_m \]

Eq. II-17 \[ C = C_m + C_s = C' m \frac{k''^2}{(k''+1)^2} \frac{d_p^2}{D_m} + C' s \frac{k''^2}{(k''+1)^2} \frac{d_s^2}{D_s} \]

Where \( d_p \) is the particle diameter, \( \lambda \) is a term dependent on the column packing, \( D_m \) is the molecular diffusion coefficient of the solute in the mobile phase, \( D_s \) corresponds to related diffusion coefficient in the stationary phase, \( d_s \) represents the migration distance within the stationary phase, \( \gamma \) is the obstruction factor of the packed bed. \( C_m, C_s, C'_m, \text{ and } C'_s \) are constants, whereby the latter two are dependent on the geometry and architecture of the column. \( k'' \) represents the zone retention factor which is closely related to the phase retention factor (\( k' \)) outlined in Eq. II-7 [19]. The corresponding plot of the plate height as a function of the linear velocity is called the Van Deemter curve Figure II-5, which depicts the minimum value of the plate height (\( H_{\text{min}} \)) at the optimal linear velocity (\( \mu_{\text{optimal}} \)). Chromatography should ideally always be performed at the minimum of this plot as it ensures obtaining the most efficient analyses. One of the most practical aspects of the Van Deemter equation is that each parameter on the equation can be linked to a phenomenon occurring in the packed column, therefore establishing a real physical meaning. As mentioned other models have been developed such as the Knox and the Giddings approach [20-22].
Figure II-5. Representative plot of the Van Deemter equation (Eq. II-13). The total plate height ($H$) is showed as a solid curve while the dashed and dotted curves represent the three contributions to the Van Deemter equation.

II.2.6 Peak capacity

A more practical way to evaluate the separation power under various combinations of experimental conditions can be performed by the determination of the peak capacity ($P_c$). The latter is a measure of the number of peaks that can fit under an elution time window ($t_1-t_2$) with a fixed resolution ($R_s = 1$) [3,23-25]. The common assumption establishes a peak spacing of four standard deviations ($4\sigma$) near the base line resolution ($\tau$), leading to the following equation:

Eq. II-18

$$P_c = 1 + \int_{t_1}^{t_2} \frac{1}{4\tau} dt$$

In the case of an isocratic elution, where $N$ is constant over the chromatogram, the peak standard deviation changes with the retention time as follows:

Eq. II-19

$$\tau = \frac{t_R}{\sqrt{N}}$$

After substituting Eq. II-19 in Eq. II-18 and integrating from the retention time of the unretained peak ($t_0$) to the last peak ($t_R$), the calculation of the peak capacity can be calculated as:
Principles of the separation techniques used in oligonucleotide analysis

Eq. II-20

\[ P_c = 1 + \frac{\sqrt{N}}{4} \ln(k' + 1) \]

The peak capacity for gradient separations is usually higher than for isocratic separations [3,26]. In the case of a gradient separation, and where every peak shows approximately the same peak width (generally occurring in most IPC, IEC and HILIC separations), the peak capacity expression can be approximated to the gradient run time \((t_G)\) as follows:

Eq. II-21

\[ P_c = 1 + \frac{t_G}{\bar{W}}; \approx \frac{t_G}{\bar{W}} \]

where \(\bar{W}\) is the average peak width at the base in the chromatographic window. However, this estimation is not always realistic as it overestimates the separation power of an actual gradient chromatogram. For instance, in the chromatogram in Figure II-6A where the average peak width is \(\bar{W} = 0.2\) min and the gradient time 10 minutes, the peak capacity according to Eq. II-21 will approximate 50 as illustrated in Figure II-6B. As can be observed in Figure II-6A, the peaks only appear between a defined time window rather than distributed over all the chromatogram, so only a fraction of the gradient is actually used. When peaks are confined within a part of the chromatogram, the effective peak capacity of the separation is less than the peak capacity value showed in Figure II-6B for the full gradient.

The separation performance can be better defined by the number of resolved peaks that can be fit between the first and the last peaks in the chromatogram and not necessarily at the beginning or at the end of the gradient:

Eq. II-22

\[ P_c = \frac{t_l - t_f}{\bar{W}} \]

where \(t_l\) and \(t_f\) are the retention times for the first and the last eluting peaks, respectively. The validity of Eq. II-21 and Eq. II-22 is limited to the case where the peak width pattern is very similar as it occurs in most IPC, IEC and HILIC gradient separations. More elaborate procedures for the calculation of the peak capacity when the peak width changes with the operating conditions can be found in literature [3,26-29].
Moreover, there are many possibilities for the determination of the average peak width $\overline{W}$. This latter is usually derived theoretically from the width at half-height ($\overline{W}_h$), from which the width corresponding to $4\sigma$ is calculated (Figure II-4) [3,23,24]. This approximation can also lead to misleading conclusions since the potential tailing and fronting effect of the peak is neglected. When analyzing complex mixtures, the determination of the peak width becomes nearly an impossible task as well resolved peaks are difficult to encounter in crowded chromatograms. Hence, it is recommend to estimate $\overline{W}$ from the chromatograms with a simple standard mixture, consisting of representative compounds with identical functionalities as the sample components, and with an elution profile that covers the entire separation space [3,26,29]. This approach was followed in this study. The peak capacity in multidimensional separations will be further discussed in section II.5.1.3.

**Figure II-6.** Peak capacity in gradient elution. A) Hypothetical chromatogram; B) representation of the peak capacity for the separation in A) according to equation (Eq. II-21); the $P_c$ approximates 50 in this case.
### II.3 Ultrahigh performance liquid chromatography

Although various strategies can be designed to increase the column efficiency such as increase in columns length and operation at more elevated temperatures, in the last decade the parameter which received the most attention has been the reduction in particle size. This approach, which allows obtaining more plates in less analysis time, is increasingly implemented as it allows faster analysis. However, the reduction of the particles size diameter is accompanied by an increased pressure drop in the pumping system as described by Eq. II-23:

\[
\Delta P = \frac{\eta F L}{K V_0 \pi r^2 d_p^2}
\]

where \( \eta \) and \( F \) are the viscosity and flow rate of the mobile phase, respectively. \( r \) is the column radius and \( K V_0 \) is the permeability which is characteristic for each column. As a consequence chromatographic systems have experienced a major modification in the last decade regarding the operational pressure limits. Conventional HPLC systems have a typical pressure limit of 400 bar, which is appropriate for most column packings (particles sizes \( \geq 3.5 \ \mu m \)) and dimensions (50-250 mm in length, 4.6 mm I.D) [1,3].

Ultrahigh performance liquid chromatography (UHPLC commercialized under the name UPLC™ by Waters) exploits the improved separation performances achievable when using particles in the 2 micrometer size range packed in 2.1 or 1 mm columns [30,31]. Operation of columns of 5 to 15 cm in length packed with such particles sizes is only possible when HPLC systems are available allowing operation at 1000 bar or more, due to relationship between the pressure and the square of the particle sizes as represented in (the Darcy equation) Eq. II-23. Although the column diameter is in theory independent of the achievable efficiencies, in practice this is not the case as at high pressures frictional heating issues leading to temperature gradients in the columns detrimentally affect the peak shapes. To allow improved heat dissipation therefore UHPLC is typically performed in narrower columns compared to HPLC. Additional benefits of the use of such narrow bore columns are reduced peak volumes and reduced solvent consumption. Extra column peak broadening phenomena are
Chapter II

However, more problematic in UHPLC due to the usage of inadequate connector and fitting dimensions [1-3].

Additionally, the reduction of the particle size permits the implementation of higher flow rates (up to system pressure limits) to speed up analyses of small molecules without changing significantly the efficiency (H<sub>min</sub>) as observed in the Van Deemter curves (full lines) in Figure II-7. The particle size reduction proves also to reduce the impact of the slow diffusion of large molecules such as oligonucleotides, as can be observed on the simulated Van Deemter curves (dashed lines) in Figure II-7. In general the mass transfer effect in the stationary phase has a mayor impact over the ON separation, and in contrast to small molecules, the separation of ONs often benefits from the usage of slower mobile phases flow rates [32].

![Figure II-7. Van Deemter curves calculated for a 50 x 4.6 mm column [32]. The parameters A, B and C were set to 1.5, 1 and 0.167, respectively. The change on the plate height as function of the mobile phase flow rate was calculated for two diffusion coefficients. (A) D<sub>m</sub> = 1.25·10^-9 m<sup>2</sup>/s; full lines and (B) D<sub>m</sub> = 1.10·10^-10 m<sup>2</sup>/s; dashed lines (representing a hypothetical oligonucleotide). The particle size of the column packing is indicated on the graph.](image)

No further emphasis will be set on UHPLC systems as the present work was performed on relatively conventional HPLC systems employing columns with particle sizes ≥ 3µm. Nevertheless, the HPLC
Principles of the separation techniques used in oligonucleotide analysis
techniques employed in this work can be implemented in UHPLC systems following straightforward adaptations to experimental parameters if required [33-35].

II.4 Principles of electrodriven separation techniques used for oligonucleotide analysis

Electrophoresis is based on the principle that under the influence of an applied electrical field, different charged species in solution will migrate at characteristic velocities depending on the ratio of their charge to the frictional drag they experience when migrating through the medium. In the simplest electrophoretic setup, a narrow plug of sample is placed in the system that contains a background electrolyte (running buffer). In such system positively charged ions migrate towards the negative electrode (cathode), while negatively charged ions migrate towards the positive electrode (anode). Moreover, as a consequence of the frictional energy released by the moving ions, the background electrolyte warms up. This phenomenon, so called Joule heating, leads to convection in the medium and is one of the most important zone broadening process in electrophoresis. Traditionally, electrophoresis has been performed on a support medium such as paper, cellulose acetate or in polymeric gel, which provided the mechanical stability for the fluidic buffer system. The latter approach is called slab gel electrophoresis and has been employed for a long time for the analysis of biomolecules such as proteins whereby the rigid gel matrix is based on a mixture of polyacrylamide/bis-acrylamide polymers. This approach also called polyacrylamide gel electrophoresis (PAGE) can also be used for the analysis of nucleic acid chains (section II.4.4.2).

In the early 80'ties Capillary electrophoresis (CE) has emerged as an alternative form of electrophoresis, where the capillary wall provides the mechanical stability for the carrier electrolyte. CE, in its current state, has been typically conducted in capillaries with inner diameters and lengths between 20-100 µm and 20-100 cm, respectively. The usage of this narrow bore tubes provides efficient removal of Joule heating by allowing this heat to be quickly dissipated to the surrounding environment. This heat removal helps to decrease the band broadening leading to more efficient and faster separations as higher electrical fields can be applied when compared to electrophoresis performed slab gels as PAGE [9,36,37]. The instrumentation needed to perform capillary
electrophoresis consists of a narrow bore capillary traditionally composed of fused silica. The ends of the capillary are placed into buffer reservoirs that contain the electrodes to make electrical contact between the high voltage supply and the capillary (Figure II-8).

CE is a widely used technique for many biological compounds that have charged or ionizable groups. Electrophoresis has been initially employed in the analysis of macromolecules such as proteins and DNA. This technique has also proven to be useful for the separations of small molecules such as amino acids, chiral drugs, vitamins, pesticides, inorganic ions, organic acids, dyes, surfactants, peptides and proteins, carbohydrates, oligonucleotides and DNA restriction fragments, and even whole cells and virus particles [36-39].

The mechanisms responsible for separation in CE are different from those in liquid chromatography, and thus can offer orthogonal analyses with complementary selectivity to LC. Moreover, sample stacking techniques have also been developed to enhance the performance of this methodology [36,40]. These aspects are used for the improved oligonucleotide analyses as described further on.

Before introducing CE for ON analysis in more detail, some principles of CE will be discussed in the following sections.

![Figure II-8. Schematic of a single capillary electrophoresis system.](image-url)
Principles of the separation techniques used in oligonucleotide analysis

II.4.1 Basic principles of capillary electrophoresis

II.4.1.1 Electrophoresis

As mentioned, the electrophoretic separation is based on differences on solute velocity under the influence on an electrical field when the solute is migrating through a media. The linear velocity of an analyte ion is given by:

Eq. II-24

\[ v_e = \mu_e E \]

Where \( v_e \) is the velocity of the ion in (cm/s), \( \mu_e \) the electrophoretic mobility (cm\(^2\)/V·s) and \( E \) is the electric field strength (V/cm). The field strength is a function of the applied voltage (V) and the total length of the capillary (L). The mobility for a given ion and medium is a constant which is characteristic of the ion. When an ion with charge (q) is subjected to an electrical field, it is attracted to the electrode charge with an electric force (\( F_E \)):

Eq. II-25

\[ F_E = qE \]

At the same time, this ion is hindered in its movement by the medium by the frictional force (\( F_F \)) [39]:

Eq. II-26

\[ F_F = -6\pi\eta rv_e \]

Where \( \eta \) is the viscosity of the solution and \( r \) is the radius of the ion. During electrophoresis, a steady state is attained when \( F_E \) equals \( F_F \) and the ion reaches a constant velocity. Combining Eq. II-24, Eq. II-25 and Eq. II-26 at this steady state, results in the following equation for the electrophoretic mobility:

Eq. II-27

\[ \mu_e = \frac{q}{6\pi\eta r} \]
Eq. II-27 explains the faster mobility of small multiply charged solutes compared to the larger solutes which migrate more slowly.

**II.4.1.2 Electroosmosis**

Next to the electrophoretic mobility of solutes in CE, another mobility phenomenon specific to the capillary format is electroosmosis [39]. Electroosmosis is the relative movement of a liquid due to the effect of the electric field on counterions adjacent to a fixed charged surface. The bulk flow generated by this process is called the electroosmotic flow (EOF). This phenomenon occurs whenever an electrical field is applied to a solution that is in close contact with a charged surface such as the walls of a fused silica capillary [36]. The magnitude of the velocity of the EOF is related to the permittivity or dielectric constant of the medium ($\varepsilon$), the zeta potential ($\zeta$) and the viscosity of the medium:

\[
\mu_{\text{eof}} = \frac{\varepsilon \zeta}{\eta}
\]

\[
\nu_{\text{eof}} = \left(\frac{\varepsilon \zeta}{\eta}\right) E = \mu_{\text{eof}} E
\]

The EOF is basically dependent on the pH and the ionic strength of the media. Moreover, the EOF leads to more efficient separations in CE compared to LC separations (pressure driven), since it establishes an uniform driving force of the flow along the capillary (Figure II-9) [36].

![Figure II-9. Flow velocity profiles. A) electrodriven system; B) pressure-driven system.](image-url)
II.4.1.3 Apparent electrophoretic mobility and migration time

In CE, the time required for a solute to migrate to the point of detection is called the migration time. The latter is depended on the magnitude of both the electroosmotic flow and the electrophoretic mobility. This relation is defined as the apparent solute mobility ($\mu_a$):

Eq. II-30

$$\mu_a = \frac{l}{t_E} = \frac{ll}{tv} = \mu_e + \mu_{eo}$$

where $l$, $L$, and $t$ correspond to the effective length and the total length of the capillary and the migration time, respectively. The effective length is the distance from the point of the injection till the detection.

II.4.2 Zone dispersion in capillary electrophoresis

When compared to liquid chromatography, capillary electrophoresis presents significantly lower dispersion and therefore an increased resolving capability. The spreading of the solute zone results from the differences in the analytes velocities in that zone. The efficiency can be expressed similarly as in LC based on the number of theoretical plates (Eq. II-31).

Eq. II-31

$$N = \left(\frac{l}{\sigma}\right)^2 = 16 \left(\frac{l}{w_b}\right)^2 = 5.54 \left(\frac{l}{w_h}\right)^2$$

The solute-zone broadening in CE under ideal conditions can be considered to be a longitudinal diffusion process along the capillary. Under those conditions and via the association of Eq. II-30, the efficiency can be related to the molecular diffusion as follows:

Eq. II-32

$$\sigma^2 = 2D_m t = \frac{2D_m ll}{\mu_e V}$$

The combination of Eq. II-31 and Eq. II-32 leads to an efficiency prediction equation in CE:

Eq. II-33

$$N = \frac{\mu_e VL}{2D_m L} = \frac{\mu_e El}{2D_m}$$
II.4.3 Resolution in CE

The resolution in CE can be calculated from the electropherogram using the following equation:

\[ R_s = \frac{\sqrt{N}}{4} \left( \frac{\Delta \mu_a}{\bar{\mu}_a} \right) \]

where \( \Delta \mu_a \) is the difference in mobilities between the two peaks \( (\Delta \mu_a = \mu_{a2} - \mu_{a1}) \) and \( \bar{\mu}_a \) is the average of those mobilities. The substitution of Eq. II-33 in Eq. II-34 yields to an equation that does not require the explicit calculation of the efficiency and also involves the effect of the EOF over the resolution:

\[ R_s = \left( \frac{1}{4\sqrt{2}} \right) \left( \frac{\Delta \mu_a}{\bar{\mu}_a} \right) \sqrt{\frac{V}{D_m(\mu_a + \mu_{eo})}} \]

II.4.4 Modes of operation in capillary electrophoresis

Numerous modes of operation have been developed in electrophoresis and in the capillary variant thereof making it a highly versatile technique. Therefore, next to the fundamentally different separation mechanisms occurring in CE compared to chromatography, the diversities in achievable selectivities in CE is further broadened due to the many possibilities in applicable separation modes. Only capillary zone electrophoresis is mentioned and more specifically capillary gel electrophoresis is discussed here. This as the former is the most basic CE mode and as the latter is the established approach for oligonucleotide analysis. The reader is referred to specialized literature for the discussion of the other CE modes which are of no relevance to this work [39].

II.4.4.1 Capillary zone electrophoresis (CZE)

CZE is essentially electrophoresis performed in a capillary. In brief, after the injection, the analytes migrate in discrete zones at different velocities in an electric field according to their charge-to-mass ratio and therefore become separated. The simultaneous separation and detection of anionic and cationic solutes is thereby possible due to the generation of the EOF, which acts a bulk flow in the system allowing elution of the species through the detector. On the other hand, neutral species do not present an intrinsic mobility and all coelute with the EOF. CZE can be performed in the positive or negative polarity mode depending on the charge of solutes in attempt to analyze them in the most
Principles of the separation techniques used in oligonucleotide analysis

efficient way. Many variants have been developed in CZE to optimize their analysis [36,41,42]. CZE is not suitable for the analysis of oligonucleotides as the charge to mass ratio of such structures do not differ much between different ONs, leading to poor separation.

II.4.4.2 Capillary gel electrophoresis (CGE)

CGE is mainly used for the size-based separation of macromolecules such as proteins, polysaccharides and nucleic acids. The separation by size of the solutes is obtained by mobilizing them through a suitable polymer acting as a molecular sieve. This CE mode is particularly useful for the separation of molecules such as DNA and denatured proteins as those cannot be separated by CZE since they have as mentioned invariable mass-to-charge ratios.

There are several practical ways in which CGE separations can be performed depending on the used sieving agent. Cross-linked rigid gel matrixes such as composed of polyacrylamide /bis-acrylamide [39] have been widely employed in slab or tube gel electrophoresis as they provide good anti-convective properties [39]. However, the limited stability of those gels under high electric fields and reproducibility issues in the cross-linking process made the usage of cross-linked gel matrixes challenging to implement on a routine basis. However, with the advent of CE and the anti-convective nature of the capillary, the need for a rigid gel that is itself anti-convective is no longer needed [9,36,37].

Another way for performing CGE includes the use of a dilute low-viscosity polymer solution that can entangle with the analytes and alter their rate of migration. Although the application of CE in entangled polymer solution is quite recent, this technique combines the advantages of free solution capillary electrophoresis (system automation, speed, reproducibility, continuously renewable media and on-line detection) with the range of application and resolving power of gel based systems [39].

In the next section, a brief description of the basic concepts from polymer physics applied to describe the relevant aspects of entangled matrixes will be presented.

II.4.4.2.1 Entangled polymer solutions

Relevant differences exist between dilute polymer solutions in which the polymer chains are hydrodynamically isolated from one another, and more concentrated solutions, in which the chains interact and overlap. Three regimes of polymer solutions can be identified: dilute, semi-dilute and
concentrated. In dilute solutions, the polymer chains are hydrodynamically isolated from each other and their properties are that of a single chain.

When the concentration of the polymer solution is increased, the transition from the dilute to the semi-dilute region occurs; consequently, the polymer chains become entangled, forming a transient network of obstacles (Figure II-10). It should be pointed out that the interactions are purely topological, in contrast to gels, where the chains are irreversibly or reversibly cross-linked [43].

II.4.4.2.1.1 Entanglement threshold

The polymer volume fraction at which the polymer chains begin to interact with each other is named the entanglement threshold \((\Phi^*)\); while the volume fraction \((\Phi)\) is defined by:

\[
\text{Eq. II-36} \quad \Phi = C \rho_p
\]

Where \(C\) and \(\rho_p\) are the polymer concentration and density, respectively. The entanglement threshold is usually determined experimentally by measuring the viscosity of the polymer solution at different concentrations and determining the point at which a deviation from the linearity occurs when plotting those variables [39,43]. A more useful approach to describe the entanglement threshold can be given by:

\[
\text{Eq. II-37} \quad \Phi^* \approx \frac{3M_W}{4\pi N_A R_g^3}
\]

where \(M_W\) and \(R_g\) are the molecular weight of the polymer and the radius of gyration, respectively; and \(N_A\) is Avogadro’s number. Eq. II-37 models the polymer chains as coils, and can be easily inferred that for sufficiently large polymers, even seemingly dilute polymer solutions can be in an entangled state. Several other models have been described in literature to approach this concept [39,43,44]; however, they will not be discussed within the framework of this thesis.

II.4.4.2.1.2 Mesh size

An entangled solution is also characterized by an average mesh size \((\xi_m)\), and can be quantitatively regarded as an average distance between polymer chains (Figure II-10D). In contrast to the
entanglement threshold, which is strongly depended on the molecular weight, the average mesh size for an entangled polymer solution ($\Phi > \Phi^*$) is only dependent on the polymer concentration:

\[
\xi = 1.43 R_g \left( \frac{\Phi}{\Phi^*} \right)^{-3/4}
\]

Thus, two polymer solutions of the same type and concentration, but differing on their molecular weight, will present the same mesh size as long as they are entangled.

In order to implement the use of entangled polymer solutions to a wide range of biopolymer separations, the ability to vary the solution mesh size is a valuable characteristic. In order to form a small mesh, the concentration must be increased; however, as this latter one is increased, so the solution viscosity. Thus, to attain a small mesh size while minimizing the viscosity of the solution, the employment of shorter polymers should be carried out. The contrary should be used if larger mesh sizes want to be established. In both cases, working near the entanglement threshold is a requirement [45].

**Figure II-10.** Schematic representation of the entanglement process from a dilute to a semi-dilute solution. A) dilute solution; B) entanglement threshold concentration; C) semi-dilute solution. $R_g =$ radius of gyration; $\xi =$ average mesh size.
II.4.4.2.1.3 Mechanisms of electrophoretic migration in entangled polymer solutions

The Ogston sieving model and the reptation model are the two main theories that describe the migration of a flexible macromolecule through a polymer network. The applicability of each depends on the size of the migrating molecule relative to the size of the network.

In the Ogston model, it is assumed that the matrix consists of a random network of interconnected pores having an average pore size \((\xi_m)\), and that the migrating solute behaves as a non-deformable spherical particle of radius \(R_g\). In this model, the smaller molecules migrate faster because they have access to a larger fraction of the available pores (Figure II-11A). The Ogston model does not take into account the facts that the migrating molecule might deform in order to squeeze through a pore as it is well known that flexible molecules such as DNA migrate even when \(R_g >> \xi_m\) [39,45]. This fact is explained by the reptation model, which assumes that the molecule instead of migrating as a non-deformable particle, the migrating coil is forced to squeeze “head first” through the tubes formed by the polymer network matrix (Figure II-11B).

Figure II-11. Schematic representation of the mechanisms of electrophoretic migration in entangled polymer solutions. A) Ogston model; B) reptation model.
When dealing with complex compound mixtures, the need for high resolution analytical techniques is essential for the adequate separation of as many components as possible in the shortest analysis time. Despite the vast effort that has been set on the optimization of the separation power in one-dimensional (1D) liquid chromatography (LC) and capillary electrophoresis (CE), these improvements are often restricted in the peak capacity or require extended analysis times or sophisticated instrumentation. For those reasons, in the last decades vast interest has been put on the development of two-dimensional (2D) separation techniques which include 2D liquid chromatography (2D-LC), 2D capillary electrophoresis (2D-CE) and the combination of them (2D-LC-CE).

In a 2D separation, the sample is subjected to two different separation mechanisms, which results in a tremendous increase of the peak capacity and a reduction in the component overlap. Basically, samples from the first dimensional analysis are taken and then subjected to a second dimensional analysis. Here, a distinction can be made between 2D on-line and 2D offline separations. The offline approach is very straightforward since after the collection of the fractions of the first dimension manually or via a fraction collector, they are re-injected on to the second dimension, with the possibility of a previous pre-concentration. On the other hand, in the on-line approach, the fractions from the first dimension are automatically collected and immediately transferred onto the second dimension via an appropriate interface.

The 2D separation methods in LC and CE can be also subdivided in sequential, heart-cutting (LC-LC, CE-CE, LC-CE) and comprehensive (LC x LC, CE x CE, LC x CE) 2D separations (Figure II-12) [46,47]. Under the on-line approach, the comprehensive and the heart-cutting analyses are the preferred configurations. The main difference between those techniques relies on the number of fractions that are transferred from the first to the second dimension. In heart-cutting separations, only few relevant fractions containing target compounds are directed to the second dimension (Figure II-12A), whereas
in the comprehensive approach, the complete first dimension is sampled and analyzed by the second dimension (Figure II-12C).

The on-line heart-cutting approach permits the resolution of a small number of components within a defined retention time window; additionally, both dimensions can be set for optimal separation conditions. On the contrary, in on-line 2D comprehensive separations, the efficiency of the second dimension is usually sacrificed since a very fast analysis is required in order to sample and analyze all the first dimension. On-line 2D separations are characterized by a higher throughput as the complete analysis is automated, resulting in a higher reproducibility. Nevertheless, the higher complexity of the system in addition to the fact that not all separation modes can be easily combined in an on-line system due to incompatibility of the separation conditions, stand out as the principal limitations of this approach.

Under the offline 2D sequential separation line (Figure II-12B), both dimensions are operated independently. Moreover, the fractions of the first dimension can be treated in order to meet compatibility requirements of the second dimension; therefore, increasing the possible combinations of separation mechanisms. Additionally, there are no restrictions in the method development for both dimensions, leading to the establishment of optimal separation conditions in each dimension, thus increasing the separation power of this methodology. The lack of automation of this technique results in a more labor intensive operation, usually accompanied by a lower reproducibility, high time consumption and possible sample contamination, loss or degradation.
Figure II-12. Schematic representation of the multidimensional separation modes of separation. A) On-line Heart-cut separation. B) Offline sequential separation. C) On-line comprehensive separation.
II.5.1 Orthogonality, sampling frequency and peak capacity

The orthogonality, the sampling frequency and their relation with the peak capacity in 2D separations will be briefly described in this section.

II.5.1.1 Orthogonality

In multidimensional separations, the orthogonality between the two dimensions is a key aspect for the development of a successful platform. An orthogonal multidimensional separation is normally attained when the separation mechanisms in the two dimensions are independent of each other [47-51]. Once different selectivities are established in each of the dimensions, the maximal separation power in a multidimensional platform is attained as the probability of coeluting compounds in both dimensions is drastically reduced. Figure II-13 represents the concept of orthogonality. In a highly orthogonal 2D methodology (Figure II-13A), the separation of the compounds will be randomly distributed over the entire multidimensional separation space. On the other hand, when a high correlation is obtained and thus low orthogonality (Figure II-13B), the separation of coeluting compounds is poorly attained therefore only complicating the setup and system.

The orthogonality also depends on the physicochemical properties of the analytes, the mobile phases, buffers, sieving media and separation conditions. Thus, those latter aspects need to be wisely chosen within the properties of the sample.

![Figure II-13](Image)

Figure II-13. Schematic representation of the orthogonality in multidimensional separations. A) High orthogonality (low correlation). B) Low orthogonality (high correlation).
Principles of the separation techniques used in oligonucleotide analysis

II.5.1.2 Sampling frequency

The number of fractions that are transferred from the first to the second dimension have a drastic effect over the peak capacity of the first dimension. If sampling periods which are significantly larger than the peak widths of the compound emerging from the first dimension are used, the resolving power of this dimension is considerably reduced as a mixing of the sample occurs prior to the transfer to the second dimension [47]. For instance, when a fraction of two compounds separated in the first dimension is collected, the compounds will remix and the attained separation in the first dimension will be lost. In order to avoid a significant loss in the first dimension separation, a minimum of four samplings across the time equivalent of an 8σ peak width of each peak should be carried out [52].

II.5.1.3 Peak capacity in 2D separations

The peak capacity was introduced already in section II.3.5. Under the 2D approach, numerous models have been described for the calculation of the peak capacity [53-55]. Basically, the peak capacity in a 2D methodology is the one resulting from the product of the individual peak capacities (Eq. II-39):

\[
\text{Eq. II-39} \quad n_{c,2D} = n_c^1 \cdot n_c^2
\]

where \(n_{c,2D}\), \(n_c^1\) and \(n_c^2\) are the peak capacities for the theoretical 2D, first and second dimension, respectively.

When the peak capacity for a 2D separation is calculated as the product of peak capacities for the two separations, it is assumed that the two separations are completely orthogonal. Furthermore, the effect of the sampling rate also needs to be examined, as in most of the cases it is impractical to collect enough fractions in the first dimension to completely avoid undersampling. Therefore in the calculation of the practical peak capacity factors such as the orthogonality, the separation time window and the sampling frequency need to be considered. One practical way to calculate the peak capacity taking into account the orthogonality was proposed based on a geometrical approach [49]. Under this model, a correlation matrix via a factor analysis is created from the analytes retention parameters. The effective area of the 2D separation space covered by the eluting peaks is used as descriptor of the orthogonality which is afterwards subtracted to the peak capacity determined by:
Eq. II-40

\[ n'_{c,2D} = n_{c,2D} - \left[ \left( \frac{2}{n_c} \right)^2 \tan(\alpha) + \left( \frac{1}{n_c} \right)^2 \tan(\gamma) \right] \]

Figure II-14 represents the geometrical approach, for which the effective area is determined by the spreading angle \( \beta_0 \) which represents the degree of similarity between the two dimensions. For a highly orthogonal system (low correlation) the angle \( \beta_0 \) can increase till equaling \( \pi/2 \); on the other hand, if the orthogonality is very low (high correlation), the angle \( \beta_0 \) will tend to equal zero and the peak capacity will not differ from the one that could be calculated under the one dimensional approach.

![Figure II-14](image)

**Figure II-14.** Geometrical model for determining the peak capacity in 2D separations taking into account the orthogonality.

The effect of undersampling over the peak capacity has been studied by several groups [53-55]. One convenient model for the evaluation of the undersampling effect consists of the determination of the first dimension broadening factor \( \langle \beta \rangle \) as function of the sampling time \( t_s \):

Eq. II-41

\[ \langle \beta \rangle \approx \sqrt{1 + 0.21 \left( \frac{t_s}{\sigma} \right)} \]

where \( \sigma \) is the standard deviation of the peaks width prior to sampling. This approach for the determination of \( \langle \beta \rangle \) is valid for \( 0.2 \leq t_s/\sigma \leq 16 \) [53]. Additionally, this model can also be
Principles of the separation techniques used in oligonucleotide analysis

implemented for low or high orthogonal separations and showed to be unaffected by the presence of non-Gaussian peaks. Finally, the effective peak capacity can be calculated as follows:

\[
Eq. \ II-42 \quad n'_{c,2D} = \frac{n_{c,2D}}{\langle \beta \rangle}
\]

The fundamental aspects regarding chromatography, electrophoresis and multidimensional separations have been introduced in this chapter to provide the reader with the necessary background information allowing the discussion of separation strategies which are typically used for the analysis of oligonucleotides. The latter aspects together with a discussion of the basic properties of oligonucleotides are described in the following chapter.

II.6 References

Principles of the separation techniques used in oligonucleotide analysis

Oligonucleotides: characteristics and contemporary analytical strategies

Summary

An overview covering the most relevant properties of oligonucleotides and of the current analytical techniques used for their analysis is provided in this chapter. More in detail, the mechanism of action of therapeutic ONs, relevance of their nuclease resistance and toxicity aspects are summarized. Additionally, electrophoretic techniques (e.g. PAGE, CGE) and chromatographic techniques such as IEC, HILIC, IPC, mixed mode and bidimensional methods for the analysis of these molecules will be reviewed.
**III.1 Introduction**

Oligonucleotides (ONs) are unmodified or chemically modified single-stranded DNA or RNA molecules. In general, they are relatively short molecules consisting in average of 13-25 nucleotides [1]. These molecules are widely used to modulate the gene expression in a various fields including research, biotechnology and therapeutic applications [2].

In the last decades, the diverse useful applications which RNA molecules play in the regulation of cellular processes have been well documented [2,3]. The role that RNA molecules perform is indispensable in the central dogma of molecular biology, which states that DNA creates RNA and RNA makes proteins (Figure III-1). This process is subdivided into the DNA replication, transcription and translation [4]. During the former, a new complementary sequence of DNA is synthetized, resulting in an exact copy of the parent DNA sequence. In the transcription process, the genetic information encoded in a section of the DNA is replicated in the form of a newly assembled piece or messenger RNA (mRNA). In eukaryotic cells the first fragment that is formed is a pre-mRNA, which must follow a series of biochemical processes to result in the formation of a mature mRNA. One relevant process in the formation of the mature mRNA from the pre-mRNA is the splicing, in which basically the introns (nucleotide sequence not present in mRNA) are removed and the exons (nucleotide sequence present in the mRNA and that encodes for a protein) are joined. Once the mRNA is formed, it exits the cell nucleus and in the cytoplasm finds its way to a ribosome where it follows the translation process. Finally, the ribosome reads the mRNA sequence and builds up the polypeptide sequence resulting in the formation of a protein.
Figure III-1. Schematic representation of the protein synthesis in eukaryotic cells.

Due to the fact that mRNAs encode all cellular proteins, oligonucleotides designed to target mRNA proved to be effective for the treatment of diseases that are not curable by current small molecule based drugs [2].

The antisense technology with oligonucleotides was first described in 1978, when a single-stranded DNA fragment was found to inhibit the mRNA translation in a cell-free system [5-7]. However, the implementation of these molecules has mainly increased since information on the human genome became available and since the synthesis of oligonucleotides has been automated. As a consequence in the last decades, a lot of research has taken place in this field for the introduction of new generation of therapeutic agents.

The fulfilment of six criteria has been established for the bioengineering of an ON and its implementation as a therapeutic agent [8]. Firstly the ONs must follow an easy synthetic pathway and in bulk. As relevant example, the development of the phosphoramidite chemistry and its
Oligonucleotides: characteristics and contemporary analytical strategies

Implementation into automated technology [9] has greatly boosted the ease of synthesis till a point where methods for large scale are currently available in the market [10]. Secondly, the ONs must show sufficient stability in vivo. For that, several chemical modifications have been investigated and will be discussed in section III.3. The other 4 criteria include their specificity to the target mRNA, high hybridization activity, their ability to enter the target cell and to be retained there. These criteria have been mainly improved via the adequate combination of chemical modifications (section III.3) the selection of the adequate sequence-length of the ON and pharmaceutical preparations for enhancing their transfection [1,11,12].

Table III-I depicts a list of representative therapeutic ONs and the corresponding drug development stage they are currently in. Currently, there are a number of therapeutic oligonucleotides which are commercially available, while more than 50 oligonucleotides are undergoing clinical trials and many others are being studied in various drug discovery programs. That number is expected to increase in the upcoming decades as the understanding of the human genome is growing [2,3,13]. This illustrates the importance these molecules are acquiring in the therapeutic field and the concomitant need for improved analytical tools for the characterization of ONs.

Table III-I. Non-exhaustive list representative therapeutic oligonucleotides which are currently under development.

<table>
<thead>
<tr>
<th>Name</th>
<th>Laboratory</th>
<th>Clinical stage</th>
<th>Therapeutic Indication</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fomivirsen</td>
<td>Pfizer</td>
<td>Commercialized</td>
<td>Cytomegalovirus retinit</td>
<td>[3]</td>
</tr>
<tr>
<td>Pegaptanib</td>
<td>Pfizer</td>
<td>Commercialized</td>
<td>Age-related macular degeneration</td>
<td>[3]</td>
</tr>
<tr>
<td>Kinamro</td>
<td>Ionis pharmaceuticals</td>
<td>Commercialized</td>
<td>Homozygous Familial Hypercholesterolemia (HoFH)</td>
<td>[14]</td>
</tr>
<tr>
<td>Alcaforsen</td>
<td>Ionis pharmaceuticals</td>
<td>Commercialized</td>
<td>Pouchitis</td>
<td>[14]</td>
</tr>
<tr>
<td>ONIS-PKKRx</td>
<td>Ionis pharmaceuticals</td>
<td>Phase 1</td>
<td>Hereditary Angioedema</td>
<td>[14]</td>
</tr>
<tr>
<td>RG-012</td>
<td>Ionis pharmaceuticals</td>
<td>Phase 1</td>
<td>Alport Syndrome</td>
<td>[14]</td>
</tr>
<tr>
<td>IONIS-GSK4-LRx</td>
<td>Ionis pharmaceuticals</td>
<td>Phase 1</td>
<td>Ocular disease</td>
<td>[14]</td>
</tr>
<tr>
<td>IONIS-HBVRx</td>
<td>Ionis pharmaceuticals</td>
<td>Phase 1</td>
<td>Hepatitis B virus infection</td>
<td>[14]</td>
</tr>
<tr>
<td>IONIS-DGAT2Rx</td>
<td>Ionis pharmaceuticals</td>
<td>Phase 1</td>
<td>NASH (Nonalcoholic Steatohepatitis)</td>
<td>[14]</td>
</tr>
<tr>
<td>ISTH0036</td>
<td>ISARNA Therapeutics</td>
<td>Phase 1</td>
<td>Ophthalmology</td>
<td>[15]</td>
</tr>
<tr>
<td>DMD EXON 45 (SRP-4045)</td>
<td>SAREPTA Therapeutics</td>
<td>Phase 1</td>
<td>Duchenne muscular dystrophy (DMD)</td>
<td>[16]</td>
</tr>
<tr>
<td>AVI-7288</td>
<td>SAREPTA Therapeutics</td>
<td>Phase 1</td>
<td>Marburg virus</td>
<td>[16]</td>
</tr>
<tr>
<td>AVI-7537</td>
<td>SAREPTA Therapeutics</td>
<td>Phase 1</td>
<td>Ebola virus</td>
<td>[16]</td>
</tr>
<tr>
<td>AVI-7100</td>
<td>SAREPTA Therapeutics</td>
<td>Phase 1</td>
<td>Influenza</td>
<td>[16]</td>
</tr>
<tr>
<td>BP1001</td>
<td>Bio-Path Holdings</td>
<td>Phase 1</td>
<td>Chronic Myeloid Leukemia (CML)Solid Tumors</td>
<td>[17]</td>
</tr>
<tr>
<td>MTL-CEBPA</td>
<td>Mina Therapeutics</td>
<td>Phase 1</td>
<td>Liver cancer</td>
<td>[18]</td>
</tr>
<tr>
<td>Name</td>
<td>Laboratory</td>
<td>Clinical stage</td>
<td>Therapeutic Indication</td>
<td>Reference</td>
</tr>
<tr>
<td>-----------------</td>
<td>---------------------</td>
<td>----------------</td>
<td>--------------------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>QR-010</td>
<td>ProQR</td>
<td>Phase 1</td>
<td>Cystic fibrosis</td>
<td>[19]</td>
</tr>
<tr>
<td>FITUSIRAN</td>
<td>Alnylam pharmaceuticals</td>
<td>Phase 1</td>
<td>Hemophilia and Rare Bleeding Disorders</td>
<td>[20]</td>
</tr>
<tr>
<td>ALN-CCS</td>
<td>Alnylam pharmaceuticals</td>
<td>Phase 1</td>
<td>Complement-Mediated Diseases</td>
<td>[20]</td>
</tr>
<tr>
<td>ALN-AS1</td>
<td>Alnylam pharmaceuticals</td>
<td>Phase 1</td>
<td>Hepatic Porphyrias</td>
<td>[20]</td>
</tr>
<tr>
<td>ALN-AAT</td>
<td>Alnylam pharmaceuticals</td>
<td>Phase 1</td>
<td>Alpha-1 Antitrypsin Deficiency</td>
<td>[20]</td>
</tr>
<tr>
<td>ALN-GO1</td>
<td>Alnylam pharmaceuticals</td>
<td>Phase 1</td>
<td>Primary Hyperoxaluria Type 1</td>
<td>[20]</td>
</tr>
<tr>
<td>ALN-TTRsc02</td>
<td>Alnylam pharmaceuticals</td>
<td>Phase 1</td>
<td>Hereditary ATTR Amyloidosis</td>
<td>[20]</td>
</tr>
<tr>
<td>ARC-521</td>
<td>Arrowhead Pharmaceuticals</td>
<td>Phase 1</td>
<td>Hepatitis B</td>
<td>[21]</td>
</tr>
<tr>
<td>ARC-AAT</td>
<td>Arrowhead Pharmaceuticals</td>
<td>Phase 1</td>
<td>Alpha-1-Antitrypsin Deficiency</td>
<td>[21]</td>
</tr>
<tr>
<td>RXI-109</td>
<td>Renova Therapeutics</td>
<td>Phase 1</td>
<td>Retinal Scarring</td>
<td>[22]</td>
</tr>
<tr>
<td>DCR-MYC</td>
<td>Dicerna Pharmaceuticals</td>
<td>Phase 1</td>
<td>Various solid tumors</td>
<td>[23]</td>
</tr>
<tr>
<td>DCR-MYC</td>
<td>Dicerna Pharmaceuticals</td>
<td>Phase 1</td>
<td>Hepatocellular Carcinoma</td>
<td>[23]</td>
</tr>
<tr>
<td>IONIS-DMPK-2.5Rx</td>
<td>Ionis pharmaceuticals</td>
<td>Phase 2</td>
<td>Myotonic Dystrophy Type 1</td>
<td>[14]</td>
</tr>
<tr>
<td>IONIS-HTTRx</td>
<td>Ionis pharmaceuticals</td>
<td>Phase 2</td>
<td>Huntington’s Disease</td>
<td>[14]</td>
</tr>
<tr>
<td>ATL1103</td>
<td>Ionis pharmaceuticals</td>
<td>Phase 2</td>
<td>Acromegaly</td>
<td>[14]</td>
</tr>
<tr>
<td>IONIS-SOD1Rx</td>
<td>Ionis pharmaceuticals</td>
<td>Phase 2</td>
<td>Amyotrophic Lateral Sclerosis</td>
<td>[14]</td>
</tr>
<tr>
<td>IONIS-FXRx</td>
<td>Ionis pharmaceuticals</td>
<td>Phase 2</td>
<td>Antithrombosis</td>
<td>[14]</td>
</tr>
<tr>
<td>IONIS-APO(a)-LRx</td>
<td>Ionis pharmaceuticals</td>
<td>Phase 2</td>
<td>High Lp(a)</td>
<td>[14]</td>
</tr>
<tr>
<td>IONIS-ANGPTL3-LRx</td>
<td>Ionis pharmaceuticals</td>
<td>Phase 2</td>
<td>Mixed Dyslipidemias</td>
<td>[14]</td>
</tr>
<tr>
<td>Apatoseren (OGX-427)</td>
<td>Ionis pharmaceuticals</td>
<td>Phase 2</td>
<td>Cancer</td>
<td>[14]</td>
</tr>
<tr>
<td>IONIS-AR-2.5Rx (AZ05312)</td>
<td>Ionis pharmaceuticals</td>
<td>Phase 2</td>
<td>Prostate Cancer</td>
<td>[14]</td>
</tr>
<tr>
<td>IONIS-STAT3-2.5Rx (AZ09150)</td>
<td>Ionis pharmaceuticals</td>
<td>Phase 2</td>
<td>Cancer</td>
<td>[14]</td>
</tr>
<tr>
<td>ATL1102</td>
<td>Ionis pharmaceuticals</td>
<td>Phase 2</td>
<td>Multiple Sclerosis</td>
<td>[14]</td>
</tr>
<tr>
<td>RG-101</td>
<td>Ionis pharmaceuticals</td>
<td>Phase 2</td>
<td>Hepatitis C</td>
<td>[14]</td>
</tr>
<tr>
<td>IONIS-GCGRRx</td>
<td>Ionis pharmaceuticals</td>
<td>Phase 2</td>
<td>Diabetes</td>
<td>[14]</td>
</tr>
<tr>
<td>IONIS-PTP1Rx</td>
<td>Ionis pharmaceuticals</td>
<td>Phase 2</td>
<td>Diabetes</td>
<td>[14]</td>
</tr>
<tr>
<td>IONIS-FGR4Rx</td>
<td>Ionis pharmaceuticals</td>
<td>Phase 2</td>
<td>Obesity</td>
<td>[14]</td>
</tr>
<tr>
<td>DMD EXON 53 (SRP-4053)</td>
<td>SAREPTA Therapeutics</td>
<td>Phase 2</td>
<td>Duchenne muscular dystrophy (DMD)</td>
<td>[16]</td>
</tr>
<tr>
<td>BP1001</td>
<td>Bio-Path Holdings</td>
<td>Phase 2</td>
<td>Acute Myeloid Leukemia (AML)</td>
<td>[17]</td>
</tr>
<tr>
<td>Apatoseren (OGX-427)</td>
<td>Oncogenex</td>
<td>Phase 2</td>
<td>Cancer</td>
<td>[24]</td>
</tr>
<tr>
<td>SB010</td>
<td>Sterna Biologicals</td>
<td>Phase 2</td>
<td>Chronic Obstructive Pulmonary Disease (COPD)</td>
<td>[25]</td>
</tr>
<tr>
<td>SB010</td>
<td>Sterna Biologicals</td>
<td>Phase 2</td>
<td>Asthma</td>
<td>[25]</td>
</tr>
<tr>
<td>SB011</td>
<td>Sterna Biologicals</td>
<td>Phase 2</td>
<td>Athopic dermatitis</td>
<td>[25]</td>
</tr>
<tr>
<td>SB012</td>
<td>Sterna Biologicals</td>
<td>Phase 2</td>
<td>Ulcerative colitis</td>
<td>[25]</td>
</tr>
<tr>
<td>PNT2258</td>
<td>ProNai</td>
<td>Phase 2</td>
<td>Cancer</td>
<td>[26]</td>
</tr>
<tr>
<td>ATU027</td>
<td>Silence Therapeutics</td>
<td>Phase 2</td>
<td>Pancreatic cancer</td>
<td>[27]</td>
</tr>
<tr>
<td>QPI 1002</td>
<td>Quark pharmaceuticals</td>
<td>Phase 2</td>
<td>Acute Kidney injury</td>
<td>[28]</td>
</tr>
<tr>
<td>QPI 1007</td>
<td>Quark pharmaceuticals</td>
<td>Phase 2</td>
<td>Acute Primary Angle Closure Glaucoma</td>
<td>[28]</td>
</tr>
<tr>
<td>PF-655</td>
<td>Quark pharmaceuticals</td>
<td>Phase 2</td>
<td>Diabetic macular Edema</td>
<td>[28]</td>
</tr>
<tr>
<td>PF-655</td>
<td>Quark pharmaceuticals</td>
<td>Phase 2</td>
<td>Wet Age-related macular degeneration</td>
<td>[28]</td>
</tr>
<tr>
<td>ALN-PCSsc</td>
<td>Alnylam pharmaceuticals</td>
<td>Phase 2</td>
<td>Hypercholesterolemia</td>
<td>[20]</td>
</tr>
<tr>
<td>ARC-520</td>
<td>Arrowhead Pharmaceuticals</td>
<td>Phase 2</td>
<td>Hepatitis B</td>
<td>[21]</td>
</tr>
<tr>
<td>RXI-109</td>
<td>Renova Therapeutics</td>
<td>Phase 2</td>
<td>Dermal Scarring</td>
<td>[22]</td>
</tr>
<tr>
<td>Samcyprome</td>
<td>Renova Therapeutics</td>
<td>Phase 2</td>
<td>Warts</td>
<td>[22]</td>
</tr>
<tr>
<td>Name</td>
<td>Laboratory</td>
<td>Clinical stage</td>
<td>Therapeutic Indication</td>
<td>Reference</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------------------------------</td>
<td>----------------</td>
<td>-------------------------------------------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>RT-100</td>
<td>Renova Therapeutics</td>
<td>Phase 2</td>
<td>Congestive heart failure</td>
<td>[22]</td>
</tr>
<tr>
<td>IONIS-TTRx</td>
<td>Ionis pharmaceuticals</td>
<td>Phase 3</td>
<td>Familial Amyloid Polyneuropathy</td>
<td>[14]</td>
</tr>
<tr>
<td>Nusinersen</td>
<td>Ionis pharmaceuticals</td>
<td>Phase 3</td>
<td>Infants with Spinal Muscular Atrophy</td>
<td>[14]</td>
</tr>
<tr>
<td>Volanesorsen</td>
<td>Ionis pharmaceuticals</td>
<td>Phase 3</td>
<td>Familial Chylomicronemia Syndrome (FCS)</td>
<td>[14]</td>
</tr>
<tr>
<td>Custirsen (OGX-011)</td>
<td>Ionis pharmaceuticals</td>
<td>Phase 3</td>
<td>Prostate / Lung Cancer</td>
<td>[14]</td>
</tr>
<tr>
<td><strong>DMD EXON 51 (ETEPLIRSEN / AVI-4658)</strong></td>
<td>SAREPTA Therapeutics</td>
<td>Phase 3</td>
<td>Duchenne muscular dystrophy (DMD)</td>
<td>[16]</td>
</tr>
<tr>
<td>Custirsen (OGX-011)</td>
<td>Oncogenex</td>
<td>Phase 3</td>
<td>Cancer</td>
<td>[19]</td>
</tr>
<tr>
<td>QPI 1002</td>
<td>Quark pharmaceuticals</td>
<td>Phase 3</td>
<td>Delayed Graft Function (DGF)</td>
<td>[28]</td>
</tr>
<tr>
<td>QPI 1007</td>
<td>Quark pharmaceuticals</td>
<td>Phase 3</td>
<td>Non Arteritic Anterior Ischemic Optic Neuropathy</td>
<td>[28]</td>
</tr>
<tr>
<td>Patisiran</td>
<td>Alnylam pharmaceuticals</td>
<td>Phase 3</td>
<td>Hereditary ATTR Amyloidosis</td>
<td>[20]</td>
</tr>
<tr>
<td>Revusiran</td>
<td>Alnylam pharmaceuticals</td>
<td>Phase 3</td>
<td>Hereditary ATTR Amyloidosis</td>
<td>[20]</td>
</tr>
<tr>
<td>BP1001</td>
<td>Bio-Path Holdings</td>
<td>Preclinical</td>
<td>Solid Tumors</td>
<td>[17]</td>
</tr>
<tr>
<td>BP1002</td>
<td>Bio-Path Holdings</td>
<td>Preclinical</td>
<td>Follicular Lymphoma</td>
<td>[17]</td>
</tr>
<tr>
<td>SB021</td>
<td>Sterna Biologicals</td>
<td>Preclinical</td>
<td>Chronic Obstructive Pulmonary Disease (COPD)</td>
<td>[25]</td>
</tr>
<tr>
<td>QP-HL2</td>
<td>Quark pharmaceuticals</td>
<td>Preclinical</td>
<td>Ménière’s Disease</td>
<td>[28]</td>
</tr>
<tr>
<td>QP-L1</td>
<td>Quark pharmaceuticals</td>
<td>Preclinical</td>
<td>Lung Transplantation, Primary Graft Dysfunction</td>
<td>[28]</td>
</tr>
<tr>
<td>ARC-LPA</td>
<td>Arrowhead Pharmaceuticals</td>
<td>Preclinical</td>
<td>Cardiovascular Disease</td>
<td>[21]</td>
</tr>
<tr>
<td>ARC-HIF2</td>
<td>Arrowhead Pharmaceuticals</td>
<td>Preclinical</td>
<td>Clear Cell, Renal Cell Carcinoma</td>
<td>[21]</td>
</tr>
<tr>
<td>Isth1106</td>
<td>ISARNA Therapeutics</td>
<td>Preclinical</td>
<td>Fibrosis</td>
<td>[15]</td>
</tr>
<tr>
<td>Isth0047</td>
<td>ISARNA Therapeutics</td>
<td>Preclinical</td>
<td>Cancer</td>
<td>[15]</td>
</tr>
<tr>
<td>Isth1047</td>
<td>ISARNA Therapeutics</td>
<td>Preclinical</td>
<td>Cancer</td>
<td>[15]</td>
</tr>
<tr>
<td><strong>DMD EXON 44</strong></td>
<td>SAREPTA Therapeutics</td>
<td>Preclinical</td>
<td>Duchenne muscular dystrophy (DMD)</td>
<td>[16]</td>
</tr>
<tr>
<td><strong>DMD EXON 52 (SRP-4052)</strong></td>
<td>SAREPTA Therapeutics</td>
<td>Preclinical</td>
<td>Duchenne muscular dystrophy (DMD)</td>
<td>[16]</td>
</tr>
<tr>
<td><strong>DMD EXON 50</strong></td>
<td>SAREPTA Therapeutics</td>
<td>Preclinical</td>
<td>Duchenne muscular dystrophy (DMD)</td>
<td>[16]</td>
</tr>
<tr>
<td><strong>DMD EXON 55</strong></td>
<td>SAREPTA Therapeutics</td>
<td>Preclinical</td>
<td>Duchenne muscular dystrophy (DMD)</td>
<td>[16]</td>
</tr>
<tr>
<td><strong>DMD EXON 8</strong></td>
<td>SAREPTA Therapeutics</td>
<td>Preclinical</td>
<td>Duchenne muscular dystrophy (DMD)</td>
<td>[16]</td>
</tr>
<tr>
<td><strong>DMD EXON 35</strong></td>
<td>SAREPTA Therapeutics</td>
<td>Preclinical</td>
<td>Duchenne muscular dystrophy (DMD)</td>
<td>[16]</td>
</tr>
<tr>
<td>OGX-225</td>
<td>Oncogenex</td>
<td>Preclinical</td>
<td>Solid tumors</td>
<td>[24]</td>
</tr>
<tr>
<td>SB020</td>
<td>Sterna Biologicals</td>
<td>Preclinical</td>
<td>Psoriasis</td>
<td>[25]</td>
</tr>
<tr>
<td>QR-110</td>
<td>ProQR</td>
<td>Preclinical</td>
<td>Leber’s Congenital Amaurosis</td>
<td>[19]</td>
</tr>
<tr>
<td>ATU111</td>
<td>Silence Therapeutics</td>
<td>Preclinical</td>
<td>Lung cancer</td>
<td>[26]</td>
</tr>
<tr>
<td>Aln-HBV</td>
<td>Alnylam pharmaceuticals</td>
<td>Preclinical</td>
<td>Hepatitis B Virus Infection</td>
<td>[20]</td>
</tr>
<tr>
<td>ARC-F12</td>
<td>Arrowhead Pharmaceuticals</td>
<td>Preclinical</td>
<td>Thrombosis and Angioedema</td>
<td>[21]</td>
</tr>
<tr>
<td>RXI-185</td>
<td>Renova Therapeutics</td>
<td>Preclinical</td>
<td>Collagenase</td>
<td>[22]</td>
</tr>
<tr>
<td>RXI-231</td>
<td>Renova Therapeutics</td>
<td>Preclinical</td>
<td>Tyrosinase</td>
<td>[22]</td>
</tr>
<tr>
<td>DCR-PH1</td>
<td>Dicerna Pharmaceuticals</td>
<td>Preclinical</td>
<td>Primary Hyperoxaluria 1</td>
<td>[23]</td>
</tr>
</tbody>
</table>
III.2 Relevant mechanisms of action of antisense oligonucleotides

Basically, antisense oligonucleotides hybridize specifically to a complementary mRNA sequence via Watson-Crick base paring [3]. Prior to that, the mRNA must be accessible to the oligonucleotide; therefore, it must be first taken up by a target cell, tissue or organ, and must be present in an adequate amount to invoke a biological response. The oligonucleotides must arrive intact at the mRNA, avoiding their degradation by the various endonucleases and exonucleases present intracellularly and within the serum and tissue [1-3,8,11,29-32]. In order to improve their stability, several chemical modifications have been carried out and they will be discussed more in detail in section III.3.

Natural oligonucleotides are usually poorly internalized by the cell, mainly because of their anionic character [2,33]. Moreover, they tend to localize in endosomes/lysosomes where they also follow degradation [3,30]. Several strategies have been investigated to facilitate the transfection of ONs through the cell membrane. The main ones include direct microinjection [30,34,35]; chemical attachment of a transfection entity such as cell penetrating peptides [36-38], and other small molecules (e.g. cholesterol) [39,40] and the usage of transfection reagents such as dendrimers [41-45], cationic polymers [33,46,47] and neutral or cationic liposomes [48-57], this last one being the most common transfection vector.

The most widely recognized mechanism of protein inhibition involves RNase H mediated cleavage of a target mRNA [1-3,11]. Once the ONs hybridize with the target mRNA, a duplex is formed, which is an adequate substrate for RNase H (Figure III-2A). This enzyme is an endonuclease that recognizes RNA/DNA duplexes and selectively cleaves the RNA strand. It has been demonstrated that once the RNA cleavage occurs, the oligonucleotide is thought to dissociate from the duplex and become available to bind a second target mRNA molecule [11]. Moreover, as many ONs modifications designed to increase the ON antisense activity and stability (section III.3) are not likely to induce the RNase H cleavage, the translation arrest and splicing inhibition are also well recognized mechanisms of action [2,3]. In the translation arrest, the ON binds near to the translation initiation region of an mRNA, and in this way impedes either the recognition, assembly, or movement of the ribosome once the assembly has taken place (Figure III-2B). In the splicing inhibition, the ON binds a specific region
Oligonucleotides: characteristics and contemporary analytical strategies

of a pre-mRNA and represses sequences or enhances splicing, resulting in exon skipping and/or inclusion of an alternatively spliced exon [2,58,59] (Figure III-2C).

Another mechanism of action includes the RNA interference pathway (iRNA), which is a natural mechanism through which natural microRNA regulates the gene expression. The gene suppression through this mechanism is very similar to the RNase H mediated mechanism; nevertheless, the main difference is that the antisense ON is bound to the cleavage enzyme prior to its hybridization with the target mRNA. More in detail, a double stranded synthetic RNA ON is complexed via a series of proteins such as DICER, Argonaute 2 (AGO2) and other proteins to form a RNA induced silencing complex (RISC). Afterwards the RISC binds to the target mRNA, allowing AGO 2 to cleave the mRNA (Figure III-3). In this mechanism, it is likely to obtain a significant number of off-target effects as the number of nucleotides that are hybridized with the target mRNA allows for a higher number of mismatches [2].

**Figure III-2.** Mechanism of action by which antisense ONs disrupt the protein synthesis. A) RNase H mediated cleavage; B) Translation arrest, the ribosomes are not able to assemble or their movement through the mRNA strand is blocked; C) Splicing inhibition, the formation a mature mRNA is impeded.
Figure III-3. RNA interference mechanism. The double stranded iRNA forms a complex with DICER and AGO 2 to form a RNA induced silencing complex (RISC) which afterwards cleaves the target mRNA.

III.3 Relevance of nuclease resistance of oligonucleotides

Unmodified DNA and RNA oligonucleotides are inherently unstable both in serum and cells as they are rapidly degraded by exonucleases and endonucleases [1-3,11,31]. In addition to that, they show poor pharmacokinetic properties as their low cell penetrating ratios and a weak interaction with plasma proteins leads to their rapid excretion from the organism by the kidney into the urine [2,30]. In order to improve their stability, various chemical modifications have been introduced into the sugar moiety, the nitrogen base and the phosphate group [60] (Figure III-4). For a more clear understanding, Figure III-4 also depicts the bases whereby the position where they bind to the sugar moiety is indicated. Additionally, Figure III-5 Error! Reference source not found. depicts the hydrogen bonding interaction that occurs between the bases (A-T, C-G for DNA and A-U, C-G for RNA).
Figure III-4. Overview of the most common chemical modifications investigated in therapeutic nucleotides. The numeration of the carbons in the sugar moiety is indicated in the central figure as well as the 5' and 3' ending positions in an ON. The bases are depicted at the bottom of the figure, where a red asterisk indicates the nitrogen which is bonded to position 1 of the sugar.

The most common chemical modification commercially implemented is the phosphorothioate, in which a non-bridging oxygen atom of the nucleotide backbone is replaced by a sulfur atom [61]. This modification encountered in the first generation of antisense ONs, fully supported RNase H activity.
and dramatically increased nuclease resistance [62]. However, this modification also reduced the binding activity towards the complementary mRNA strand. The first approved ON drug Vitravene (Formivirsem), used for the treatment of retinal inflammation caused by the cytomegalovirus (CMV) in AIDS patients, contained this type of phosphate backbone modification. Since then, the majority of the ONs currently following clinical trials are based on this kind of chemistry. Other common modifications in the backbone include the phosphoramidate [9,63,64] and the methylphosphonate [65], in which the oxygen in position 3 in the sugar is substituted with a 3-amino group for the first one and a methyl group is placed instead of an oxygen in the phosphate group for the second one, respectively. Furthermore, the thiophosphoramidate modification which combines the phosphoramidate and the phosphorothioate modifications was also developed [66] and demonstrated good hybridization activity although this did not activate the RNase H mechanism.

Several other substitutions of the phosphodiester groups [2,60,67] such as acetals [66], hydroxylamines [68] and amides [69] have been described for the replacement of the phosphorous atom; however, improvements over the phosphorothioate backbone have not been attained yet [2].

![Figure III-5. Base paring in DNA and RNA oligonucleotides. The arrows indicate the position where the base forms a covalent bond with the sugar moiety. The hydrogen bonding is indicated with the green dotted lines.](image)

The second generation of therapeutic ONs is limited to RNA based ONs as modifications with 2'-O-methyl, 2'-O-methoxyethil, or 2'-O-aminopropyl group at the 2’ position of the ribose moiety have
Oligonucleotides: characteristics and contemporary analytical strategies

been described [70]. These modifications improve the hybridization with the complementary mRNA, and the proximity of the 2’ substituent to the 3’ phosphate group in an ON, usually demonstrated increased nuclease resistance. Nevertheless, these modifications form duplexes with the mRNA that are practically not cleaved by the RNase H. This limitation has been successfully overcome by the combination of modifications, such as ONs with phosphorothioate modifications in the central part and 2’-O-methylribonucleosides at the ends. These terminal nucleotides increased the affinity and nuclease resistance, whereas the central region of the ON allowed for an effective RNase H mediated cleavage.

Locked nucleic acids (LNA), morpholino and peptide nucleic acids (PNA) are part of the third generation of therapeutic ONs. All of them demonstrated improved resistance towards nucleases and good hybridization activities [71-75]. The LNA ONs are considered as analogs of the 2’-O-methyl modification as the 2’-substituent is bounded to the 4’-C atom in the ribose, forming a bicyclic system. Uniform LNA ONs do not support the RNase H cleavage [76] and therefore this modification is encountered in combination with other ones in a similar manner as the 2’-O-methyl ONs [73].

In the morpholino ONs the sugar moiety is replaced by a morpholine ring and a phosphorodiamidate linkage connecting the morpholine nitrogen atom with the hydroxyl group of the 3’-side residue is used for connecting the nucleotide units. Despite good properties in nuclease resistivity and hybridization activity, these ONs do not activate RNase H and are primarily being explored in the framework of translation arrest and splicing inhibition mechanisms [75].

The PNA ONs are characterized by the replacement of the sugar phosphate backbone by a peptide. These ONs are highly resistant to degradation by nucleases and proteases and demonstrated high hybridization activities [77-80]. However, similarly to morpholino ONs they have been used primarily in the splicing inhibition and translation arrest mechanisms.

III.4 Toxicity aspects of antisense oligonucleotides

In similar way as for the small molecule based therapeutics, antisense ONs also exhibit dose dependent toxicities. The potential toxicological effects of these drugs are usually sub-classified as
hybridization dependent or not [2,30,61,81]. The hybridization dependent toxicity is generated by the over inducement of the pharmacological effect and also by the hybridization to non-targeted RNAs. Most of the antisense ONs have a length between 15-20 bases. That length, theoretically should hybridize specifically with one unique complementary mRNA encoded by the human genome [82]. ONs with this length range have proved to successfully discriminate between two genes that differ by a mutation on a single base [11,83]; however, the ON will partially complement several other genes. Longer ON sequences (> 30 nucleotides) might increase the risk of non-sequence specific mRNA cleavage as the probabilities for the hybridization to other mRNA sites increases. These effects are minimized by the adequate sequence-length bioengineering that will allow for a specific hybridization or a maximum of three mismatches (number of genes to which the ON is predicted to hybridize directly) [2]. Shorter ONs (< 13 nucleotides) generally do not show sufficient affinity to result in adequate potency. Moreover, for ONs that suppress the protein formation RNA interference (iRNA) represents a significant challenge as only six to eight nucleotides may be required for activity, which can result in hundreds to thousands of potential non target interactions [2].

The non-hybridization toxicity results from the interaction of the ON with proteins. This toxicity mainly depends on the chemistry of the ON and the way they interact with a specific protein. Effects on the immune cell activation [84,85], complement activation [86] and coagulation [87] have been mainly reported. Furthermore, this toxicity is also influenced by the dose; therefore, similar studies and precautions such as for the small molecule therapeutic drugs need to be carried out. Therefore having good purity indicating methods for the analysis of ONs is an essential control parameter.

### Ill.5 Analytical separation methods for the analysis of oligonucleotides

After reviewing the mechanisms of action of antisense ONs, the vast number of chemical modifications and combinations and possible toxic effects, it is clear that the need for specific and sensible analytical methodologies is required for pharmaceutical research and quality control purposes. The analytical methods must be able to accurately quantify, separate and characterize
these therapeutic molecules, their metabolites and degradation products in pharmaceutical preparations, and in vitro and in vivo studies.

The main synthetic impurities and degradation products encountered in these molecules are very closely related to each other, and include mainly adduct sequences (\(n+1\), \(n+2\)...) or sequence deletions (\(n-1\), \(n-2\)...). In addition, other impurities resulting from oxidation, depurination and other bases are also included in a family of closely related impurities that increase the challenges in the resolution of the main product. Furthermore, the number of impurities can drastically increase when dealing with double stranded iRNAs as each strand carries its own set of impurities. In this duplex, each strand has a 5’ and 3’ phosphate and hydroxyl group, respectively [2,3]. Also, they can follow metabolism by nucleases either on the 3’ or the 5’ position, leading to the introduction of additional impurities by each single strand. Moreover, the analytical methodologies also need to be capable to separate ONs of the same length but different base compositions, which extends their applicability as primers in the polymerase chain reaction and microarray analysis.

In this section the separation analytical methodologies employed thus far for the analysis of the oligonucleotides will be reviewed. Electrophoretic and chromatographic methods employing ultraviolet (UV), fluorescence and mass spectrometric detection (MS) will be discussed.

**III.5.1 Electrophoretic methods for the analysis of oligonucleotides**

The electrophoretic methodologies have become essential in the analysis of biopolymers. Since the first electrophoretic separations reported by Tiselius in the 1930s [88], electrophoretic separations have been under constant evolution. The first separations of nucleic acid polymers were carried on slab gels matrixes, mainly consisting of agarose and polyacrylamide. Agarose gels were mainly used for the separation of DNA due to their larger pore structure while polyacrylamide was primarily employed for a wide variety of biopolymers including proteins and oligonucleotides [89]. Polyacrylamide gel electrophoresis (PAGE) has been widely employed for the stability assessments of oligonucleotides [90-96], where fluorescent or radiolabeled ONs were employed (Figure III-6A). The usage of slab gels provided a great separation power, where ONs of 600 bases could be separated by a nucleotide difference [97]. Nevertheless, the slab gel approach presents the disadvantage of being slow, difficult to quantitate and consisting of many time consuming manual steps. Moreover,
the detection requires the usage of toxic agents and exposure to UV radiation, therefore this technique is currently mainly reserved for purification purposes. The advances in instrumentation technology led to the establishment of gel electrophoresis in the capillary format, which offered the possibility to transfer the routine slab gel electrophoretic methods into an automated capillary format. Additionally, with the implementation of capillary electrophoresis, the diverse modes of operation it presents (section II.4.5) greatly expanded its applicability to diverse type of molecules. Nevertheless, the application of CE for the separation of therapeutic ONs remains largely limited to CZE and CGE.

Capillary zone electrophoresis (CZE) only proved to be useful for the separation of nucleotides [98-100] and oligonucleotides up to 13 bases in length [98] since the variability in the mass-to-charge ratio of ONs is greatly diminished as the length increases. CZE-MS analysis has been used to characterize ONs [99,101,102]; however, a separation between larger ONs is challenging to attain with this technique. Micellar electrokinetic chromatography, which is a variant of CZE containing charged micelles and typically is used for the analysis of neutral solutes, has thus far only been used for the analysis of nucleotides [103,104].

Capillary gel electrophoresis (CGE) has been the technique of choice in the early years of ON pharmaceutical development [105]. The separation of ONs occurs via a sieving mechanism (section II.4.5.5.1.3). However, the exploitation of hydrophobic interactions in this mode has also been proposed with the usage of poly(N-vinylpyrrolidone) as a pseudophase [106]. In general the CGE methodologies of ONs can be differentiated mainly by the gel matrix they employ.

Rigid gel matrixes are chemically attached to the surface of the capillary column, and in the same way as for slab gel electrophoresis, a chemical cross-linker is used to form the gel and tune the matrix pore size and viscosity. The higher the concentrations of the cross-linker and the monomer, the smaller the pore size of the gel. The first separations of ONs on single base of a dA40-60 DNA ON ladder proved the value of this methodology [107] (Figure III-6B). Nevertheless, the problems associated with gel matrixes such as their sensitivity to temperature, changes in pH and high voltage, usually lead to the formation of a bubble and as a consequence the interruption of the current flow. Additionally, the low reproducibility between batches, the extended preparation time and the short life time of this gel filled capillaries, limited the development and validation of these methods for quantitative analysis.
The replaceable gel matrixes greatly increased the reproducibility and the ease of use of CGE. These matrixes consisting of entangled polymer solutions (section II.4.5.5.1) such as linear polyacrylamide [108-110], alkylcellulose [110,111], and low melting agarose [112,113] have been implemented for the separation of double and single stranded DNA , ranging from few nucleotides till thousands of base pairs [114]. Matrixes consisting of hydroxyethyl cellulose [115,116], polyethylene glycol [110,117,118] and many mixtures and derivatives of them such as polyvinylpyrrolidone and poly-TrisA [119] have been used for DNA sequencing. In the last decade, polymer solutions capable to separate single stranded ONs by a single base resolution have been reported [120]. The usage of denaturing conditions such as 7M urea [89,109-111,121] and organic modifiers (e.g. acetonitrile) [117,122] has also been reported in combination with replaceable polymers and attaining a single base resolution. This greatly extended the applicability to the analysis of therapeutic iRNA ONs and to other hybridization studies. Furthermore, novel thermoresponsive copolymers consisting of hydrophobic and hydrophilic blocks such poly-(ethylene oxide)/poly(propylene oxide) (pluronic gels) [113,123-125] (Figure III-6C), Poly(N-isopropylacrylamide) [126], N,N-diethylacrylamide/N,N-dimethylacrylamide [127], polyacrylamide/poly(propylene oxide) (PPO) [128] and poly-N-alkoxyalkylacrylamide [129] have been employed for the separation of DNA and ONs and were capable to achieve a single base resolution.

Figure III-6. A) Slab PAGE of a 25pb DNA ladder [130]; B) CGE on a rigid gel matrix (acrylamide/bis-acrylamide) of an A40-mer till A60-mer ON mixture [107]; C) CGE of a T19-T24 ON mixture using liquid crystalline Pluronic F127 as sieving media [123].
CGE has proven to be a viable technique in the stability tests of ONs [116,121,131-133] and in pharmacokinetic and metabolism studies of ONs in biological matrixes in combination with diverse extraction and sample preparation procedures [111,134-137]; however, improvements of the reduction the sample preparations steps, and research into different gel matrixes can be carried out to increase the ease of use of this technique and to improve the robustness of the method to different sample matrixes and injection modes.

III.5.2 Liquid chromatography methods for the analysis of oligonucleotides

The higher robustness of liquid chromatography (LC) when compared with CE makes this technique ideal for the analysis of ONs, as the sample matrix effects over the sample introduction and separation are somewhat diminished. The first attempts for the substitution of slab gel electrophoresis with LC using porous stationary phases did not yield to the expected enhancement on the resolution and analysis time due to the lower inter-particle distribution of these molecules [138]. Nevertheless, since the introduction of non-porous micropellicular packings [139-142], the chromatographic analyses of nucleic acid polymers can be obtained within minutes due to an improved mass transfer kinetics, increased surface accessibility and fast column regeneration.

The LC analytical methodologies employed thus far for the analysis of ONs mainly include ion-exchange chromatography (IEC), ion-pair chromatography (IPC), hydrophilic interaction liquid chromatography (HILIC) and size exclusion chromatography (SEC). These techniques will be briefly reviewed in the following sections.

III.5.2.1 Ion-exchange chromatography for the analysis of oligonucleotides

Ion-exchange chromatography (IEC) has evolved to become a true high performance method for the separation of charged molecules (section II.2.1) which proves particularly suitable for the analysis of biomolecules including oligonucleotides. The ability of IEC for the analysis of ONs in biological matrixes has proven its versatility as it allows separation of ONs from other polar molecules in complex matrixes. Additionally the significant tolerance for salts it presents allows the use of simplified sample preparation procedures when compared to other analytical techniques such as
capillary electrophoresis. IEC has been particularly useful for quality control purposes and purification of ONs; however, its ability to detect changes in the nucleotide composition is somehow limited. IEC separations allow the implementation of UV and fluorescence detection of ONs.

Both strong and weak anion exchanges have been successfully implemented for the separation of DNA and RNA nucleic acid polymers and ONs [143-160] (Figure III-7A). In the last years, several improvements on the stationary phases of IEC columns have been carried out, in particular for strong anion exchangers. Porous silica alkylamine derivatized [154] and polyethyleneimine (PEI) [155,157] coated supports were developed for the analysis of ONs. Further optimizations on the PEI chemistry which included the establishment of a quaternary amine on the ion exchanger, improved the separation of ONs up to 50 bases in length [152].

The mechanism of separation of ONs by IEC is not purely ionic, and a complex set of secondary interactions might occur. The incomplete coverage of the support particles by the ammonium groups in the stationary phase also leads to the possibility of non-ionic interactions between the analyte and the stationary phase. This effect is particularly dramatic in the separation of phosphorothioate ONs, where a chiral center is formed after the substitution of an oxygen by the sulfur atom (section III.3) creating a mixture of disastereomers with a range of hydrophobicities [13,156]. This effect usually causes peak broadening and a diminishment over the resolution. Therefore, IEC has certain limitation in the separation of racemic mixtures of these molecules [13], although the separation of two identical diastereoisomeric ONs containing two internal phosphorothioate linkages has been reported [160] (Figure III-7B).

Moreover, methacrylate-based (polymeric) stationary phases have also been developed and successfully applied to ON separations [143,144]. The first glycidyl methacrylate polymer based columns have been enhanced with a glycidoxyethyl methacrylate [143] (Figure II-7C). These columns proved to be more stable at higher temperatures and pH > 11 and exhibited a improved life time. This is particularly useful in the analysis of primers and iRNA ONs as single stranded and double stranded ONs can be analyzed by just establishing or not denaturing conditions (high temperature and/or pH>10.5). Additionally, it has been also demonstrated that the orientation of the bases in the ON and their basicity influence ion-exchange retention, and that certain ONs with the same length and differing in base composition were able to be resolved [144,151]. This due to the fact that the negatively charged phosphate groups are attracted to the stationary phase, while the positively
charged bases are repelled. At pH values above 10, the deprotonation of thymine and guanine will contribute to the overall negative charge of the ON, while secondary conformational changes caused by Watson-Crick interactions also influence the retention. Therefore the capability to selectively modify the retention with the aid of the pH, proves its value for the separation of mixtures of ONs.

Finally, the introduction of a novel monolithic stationary phase with porous ion-exchange “nanobeads” has been carried out for ON purification purposes [153,158]. This stationary phase employs porous ion-exchange “nanobeads” that are attached to a monolithic polymer column. Under this approach the pore size was optimized to allow for mass transfer through convection rather than diffusion to improve the column efficiency.

**III.5.2.2 Ion-pair chromatography for the analysis of oligonucleotides**

Ion-pair chromatography (IPC) (section II.2.2) has become one of the most versatile LC techniques for the analysis and purification of ONs, as quite robust separations can be attained regardless of the matrix effects, in addition to a more friendly mobile phase with mass spectrometry detectors. Briefly, the negative charge of the ON interacts with a positive charge of the ion-pair reagent, allowing for a
chromatographic separation of these analytes. The number of available charges in the ON as well as the secondary structure that it presents, determines their interaction with the IP reagent. Therefore, the separation occurs not only based on the length of the ON but also on its nucleotide composition. It has been observed that for DNA ONs, the hydrophobicity of the bases follows the order C < G < A < T [161-163]. This is particularly useful for quality control purposes, as with this technique, the separation of impurities derived from changes in the ON sequence such as depurinations and other chemical modifications on the bases can be attained. Since the first successful separation of ONs in the later 70s [164], the main research goals have focused on the ion-pairing reagent, allowing the hyphenation of this technique with MS, but also permitting enough selective separations based on the ON length and base composition.

Triethylammonium acetate (TEAA) in combination with acetonitrile as organic modifier has become one of the most common mobile phases for the separation of ONs in IPC [13,165-169]. This buffer has been employed in the separation of single stranded ONs from 2-30 mer [161,164,170] till 450-mer [171] and also double stranded ONs [163,172,173] (Figure III-8A). Furthermore, its applicability has also been extended to RNA ONs with a particular application to the analysis micro iRNAs [174,175]. Moreover, retention prediction models of heterooligonucleotides using this ion-pairing reagent have been broadly investigated [161,162,164,171]. Triethylammonium bicarbonate (TEAB) has also been introduced due to its higher compatibility with MS [173,176,177]. When compared to TEAA, improvement on the separation of ONs < 13-mer has been observed; however, no changes were detected for larger ONs. Butyltrimethylammonium (BDMA), a less polar IP reagent when compared to TEAA and TEAB has been employed with no significant improvement over the separation of ONs. Nevertheless, it allows for higher concentrations of organic modifier (acetonitrile), thus improving the MS detection [178,179]. Other IP reagents such as diisopropylammonium acetate [180,181] and Hexylammonium acetate [182] have also been investigated; however, non-substantial improvements were obtained when compared to other IPC reagents. The combination of hexafluoroisopropanol (HFIP) and TEA demonstrated a tremendous improvement particularly in the MS detection of ONs [183-185] (Figure III-8B). In this buffer system, methanol needs to be used as organic modifier as HFIP is insoluble in acetonitrile; moreover, the solubility of TEA is also reduced in the presence of HFIP, consequently, TEA ions tend to stick more to the stationary phase allowing a more stable layer of the IP reagent. These allows for a predominantly ionic interaction mechanism [171], which improves the separation of phosphorothioate ONs where the peak broadening caused
by the different hydrophobicities of the phosphorothioate enantiomers is significantly diminished [186-188].

In the separation of ONs and other relatively large molecules, the mass transfer (section II.3.4.1) effect becomes one of the major contributors to the peak broadening on porous C18 stationary phases [171,189]. This limitation has been diminished by the usage of smaller particle sizes (Figure III-8C), as the diffusion path is shortened [163,190,191]. Furthermore, the usage of core-shell C18 particles has significantly improved the separation of ONs [141,192,193]. Finally, pellicular or monolithic poly(styrene-divinylbenzene) (PS-DVB) phases have also been used for IPC of ONs, demonstrating an outstanding sensitivity in UV and MS detection with the usage of capillary columns of these phases [138,161,172,176,194-196].

![Figure III-8](image)

**Figure III-8.** A) IPC analysis of a double stranded DNA fragments from a plasmid digest [163]; B) IPC of a poly T ladder using HPIF/TEA buffer [183]; C) IPC-ESI-MS analysis of a 21-mer ON performed on a C18 column of 1.7 µm core shell particles [191].

### III.5.2.3 Hydrophilic interaction liquid chromatography for the analysis of Oligonucleotides

Hydrophilic interaction liquid chromatography (HILIC) (section II.2.3) has become a popular chromatographic mode for the separation of polar hydrophilic and ionic compounds. Moreover, the high organic content of the mobile phase makes this technique perfectly suitable with mass spectrometry detectors.

The simplest and first stationary phase used in HILIC separations consisted of bare silica [197]. In a typical separation, a high content of an organic solvent such as acetonitrile is employed while the strong eluting solvent is water. The substitution of water by solvents such as methanol (MeOH),
Oligonucleotides: characteristics and contemporary analytical strategies

ethanol (EtOH) and ethylene glycol (EtOH$_2$) has also been reported. The eutrophic strength of these protic modifiers has been established as EtOH < MeOH < Et(OH)$_2$ < H$_2$O [197,198]. HILIC separations have been able to successfully resolve nucleotides and other small related molecules; however, its applicability for the separation of therapeutic ONs is limited as only a base resolution can be achieved for ONs < 10-mer [199].

The low selectivity and retentions shown by HILIC columns made of bare silica have been improved by the introduction of polar functional groups such as amide, amino, diol, cyano, zwitterionic and other charged and non-charged groups [197]. One of the simplest modifications consists of introducing alkyl chains with hydroxyl functional groups. The separation mechanism is based on the same principle as for bare silica columns (partition equilibrium), although the presence of hydroxyl groups allows increased interactions via hydrogen bridges, leading to an increase in selectivity. The development of a hydroxymethyl methacrylate based monolithic column using HILIC has been reported for the separation of three ONs 15-mer, 19-mer and 20-mer [200] (Figure III-9A).

![Figure III-9. A) HILIC of a 15-mer, 19-mer and 20-mer ONs on a hydroxymethyl methacrylate based monolithic column using HILIC conditions [200]; B) HILIC separation of a T15-T30 ON ladder on a PEEK ZIC-HILIC column, the effect of the increase in concentration of ammonium acetate in the mobile phase is represented [201].](image)

Another common modification consists in the derivatization with amine and amide terminal groups. In the same way as in the stationary phases modified by hydroxylation, the separation mechanism
involve a partition equilibrium with the additional hydrophilic interactions, which lead to an increased selectivity. Furthermore, these stationary phases allow interactions by hydrogen bonding as donor and acceptor [198]. Additionally, an ionic interaction mechanism may take place depending on the pH of the mobile phase, as the amine groups may be positively charged. When compared to the hydroxyl columns, these amino base columns have demonstrated an improved separation of ONs, although a base resolution between ONs > 15-mer has not been attained yet [202]. These observations lead to the development of HILIC charged stationary phases. In the simplest columns, the silica has been modified with different alkyl chains containing residues with positive and/or negative charged groups. The most widely studied charged stationary phase in the separation of polar compounds is ZIC-HILIC. This zwitterionic stationary phase consists of quaternary ammonium groups and sulphonic groups [197,198]. The use of polymer based zwitterionic columns (ZIC-pHILIC) has also been introduced as this allows the usage of mobile phases with pH values over the operational pH values of silica based columns. A ZIC-HILIC column made of from poly(etherether ketone) (PEEK) has been applied for the ON analysis using mass spectrometric detection [201] (Figure III-9B); nevertheless, the separation performance when compared to ion-pair chromatography and ion-exchange chromatography, has not demonstrated a significant enhancement.

**III.5.2.4 Mixed mode, affinity and bidimensional separation methods for the analysis of oligonucleotides**

The successful separation methodologies of ONs in IPC, IEC and HILIC do not always follow exclusively one mechanism of separation. Under those approaches somewhat complementary separation mechanisms are being exploited, especially between IPC and IEC. On the other hand, HILIC shares the IEC mechanism in its newly implemented ZIC-HILIC columns. Based on these observations, mixed-mode chromatographic columns possessing both reversed-phase and ion-exchange properties have been developed for the separation of ONs [166]. In mixed-mode columns, the ONs are subjected to ionic and hydrophobic interactions simultaneously, and the predominance of certain mode can be established by choosing the adequate mobile phase conditions. Few years ago, a comparative study of different commercialized mixed-mode columns (Scherzo C18, Imtakt USA) has demonstrated the potential use of these stationary phases for the separation of ONs [203].
Oligonucleotides: characteristics and contemporary analytical strategies

Those columns, depict both hydrophobic retention obtained via C18 groups and ionic interactions from ionizable groups which have also be anchored to the same supporting material. These stationary phases proved to be adequate for the separation of typical ON impurities such as sequence deletions n-x. Additionally, the separation of isomeric ONs in which one single base was inversed with a neighboring base proved possible with such mixed mode columns, while the corresponding separation proved more challenging via pure IPC and IEC (Figure III-10A). Recently a surface bonded N-11-undecenyl-3-aminoquinuclidine based on 5 µm thiol-silica column was developed, proving adequate results for the separation of isomeric ONs [204]. The future trends in the development of these stationary phases include the establishment of functional groups that will allow the use of more MS compatible conditions.

The selectivity of the stationary phases for the analysis of ONs has also been modified by the attachment of nucleic bases. A limited number of reports have described the affinity LC variant, in which adenine or thymidine has been immobilized on silica supports [205,206]. One support employing a propylamidopropyl spacer group between the thymine and the silica support has resulted in a mixed mode phase capable to separate ONs arguably based on affinity, weak anion-exchange and reverse phase mechanisms [206]. Although this stationary phase allowed for MS compatible conditions, excessive peak broadening for certain ONs and a diminished resolution between ONs >15-mer have been observed (Figure III-10B).

![Figure III-10](image)

**Figure III-10.** A) Separation of positional isomers of a 21-mer ON using a mixed mode IPC-IEC column [203]; B) separation of A12-A18 ONs on thymidine immobilized column [206].
Despite the progresses in column technology and in the applied methodologies for the separation of ONs under the different modes covered in this chapter, the resolution of a complex mixture of ONs under the unidimensional approach remains extremely challenging. Therefore, the exploration into two dimensional separation techniques has also been carried out to improve peak capacity. Considerable progress has been made in the last few years in the development of comprehensive multidimensional LC methodologies for various applications [207-210], and few years ago a comprehensive HILIC x IPC method for ONs has been reported [199]. This methodology was able to successfully separate a mixture of 27 ONs up to 10 bases in length (Figure III-11) nevertheless, most of the therapeutic active ONs are on average double that length, limiting the applicability of this platform for therapeutic ONs.

Figure III-11. Comprehensive HILIC x IPC separation of mixture of 27 ONs (< 10-mer) [199].
Consequently, the development of alternative improved separation platforms for the analysis and characterization of complex mixtures of ONs is carried out in the framework of this thesis as will be covered in detail in the following chapters.

III.6 References

Oligonucleotides: characteristics and contemporary analytical strategies

Oligonucleotides: characteristics and contemporary analytical strategies


Oligonucleotides: characteristics and contemporary analytical strategies


Combining liquid chromatography with multiplexed capillary gel electrophoresis for offline comprehensive analysis of complex oligonucleotide samples

Summary

Two offline, two dimensional liquid chromatography (LC) × capillary gel electrophoresis (CGE) and LC × (24) multiplexed-CGE, methodologies were developed for the separation of oligonucleotides of therapeutic size. Both ion-pair chromatography (IPC) and ion-exchange chromatography (IEC) were studied as methods for the separation in the first dimension and single and multiplexed capillary electrophoresis methods in entangled polymer solutions were used for the second dimension analysis. Both, electrokinetic and pressure injection were evaluated for the investigation of the collected LC fractions in the CE capillary. The comprehensive separation was optimized with standard mixtures of poly adenosine, thymidine, cytosine and uracil homodeoxyoligonucleotides up to 35 bases long. Highly orthogonal methodologies and overall peak capacities of 6435 and 6993 for IPC × CGE and IEC × CGE, respectively, were obtained within a few hours analysis time.
Combining liquid chromatography with multiplexed capillary gel electrophoresis for offline comprehensive analysis of complex oligonucleotide samples

Published as:


IV.1 Introduction

Synthetic oligonucleotides (ONs) typically up to 35 bases in length are increasingly being used in a variety of biological and therapeutic applications including DNA sequencing, genotyping, as primers for polymerase chain reaction and in clinical diagnosis. As quite a few therapeutically active oligonucleotides are today in the clinical trial stage [1], there is an increasing need for improved separation protocols, especially for e.g. drug purity assessment testing. Because of the nature of ON synthesis, the synthetic impurities often include related families of fail (N-x) and adduct (N+x) sequences next to a large number of possible degradation products that can be generated under stress conditions. Moreover, the development of many additional ONs impurities is also experienced during chemical and enzymatic digestions in the framework of pharmacokinetic and pharmacodynamic studies [2,3] and the use of combinatorial therapies can also increase the sample complexity. Because of the rapidly increasing number of structural and chemical variations in ONs of certain chain length, conventional 1-D based separations are therefore insufficient for adequate mapping of all impurities and degradants which can be generated. Therefore, the development of 2D separation methods capable to resolve this increasing sample complexity is imposing itself. Up to now, high performance liquid chromatography (HPLC) and capillary electrophoresis (CE) have primarily been used independently as separation techniques for resolving active ON strands from the impurities. Ion-exchange (IE) and Ion-pair mechanisms (IP) have mostly been used for the separation of ONs by liquid chromatography approaches [1]. The former ion-exchange chromatography (IEC) approach has thereby been favored as, when compared to the other applicable LC techniques, somewhat higher peak capacities can thereby be obtained while involving relatively simple sample preparation procedures [1,4,5]. Strong over weak ion exchangers are thereby preferred as they allow
selectivity control over a broader pH range [1]. The IEC separation is essentially governed by the number of charges of the ONs at a particular pH. Those charges are mainly established by the phosphate backbone and are proportional to the length of the ON. However, the secondary structure and the hydrophobic interactions of the bases with the stationary phase can significantly alter the eulotropic behavior of the ONs [4,6,7]. Also in IEC, the selectivity can be tuned with the pH, the choice of counter anions and organic modifiers, and by temperature control [6,8]. For the separation of samples containing multiple and diverse ONs sequences, the suppression of intramolecular and intermolecular interactions is essential for obtaining reproducible results. Elevated temperatures, high pH (> 11.0) and chaotropic agents are used for this purpose [5,9,10]. IEC has demonstrated capabilities for the separation of single nucleotide addition and deletion sequences. However, the resolution tends to decrease for the separation of ONs larger than 30 bases. Also due to the high ionic strength of non-volatile salts in the mobile phase, coupling IEC to mass spectrometry (MS) is considered too challenging for routine implementation.

The alternative LC approach for ONs analysis is ion-pair chromatography (IPC). This technique uses a mobile phase typically containing a triethylammonium acetate (TEAA) buffer and acetonitrile as organic modifier [11]. The number of available charges and the secondary structure of the ONs govern the interaction with the ion-pair reagent allowing hydrophobic retention based on the length of the ONs. Next to the phosphate-amine interaction, the hydrophobicity of the individual bases thereby also affects the separation. TEAA when compared to other ion-pair reagents, allows for excellent separations, is inexpensive, volatile, and allows coupling to MS [11-17]. However, when compared to IEC, lower peak capacities are typically obtained. The loss in resolution with the increase in the length of the ONs is comparable to what is observed in IEC.

Next to the LC approaches, since the emergence of high efficiency capillary variants of electrophoresis, these electrodriven methods have been increasingly used for the separation of complex ONs mixtures. As such, capillary zone electrophoresis (CZE) cannot be used for the separation of ONs larger than 10 bases long because of the almost identical mass to charge ratio observed for ONs of further increasing chain length. Capillary gel electrophoresis, on the other hand, has demonstrated the capability to separate ONs differing in length by one nucleotide from 10 to 600 bases long involving a separation by size mechanism [18-24]. The cross-linked polyacrylamide gel capillaries which were initially introduced for this purpose depict some drawbacks, such as limited gel life time and irreproducibility of the polymerization in the capillary. The usage of entangled polymer solutions as sieving media has therefore become more popular as it overcomes those
Combining liquid chromatography with multiplexed capillary gel electrophoresis for offline comprehensive analysis of complex oligonucleotide samples

limitations [8,20,22,23,25,26]. The polymer solution can thereby be replaced after each analysis, avoiding sample material remaining in the capillary; additionally, on-capillary electroextraction and concentration [27,28] and hydrodynamic injection becomes possible with this approach. Nevertheless, a separation by nucleotide composition of ONs of identical length, only through sieving polymers cannot be performed. As the overall complexity of many ONs samples exceeds the individual peak capacity of LC and CGE, this problem can be solved by the combination of orthogonal techniques. In this way a combined methodology with a significantly increased resolving power can be obtained.

The first implementations of 2D separations, mainly for biomolecules, involved 2D slab gel electrophoresis. This methodology has been successfully applied for the separation of peptides and proteins involving a separation by size and isoelectric point [29]. The benefits of these methodologies for the separation of ONs are therefore limited, as only a sieving mechanism can effectively separate these molecules. Among the 2D methodologies for the analysis of biomolecules, until now, only a comprehensive LC x LC method using hydrophilic interaction liquid chromatography x reversed phase liquid chromatography (HILIC x IPC) for ONs up to 10 bases long has been developed [30]. Thus far, the use of on-line and offline LC x CGE methodologies for ONs has not been reported. Some methodologies for hyphenation of LC with CE have been described for the analysis of peptides and proteins [29,31,32].

Coupling LC to CE involves a number of additional interfacing challenges compared to LC x LC, for which the solvent incompatibility, the excessive band broadening due to column and interface configuration, and a faster second dimension in comprehensive separations have been described. In LC x CE methodologies the difference in peak volume of LC fraction and the injection volume in CE, the ionic strength of the LC mobile phase and the way the electrode is used at the coupling side of CE need to be considered [29,31-33]. For those technical reasons, CZE is the most viable technique for performing an on-line LC x CE separation. On the contrary, CGE in entangled polymer solution is challenging due to the fairly long capillary regeneration time required in an on-line setup. The appearance of multiplexed capillary electrophoresis systems [34-36] offers an alternative solution to this problem by allowing simultaneous fast off-line analysis of all or of a segment of the eluting fractions from the LC column. Therefore in this study, offline IEC x CE, IPC x CE and IPC x multiplexed CE (IPC x 24CE) methodologies were developed and evaluated for the separation of
poly-adenosine (A), thymidine (T), cytosine (C) and uracil (U) homodeoxyoligonucleotide ladders up to 35 bases long.

### IV.2 Experimental

#### IV.2.1 Chemicals

Triethylamine (TEA), acetic acid, sodium chloride, sodium hydroxide, Tris, EDTA, adenosine monophosphate (AMP), hydrochloric acid and (Hydroxypropyl)methyl cellulose (HPMC) viscosity 40-60 cP, 2% in H₂O (20°C) were obtained from Sigma-Aldrich (Steinheim, Germany). Tricine was supplied by Jansen Chimica (Beerse, Belgium). HPLC grade acetonitrile (ACN) from Fischer Scientific (Loughborough, U.K.) and Milli-Q water (Milipore, Milford, MA) were used. All the mobile phases were filtered through 0.22 μm nylon filters (Grace Davison Discovery Sciences, Lokeren, Belgium). Homo-oligonucleotides of deoxycytidine (dC), deoxythymide (dT) and deoxyadenosine (dA) bases with a length between 5-mer and 35-mer (increasing step by 5-mer) and deoxyuracil (dU) 5-mer and 15-mer were purchased from Eurogentec (Liege, Belgium). The lyophilized ONs were dissolved in an appropriate volume of 10 mM Tris–HCl, 1 mM EDTA (pH 8.0) in order to obtain 100 mM stock solutions. The final concentration of the analytes in the mixture solution was adjusted for achieving similar peak heights in the plots for visualization purposes, and their concentration ranged between 8 and 0.8 mM. During the preparation of these mixture solutions, the solvent was evaporated to dryness by a gentle N₂ flow and then redissolved in order to obtain consistent concentrations during the method development.

#### IV.2.2 Instrumentation

An Agilent 1200 series HPLC composed of a binary pump and an UV detector set at 254 nm was used as the first dimension (Agilent Technologies, Waldbronn, Germany). LC fractions were collected with a Waters Fraction Collector II (Waters Corporation, Milford, MA, USA) into 1 mL vials. For the second dimension a Beckman P/ACE 5510 single capillary (Beckman Coulter Inc., Fullerton, CA, USA) and a CePRO 9600 (CombiSep, Ames, IA, USA) multiplexed capillary electrophoresis instrument were used, controlled by the P/ACE Station version 1.21 and the CePRO Manager version 6.01 software,
Combining liquid chromatography with multiplexed capillary gel electrophoresis for offline comprehensive analysis of complex oligonucleotide samples

respectively. The detector wavelength and sampling frequency was set at 254 nm at 10 Hz and 214 nm at 8 Hz for the single and multiplexed capillary systems, respectively. The contour plots were constructed using the GC Image software version 1.9b7 (University of Nebraska, Lincoln, USA).

IV.2.3 Chromatographic conditions

IPC was performed on a 50 x 4.6 mm I.D. x 3 µm XBridge C18 column ( Waters, Zellik, Belgium). The mobile phase was composed of 100 mM TEA (pH 5.5 adjusted with acetic acid) (A) and ACN (B). A flow rate of 0.3 mL/min was used with the following gradient program: 0-40 min, 2-14%B. The column temperature was set at 60 °C. The IEC experiments were performed on a polymeric based analytical column 4 x 250 mm DNAPac PA200 protected with a 4 x 50 mm DNAPac PA200 precolumn (Thermo Scientific, Erembodegem-Aalst, Belgium). The mobile phase consisted of 20/80 ACN/water (pH 11.5 adjusted with NaOH) (A) and 20/80 ACN/water 1.25 M NaCl (pH 11.5) (B). A flow rate of 0.3 mL/min was used and the gradient profile was 0-75 min, 0-75 %B. The column temperature was set at 25 °C. In all the experiments the injection volume was maintained at 35 µL. Fractions from the first dimension were collected every 60s for subsequent CE analysis.

IV.2.4 Electrophoretic conditions

IV.2.4.1 Single capillary system

Polyimide coated fused-silica tubing (Polymicro Technologies, Phoenix, AZ) of 100 µm I.D. (360 µm O.D.), 26 cm of total length (L) and 19.5 cm effective length (l) was used for the Beckman P/ACE 5510 system. The capillary temperature was set at 20 °C. The buffer consisted of 50 mM Tris-Tricine (pH 8.1) filtered through a 0.22 µm PVDF syringe filter. The replaceable polymer solution was prepared by dissolving an appropriate amount of HPMC in the Tris-Tricine buffer in order to obtain a 5% m/v solution. The polymer solution was filtered the same way as the buffer solution and was then placed in an ultrasonic bath for 60 min in order to eliminate air bubbles. The capillary was conditioned by rinsing it (20 psi) for 2, 4 and 2 min with NaOH 0.1M, water and Tris-Tricine buffer, respectively. The polymer solution was subsequently pressurized through the capillary (20 psi) for 20 minutes, then the inlet vial containing HPMC was exchanged with a vial containing water and pressure (20 psi) was applied for 0.2 min. The capillary was then ready for analysis. A pressure injection was performed (20 psi) for 0.5 min and the separation was performed at -10 kV (-385 V/cm) for 12 min. During the
injection and separation process a vial containing the polymer solution was placed in the outlet position and meanwhile a vial with Tris-Tricine buffer was placed at the inlet position at the instant of the separation process. Between each run the capillary was rinsed (20 psi) for 8 and 1 minutes with water and Tris-Tricine buffer, respectively.

IV.2.4.2 Multiplexed capillary system
A 24/96 capillary array 75 µm I.D. (200 µm O.D.), L = 80 cm and l = 55 cm was used (Figure IV-1). The capillaries were washed (50 psi) for 30, 60 and 30 min with NaOH 0.1M, water and Tris-Tricine buffer, respectively. The capillaries were filled with the HPMC solution for 120 min (70 psi) using an external nitrogen pressurizing system. Before injection, the capillary array was conditioned at -15 kV (187 V/cm) for 60 minutes, followed by an electrokinetic injection at -10 kV for 60 s. The separation took place at -15 kV for 60 min.

Figure IV-1. Schematic of the multiplexed capillary electrophoresis array.
Combining liquid chromatography with multiplexed capillary gel electrophoresis for offline comprehensive analysis of complex oligonucleotide samples

IV.3 Results and discussion

IV.3.1 Offline comprehensive ON analysis by IEC x CGE

Classically, in on-line comprehensive methodologies the peak capacity of the second dimension is sacrificed in order to establish an adequate sampling frequency on the first dimension [33]. Under this approach, the maximal peak capacity that a combination of two dimensions can theoretically generate, is not reached in many cases. By contrast, in offline comprehensive methodologies, optimal peak capacities for each dimension can be approached in an easier way as no limitations due to real time sampling frequency in the first dimensions are involved. However, under those conditions and if the second dimension is performed sequentially, the analysis time is drastically extended making the approach poorly applicable. In this study, the conditions of each dimension were adjusted for achieving the highest peak capacity, and subsequently the best performing methodology was tested towards use with multiplexed CE for faster offline comprehensive analysis.

As IEC has been reported as a highly efficient technique for ON separations, this LC mode was selected first for the development of this comprehensive approach. A strong ion exchange column was used for this purpose [5]. An aqueous mobile phase of pH 11.5 containing 20% of ACN was used, whereby the ONs eluted with a gradient of increasing NaCl concentration. The optimal separation was reached with a gradient slope of 1% B/min, although a complete separation of various combined ladders was not possible under those conditions (Figure IV-2A). Due to the polymeric nature of the IEC column, the implementation of denaturing conditions as elevated temperature and pH was not a limitation [5]. The high pH was necessary, as lower pH (< 10.5) did not overcome the hydrogen bonding between the various ONs analytes, as has been well reported before [5,6]. Raising the column temperature to 80 °C, suppressed partially this intermolecular interaction. However, increasing the pH to 11.5, as such at 25 °C, demonstrated to be the most effective solution as the combination of both high pH and temperature lowers the column life time [5]. The usage of chaotropic agents was avoided, as their implementation can have detrimental effects in the operation of IEC and towards the column lifetime.
Because of the increasing applications of filled gels in CGE, in this work replaceable polymers were used for ONs separation and all the experiments were performed in untreated fused silica capillaries because the used multiplexed arrays are currently only available with this type of capillaries. In order to control the separation of ONs in CGE, the polymer concentration and viscosity are thereby the most relevant parameters to adjust. When the former is increased, a transition from a dilute to a semi-dilute regime occurs and the polymer chains become entangled, forming a transient network of obstacles. Once the entanglement occurs the mesh size is only depended on the polymer concentration. Polymers of high molecular weight are thereby preferred as they can form a more rigid network [7,8,19].

Amongst the different polymers tested for the separation of DNA, cellulose polymers assume in water a stiff and extended conformation and become entangled in a more easy way compared to flexible random coil polymers. Hence the working concentration of cellulose based polymers is therefore lower than for polyacrylamides [8,19]. A solution of 5% HPMC generated the best separation for the studied ONs mixture (Figure IV-2B) whereby at a concentration below 3.5% the resolution was lost. On the other hand no incremental benefit was gained when raising the concentrations up to 7.5%, a polymer concentration at which solution viscosity is much raised severely lengthening the capillary filling time.
Combining liquid chromatography with multiplexed capillary gel electrophoresis for offline comprehensive analysis of complex oligonucleotide samples

Figure IV-2. IEC chromatogram of the 23 ONs standard mixture (A). Electropherogram of the 58th IEC fraction (B). IEC x CGE contour plot (C). The gradient profile of the LC separation is displayed. For conditions, refer to sections IV.2.3 and IV.2.4.1

Efficient transfer of the IEC fraction into CGE was complicated by the NaCl content. As discrimination free electrokinetic injections appeared not possible due to a too high dependency on the salt content in the LC phase, pressure injection was selected for the CGE mode. However, this prohibited implementation of the used multiplexed capillary system as a second dimension. Although the used multiplexed instrumentation allowed sample loading through the application of a vacuum at the
capillary outlets, the applied vacuum of -1 psi, was insufficient for introducing a detectable amount of sample in a reasonable time under the used CGE conditions. Additionally, the increasing NaCl content between the earlier and later eluting fractions in IEC leads to changes in the migration time window of the ONs and in the measured current profiles during the respective CGE separations (Figure IV-3). As the mobility of an ion is also dependent on the ionic strength of the buffering medium, the increasing (and high) ionic strength of the IEC fractions can easily lead to shifts in migration times of the ONs in CGE. Also note that higher observed currents can lead increased Joule heating and to corresponding losses in separation performance.

![Graph of current profiles](image)

**Figure IV-3.** Current profiles in the single CE system after performing pressure injection for an IPC fraction and for the 1st, 25th and 50th fraction of the IEC analysis. For conditions, refer to sections IV.2.3 and IV.2.4.1

When injecting an ON ladder mixture dissolved in a mobile phase composition corresponding to the first and last IEC fraction, a variation on the migration time window of around 20% was determined with the CGE dimension. Nevertheless, the migration time shifts between the ONs remained practically constant as their separation profile is mainly established by the sieving media. In order to correct the migration time window differences, 5 μL of an AMP solution (150 μg/mL) was added to each fraction before the injection. The AMP migration time was subsequently used to align the electropherograms with a straightforward but automated in-house developed algorithm.
Combining liquid chromatography with multiplexed capillary gel electrophoresis for offline comprehensive analysis of complex oligonucleotide samples

The much improved separation performance achieved by combining the different separation mechanisms can be appreciated on the contour plot in Figure IV-2C. The IEC method essentially separates the ONs on base composition leading to the appearance of zones in the LC chromatogram for each type of ONs ladder; however, as can also be seen in Figure IV-2C some of the ONs ladders show significant overlap. By the implementation of CGE as second dimension, an orthogonal separation mechanism that separates the ONs by their size is capable to completely resolve the ONs eluting from the first dimension. In the LC coelution region of the adenosine and cytosine ONs, the presence of additional compounds that appear, correspond to fail n-1, n-2, n-n, sequences, present as impurities in the poly-A and poly-C standards. A proper identification of those compounds would not be achievable by a one dimensional analysis. This example demonstrates some of the benefits of this comprehensive approach for the separation of complex mixtures.

Although satisfactory separations are obtained in this way, the described hyphenation problem with multiplexed CGE led to the need for investigation of the combination of IPC with CGE.

**IV.3.2 Offline comprehensive ON analysis by IPC x CGE**

Despite the somewhat lower resolution that can be achieved for the separation of ONs in IPC when compared to IEC [1], the much lower ionic strength of the mobile phase in IPC allows it to be more easily implemented as first dimension in the LC x CGE methodology. When employing a weak ion-pair agent, such as TEA, a small hydrophobic interaction of the ONs bases as such with the stationary phase is also still emerging. In the separation of ONs shorter than 10-mer, this interaction dominates the separation over the ion-pair mechanism. However the relative importance of the hydrophobic effect is decreasing when the chain length increases, up to a point when the ion-pair mechanism gains control over the separation [1]. Under a gradient of ACN and a TEAA buffer on a C18 column, the elution order of the ONs bases T > A > C > U was confirmed [12,37]. Under the selected chromatographic conditions a complete separation of the 23 ONs can also not be achieved (Figure IV-4). Also the loss in resolution between some families of ONs as consequence of hydrogen bond interactions between the ONs can be exemplified with the cytosine ONs larger than 20-mer, where no separation can be attained (Figure IV-5). The concentration of TEAA, the pH and the gradient profile, have the highest effect over the separation. In the IPC approach, concentrations higher than
100 mM of TEAA did not improve the separation and at lower concentrations the resolution is decreased [12-15]. In IPC of ONs, a pH between 3.5 and 6.5 should be maintained for an optimal separation as the ionization of the TEA and the phosphate backbone of the ONs allowing the establishment of an additional anion exchange mechanism that contributes to the separation [38]. Lowering the gradient steepness leads to a partial separation between the coeluting A10-C15, A15-T5, A20-A25 and A30-A35 ONs; however, the separation profile of the cytosine ONs was thereby not improved (Figure IV-4).

**Figure IV-4.** IPC chromatogram of the 23 ONs standard mixture (A) and the corresponding IPC x CGE contour plot (B). The gradient profile of the LC separation is displayed. For conditions, refer to sections IV.2.3 and IV.2.4
Combining liquid chromatography with multiplexed capillary gel electrophoresis for offline comprehensive analysis of complex oligonucleotide samples

Figure IV-5. IPC chromatogram of the cytosine ONs standard mixture (A) and the corresponding IPC x CGE contour plot (B). The gradient profile of the LC separation is displayed. For conditions, refer to sections IV.2.3 and IV.2.4

Both, pressure and electrokinetic injections of the fractions could be performed, and only slight differences over the separation performance were observed between both injection modes. The pressure injection generated relatively more reproducible peak areas as the differences in the composition of the fractions presented a diminished effect over the loadability of the ONs into the capillary. An electrokinetic injection up to 60s at -5kV could be performed without affecting the performance of the separation; however, the alignment procedure using AMP as internal standard was required for improving the reproducibility. As can be observed in the contour plots of Figure
IV-5B and Figure IV-4B, this methodology completely resolves the components of the mixture despite the coelution of some ONs in the first dimension (Figure IV-5A and Figure IV-4A). The possibility to perform electrokinetic injections with these fractions allowed efficient implementation of multiplexed capillary electrophoresis as second dimension in an easy way, where a 24 capillary array was employed (Figure IV-6). The usage of larger capillaries and of a lower electric field (the output voltage of the instrument is limited to ±16 kV) on the multiplexed capillary system improved the resolution, although broader peaks where thereby obtained. The higher throughput allowed much shorter fractionation periods without compromising the overall analysis time excessively. Capillary arrays up to 96 capillaries are available for this system.

![Figure IV-6. IPC x CGE contour plot for the 23 ONs mixture, determined using the multiplexed (x24) capillary electrophoresis system. For conditions, refer to sections IV.2.3 and IV.2.4.2](image)

**IV.3.3 Peak capacity of IEC x CGE and IPC x CGE approaches**

The degree of orthogonality of the LC x CGE methodologies is reflected in the low correlation factor value $r^2$ of 0.52 (Table IV-1) calculated after plotting the normalized retention data [39]. Lower values were not obtained since the separation of homologue ON in LC is also partially done by size. The peak capacities in the individual dimensions were calculated on completely resolved ONs with a symmetric
Combining liquid chromatography with multiplexed capillary gel electrophoresis for offline comprehensive analysis of complex oligonucleotide samples

Peak shape, whereby the peak width at half height was used to calculate the peak width at 4σ for the peak capacity calculation. The average free mobility ($\mu_0$) of ONs was calculated in order to determine the peak capacity in CGE. Standard mixtures of T5-T35 ONs were used, each dissolved in the respective mobile phase composition according to the gradient program at their retention time. The $\mu_0$ was calculated by constructing a plot of log $\mu$ versus %HPMC, generating a linear relationship in which the intercept is equal to $\mu_0$ [19]. The calculated theoretical peak capacities were 6435 and 6993 for IPC x CGE and IEC x CGE, respectively (Table IV-1).

Table IV-1. Summary of the peak capacities in each dimension of IPC x CE and IEC x CE together with the correlation coefficients between the individual dimensions and the repeatability after performing three consecutive contour plots. Refer to sections IV.3.3 and IV.3.4

<table>
<thead>
<tr>
<th></th>
<th>IPC x CE</th>
<th>IEC x CE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation ($r^2$)$^a$</td>
<td>0.52</td>
<td>0.521</td>
</tr>
<tr>
<td>Peak capacity first dimension$^b$</td>
<td>165</td>
<td>189</td>
</tr>
<tr>
<td>Peak capacity second dimension$^b$</td>
<td>39</td>
<td>37</td>
</tr>
<tr>
<td>Theoretical peak capacity</td>
<td>6435</td>
<td>6993</td>
</tr>
<tr>
<td>Effective peak capacity$^c$</td>
<td>852</td>
<td>1474</td>
</tr>
<tr>
<td>Average %RSD on migration time$^d$</td>
<td>0.33 (SCS), 1.14 (MCS)</td>
<td>0.56</td>
</tr>
</tbody>
</table>

$^a$ Calculated from the normalized retention factors in each dimension [39].
$^b$ Value calculated from the average 4σ peak widths of the non coeluting A, T and U oligonucleotides.
$^c$ Calculated according to Davis et al. [40]

SCS: single capillary system; MCS: multiplexed capillary system.

The effective peak capacity, a parameter established for the estimation of the peak capacity in a comprehensive separation considering the undersampling of the first dimension, which leads to loss on the primary resolution, was also calculated [40], and resulted in 852 and 1474 for IPC x CGE and IEC x CGE, respectively. This parameter has been widely studied concluding that a minimum of two or three fractions per first dimension peak should be sampled [40-43]. Nevertheless, the implications on the analysis time in this LC x CGE methodology when employing a single capillary system becomes a parameter of concern as challenging to realize. The greater peak capacity obtained in the IEC x CGE
methodology was attributed to the higher efficiency of the IEC column; additionally, this LC methodology has also demonstrated more homogeneous peak shapes for all the studied ONs.

### IV.3.4 Figures of merit of the developed methodology

In order to compare the repeatability between IEC x CGE, IPC x CGE and IPC x 24CGE, the relative standard deviations (RSD) of the elution and migration time were determined. The highest variability was found in the CE separation as a consequence of the varying composition of the fractions and because of the nature of electrodriven techniques itself. The RSDs on the single and multiplexed capillary systems were determined using the standard mixture of thymidine ONs. The measured %RSDs without applying the alignment procedure were 3.6% (n = 16) and 8.5% (n = 52) for the single and the multiplexed capillary systems, respectively. This has repercussions for the construction of the contour plot, as no contour spots belonging to the same compound could be joined when a peak is split in two or more fractions, leading to erroneous assumptions on the number of compounds and in the peak assignments. The implementation of the alignment procedure improved the RSDs to values below 1.2% as can be observed in Table IV-1, where the %RSD after the construction of three consecutive contour plots is displayed. The variable electric field in the capillary, established after the injection of the IEC fractions, slightly changed the mobility of the ONs. This variability was reflected in the higher %RSDs observed in the IEC x CE methodology when compared to the IPC x CGE method in the single capillary system. A higher variability for the multiplexed capillary system was expected because some differences between capillaries were observed. For quantification purposes, the area response should also be corrected with an internal standard. The %RSD’s of the peak areas were 26.6% (n = 52) and 4.2% (n = 16) for the multiplexed capillary system and for the single capillary system, respectively. This higher variability on the area observed in the multiplexed capillary system was mainly derived from the variation between the capillaries.

In order to obtain a better estimation of the potential performance for e.g. purity assessments, the limit of detection (LOD) and the limit of quantification (LOQ) for a number ONs were calculated. Usually, the LOD and LOQ are defined as the concentrations giving rise to a signal corresponding to three and ten times, respectively, the intensity of the blank signal. Other approach to determine the LOD and LOQ consists in employing the following equations: \( \text{LOD} = Y_B + 3S_B \); \( \text{LOQ} = Y_B + 10S_B \); where \( Y_B \) is the blank signal and \( S_B \) the standard deviation of the blank [44-46]. In practice, \( Y_B \) and \( S_B \) were determined from a calibration line adjusted with the least-square method, in which the statistic \( S_y/x \).
Combining liquid chromatography with multiplexed capillary gel electrophoresis for offline comprehensive analysis of complex oligonucleotide samples

which estimates the random errors in the Y direction was used for calculating $S_{b}$, and the intercept was used as an estimation of $Y_{B}$ [44]. Individual injections of the selected ONs dissolved in the corresponding mobile phase were performed in CGE for this calculation. The lowest LOD (0.01 µg/mL) was obtained with electrokinetic injection of IPC fractions in the single capillary system (Table IV-2). This injection mode, when compared to the pressure injection presented average LODs and LOQs which were about 15 times lower, exhibiting the capabilities of in-capillary concentration of ONs when long injection periods combined with simple stacking procedures are performed. The effect of strong ionic mobile phases over the sensitivity in CGE is also reflected in the higher LOD and LOQ values of the IEC x CGE method when compared to the IPC x CGE method, both employing the pressure injection mode. These values are similar to the ones obtained with the multiplexed capillary system, in which its lower sensitivity is mainly a consequence of the instrumental configuration. Depending on the selected LC technique in the first dimension and when low concentrations of ONs are expected in the given samples, an adequate sample concentration procedure should also be considered. Note that significant improvements are conceivable upon evaporation and concentration of the collected LC fractions.

<table>
<thead>
<tr>
<th>Compounds</th>
<th>IPC x CGE$^a$</th>
<th>IEC x 24CGE$^b$</th>
<th>IEC x CGE$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pressure injection</td>
<td>Electrokinetic injection</td>
<td>Pressure injection</td>
</tr>
<tr>
<td></td>
<td>LOD</td>
<td>LOQ</td>
<td>LOD</td>
</tr>
<tr>
<td>T5</td>
<td>0.07</td>
<td>0.21</td>
<td>0.01</td>
</tr>
<tr>
<td>T15</td>
<td>0.10</td>
<td>0.30</td>
<td>0.01</td>
</tr>
<tr>
<td>T25</td>
<td>0.20</td>
<td>0.79</td>
<td>0.01</td>
</tr>
<tr>
<td>T35</td>
<td>0.23</td>
<td>0.80</td>
<td>0.01</td>
</tr>
</tbody>
</table>

The values are given in µg/mL
IV.4 Conclusions

Two new offline two-dimensional comprehensive LC x CGE and LC x multiplexed CGE platforms for the analysis of ONs were presented. The combination of IPC and IEC with CGE in entangled polymer solutions generated orthogonal methodologies with increased peak capacities, capable to separate ONs based on their size and nucleotide composition. These methodologies show promising results for their implementation in the analysis of complex biological matrixes involving the presence of therapeutic ONs and double and single DNA strands. Multiplexed capillary electrophoresis has proven to be an applicable and high throughput technique to be used as second dimension in LC x CGE analysis.

IV.5 References

Combining liquid chromatography with multiplexed capillary gel electrophoresis for offline comprehensive analysis of complex oligonucleotide samples

Possibilities and limitations of spectral deconvolution in photodiode array detection separations of oligonucleotides

Summary

Liquid chromatography (LC) and capillary gel electrophoresis (CGE) stand out among the most widely implemented techniques for the analysis of oligonucleotides. In both approaches, the usage of a photodiode array detection (PDA) system is commonly preferred as the UV spectra can be simultaneously recorded. Although both HPLC and CE depict satisfactory separation capabilities for the analysis of a variety of ONs mixtures, in many cases the separation of ONs by these techniques or through their combination is incomplete. The chemometrical deconvolution method discussed in this chapter provides a fast additional in silico tool for the separation of partially overlapping ON signals that present certain UV spectral differences, and it might be easily implemented as an additional feature in LC-PDA, CGE-PDA and LCxCGE-PDA ON analyses. The in-house-developed MATLAB script differentiates between the ONs based on their dissimilarity in the UV-spectrum and retention time, using information from the retention time and UV spectral analysis. Consequently, the script generates individual chromatograms and UV-spectra for each of the compounds. In this work the possibilities and limitations of the deconvolution script in the separation of partially overlapping ONs were explored. Homooligonucleotides and ONs with random base composition were used to test this computational tool.
Possibilities and limitations of spectral deconvolution in photodiode array detection separations of oligonucleotides

V.1 Introduction

The separation methodologies for the characterization of oligonucleotides (ONs) have evolved from the implementation of one dimensional approaches (Chapter II) till the development of comprehensive bidimensional separation methods capable to resolve complex mixtures based on their size and base composition (Chapter III) [1,2]. The separation from base to base of ON mixtures using one dimensional methods is often rather difficult to obtain. Moreover, even with the implementation of multidimensional approaches, a complete baseline separation might not be attained for all the ONs present in complex mixtures [2]. The need for further characterization tools of complex mixtures has oriented the research into post-column or post-capillary separation strategies that are not necessarily chromatographically or electrophoretically based. Such approaches have been conceived by using detection systems such as mass spectrometers and photodiode array detectors (PDA) mainly, which in combination with chromatographic methods are capable to generate a bidimensional set of data. In a typical chromatographic analysis using either of those detectors, the first dimensional data set of a chromatographic peak is correspondent to the elution time versus the signal intensity, and for each time point, a second dimensional data is obtained based on the measured mass to charge ratios (m/z) or the wavelengths with a given signal intensity in mass spectrometry or spectroscopy, respectively.

The analysis of ONs using mass spectrometry has extensively been described by using direct infusion electrospray ionization (ESI-MS) [3], hyphenated to liquid chromatography (LC-ESI-MS) [4-12] or by employing matrix-assisted laser desorption/ionization (MALDI-MS) [10,13-18]. Several of those techniques may permit performing a mass spectrometric signal deconvolution of coeluting compounds if particular ions are generated for each of the coeluting compounds, also excluding in this case the effects of the ionization suppression that might occur in mixtures of solutes. Direct infusion ESI-MS has been rather used for the characterization of pure ONs; on the other hand, the usage of offline MALDI-MS after the LC separation or the on-line hyphenation of LC with ESI-MS has proved to be possible for the separation of mixtures of ONs. Although the applicability, sensitivity and the resolution of mass spectrometers have greatly improved in the last decades, the mass spectrometric deconvolution of overlapping oligonucleotide signals is not a straightforward approach.
with broad applicability, and is rather time consuming as the structural similarities between ONs and their polyanionic nature generate rich MS spectra consisting of multiple m/z ions. LC-ESI-MS methodologies have been limited to ion-pair chromatography (IPC) [13-18] and hydrophilic interaction liquid chromatography (HILIC) [10]. Though HILIC presents better compatibility than IPC with ESI-MS detection, the lower peak capacities of this methodology discard this technique as first choice for ON analysis. IPC on the other hand, allows for an improved separation; nevertheless, the signal suppression is higher, which results in a diminished sensitivity of the methodology. Moreover, due to the polyanionic character of ONs, the formation of adducts with alkaline cations predominantly is a common problem that complicates MS analysis and deconvolution of overlapping signals. Multiple efforts have been described for suppressing cation adduction [4,5,7,9,13,17]; however, the strategies complicate the applicability of the methodologies for their routine implementation. Furthermore, the fractionation of the eluent for offline MALDI-MS analysis in addition to the cation adduction suppression steps, causes significant loss of data and the discrete nature of MS-spectra hinders full deconvolution of disperse peaks [15,19]. Finally, it is relevant to mention that the hyphenation of MS with ion exchange chromatography (IEC) and capillary gel electrophoresis (CGE) is challenging, therefore impeding the implementation of MS deconvolution of ONs.

The spectroscopic detection systems that generate continuous spectra such as in UV or fluorescence detectors offer possibilities which are not available with spectrometric detections such as MS. These aspects are mainly the easier implementation in laboratories (infrastructure and budged wise) and higher robustness toward changes in mobile phase composition and in flow rates. Additionally UV detectors are fully compatible with IEC mobile phases, which would present an additional challenge if coupled to MS due to the elevated concentration of non-volatile salts which are thereby used. Hence, the study of the potential of the development of improved deconvolution tools for continuous detections such as PDA is justified by the challenges encountered in mass spectrometry and by the ease of operation, the high robustness, the limited cost and by the universal compatibility of UV detection with all fluid based separation techniques. This study has been limited to address those main problems for which no satisfactory solution can be found via the straightforward choice of a selective detection wavelength. The studied situation often arises when several species overlap which are also present in varying concentrations.
Possibilities and limitations of spectral deconvolution in photodiode array detection separations of oligonucleotides

The usage of mathematical methods to process analytical data has rapidly increased in the last decades due to the advances in hardware and computational software tools. As already mentioned, PDA detectors when coupled to LC generate a bilinear or convoluted data matrix. As can be observed in Figure V-1, the data matrix obtained from a typical IPC-PDA, IEC-PDA or CGE-PDA analysis, consists of a retention time dimension (chromatographic/electrophoretic profile) and a wavelength dimension (spectroscopic profile). In contrast to data acquisition, deconvolution is mathematical process that generates both a pure chromatographic and spectroscopic profile from the bilinear data matrix \( D \), by the application of certain constrains during the optimization process. In this study the multivariate curve resolution-alternating least squares algorithm (MCR-ALS) was employed (Eq. V-1).

\[
D = C_{nc,n} \cdot S^T_{ns,n} + E_{nc,ns}
\]

where \( D \) represents the convoluted data matrix; \( C \) and \( S \) represent the pure chromatographic and spectroscopic profiles, respectively. \( E \) is the residual matrix and the integral numbers of \( nc \) and \( ns \) are the amount of data points in both dimensions, respectively. Finally \( n \) stands for the number of individual signals in the convolution, which corresponds to the number of ONs in the sample.

The MCR-ALS algorithm was first proposed in 1995 [20] and since then it has been increasingly applied to process bilinear data generated by LC-PDA [21-25], multidimensional techniques [26] and for monitoring chemical processes [27]. As depicted in the algorithm workflow in Figure V-2, MCR-ALS begins with a single value decomposition (SVD) or principal component analysis (PCA) to estimate the number of compounds in the sample data matrix \( D \). In our case, as this information is already known before the data processing, this process is skipped in the present approach. Afterwards, a threshold is set for the cumulative percentage of the total variance. Next, the initial chromatograms and spectra for the whole set of compounds are estimated by detecting the signal purity along the chromatographic profile. Subsequently, the main process of the ALS algorithm takes place, which is executed by iterations that minimize the least squares of the residual signals in both chromatographic and spectrometric profiles.
Based on the understanding of chemical and physical properties of the analytes and instrumental technique, certain constraints are implemented in the ALS process, for instance non negativity constrains for spectral absorbance and unimodality constrains for chromatographic peaks. Once a preset number of iterations is performed or when the optimization reaches convergence, the process stops generating the optimal pure chromatograms ($C$) and spectra ($S$) for all the defined compounds as well as the estimated residual matrix $E$ [29]. These basic deconvolution principles have been implemented in a variety of applications such as in the deconvolution of spectroscopic methods including chiral and size exclusion separations [19,25,30], isotopic patterns [31] and image restoration [32]. So far, research on the MRC-ALS approach for the PDA spectral deconvolution of ONs has not been explored yet. In this study the possibilities of the implementation of the MCR-ALS algorithm for the deconvolution of ON overlapping signals were investigated. This computational tool provides a possible extra in silico separation feature that can be easily added to extend the separation power of LC, CGE and LCxCGE methodologies for the characterization of complex mixtures of ONs.
Possibilities and limitations of spectral deconvolution in photodiode array detection separations of oligonucleotides

Figure V-2. Work flow of the MCR-ALS deconvolution algorithm.

V.2 Experimental

V.2.1 Chemicals

Acetic acid and trimethylamine (TEA) were purchased from Sigma-Aldrich (Steinheim, Germany). Acetonitrile (ACN) HPLC grade from Fischer Scientific (Loughborough, U.K.) and Milli-Q water (Millipore, Milford, MA) were used. All the mobile phases were filtered through 0.22 µm nylon filters (Grace Davison Discovery Sciences, Lokeren, Belgium). Deoxyoligonucleotides differing in length and base composition (the specific sequences are indicated the results and discussion section V.3 ) were purchased from Eurogentec (Liege, Belgium). The lyophilized ONs were dissolved in an appropriate volume of 10 mM Tris–HCl, 1 mM EDTA (pH 8.0) in order to obtain 100 mM stock solutions. The final concentration of the analytes in the mixture solution was adjusted for achieving similar peak heights in the plots for visualization purposes, and their concentration ranged between 8 and 0.8 mM.

V.2.2 Instrumentation and conditions

All the analyses were performed on an Agilent 1100 HPLC system, equipped with a binary pump and a PDA detector recording the spectra from 190 to 400 nm. The sampling frequency of the detector was set to 10 Hz and a slit width of 1 nm was employed. The chromatograms were acquired at 260 nm using a reference wavelength at 380 nm with a width of 20 nm. The separation was performed isothermally at 40 °C on a 50 x 4.6 mm I.D. x 3 µm XBridge C18 column (Waters, Zellik, Belgium). The
mobile phase was composed of 100 mM TEA (pH 5.5 adjusted with acetic acid) (A) and ACN (B). The flow rate was maintained at 0.5 mL/min while the gradient program was deliberately varied to modulate the separation between the different ONs mixtures, and it ranged between 0% and 16% of solvent B in a time window between 0 and 10 minutes.

### V.2.3 Data treatment

Once the chromatographic and spectroscopic data were acquired, they were exported from Chemstation (version B 04.03) in a wavelength range from 190 till 360 nm, from the beginning of the analysis time till elution of last studied peaks. The reduction of the time and wavelength window was necessary to avoid inadequate deconvolution results as already reported in a previous study [25]. The exported data files in comma separated values (CSV) format were afterwards imported in MATLAB, where the deconvolution algorithm was performed. All the chromatograms were automatically normalized (the normalization was encoded in the deconvolution script) to the deconvoluted spectrum of their corresponding component. For the successfully deconvoluted peaks, the peak area was calculated for further quantitative pre-validation purposes.

### V.3 Results and discussion

The potential of the MRC-ALS algorithm for qualitative and quantitative deconvolution purposes of complex mixtures of racemates [28] and polymers [25] has been recently demonstrated. The main objective of this work was to explore the potential of the algorithm for the deconvolution of ON signals that could not be successfully separated by LC or CGE. At this point and for the ease of this study, the overlap of signals was only explored using IPC. Nevertheless, its implementation in other HPLC modes and in CGE can be easily carried out, as the differences in the UV spectra should be minimal because the separation media should not generate absorbance at wavelengths above 230 nm. Consequently, for the accurate deconvolution of the signals, the UV signals below 230 nm were discarded. When using shorter wavelengths, poorer deconvolution and a much increased noise signals were observed. Additionally, shorter wavelengths produced UV spectra that were not
Possibilities and limitations of spectral deconvolution in photodiode array detection separations of oligonucleotides

independent of the mobile phase composition. Therefore, shorter wavelengths limit and complicate the implementation of the algorithm in broader applications.

Table V-1. Sequence and molecular weight of the studied ONs.

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>Molecular weight (g/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A20</td>
<td>5’-AAA-AAA-AAA-AAA-AAA-3’</td>
<td>4636.2</td>
</tr>
<tr>
<td>GC15</td>
<td>5’-GGG-CCC-GGG-CCC-GGG-3’</td>
<td>4636.0</td>
</tr>
<tr>
<td>T10</td>
<td>5’-TTT-TTT-TTT-TTT-3’</td>
<td>4501.0</td>
</tr>
<tr>
<td>U15</td>
<td>5’-UUU-UUU-UUU-UUU-UUU-3’</td>
<td>4290.6</td>
</tr>
<tr>
<td>AC15</td>
<td>5’-AAA-CCC-AAA-CCC-AAA-3’</td>
<td>4492.0</td>
</tr>
<tr>
<td>R41</td>
<td>5’-GTT-GGA-TTA-AAC-AAC-CGT-TCC-CGT-CTC-TAT-CAG-CTT-AGT-GT-3’</td>
<td>12541.2</td>
</tr>
</tbody>
</table>

Table V-1 indicates the base composition of the studied ONs. The preferred wavelength for analysis of ONs is around 260 nm as these molecules will generate the highest intensity signal independently of the composition of the common separation media used in the mentioned instrumental techniques. As a first example, a mixture of two ONs with significant differences in base composition (purine and pyrimidine nucleotides) was explored. Figure V-3A depicts the IPC chromatogram of a mixture of an A20 and a GC15 ONs at two different wavelengths. As can be observed, the ON signals are not fully resolved and only differences on the signal intensity of each analyte can be distinguished. The UV spectra of each of those pure ONs is displayed in Figure V-3B, were it can be observed that both molecules present their maximal absorptivity around the region of 260 nm, and by selecting different wavelengths such as the one at 235 nm only has an effect on the relative peak intensities but not on the separation of the signals. It also must be noticed that the UV spectra are not identical, which presents an advantage as the deconvolution of the signals can therefore be performed smoothly as depicted in Figure V-3C.
Figure V-3. A) Chromatogram of an A20 and a GC15 ON mixture recorded at 260 and 235 nm; B) UV spectra of the GC15 and A20 ON; C) Overlay of the deconvoluted chromatograms for the A20 and GC15 ONs with the original chromatogram recorded at 260 nm.

Moreover, it is important to notice that in order to facilitate the deconvolution process, the operations set was performed in the time regions of interest in the chromatogram. A previous work based on SEC with PDA detection revealed that artifact peaks tend to appear in the total exclusion and permeation region [25]. In the current IPC case, similar artifacts can occur close to the void time and at the top of the gradient profile.

As the similarity between the UV spectra of the ONs increases, the success ratio for a deconvolution decreases significantly. In Figure V-4A the original and deconvoluted chromatogram of a T10 and U15 ON is depicted. As can be observed on the deconvoluted signals, shoulders corresponding to the adjacent ON are obtained for each signal. Although the U15 ON presents a clear bathochromic shift compared to the T10 ON (Figure V-4B), a considerable part of the UV spectra shows no significant difference leading to errors in the deconvolution of the peaks. In the case presented in Figure V-4, both ON sequences are totally different in base composition but the nucleotides are based on pyrimidine rings leading to similar UV spectra. For this latter case, an error in the deconvolution of around 15% in the deconvoluted peak area can be estimated due to the appearance of the shoulders.
These observations demonstrate that performing quantitative analysis in this way could lead to false signals generating the mentioned errors. Nevertheless, if effort is set for increasing the resolution between those peaks, the deconvoluted profile might present acceptable results for quantitative purposes, although it compromises the merits of the algorithm in the process.

**Figure V-4.** A) Chromatogram of a T10 and an U15 ON mixture recorded at 260 and the deconvoluted chromatograms. An estimation of error for each peak after the deconvolution in terms of peak area is indicated; B) Corresponding UV spectra of the T10 and U15 ONs.

The cases presented in Figure V-3 and above (Figure V-4), in which homooligonucleotides are tested, demonstrate positive results for the deconvolution of partially overlapping signals. Nevertheless, therapeutically active ONs present random combination of the nucleic bases, which results in similar UV spectra that do not allow much deconvolution of the signals. Figure V-5 exemplifies this case, where a 41-mer ON with a nucleotide random sequence (R41) is partially overlapped with an AC15 ON. As can be observed on the deconvoluted signals (Figure V-5A), the presence of shoulders on each peak will not allow for a proper quantification even if the peaks are somewhat much more resolved than the peaks presented in Figure V-3. Note also that the UV spectra (Figure V-5B) does not present significant differences that would allow a proper deconvolution process.
The few cases in which the deconvolution process might be successfully implemented are mainly reduced to ONs that present significant differences between their nucleotide base compositions. While emphasis on the differences in the aromatic base structure of the nucleotides (purinic/pirimidinic). Moreover, for a successful deconvolution, the resolution between the signals is also strictly depending on the similarity of the UV spectra that the signals present, requiring larger resolutions for ONs that present UV spectra with less dissimilarities. Likewise, the concentration ratio in which each ON is present with respect to each other plays also an important role over the success of this deconvolution process. Taking as a point of reference the ON mixture of A20 and GC15 (Figure V-3), where both ONs are present in similar molar concentrations (~50 mM each), the deconvolution can take place successfully. Nevertheless, this might not be the case if the ON concentrations are not equimolar. Figure V-6 shows an overlay of the UV spectra of the A20 ON maintaining a constant concentration at 50 mM while the concentration of the G15 ON was set to 50 mM, 25 mM and 5 mM, respectively. It can be noticed that when decreasing the concentration of one ON with respect to the other, the UV spectra of the most concentrated ON masks the UV spectra of the less concentrated one, impeding the spectral deconvolution of the overlapping signals.

Finally, for the validation of quantitative purposes through the usage of deconvoluted chromatograms, individual calibration curves must be constructed for each of the ONs. It is hereby of great importance to mention that the peak areas measured in the deconvoluted chromatograms are not equal to the peak areas of an identical amount of ONs measured at any wavelength in a 2D signal. The deconvoluted chromatograms are constructed and based on all available spectral and
temporal information in the data matrix. As first assumption, this data treatment could be considered unnecessary for the calibration data to isolate the signal and thereby the peak areas, as the numeric data transformation of the deconvolution script implies that deconvolution is also necessary and therefore must be performed on all the calibration data as well.

Figure V-6. UV Spectra overlay of an A20 ON and GC15 ON at different mM concentrations.

V.4 Conclusions

The spectral deconvolution of overlapping oligonucleotide signals with slightly dissimilar UV spectra has been demonstrated. Although the usage the MCR-ALS algorithm may be limited to the separation of homooligonucleotides with significant differences in their purinic and pirimidinic composition, it offers an additional tool for compositional analysis of ON samples. Currently its applicability for the separation of therapeutically active ONs and related impurities appears limited. Nevertheless, possible future applications may not be discarded with the new structural modifications that ONs are undergoing. The latter ones might generate molecules and impurities in which the UV spectra might slightly differ, but not to the extent in which a specific wavelength for each ON is present. In those cases, the MRC-ALS algorithm might find its place as valuable addition to the current separation methodologies. Additionally, this algorithm has been developed as an easy-to-operate method not requiring vast mathematical knowledge for its processing. Therefore, its implementation in data acquisition/data treatment software packages could be explored as it could be easily linked with the
common multiple-wavelength-PDA detection systems used in HPLC and CGE that are widely used in the analysis of ONs.

V.5 References

Possibilities and limitations of spectral deconvolution in photodiode array detection separations of oligonucleotides

Overview of enzyme immobilization approaches for on-line coupling to HPLC

Summary

In this chapter an overview of the main aspects related to the manufacturing and characterization of immobilized enzyme reactors (IMERs) for hyphenation with LC is covered to allow the reader to better understand the new proposed methodologies described in chapters VII and VIII. Emphasis is thereby set on typically used and manufacturing procedures relevant to this work and on providing an introduction in kinetic studies required for characterization of IMER performance.
Overview of enzyme immobilization approaches for on-line coupling to HPLC

VI.1 Introduction

Since the first reported enzyme immobilization almost 100 years ago, research has been carried out to continuously develop innovative modes of insoluble biocatalysts [1]. The anchoring of enzymes on supporting materials has been described for a variety of applications including bioprocessing, affinity chromatography, biotransformation and bioanalysis [2]. This concept once manufactured into column or cartridge is better described as an immobilized enzyme reactor (IMER), which presents several advantages over enzymatic reactions carried out in solution. First, the elevated cost of enzyme production and purification requires the usage of low concentrations of enzymes making processes more time consuming and less efficient. Next, the automatization of processes with soluble enzymes is somewhat difficult to achieve, as well as the recycle of the biocatalyst. The implementation of IMERs can easily overcome those disadvantages, increasing the quality of the enzymatic products and decreasing the operational costs.

In the following sections, the main aspects with respect to enzyme immobilization, the influence of the packing materials and some relevant points related to enzyme kinetics are discussed to allow better understanding of the new developments proposed in the subsequent chapters (Chapter VII and Chapter VIII).

VI.2 Enzyme immobilization

One key aspect in the manufacturing of the IMERs includes the structure of the support, since it determines the accessibility of the substrate to the active sites. An overview of the most common supporting materials will be covered in section VI.3 In this section, the most popular methods for enzyme immobilization will be discussed. The process can be classified into three groups: covalent immobilization, physical adsorption and encapsulation using sol-gel methods [1]. The selection of the appropriate support and immobilization technique is therefore a critical aspect in the performance and applicability of the IMERs.
VI.2.1 Covalent immobilization

The covalent immobilization has established itself as the most used method mainly because it avoids the leakage of the enzyme. This can greatly extend the IMERs lifetime and counterbalances the decrease in the enzymatic activity due to the alteration of the native enzyme structure. Additionally, the covalent binding also confers an extended thermal stability to the enzyme, since the strong interaction of the enzyme to the support grants rigidity to the protein structure, and consequently limits its movement under higher temperatures.

Silica based derivatized matrices are the most commonly used supporting materials in this approach. The functional groups that are usually used for the attachment include amino, epoxy, carbonyl, diol and phenolic groups, though the most widely used are the former two. The same chemistries can be applied to functionalize monolithic supports and silica beads, and basically the differences lay on the activation and immobilization procedures. The activation of silica beads may occur via a static or dynamic method, while the dynamic method is usually preferred for monoliths. The latter refers in general to the functionalization and immobilization by flowing the reactants through the monolithic support or through a cartridge packed with silica beads for instance.

The immobilization via epoxy groups consists of a single step immobilization, as after the functionalization of the silica support, the epoxy groups are ready to be reacted with nucleophilic groups present in the enzyme (e.g. amino, thiol or hydroxyl), resulting in the covalent attachment of the enzyme to the support (Figure VI-1A). At the end of the immobilization process, the unreacted epoxy groups can be deactivated by reaction with different thiol or amine compounds, thus preventing further uncontrolled reactions between the support and the enzyme that could reduce the enzyme stability. Another direct approach for the covalent immobilization includes the attachment via aldehyde groups, which also can be carried out with mild conditions, but in a shorter time (Figure VI-1B). However, the functionalization of the silica with aldehyde groups is more challenging.

The functionalization of silica with reactive carbamate groups via 1,1-carbonyldiimidazole (CDI), also allows for a direct reaction with nucleophiles present in the enzyme, resulting in the formation of an uncharged urethane-like bond stable against hydrolysis and proteolysis (Figure VI-1C). The deactivation of the imidazole-carbamate groups can be carried out using a solution of high pH, generating the original hydroxyl groups followed by the release of imidazole and CO₂.
The reaction of activated esters with amines provides the formation of stable amide bonds (Figure VI-1D). This process occurs almost quantitatively without side effects and under mild conditions. Two strategies have been applied in this kind of chemistry. The first consists in the activation of the support while the second one in the activation of the enzyme. Under the first strategy, the practicality of the activation procedure compared to other methods results very low. On the other hand, the second strategy has limited application for enzymes as usually their activity is diminished.

Figure VI-1. Enzyme immobilization via a direct reaction approach. A) epoxy; B) aldehyde; C) carbamate; and D) esters reactive groups.
Other immobilization procedures consisting of multistep approaches have been described [1,3,4], although their implementation is less practical compared to the direct immobilization chemistries. One of the most representative procedures under this approach consists in employing amino activated silica, which is subsequently reacted with glutaraldehyde (Figure VI-2). This functional crosslinking agent leads to high surface coverage while also acting as a spacer. In general, the amino-bearing ligands via aldehyde groups result in the formation of an imine bond (-C=N-) which is not very stable and undergoes hydrolysis liberating the protein [5]. As a consequence, an additional hydrogenation step is required to convert the imine bonds to stable secondary amine bonds.

![Figure VI-2. Enzyme immobilization via glutaraldehyde chemistry.](image)

Another common procedure involving amino functionalized surfaces consists in their activation via disuccinimidilsuberate (DSS) (Figure VI-3). This other bifunctional cross-linking agent contains two N-hydroxysuccinimide esters that are reactive primarily towards amines.

Finally, in order to provide flexibility to the enzyme and to minimize the effect of the solid support on the enzymatic activity, a spacer can be introduced. This moiety basically distances the enzyme from the solid phase surface and is generally recommended with bead based supports. To a certain extent, the activation via the above described functional groups such as DSS, glutaraldehyde or 3-glycidoxypropyltrimethoxysilane also can be considered as spacers [6,7]. Many procedures in relation to the introduction of spacers have been reported in the literature [1,8,9] and will not be
Enzyme immobilization approaches for on-line coupling to HPLC

covered in detail in this chapter, as this approach is not specifically useful for the subsequent chapters.

**Figure VI-3.** Enzyme immobilization via disuccinimidilsuberate chemistry.

## VI.2.2 Physical adsorption

Physical adsorption of enzymes represents one of the simplest immobilization methods. This immobilization is based on weak non-covalent binding between enzymes and the supporting material. The encountered binding forces include van der Waals force, ionic and hydrogen bond type interactions. The commonly used carriers are controlled pore glass (CPG), activated carbon, bleached earth, alumina, silica, calcium phosphate, metal oxide and other organic carriers, such as starch, albumin and natural macromolecular carriers [10].

One adsorption strategy consists in covering the silica surface with polyethylenimine (PEI) followed by enzyme immobilization via electrostatic attraction between enzymes and the PEI chains [11]. A more popular method consists of the usage of metal-ion chelated supports [12,13], which provides the possibility to regenerate the IMER via reversible enzyme immobilization. In this method, the surface is bonded to iminodiacetic acid moieties followed by the treatment with Cu²⁺ and the enzyme in solution. Consequently, desorption of the enzyme can be easily carried out using ethylenediaminetetraacetic acid (EDTA).
Moreover, under this immobilization approach, the perturbation to the native structure of the protein is minimal when compared to the covalent immobilization. As a consequence, typically higher enzymatic activities are obtained in this way. Nevertheless, the leakage of the protein from the support during its use presents a serious disadvantage. This can occur especially under mild changes in temperature, pH, and ionic strength or by the simple presence of the substrate. This phenomenon is exploited in regenerative biocatalytic systems [1,11]. However, for many enzymes and applications, the costs of regeneration are the main limitation of this immobilization strategy.

**VI.2.3 Enzyme entrapment**

Enzyme entrapment can occur by embedding it into a network of polymer in gel or into a semi-permeable membrane. In this immobilization approach, the amino acid residues of the enzyme usually do no react, therefore the original enzymatic structure is rarely changed. Because only smaller molecules can be diffused into the pore network for polymer gel, the diffusion resistance might cause changes in the kinetic behavior of the immobilized enzyme reducing its activity. Thus, the embedding method is mainly suitable for small substrates and products, limiting its application to study the activity of the enzymes on oligonucleotides as describe in the further chapters.

The entrapment of the enzyme is carried out in practice by mixing a catalyst with the gel formation ingredients in conjunction with the enzyme. Thereafter, once the gel is formed, the enzyme remains trapped in the matrix. This process can be carried on inorganic or organic polymeric supports. In the first case, the enzyme can be entrapped into a monolithic silica matrix, obtained via the polycondensation of alkoxysilanes [14]. The enzyme embedment into organic polymer matrixes has also been explored using polyacrylamide, although the toxicity of this gel limits the application of IMERs in food and pharmaceutical processes [4].

An alternative kind of embedding has been more recently introduced through the usage of immobilized artificial membrane (IAM) stationary phases. Such materials are obtained via the covalent attachment of phosphatidylcholine to aminopropyl silica via a terminal amide linkage [3,4,15,16]. As a result, the phospholipid head groups from the surface of the support and the hydrocarbon side chains produce hydrophobic cavities that extend from the charged head groups to the surface of the aminopropyl silica. This structure limits the access to the unbound amine groups
Overview of enzyme immobilization approaches for on-line coupling to HPLC

and allows the enzyme to interact with any combinations of polar head groups and hydrophobic chains in a way which is similar to the physical adsorption process.

VI.3 Immobilization supports

An ideal support must be inert, stable and resistant to mechanical strength, allowing the enzyme to maintain its ternary structure and ensuring interaction of the substrate with the active site of the enzyme. Moreover, properties such as porosity, pore size distribution and charges are also essential as they influence directly the kinetic process [4]. Various types of supporting materials have been developed for the manufacturing of IMERs, and can be summarized into silica based derivatized matrixes, monolithic chromatographic supports and immobilized artificial membrane stationary phases [3]. Although there is no universal support that might cover the needs of all applications and enzymes, many desirable characteristics are common for the immobilization of enzymes. These support materials have been primarily classified into inorganic and organic materials.

VI.3.1 Inorganic materials

Inorganic materials present some advantages over polymer organic constituents such as a higher stability and mechanical strength, large surface areas, less propensity towards microbial decomposition and a relative lower cost. The mainly used supports in the manufacturing of IMERs include controlled pore glass beads (CPG) and silica [10]. CPG has been used broadly in the last decade because their large surface area, narrow pore size distribution, mechanical strength, rigid structure and chemical inertness satisfy the requirements for on-line HPLC applications of IMERs. On the other hand, silica based materials provide a larger surface area and good mechanical properties. Additionally, they can be functionalized in many ways broadening thus the immobilization possibilities and applications.

In the last years the usage of monolithic columns in HPLC allowed obtaining increasingly efficient separations while depicting excellent column permeability as a consequence of the small skeleton sizes and due to wider overall porosities of such columns, respectively [17]. Therefore, the dimensions of an IMER, based on monoliths over the pressure drop in an on-line system does not
become a limitation. Due to the improved diffusivity of the molecules through monoliths, they have been widely employed for the development of IMERs. However, research on silica monoliths is tightly restricted by the lack of availability of monoliths having different pore and domain size distributions in addition to a poorer batch to batch reproducibility. Thus, further improvement of the IMER supports might be focused on increasing the structural uniformity and by decreasing the domain sizes.

VI.3.2 Organic materials

Organic materials include natural and synthetic supports. Natural polymers such as glucan, agarose, chitosan or fibrin polysaccharide matrices present a high biocompatibility; nevertheless, they also have weak mechanical properties. Natural polysaccharides are particularly suitable as most of them are non-toxic and present an excellent mass transfer performance. Chitosan is one of the most popular supports in this class, which has been used particularly to support enzymes and cells [4,18]; however, chitosan swells unsteadily in water in addition to its weak mechanical resistance.

Synthetic polymer materials include polystyrene, methyl methacrylate and divinyl benzene resins, among others. Their physicochemical properties and microbial corrosion resistance, allows their implementation as supports for practically any enzyme. Some relevant supports that have been used in the last decade include poly(glycidylmethacrylate-co-acrylamide-co-ethylene glycoldimethacrylate)monolith [19], Eupergit® C [20,21] and monolithic convective interaction media (CIM) Disks [20,22,23], and commercial immobilized artificial membrane (IAM) (section VI.2.3). Among this group, one versatile support includes the Convective Interaction Media (CIM), which is the trade name of a well characterized methacrylate polymer monolith. This monolith has a defined pore size distribution that allows low back pressures in flow through applications. Its versatility and simplicity of use has been proven in various applications [22], as few activation chemistries are commercially available. EDA-CIM Disk is an amine activated monolithic support obtained from the native epoxy groups with a convenient ethylenediamine spacer. EDA monolithic disks have been used for bioconversion after coupling proteins, peptides or other ligands through a cross-linking reaction with a suitable bifunctional reagent [10].

Finally, magnetic polymer microspheres have also been described [4,24,25]. This material is formed by the coverage of a kind of magnetic metal such as Fe₃O₄ with a polymer containing active
Overview of enzyme immobilization approaches for on-line coupling to HPLC

immobilization groups. Although their implementation in on-line systems with LC is limited, their applicability in batch reactions is wider as a facilitated separation and recycling of the immobilized enzyme can take place by the application of a magnetic field.

VI.4 Kinetics of immobilized enzymes

In sections VI.2 and VI.3 several supports and enzyme immobilization mechanisms have been briefly described. Their characteristics play a crucial role over the enzyme kinetics as they determine the conformational changes of the enzyme, exposure to its active sites and the interaction with the substrate.

The recognition of substrates by enzymes is represented by a dynamic process which is accompanied by conformational changes at the active site of the enzyme. After that substrate is transformed into products through a series of steps that are particular for each enzyme and define under the enzymatic mechanism. One of the simplest and most widely used models to describe the kinetic properties of an enzyme is based on the one Michaelis-Menten kinetics [26]. According to this model, the enzyme (E) associates with a substrate (S) to form an enzyme-substrate complex, which afterwards dissociates into the native enzyme and the product (P):

\[ E + S \xrightarrow{k_1/k_2} ES \xrightarrow{k_3} E + P \]

The enzyme acts as a catalyst and therefore it cannot alter the equilibrium of the chemical reaction. In the steady state, the concentrations of intermediates remain constant while the concentrations of starting compounds and products change. This occurs when the rates of breakdown and formation of the ES complex are equal. Under those conditions the enzymatic reaction constant \( K_M \), known as Michaelis–Menten constant is represented by:

\[ K_M = \frac{k_2+k_3}{k_1} \]
Under certain assumptions such as the enzyme concentration being much less than the substrate concentrations the enzymatic reaction rate $V$ is given by:

\[ V = V_{max} \frac{[S]}{[S] + K_M} \]  

The experimental determination of $K_M$ and $V_{max}$ is derived from measuring the reaction rate at different substrate concentrations. The rate of enzymatic reaction $V$ is calculated as:

\[ V = \frac{[P]}{t} = -\frac{[S]}{t} \]

Where $t$ is the reaction time. Eq. VI-3 predicts the reaction velocity versus the substrate concentration $[S]$. At very low substrate concentrations, when $[S] << K_M$, the rate becomes directly proportional to the substrate concentration, leading to:

\[ V = V_{max} \frac{[S]}{K_M} \]

On the other case, at high concentrations of $S$, when $[S] >> K_M$, the rate becomes maximal and independent of the substrate concentration ($V = V_{max}$). When $[S] = K_M$, then $V = V_{max}/2$; consequently, $K_M$ is equal to the substrate concentration at which the reaction rate is half of its maximal value, or in other words, at the rate at which half of the active sites of the enzyme are filled. Under those conditions $k_2 >> k_3$, meaning that the dissociation of the ES complex to $E$ and $S$ is much more fast than the release of the enzyme and the product formation, and $K_M$ is equal to:
Under this assumption, $K_M$ is equal to the dissociation constant of the $ES$ complex. When this occurs, the value of $K_M$ represents a measure of the affinity of the $ES$ pair. More in detail, high $K_M$ values indicate a weak binding, while low $K_M$ values designate a strong interaction. Most of the enzyme kinetic properties can be described under this model; nevertheless, allostERIC enzymes or enzymes that possess multiple active sites might be described by other models [26] and will not be covered under the framework of this thesis.

In the case of flow through systems such as in IMERS, the reaction time is defined by the residence time of the product in the reactor, and depends directly on the applied flow rate. By plotting the $V$ versus the substrate concentration $[S]$, the so called Michaelis-Menten curve is constructed (Figure VI-4). The graphical evaluation of this nonlinear plot to obtain the kinetic parameters relies on accurate fitting. Few linearization methods such as the Hanes–Woolf plot and Lineweaver–Burk plot have been employed. However, in most of the cases the linearization via those method generates errors in addition to inaccuracy to the determination of the constants. With the available computational tool nowadays, the curve fitting via nonlinear equation generates so far the most accurate results with a straight forward approach.

![Michaelis-Menten saturation curve](image)

**Figure VI-4.** Michaelis–Menten saturation curve for an enzyme reaction showing the relation between the substrate concentration and reaction rate.
Besides $K_M$ and $V_{max}$, another valuable parameter is the turnover number, which is the maximum number of substrate molecules converted to product per enzyme molecule per time unit. The turnover number is related with the first order kinetic constant $k_3$ (Eq. VI-1) also named in the literature as catalytic constant ($k_{cat}$). The larger $k_3$, the more rapid the reaction takes place. This constant is calculated as:

$$k_3 = \frac{V_{max}}{[E]}$$

Eq. VI-7

The enzymatic activity is expressed as the moles of substrate converted per time unit, and is calculated from $V_{max}$ and the reaction volume ($\phi$):

$$U = V_{max} \cdot \phi$$

Eq. VI-8

The specific activity ($A_{sp}$) is the activity related to the milligrams of enzyme present in solution or immobilized ($m$), and is calculated as:

$$A_{sp} = \frac{U}{m}$$

Eq. VI-9

For proper description purposes, after the immobilization, usually the immobilization yield and enzymatic activity should be determined. The amount of immobilized enzyme does not necessary correspond to the active immobilized enzyme. One feature determines the yield of immobilization in terms of protein bound to the matrix while the other represents the number of units that have remained active.
Finally, it is well known that the enzymatic activity strongly depends on the environmental conditions. Different factors that affect enzymes in solution will affect immobilized enzymes. Therefore, care must be taken in the control of temperature, ionic strength, pH, flow rate, organic modifier, the enzyme concentration, and support and immobilization chemistry. The influence of those parameters over the enzymatic activity can usually be predicted for the immobilized enzyme based on the observation of the enzyme in solution; however, no accurate prediction can be made and most of the times the effects drastically change, therefore the experimental determination must take place.

VI.5 References

Chapter VII

Hyphenation of a Deoxyribonuclease I immobilized enzyme reactor with liquid chromatography for the on-line stability evaluation of oligonucleotides

Summary

The stability of antisense oligonucleotides (ONs) towards nucleases is a key aspect for their possible implementation as therapeutic agents. Typically, ON stability studies are performed off-line, where the ONs are incubated with nucleases in solution, followed by their analysis. The problematics of off-line processing render the detailed comparison of relative ON stability quite challenging. Therefore, the development of an on-line platform based on an immobilized enzyme reactor (IMER) coupled to liquid chromatography (LC) was developed as an alternative for improved ON stability testing. More in detail, Deoxyribonuclease I (DNase I) was immobilized on epoxy-silica particles of different pore
sizes and packed into a column for the construction of an IMER. Subsequently, the hyphenation of the IMER with ion-pair chromatography (IPC) and ion-exchange chromatography (IEC) was evaluated, leading to the successful development of two on-line methodologies: IMER-IPC and IMER-IEC. More specifically, natural and modified DNA and RNA oligonucleotides were used for testing the performance of the methodologies. Both methodologies proved to be simple, automatable, fast and highly reproducible for the quantitative and qualitative evaluation of ON degradation. In addition, the extended IMER life time in combination with a more straightforward control of the reaction kinetics substantiate the applicability of the IMER-LC platform for ON stability tests and its implementation in routine and research laboratories.

Published as:


VII.1 Introduction

In the last two decades the therapeutic potential of antisense oligonucleotides (ONs) has been firmly recognized. These single-stranded short nucleic acid chains are designed to hybridize complementary sequences in DNA, pre-mRNA and mRNA, and in this way, block the expression of a target gene [1-3]. The main problems related with the development and implementation of therapeutic ONs include their intracellular delivery, transfection across the membrane and stability against nucleases [2,3]. In an attempt to improve upon the latter, many modifications on the phosphodiester backbone, the sugar or the base moiety have been described for increasing nuclease resistance [2,3]. For instance, the phosphorothioate modification is one of the simplest chemical modifications which can be carried out in the ON [3,4]. Hereby a non-bridging oxygen of the phosphorus is substituted by a sulfur atom. Phosphorothioates have demonstrated high resistance towards nucleases; however, their biggest drawback is the reduced ON hybridizing ability towards complementary strands [3,4].
Therefore, many other chemical modifications and their combinations have been considered in order to seek for an optimal balance between the hybridization ability of the ONs and their resistance towards nucleases [3-5].

Stability studies of ON have been mainly performed in vitro [4,6-24]. For this purpose ONs are typically incubated with a specific nuclease, mixture of nucleases, biological fluid or tissue homogenates, followed by their analysis in various ways. The initially developed analysis techniques required the usage of fluorescent or radiolabeled ONs in combination with polyacrylamide gel electrophoresis (PAGE) [6-8,10,14,15,25] or thin layer chromatography (TLC) [6]. However, those procedures were laborious, often lacking good sensitivity, and attaining accurate and precise quantitation results remains challenging. Afterwards, with the increasing use of capillary electrophoresis (CE) [23,24,26-28] and liquid chromatography (LC) [6,29], labeled ONs were no longer required. The usage of CE has proven to deliver a base to base resolution for ONs differing by one nucleotide in length. Moreover, the substitution of polyacrylamide filled capillaries by entangled polymer solutions drastically improved the reproducibility of the methodology and facilitated its implementation [27,30]. Nevertheless, the high ionic strength of the enzymatic incubation media, requires a desalting step of the sample for an adequate CE analysis [23,24,26,27,31], complicating in this way the sample manipulation procedure, affecting the analysis time and reproducibility of the methodology. In contrast, the greater robustness of LC, in addition to recent advances in the development of higher efficiency columns which are able to provide a high resolution separation of ONs in the range of 10-mer to 30-mer [29,30,32], made this technique the preferred candidate for ON stability studies in pharmaceutical environments [9,11,13,16,18,20-22]. Among the existing LC techniques, ion-pair chromatography (IPC) and ion-exchange chromatography (IEC) have proven to be suitable for this purpose [20,24,29,30,33]. Additionally, ON separations by IPC and IEC (with a polymeric support), do not occur exclusively based on the length of the ON, but also depend on the base composition [32]. Thus, discerning between two ON product fragments which have the same length but a significantly different base composition can be efficiently achieved using either of these two techniques [30]. Analytical procedures employed thus far for ON stability share in common the presence of enzymes in solution, for which the reaction is performed by batch incubation [4,6,9-11,13-16,18,19,25-27] or in on-line way [34]. Incubation can be performed using either the enzyme present in solution or immobilized on a solid support. Enzyme immobilization offers several advantages when compared to enzymatic reactions carried out in solution. In addition to an easier
Hyphenation of a Deoxyribonuclease I immobilized enzyme reactor with liquid chromatography for the on-line stability evaluation of oligonucleotides

Separation of the enzyme from the reaction products and/or substrate, the immobilization of enzymes also confers them an increased stability towards mechanical stress, pH, heat, ionic strength and organic solvents [35-41]. Moreover, once packed into the columns to be used as immobilized enzyme reactors (IMERs), they can be employed in a continuous operation mode, reducing in this way the sample manipulation steps, increasing the reproducibility and allowing their reutilization; therefore, reducing costs and analysis time.

The immobilization of enzymes on solid supports has been widely studied and numerous immobilization methods and supports which are commonly employed have been reported in literature [35,36,38,40-43]. Several nucleases are commercially available, and a broad variety of combinations can be tested based on the type of substrate they hydrolyze (DNA or RNA), the type of nucleophilic attack (exonuclease and/or endonuclease), the nature of the hydrolytic products (mono or oligonucleotides) and the nature of the bond which is hydrolyzed [44,45]. Furthermore, factors such as their biological relevance, kinetic performance and cost, also need to be considered for their application in IMERs.

Deoxyribonuclease I (DNase I), is a 29.1 kDa endonuclease which cleaves the phosphodiester backbone of double and single stranded DNA, requiring divalent cations as cofactors [46,47]. The enzyme cleaves the P-O3’ bond of the DNA backbone, yielding 5’-phosphate oligonucleotides [47-49]. Moreover, DNase I has been used as a powerful footprinting agent and for DNA probing [50,51]. The immobilization of DNase I on monolithic supports [52,53], polymers [54] and magnetic particles [55] has been reported. Furthermore, the good catalytic performance of the immobilized endonuclease and its low cost, makes DNase I a good candidate for its implementation in IMERs.

In this study, we describe the construction of an IMER with bovine DNase I, and the evaluation of its hyphenation with IPC and IEC using natural and modified DNA and RNA ONs. Two on-line platforms (IMER-IPC and IMER-IEC) with potential applications in ON stability studies were successfully developed. To the best of our knowledge, the development of on-line IMER-LC methodologies for stability testing of ONs had not been reported yet, and could be of high relevance in the development of improved antisense ON therapies.
VII.2 Experimental

VII.2.1 Chemicals

Triethylamine (TEA), acetic acid, adenosine monophosphate (AMP), hydrochloric acid, chloroform, Tris, NaCl, acetone (HPLC grade), Na2HPO4, EDTA, CaCl2, MgCl2, glycine, (3-Glycidyloxypropyl) trimethoxysilane (GPTMS) and Deoxyribonuclease I (DNase I) from bovine pancreas (≥400 Kunitz units/mg of protein) were purchased from Sigma-Aldrich (Steinheim, Germany). Acetonitrile (ACN) HPLC grade from Fischer Scientific (Loughborough, U.K.) and Milli-Q water (Millipore, Milford, MA) were used. All the mobile phases were filtered through 0.22 µm nylon membrane filters (Grace Davison Discovery Sciences, Lokeren, Belgium). Homo-oligonucleotides of deoxythymidine 15-mer (T15), 20-mer (T20) and 30-mer (T30), deoxyadenosine 30-mer (A30), 12-mer ON (5’-GCA-CAC-CGT-CAG-3’) and a 41-mer ON (5’-GTT-GGA-TTA-AAC-CGT-TCC-CGT-CTC-TAT-CAG-CTT-AGT-GT-3’) all based on the phosphodiestere backbone; 15-mer (5’-T*T*T*-T*T*T*-T*T*T*-T*T*T*-3’) with a phosphorothioate backbone, and the 2’ O-methyl RNA 12-mer ON (5’-GCA-CAC-CGU-CAG-3’) were purchased from Eurogentec (Liege, Belgium). The lyophilized ONs were dissolved in an appropriate volume of Milli-Q water in order to obtain 100 µM stock solutions.

VII.2.2 Instrumentation

An Agilent 1100 series HPLC composed of a binary pump and a single wave length detector was employed for the delivery of the chromatographic mobile phases, while the enzymatic reaction buffer was delivered to the reactor by a HP 1050 quaternary pump (Agilent Technologies, Waldbronn, Germany). A CTO-20AC prominence HPLC column oven (Shimadzu, Kyoto, Japan) was used for controlling the temperature of the IMER.
Hyphenation of a Deoxyribonuclease I immobilized enzyme reactor with liquid chromatography for the on-line stability evaluation of oligonucleotides

**VII.2.3 Preparation of epoxy-functionalized silica**

Nucleosil 5 µm spherical silica particles (1.2 g) of 300, 1000 and 4000 Angstrom from Macherey-Nagel (Duren, Germany) were used. The functionalization of the silica particles was carried out by suspending them in a 10% (v/v) GPTMS solution in acetone at room temperature for 2 h under intense stirring. Afterwards, the particles were filtered and thoroughly washed with acetone, followed by phosphate buffer (pH 7.0; 15 mM) containing 0.1 M NaCl and finally with phosphate buffer [56]. The particles were dried at 37 °C and then stored at 4 °C until the enzyme immobilization was carried out.

**VII.2.4 Immobilization of DNase I**

Three mL of a 5 mg/mL solution of DNase I in phosphate buffer (pH 7.0; 15 mM) was mixed with 1 g of the functionalized silica particles and stirred at 3000 rpm for 30 min. Subsequently, the mixture was set in the fridge (4 °C) for 12 h. When the immobilization was completed, the enzyme modified silica was filtered and washed with phosphate buffer. Afterwards, deactivation of the unreacted epoxy groups was carried out by mixing the particles with a 0.2% (m/v) glycine solution in phosphate buffer and stirring for 2 h at room temperature. Once the deactivation was completed, the particles were filtered and washed with phosphate buffer containing 0.1 M NaCl, followed by phosphate buffer and then by the reaction buffer (pH 7.5; 50 mM Tris, 5 mM CaCl2 and 5 mM MgCl2). Finally, the enzyme modified silica particles were dried at 37 °C and stored at 4 °C before use. The amount of enzyme bound to the silica particles was determined from the concentration difference of the DNase I solution before and after the immobilization using flow injection analysis at 280 nm.

**VII.2.5 Preparation of the IMERs**

150 mm x 2.1 mm I.D. stainless steel columns were packed using 5 mL slurry containing 1g of enzyme-silica particles in phosphate buffer (pH 7.0; 15 mM). The slurries where previously homogenized for 10 minutes in an ultrasonic bath and consequently packed in the columns at a constant pressure of 400 bar for 2h using the phosphate buffer as packing solvent.
Chapter VII

VII.2.6 Kinetic determinations

The biological activity of DNase I was determined by measuring the degradation that a deoxythymidine 30-mer oligonucleotide (T30) experienced under the incubation with this enzyme. For this, Ion-pair chromatography (IPC) using a XBridge C18 column 50 mm × 4.6 mm I.D. × 3 µm column (Waters, Zellik, Belgium) was used, and the separation temperature was maintained at 60 °C. The IP mobile phase was composed of 100 mM TEA (pH 5.5 adjusted with acetic acid) (A) and ACN (B). The gradient profile consisted of 0-50 min, 0-16B% and the detection was performed at 260 nm. A calibration curve for the T30 ON (0.05-60 µM) was constructed for calculating the amount of T30 in the blank and in the test samples. The enzyme activity was expressed as the pmol of T30/min degraded under the assay conditions described in sections VII.2.6.1 and VII.2.6.2.

VII.2.6.1 Activity of free DNase I

A T30 ON (final concentration 1-50 µM) was incubated with DNase I (0.01-0.0005 mg/mL) in 50 mM Tris buffer pH 7.5, 5 mM CaCl₂ and 5 mM MgCl₂ at 37 °C (final volume of 100 µL) and stirred at 2000 rpm. The incubation time ranged from 0.5-5 min and the reaction was stopped by adding 10 µL of 0.5 M EDTA solution. Afterwards, the sample was analyzed by IPC as described in section VII.2.6. The reaction rate (µM T30/min) was calculated at the beginning of the reaction in the linear part, and the results were used to construct the Michaelis–Menten plots [37]. The specific activity was calculated by dividing the enzyme activity by the amount of dissolved enzyme. Non-linear regression analysis was performed for calculating the kinetic parameters.

VII.2.6.2 Activity of immobilized DNase I

The activity of the immobilized DNase I was determined in batch in order to attain similar reaction conditions to the assay in section VII.2.6.1. The immobilized DNase I silica particles (10-1 mg) were incubated with a T30 ON (final concentration 1-50 µM) in 50 mM Tris buffer pH 7.5, 5 mM CaCl₂ and 5 mM MgCl₂, at 37 °C (100 µL) and stirred at 2000 rpm. The vials were incubated for a period ranging from 5-90 min followed by the addition of 10 µL of 0.5 M EDTA solution in order to stop the reaction. Afterwards, the solutions were filtered through 0.45 µm syringe filters before their instrumental
Hyphenation of a Deoxyribonuclease I immobilized enzyme reactor with liquid chromatography for the on-line stability evaluation of oligonucleotides

analyses. The kinetic parameters were determined as described in section VII.2.6.1. The specific activity was calculated by dividing the enzyme activity by the amount of immobilized enzyme.

VII.2.7 On-line IMER-LC setup

The employed instrumental setup is depicted in Figure VII-1. This fully automated analytical procedure can be subdivided in two main events: the reaction and focusing of the analytes (i) and the separation (ii). Table VII-1 resumes the chromatographic conditions of the integrated IMER-LC system.

(i) Pump 1, which delivered the reaction buffer, was directly connected to the HP 1100 auto-injector (1-5 µL of injection volumes were used), followed by the IMER which was kept at 37 °C during all the experiments. By changing the flow rate passing through the reactor, the residence time of the sample could be tuned in order to attain the desired degradation rate for the ONs. The system’s and reactor’s dead volumes were determined by injecting a 1 mM caffeine solution. The effluent from the bioreactor was mixed with the LC mobile phase (pump 2) with the aid of a T-piece. In this way, both the initial chromatographic mobile phase and the reaction buffer, were continuously delivered to the LC column. This composition and flow rates were kept constant for sufficient time in order to assure that all the sample plug exited the reactor and focused on the head of the LC column; for instance: for a reaction buffer flow rate of 50 µL/min (pump 1), the LC separation mobile phase was kept at 500 µL/min for 40 min, exceeding by this, on average, 4 times the reactor’s dead volume (Table VII-1).

(ii) Two separation modes were explored: ion-pair chromatography (IPC) and ion-exchange chromatography (IEC), and the separation conditions were based on a previously described methodology [30]. The detection wavelength was fixed to 260 nm and the flow rate was set to 500 µL/min for both separation modes. For IPC, the same separation conditions were used as in section VII.2.6. During the reaction-focusing step, the IPC mobile phase composition was kept at 100 % of A; after which, the separation process took place, for which the reaction buffer flow rate (pump 1) was automatically lowered to
1 µL/min and the IPC gradient was started (pump 2). The IEC experiments were performed on a polymeric based analytical column 4 mm × 250 mm DNAPac PA200 protected with a 4 mm × 50 mm DNAPac PA200 precolumn (Thermo Scientific, Erembodegem-Aalst, Belgium). The column temperature was set to 25 ℃. The mobile phase consisted of water at pH 11.5 (adjusted with NaOH) (A) and water with 1.25 M NaCl at pH 11.5 (B). The gradient profile was 0–90 min, 0–90%B. The same conditions as in the IPC methodology were set for pump 1.

It must be pointed out that each pump functioned accordingly to its own elution program and that they were interconnected via an analog interface for their synchronization; thus, the start of the analysis was controlled by the ChemStation software from the Agilent 1100 instrument.

Figure VII-1. Schematic representation of the IMER-LC setup.
Hyphenation of a Deoxyribonuclease I immobilized enzyme reactor with liquid chromatography for the on-line stability evaluation of oligonucleotides

Table VII-1. Chromatographic conditions of the IMER-LC setup.

<table>
<thead>
<tr>
<th>Event</th>
<th>Flow rate (µL/min)</th>
<th>Time (min)</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) Elution of the analytes through</td>
<td>Pump 1 50 500</td>
<td>0.00-40.00</td>
<td>(i) Elution of the analytes through the IMER and focusing in the analytical column</td>
</tr>
<tr>
<td>the IMER and focusing in the</td>
<td>Pump 2 (Eluent)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>analytical column</td>
<td>50 (100% A)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(ii) Separation of the analytes by</td>
<td>1 500 500</td>
<td>^a 40.01-90.00</td>
<td>(ii) Separation of the analytes by the analytical column</td>
</tr>
<tr>
<td>the analytical column</td>
<td>(Gradient A-B)</td>
<td>^b 40.01-130.00</td>
<td></td>
</tr>
</tbody>
</table>

^a Ion-pair chromatography (A: 100 mM TEA, pH 5.5; B: ACN)
^b Ion-exchange chromatography (A: H₂O, pH 11.5; B: 1.25 NaCl, pH 11.5)

VII.3 Results and discussion

VII.3.1 Immobilization of DNase I

The first goal was to develop a reactor allowing maximal enzymatic activity while maintaining the residence time to a minimum. Several variables may influence the IMER performance; among them, the choice of the supporting material and the immobilization chemistry of the enzyme to the support, stand out as the most critical. An ideal support should be chemically inert to the reaction conditions and samples. In addition, its physical properties such as porosity, pore size distribution, shape, swelling capacity and charges, play also an important role in the kinetics of the process [35-38]. Spherical silica particles were chosen as support for the enzyme as a variety of pore diameters and particle sizes are available, allowing for example, the optimization of the diffusion process and the pressure drop in the IMER. In addition, their increased resistance to degradation by contaminating microorganisms and a stable morphology towards different solvent conditions, distinct as relevant factors for the development of an IMER [38,41,57,58].

Many strategies for immobilizing enzymes have been described, and among the most relevant stand the adsorption and the covalent immobilization [35,36,38,41]. Though covalent immobilization of the enzyme ensures an increased ruggedness and life time of the reactor, one should take into account that, due to the drastic reaction conditions used during the covalent immobilization process, enzymatic activity can be somewhat reduced.
A common, if not the most common crosslinking agent for enzyme immobilization is glutaraldehyde [35,36,38]; however, amino-propyl silica is thereby required as starting material, presenting the disadvantage that ONs strongly adsorb and tail to the unreacted silanol amino functions, rendering this activation process inadequate for the purpose of the here developed methodology. Therefore in this work, the immobilization of DNase I was carried out via nucleophilic attack of the amine groups of the enzyme on the immobilized epoxide groups in silica. Next to the simplicity of this type of silica activation, this procedure results in quasi inert and stable covalent bonds [38,41,56,59]. In addition, no adsorption of the ONs onto the support material could be observed when the glycidyloxypropyl bond chemistry was used as an alternative.

As the accessibility of the substrate to the active sites and its diffusivity in the IMER play an important role in kinetic reactions, immobilization on 5 µm silica particles with a pore diameter of 300 Å, 1000 Å and 4000 Å was evaluated. Table VII-2 shows the amount of enzyme immobilized on the silica supports. It can be noticed that particles with larger pore diameters attach less enzyme in direct correlation with their decreased surface area [60]. This also demonstrates that under the immobilization process conditions, the larger fraction of the enzyme is expected to immobilize close to the pore mouth and on the external surface of the particle rather than deep in the pores [61].

Table VII-2. Properties of the immobilized DNase I silica particles

<table>
<thead>
<tr>
<th>Silica pore size</th>
<th>Pore volume (mL/g)</th>
<th>Surface area (m²/g)</th>
<th>Immobilized DNase I (mg/g)</th>
<th>Activity/support (pmol T30/min·g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>300 Å</td>
<td>0.8</td>
<td>100</td>
<td>8.2 ± 0.3</td>
<td>123.8 ± 42.5</td>
</tr>
<tr>
<td>1000 Å</td>
<td>0.8</td>
<td>25</td>
<td>6.1 ± 0.2</td>
<td>643.3 ± 156.0</td>
</tr>
<tr>
<td>4000 Å</td>
<td>0.7</td>
<td>10</td>
<td>4.9 ± 0.1</td>
<td>361.8 ± 23</td>
</tr>
</tbody>
</table>

5 µm silica particles were employed. The free DNase I used for immobilization has a specific activity of 6.09 x 10⁵ pmol T30/min·g. Result expressed as the mean ± CI (α = 0.05) (n = 3).

VII.3.2 Activity of free and immobilized DNase

Typically the activity of DNase is determined based on the Kunitz hyperchromicity assay [62], in which the activity of the DNase is manifested as the increase on absorption at 260 nm as a DNA solution is degraded. The substrate used for activity tests has significative effect over the activity of DNase to cleave the phosphodiester backbone. DNase is able to cleave both, single and double stranded DNA,
being more active towards this latter one [46]. It has been documented that the sequence also plays an important role, which has been attributed mainly to the flexibility of the DNA [48]. DNA regions rich in A/T nucleotides are rather less flexible and therefore a diminished cleavage is expected [63]. In this study we used a homo-oligonucleotide of deoxythymidine 30-mer (T30) as substrate, as it will provide a more realistic approach of the DNase activity towards the goals of the developed methodology, which is to evaluate the stability single stranded oligonucleotides.

The activity of DNase was calculated as the picomoles of T30 degraded per minute. The determination was performed in a batch sequence for both the free enzyme and the immobilized enzyme, to maintain the same incubation conditions, such as the same agitation speed. The activity of the free enzyme was determined as $6.09 \times 10^5$ pmol T30/min while the retained activity percentage of the immobilize enzyme was around 0.1% (Table VII-2). The activity decrease upon immobilization is expected, as once immobilized, the enzyme is somewhat restricted and subjected to conformational adaptations in addition to a limited access of the substrate to the active sites. Note that usage of spacers could be one strategy allowing addressing this loss in activity. Moreover, it needs to be considered that the covalent immobilization occurs on the multiple regions of the enzyme and that the exposure of the active site might be impeded. Therefore the calculated specific activity values are inherently reduced and might be not directly correlated with the immobilized amount. Moreover, additional supports and immobilization techniques might also be explored, what could result in different activities. Another strategy to obtain increased reactor activity could be found in the usage of more active batches of DNase I, however with the accompanied drawback of increased cost. However, it is important to point out at this point that by adjusting the flow rate through the IMER, differences in the degree of degradation can be controlled.

By comparing the activity of the immobilized DNase I in the 5 µm silica particles of 300 Å, 1000 Å, and 4000 Å it can be noticed that the most active support was the one with an average pore size of 1000 Å (Table VII-2). On the other hand, the least active support was the one with the smaller pore sizes, despite of a larger amount of immobilized enzyme. This suggests that the larger pore sizes allow for a facilitated interaction of the substrate with active site of the enzyme. Large molecules such as ONs have large hydrodynamic volumes. Considering a T30 ON (~9000 Da) as a sphere in solution, its hydrodynamic diameter is estimated to be around 100 Å (based on a correlation to a protein MW). Therefore, in order to avoid restriction through the pores in the particles (a cause of lower activities),
a pore diameter at least 3 times bigger than the one of the analyte hydrodynamic volume is required (e.g. 300 Å). When larger pores are used (e.g. 1000 Å) the enzyme and the ON can easily interact with flexibility and the interaction does not occur mainly at the head of the pores. However, when too large pores are employed (e.g. 4000 Å) the interaction between the ON and the immobilized enzyme might not occur so frequently as the ON might pass freely through the pores, in addition to a lower enzyme load due to the decrease in the surface area with larger pores.

This observation is corresponding to the degradation profiles of the T30 ON encountered in the IMERs (Figure VII-2) tested in the developed methodology (section VII.3.3) when residence times below 10 min are used, where also the 1000 Å support stands as the most active. It is important to notice that that the diffusivity that ONs experience through the pores in a flow through system are different than the ones experienced in batch. At higher flow rates, and due to a reduced diffusivity of ONs, large enough ONs experience diffusion restriction with smaller pore sizes, such as the case observed for the T30 ON (Figure VII-2). By contrast, when longer residence times are used (lower flow rates through the IMER), the T30 ON is able to penetrate into the 300 Å pores more effectively when compared to shorter residence times (faster flow rates through the IMER). Therefore the highest activity can be attributed to the support consisting of the 300 Å, this also in correlation with the amount of loaded enzyme. This illustrates that factors such as the analyte diffusivity and the hydrodynamic volume of ONs play an important role in the activity of IMERs. A right balance needs thereby to be found between the pore size of the supporting material and the achievable enzyme load.

Furthermore, in Figure VII-2, it can be noticed that the degradation of the 41-mer ON increases as the pore diameter increments, as can be observed in curves A and B in Figure VII-2 corresponding to the IMERs with 1000 Å and the 300 Å pore diameter particles, respectively. The curve of the 4000 Å IMER for the 41-mer ON is not presented, as this ON practically degrades in its totality at the fastest residence time tested in our method (~4 min). The higher degradation rate observed for 41-mer ON when compared to the T30 ON is related to secondary structure formation resulting in the presence of a certain degree of duplex structures and also to a mixed nucleotide sequence towards which DNase I is more active [46]. Additionally, a facilitated diffusion of the larger ON through the IMER with the larger pore diameters, also explains this significant differences as already discussed for T30. Most importantly, the differences in degradation between the ONs are also derived from the DNase I cleavage dependence on the ON sequence [48,49,63]. Although further studies on a broader set of
Hyphenation of a Deoxyribonuclease I immobilized enzyme reactor with liquid chromatography for the on-line stability evaluation of oligonucleotides

ONs could reveal additional factors, at this time it is important to note that this setup satisfactorily allows the comparison and modification of the residence times by adjusting the flow rate.

Figure VII-2. Oligonucleotide degradation in the the 300 Å, 1000 Å, 4000 Å IMERs. A T30 ON and the 41-mer ON are compared. The degradation percentage is plotted against de residence time in the IMER. The confidence interval (α = 0.05, n-1 d.f.) of the mean is plotted for each point.

VII.3.3 On-line IMER-LC

The ease of operation, automation and reproducibility of LC when combined with the specificity of an IMER, allows for a highly sensitive methodology to study ON degradation. As in the developed methodologies, the IMER was assembled in the pre-column configuration, the LC separation conditions which could in principle cause possible enzymatic denaturation due to organic solvents, salts and pH of mobile phase, had no effect on the IMER performance. Vice versa, in this setup, the enzymatic reaction conditions should also not alter the performance of the LC separation.

Typically, in pre-column IMER assemblies, both LC chromatographic pumps run independently by the aid of a switching valve [36,39-41,64]; nevertheless, in the proposed methodologies, the pumping systems were harmonized and the use of a switching valve was substituted by a T-piece. In this
configuration, the IMER effluent was also continuously percolated through the LC column. This increases the simplicity of the methodology; however, the IMER reaction conditions may have an impact over the LC separation performance. Therefore, as first criterion, the IMER reaction conditions must be compatible with the LC methodology. A second possible drawback of this approach is the band broadening, which might be caused due to long residence times in the IMER and non-specific interactions of the analytes with the enzyme and the support material.

As it was required to achieve a high resolution separation of complex ONs mixtures after the reaction, it was critical that the analytes properly focused at the head of the separation column once they emerged from the IMER. As this methodology avoids a switching valve which might allow coupling a trapping column, the LC mobile phase must allow for a proper focusing of the analytes as such.

Ion-pair chromatography (IPC) and ion-exchange chromatography (EIC) were investigated as those LC methodologies have demonstrated to be suitable for a high resolution separation of therapeutic ONs [29,30]. In addition, the IMER reaction buffer (Tris 50 mM, pH 7.5, CaCl$_2$ 5 mM, MgCl$_2$ 5 mM) was compatible with those LC mobile and stationary phases.

**VII.3.3.1 On-line IMER-IPC**

IPC has been widely used to purify ONs; therefore, the efficient focusing effect of the hydrophobic ion-pair formed between the negatively charged ON and triethylamine (TEA) to an octadecyl silane (C18) stationary phase has been documented [29,33].

Firstly, the focusing of the ONs in IPC under different mixtures ratios of IPC initial mobile phase (TEA 100 mM, pH 5.5) with the IMER reaction buffer was explored in the absence of the IMER and by injecting a poly-T ON mixture under a defined flow rate in pump 1 and pump 2. The conditions were maintained for a defined time, which was established as the time necessary to elute at least 2 to 5 IMER volumes by the flow rate of P1. The tested flow rate ratio (pump 2/pump 1) varied between 2.5 and 20, corresponding to a flow rate ratio pump 2/pump 1 of 500/200 µL/min to 500/25 µL/min, respectively. Those flow rates were mainly dependent on the operational pressure limits in pumps 1 and 2 and the fact that an adequate elution flow rate for the LC separation must be set. Therefore, the flow rate in pump 2 was maintained at 500 µL/min during all the experiments. It must be pointed out that the pressure drop observed in pump 1 with the IMER at a flow rate of 200 µL/min was around
Hyphenation of a Deoxyribonuclease I immobilized enzyme reactor with liquid chromatography for the on-line stability evaluation of oligonucleotides

250 bar. In addition, the pressure experienced by pump 1 was also influenced by the pressure drop in pump 2. As consequence, pump 1 was operating near its pressure limit (400 bar) when a flow rate of 500 and 200 µL/min were used in pump 2 and pump 1, respectively. Under the applied conditions before the start of the gradient in pump 2, no breakthrough and satisfactory peak focusing of the ONs was observed. The gradient elution was achieved by programming the pumps in such way that once the focusing was performed, pump 1 automatically reduced the flow rate to 1 µL/min and pump 2 started the gradient program. In this way the influence of the reaction media over the separation profile was diminished. This was possible as both pumps were programmed independently and as they were interconnected by an analog interface for their synchronization with the LC injector.

Finally, the peak shapes and efficiencies were compared, and no significant loss in resolution was observed down to a pump 2/pump 1 flow ratio of 5. If higher flow rates are required in pump 1 to establish shorter residence times in the IMER, the flow rate in pump 2 must be concomitantly increased in order to maintain a pump 2/pump 1 flow rate proportion of at least 5 for a subsequent highly efficient separation. Note that the system pressure limit becomes the boundary condition in this set up, which can in principle be easily solved by reducing the pressure drop in pump 1 through the use of larger particles, shorter columns or broader columns in the IMER configuration.

Figure VII-3A and Figure VII-3B show the chromatograms of the 41-mer ON using a blank reactor (epoxy-silica deactivated with glycine) and the IMER, respectively. A flow rate of 50 µL/min was used in pump 1, and was maintained for 40 min in order to elute at least 4 IMER volumes. The effect of the IMER on the ON can be clearly observed, and the main degraded fragments can be resolved using this chromatographic technique.

Subsequently, the usage of an internal standard was explored as one potential application of this methodology is to perform quantitative degradation studies. Caffeine (1 mM) was used as internal standard (IS) as it demonstrated not to interact with the IMER, and also presented sufficient focusing in the methodology. However, the variation of the peak shape of the IS was noticeable at different residence times, and was not suitable for a focusing period which exceeded 6 times the IMER volume. By exceeding 4 times the IMER volume, it could be ensured that all the sample bands eluted from the IMER. On the other hand, this drastically extended the analysis time, particularly when long residence times were used. However, by comparing the area of the internal standard at different focusing times
and IMER residence periods, no significant differences on the peak areas were observed up to 1.5 IMER volumes. To deliver degradation profiles that do not only focus on the degradation of the main compound, the residence time was tuned to allow the proper detection of the degraded peaks which might not be detected if the activity of the reactor is too high as they also subjected to degradation. The 300 Å IMER was therefore employed for attaining the uncomplete degradations profiles presented in this paper, as if more active reactors are used the flow rate through the IMER must be increased, what increases the system pressure and mobile phase consumptions.

**Figure VII-3.** IMER-IPC chromatograms of a 41-mer ON. A) blank IMER; B) DNase I IMER. ~10 min of residence time in the 300 Å IMER (Flow rate of pump 1 = 50 µL/min). The sample was allowed to focus for 40 min before the LC gradient start. IS: internal standard. Injection: 1 µL of a 100 µM ON solution.

**VII.3.3.2 On-line IMER-IEC**

It is known that IEC generates higher peak capacities and similar peak shapes for A, T, C , G and U homo-oligonucleotides in comparison to IPC, were ONs rich in C and G are poorly separated [29,30]. Therefore, the hyphenation of the IMER with IEC was also investigated. The same experiments as in the IPC approach were performed in order to explore the focusing of the ONs in this IEC variant. The ONs focused without any problem, and flow rates ratio pump 2/pump 1 up to 2.5 were successfully tested. Higher flow rates (> 200 µL/min) could be used in pump 1 without any consequences on the LC performance; nevertheless, they were not fully explored as the pressure drop in pump 1 almost reached its operational limits with the IMERs. It is important to point out that the pressure
Hyphenation of a Deoxyribonuclease I immobilized enzyme reactor with liquid chromatography for the on-line stability evaluation of oligonucleotides

experienced by pump 2 with the 25 cm IEC column at 500 µL/min did not exceed 100 bar; however, the operational pressure limit of this column is 250 bar.

Figure VII-4A and Figure VII-4B, depict the chromatograms of the 41-mer ON using a blank reactor and the IMER respectively. The same instrumental conditions were used as for the chromatograms in Figure VII-3 with the exception for the injection volume, which was increased to allow for an improved detection. Adenosine monophosphate (AMP) was employed as IS (1 mM), and it demonstrated to focus in an excellent way, in the same extent as the ONs. The higher peak capacity in the IEC methodology can be assessed after comparing the chromatograms of Figure VII-3 and Figure VII-4. An important aspect to emphasize in the IEC methodology is that the separation must be performed under denaturing conditions (pH > 10.5) [30,32], as the reaction fragments can experience hydrogen bond interactions between them, which would lead to unreproducible degradation profiles if the separation was to be performed at lower pH values.

Figure VII-4. IMER-IEC chromatograms of a 41-mer ON. A) blank IMER; B) DNase I IMER. ~10 min of residence time in the 300 Å IMER (Flow rate of pump 1 = 50 µL/min). The sample was allowed to focus for 40 min before the LC gradient start. IS: internal standard, adenosine monophosphate (AMP). Injection: 3 µL of a 100 µM ON solution.
**VII.3.4 Stability evaluation of antisense ONs**

A number of natural and modified ONs were subsequently tested with the developed IMER-LC methodologies for the evaluation of the performance of the method in stability assessments. The degradation profiles were recorded for both methodologies (IMER-IPC and IMER-IEC) and comparisons with the respective blanks were carried out.

The improved, but not complete resistance towards DNase I of the phosphorothioate modification is visualized in Figure VII-5, where plots A and B correspond to a T15 phosphodiester ON, while plots C and D to a the phosphorothioate ON. It can be observed that after ~50 min of residence time in the IMER, the T15 ON is practically completely degraded (plot B compared to its blank plot A), while the T15- phosphorothioate ON only slightly degraded (plot D compared to its blank plot C). In the plots E to H the specificity of the IMER towards DNA ONs is demonstrated, as the RNA ON passes intact through the IMER (plots G and H).

Changes in the degradation profiles when injecting a mixture of ONs demonstrate the applicability of the methodology for footprinting analysis and/or cleavage protection carried out by a hybridization process for instance. Hybridization between the A30 and T30 ONs was tested for that purpose. The differences in the degradation profiles are visualized in Figure VII-6, where plot A corresponds to the analysis of a (T30, A30) ON mixture in the blank IMER, while plots B, C and D depict the degradation profiles of a single stranded A30 ON, a single stranded T30 ON and of the (A30, T30) ON duplex, respectively.

Moreover, similar ON degradation profiles consisting mainly of nucleotide deletions form the main compound as observed in Figure VII-6 have been observed in vivo [31] and in vitro metabolism studies of ONs [23,24,31] as well as for small interfering RNA (siRNA) [20-22]. This proves the utility of the developed methodologies as no sample preparation methodologies are required previous to the instrumental analysis. Furthermore, under this approach the usage of only one enzyme can be extended to multiple enzymes to deliver more representative information on the metabolism of these molecules.

One of the advantages of the usage of IMERs is their stability and extended life time. Indeed, the IMERs demonstrated to be highly active even up to 3 months after their construction, a period during which they were subjected to analysis and storage conditions (4 °C). Figure VII-7 shows the
Hyphenation of a Deoxyribonuclease I immobilized enzyme reactor with liquid chromatography for the on-line stability evaluation of oligonucleotides

degradation profile of a (T30, A30) ON mixture tested on the 300 Å IMER after 3 months from its manufacturing. Furthermore, the high reproducibility of the methodology can also be visualized in Figure VII-7, where the degradation profiles obtained during three consecutive days are presented. Additionally, the relative standard deviation (RSD) on the area of selected product peaks is also displayed. The %RSD did not exceed 15%, and such high values, have only been observed for the smallest product peaks, while for larger product fragments the RSDs were around 5%. The average %RSD in all cases was below 10%, which is exceptionally good as it was measured on the generated degradation fragments, demonstrating the potential of the methodology for performing on-line quantitative degradation studies. Note that the %RDS on the retention times were below 0.1% for all the cases.

**Figure VII-5.** IMER-IPC chromatograms. The analyzed ON sample is indicated for each chromatogram. A), C), E) and G) correspond to a blank IMER, while B), D), F) and H) to the DNase I IMER. ~50 min of residence time in the 300 Å IMER (Flow rate of pump 1 = 10 µL/min). The sample was allowed to focus for 80 min before the LC gradient start. IS: internal standard, caffeine 1 mM. Injection: 1 µL of a 100 µM ON solution.
Figure VII-6. IMER-IEC chromatograms. The analyzed ON sample is indicated for each chromatogram. A) corresponds to a blank IMER, while B), C) and D) to the DNase I IMER. ~50 min of residence time in the 300 Å IMER (Flow rate of pump 1 = 10 µL/min). The sample was allowed to focus for 80 min before the LC gradient start. Injection: 10 µL of a 50 µM ON solution. Each ON is present at the same concentration. No internal standard was employed.

The limits of detection (LOD) of the IMER-LC methodologies are determined by the signal to noise ratio (S/N = 3) [64]. The calculated LOD values for the 12-mer and the 41-mer ONs are shown in Table VII-3. Two LOD values were determined, one corresponding to the instrumental LOD, determined by injection of the ONs through a blank IMER, and the IMER-LC LOD, which was determined using the
Hyphenation of a Deoxyribonuclease I immobilized enzyme reactor with liquid chromatography for the on-line stability evaluation of oligonucleotides

IMER. This last LOD, was calculated based on the product fragments rather than by the main peak, and was defined as the concentration where at least 80% of the fragments in a degradation profile were detected with a S/N=3. This LOD is dependent on the residence time, as the number and intensity of the generated fragments are strongly influenced by it. Nevertheless, this LOD provides a useful insight for setting on analyses with this setup. In general, somewhat higher LODs were observed for the IMER-IEC methodology when compared to the IMER-IPC, and were attributed to the higher signal noise ratio generated by the IEC mobile phase containing 1.25 M NaCl.

**Figure VII-7.** Three IMER-IEC chromatograms of a (T30, A30) ON mixture analyzed during 3 consecutive days in a 3 months old 300 Å IMER. The %RSD on the area is indicated for few peaks. ~50 min of residence time in the IMER (Flow rate of pump 1 = 10 µL/min). The sample was allowed to focus for 80 min before the LC gradient start. Injection: 5 µL of a 50 µM ON solution. No internal standard was employed.

<table>
<thead>
<tr>
<th>Compounds</th>
<th>LOD Main peak</th>
<th>LOD Product fragments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IPC</td>
<td>IEC</td>
</tr>
<tr>
<td>12-mer ON</td>
<td>0.01</td>
<td>0.11</td>
</tr>
<tr>
<td>41-mer ON</td>
<td>0.03</td>
<td>0.12</td>
</tr>
</tbody>
</table>

The values are given in µg/µL concentrations and considering an injection volume of 1 µL. Only DNA based ON were tested.

*a* Determined with a blank IMER, S/N = 3.

*b* Determined under a flow rate of 50 µL/min through the 300 Å IMER and considering a detection of 80% of the product fragments with a S/N = 3.
Chapter VII

VII.4 Conclusion

An on-line IMER-LC platform for the stability evaluation of antisense DNA ONs was developed. IPC and IEC were successfully coupled with a DNase I IMER. More in detail, the IMER-IEC method presented a higher resolving power for the ON degradation products when compared to the IMER-IPC method. On the other hand, the latter allows for easier hyphenation with mass spectrometry [29,65]. In addition to the extended life time of the IMER, both methodologies showed a high reproducibility in area and retention time. This platform demonstrates potential for application in laboratories where stability tests of ONs are often performed. Moreover, the construction of IMERs with other endonucleases and exonucleases, specific or not to the sugar and their coupling to LC should also be explored, as this would cover stability assessments of a much wider range of ON modifications. Furthermore, the construction of an IMER containing multiple nucleases, based either on the simultaneous immobilization of various nucleases to a support or by the interconnection of IMER subunits is promising, as many nucleases share the same activation conditions by divalent cations such as $\text{Ca}^{2+}$ and $\text{Mg}^{2+}$, allowing for an easy hyphenation with LC. Implementation of the here described methodology for nuclease resistance assessment will allow generation of uniform and reliable data, finally enabling reliable comparison of different approaches towards enhancement of nucleic acid stability, which is of prime importance in the fast developing field of ON based drug development.

VII.5 References

Hyphenation of a Deoxyribonuclease I immobilized enzyme reactor with liquid chromatography for the on-line stability evaluation of oligonucleotides


[64] ICH harmonized tripartite guideline: validation of analytical procedures: text and methodology Q2(R1), International Conference of harmonization of technical requirements for registration of pharmaceuticals for human use, 2005.

Offline comprehensive liquid chromatography in combination with a Deoxyribonuclease I immobilized enzymatic reactor for selective screening of oligonucleotide mixtures

Summary

The development of a comprehensive ion-pair chromatography-immobilized enzyme reactor x ion-pair chromatography (IPC-IMER x IPC) methodology for the advanced characterization of DNA/RNA oligonucleotides (ON) mixtures has been carried out. More in detail, a DNase I IMER has been coupled to IPC in the post column configuration, followed by the collection of the eluting fractions and reanalysis by IPC. The effect of the mobile phase over the IMER activity was qualitatively evaluated. The methodology proved to generate relevant ON degradation profiles that might be correlated with the ON stability towards nucleases. Moreover, this platform shows potential for its further implementation in selective analysis of ON mixtures and in mapping studies.
Offline comprehensive liquid chromatography in combination with a Deoxyribonuclease I immobilized enzymatic reactor for selective screening of oligonucleotide mixtures

Published as:

1.1 Introduction

In the last decades, the usage of oligonucleotides (ONs) for the antisense technology has significantly increased. These molecules are essentially unmodified or chemically modified single-stranded DNA or RNA entities. In general, they are relatively short molecules consisting in average of 13-25 nucleotides [1]. They are widely used to modulate the gene expression in various fields including research, biotechnology and therapeutic applications [2]. Basically, antisense oligonucleotides hybridize specifically to a complementary DNA, pre-mRNA or mRNA via Watson-Crick base pairing [3], and in this way, they block the expression of a target gene [1,4,5]. Unmodified DNA and RNA oligonucleotides are inherently unstable both in serum and in cells as they are rapidly degraded by exonucleases and endonucleases [1-4,6]. Several modifications have been implemented to increase the nuclease resistance of ONs including modifications of the sugar moiety, the nitrogen base and of the phosphate group [7]. Although these alterations significantly reduce their predisposition to nuclease degradation, they do not completely inhibit it. As additionally, many of such modifications of the ON backbone also reduce the corresponding hybridization activity necessary for the therapeutic action, an adequate balance between nuclease resistivity and activity needs to be found [1,8].

The main synthetic impurities and degradation products encountered in these molecules are very closely related to each other and include mainly adduct sequences (n+1, n+2...) or sequence deletions (n-1, n-2...). In addition, other impurities resulting from oxidation, depurination and other bases are also included in a family of closely related impurities that increase the challenges in the resolution of the main product. Furthermore, the number of impurities can drastically increase when dealing with double stranded interfering RNA (iRNA) as each strand carries its own set of impurities [2,3].
Due to the vast number of structural and chemical variations available for ONs, conventional 1D based separation techniques are too limited for adequate mapping of all impurities and degradants which can be generated [9-11]. Therefore, the development of comprehensive 2D separation methods capable to resolve this increasing sample complexity imposes itself. The approaches developed thus far for the analysis of ONs include a comprehensive liquid chromatography (LC x LC) method using hydrophilic interaction liquid chromatography x ion-pair chromatography (HILIC x IPC) for ONs up to 10 bases in length [12], and a comprehensive liquid chromatography x capillary gel electrophoresis (LC x CGE) method based on ion-pair chromatography (IPC) and ion exchange chromatography (IEC) as LC techniques in combination with CGE using entangled polymer solutions [13]. This latter LC x CGE platform, proved to deliver increased peak capacities compared to the 1D separation techniques for the separation of ONs of therapeutic sizes.

The combination of separation techniques with immobilized enzyme reactors (IMERs) leads to improved characterization of complex mixtures by allowing prior isolation of the ONs prone to the enzymatic degradation. Recently, a methodology combining a DNase I IMER coupled in the pre-column configuration to LC was proposed for ON stability assessments [14]. Pure ON samples were thereby subjected to analysis in order to generate a characteristic degradation profile. However, in the presence of samples containing multiple ONs in combination with other molecules, a characteristic ON degradation profile cannot be achieved independently for each entity.

By merging the concept of comprehensive 2D separations with the use of in-line enzymatic reactors, the development of a methodology that is more selective, which depicts more peak capacity and which allows the generation of specific ON stability information becomes possible. Similar platforms (LC-IMER x LC) have been implemented before for the advanced characterization and/or sample preparation of mixtures of peptides and proteins [15-21].

In this study, we present the highlights from the development of a comprehensive LC-IMER x LC methodology based on ion pair chromatography for separations and applying reactors with immobilized Deoxyribonuclease I. It is illustrated that this tool can be used for comparative assessment of the stability of particular ONs in mixtures of DNA and RNA. To the best of our knowledge, the development of a comprehensive LC-(DNase I) IMER x LC platform applied to ONs had not been developed yet, and could be of relevance as novel tool in the development of antisense ON therapies.
Offline comprehensive liquid chromatography in combination with a Deoxyribonuclease I immobilized enzymatic reactor for selective screening of oligonucleotide mixtures

VIII.2 Experimental

VIII.2.1 Chemicals

Triethylamine (TEA), acetic acid, hydrochloric acid, Tris, CaCl₂, MgCl₂, glycine, (3-Glycidyloxypropyl) trimethoxysilane (GPTMS) and Deoxyribonuclease I (DNase I) from bovine pancreas (≥400 Kunitz units/mg of protein) were purchased from Sigma-Aldrich (Steinheim, Germany). Acetonitrile (ACN) HPLC grade from Fischer Scientific (Loughborough, U.K.) and Milli-Q water (Millipore, Milford, MA) were used. All the mobile phases were filtered through 0.22 µm nylon membrane filters (Grace Davison Discovery Sciences, Lokeren, Belgium). The oligonucleotides: (DNA 12) 12-mer DNA ON (5’-GCA-CAC-CGT-CAG-3’), (AC 15) 15-mer DNA ON (5’-AAA-CCC-AAA-CCC-AAA-3’), (DNA 20) 20-mer DNA ON (5’-AGC-GAT-AAG-ATT-CAT-ATA-TC-3’), 30-mer DNA ON of Homo-deoxythymidine (T30), (R41) 41-mer DNA ON (5’-GTT-GGA-TTA-AAC-ACC-CGT-TCC-CGT-CTC-TAT-CAG-CTT-AGT-GT-3’) and the 2’-OMe RNA 12-mer ON (5’-GCA-CAC-CGU-CAG-3’) were purchased from Eurogentec (Liege, Belgium). The lyophilized ONs were dissolved in an appropriate volume of Milli-Q water in order to obtain 100 µM stock solutions.

VIII.2.2 Instrumentation

An Agilent 1200 series HPLC capillary system composed of a binary pump and a single wavelength detector was employed for the delivery of the chromatographic mobile phases, while the enzymatic reaction buffer was delivered to the reactor by a HP 1050 quaternary pump (Agilent Technologies, Waldbronn, Germany). A CTO-20AC prominence HPLC column oven (Shimadzu, Kyoto, Japan) was used for controlling the temperature of the IMER. The collection of the fractions was carried out using a micro fraction collector Agilent 1200 series. The detector wavelength and sampling frequency were set at 260 nm and 10 Hz, respectively. The contour plots were constructed using MATLAB (Natick, MA, USA).
Chapter VIII

VIII.2.3 Preparation of the IMER

An IMER manufactured and characterized in a previous study has been used in the development of this comprehensive platform [14]. In brief, DNase I was immobilized on epoxy activated silica particles (Nucleosil 5 µm, 4000 Angstrom pore size from Macherey-Nagel, Duren, Germany) using GPTMS as activator, while glycine was used for deactivating the unreacted epoxy groups. The DNase I immobilized silica particles (4.9 ± 0.1 mg of DNase I /g of silica) presented a specific activity of 361.8 ± 23 pmol T30/min·g. This activity was determined using a T30 ON as a substrate in a medium consisting of 50 mM Tris buffer pH 7.5, 5 mM CaCl$_2$ and 5 mM MgCl$_2$ at 37 °C. Afterwards, the enzyme-silica particles were slurry packed into 150 mm x 2.1 mm I.D. stainless steel columns at a constant pressure of 400 bar.

VIII.2.4 Comprehensive IPC-IMER-IPC setup

VIII.2.4.1 First dimension, IPC-IMER

The employed instrumental setup is depicted in Figure VIII-1A. This fully automated procedure can be subdivided into three main events: the chromatographic separation of the ON mixture (i), reaction of the analytes coming from the chromatographic effluent in the IMER (ii), and the collection of fractions eluting from the IMER (iii).

(i) IPC was employed as separation technique in the first dimension. For this, a 100 mm x 1 mm I.D. x 3.5 µm Xbridge C18 column (Waters, Zellik, Belgium) was used. The column temperature was set at 60 °C. The mobile phase was composed of 100 mM TEA (pH 5.5 adjusted with acetic acid) (A) and ACN (B). A flow rate of 20 µL/min was used with the following gradient program: 0–300 min, 0–19.2%B. Pump 1, which delivered the chromatographic separation conditions, was directly connected to the HP 1200 autoinjector (10 µL of injection volume) followed by the chromatographic column.

(ii) Consequently, the effluent of the chromatographic column was mixed with a concentrated reaction buffer solution (pH 7.5; 1.25 M Tris, 55 mM CaCl$_2$ and 55 mM MgCl$_2$) delivered by Pump 2 with the aid of a T-piece followed by a 2 µL mixing loop (0.127 mm I.D.). The flow rate delivered by Pump 2 was set to 2 µL/min in order to approximate a 1/10 dilution of the reaction buffer with the chromatographic mobile phase. The
resulting concentration of the enzyme activating cations (Ca$^{2+}$ and Mg$^{2+}$) and pH in the mixture mobile phase-buffer were the same as to the one tested in the activity of the IMER (Section VIII.2.3). Subsequently, this effluent was percolated at a total flow rate of 22 µL/min through the IMER kept at 37 °C. The system’s and reactor’s dead volumes were determined by injecting a 1 mM caffeine solution. This allowed the estimation of the residence time that the analytes experienced in the IMER (~20 min).

Figure VIII-1. Schematic representation of the LC-IMER-LC setup. A) First dimension LC-IMER, B) second dimension LC. In the first dimension, an ON mixture is separated by IPC and the effluent is subjected to a reaction in the IMER. The IMER effluent is thereafter connected to a fraction collector. The fractions collected from the first dimension are subsequently directly injected into the second dimension and analyzed.
Thereafter, the effluent from the IMER was connected to a UV detector followed by a fraction collector. Finally, fractions of this first dimension were collected every 120 s into 100 µL vial inserts.

**VIII.2.4.2 Second dimension, IPC**

The fractions collected from the first dimension (section VIII.2.4.1) were injected (35 µL injection volume) into a conventional HPLC system setup employing IPC as separation technique (Figure VIII-1B). The separation was carried out in a 50 mm × 4.6 mm I.D. × 3 µm XBridge C18 column (Waters, Zellik, Belgium). The column temperature was maintained at 60 °C. The mobile phase consisted of 100 mM TEA (pH 5.5 adjusted with acetic acid) (A) and ACN (B). A flow rate of 0.5 mL/min was used with the following gradient program: 0-50 min, 0-16%B.

**VIII.3 Results and discussion**

Despite the unavoidable obtained reduction in enzymatic activity after immobilization on a solid support, the obtained gain in mechanical and chemical resistance of the enzyme is one of the most important features of this process when comparing to free enzymes in solution. The employed DNase I IMER proved to have an exceptional life time with intrinsic activities capable to generate degradation profiles of DNA based ONs in a reproducible way when combined with liquid chromatography in the pre-column configuration [14]. In the present study, concept of employing the Deoxyribonuclease based IMERs in the post-column configuration is introduced and evaluated. For this, the preservation or suppression of the enzymatic activity was the key evaluated aspect.

In a previous study, the hyphenation of both ion-pair chromatography (IPC) and ion exchange chromatography (IEC) in an IMER-LC setup proved to be suitable [14]. Those LC techniques are also among the preferred ones when seeking for robust high performance separations of ONs, and were therefore evaluated as candidates for the LC-IMER setup. The harsh properties of the mobile phases of those LC techniques (IPC: TEA 100 mM pH 5.5; IEC: ~1.2 M NaCl pH >10 [14]) are in first instance inadequate for the activity of DNase I. Nevertheless, when mixed or diluted with a buffer of satisfactory ionic strength and buffer capacity, suitable conditions for the enzymatic activity may be
Offline comprehensive liquid chromatography in combination with a Deoxyribonuclease I immobilized enzymatic reactor for selective screening of oligonucleotide mixtures

obtained. In an online LC-IMER setup, the establishment of appropriate conditions for the enzymatic activity may be attained by two strategies: (i) dilution of the mobile phase in the activity buffer, (ii) addition of concentrated activity buffer to the mobile phase. Both possibilities were explored by incubating immobilized DNase I particles in individual vials with the appropriated solvent compositions. The DNase I silica particles were thereby mixed with LC mobile phase compositions which were either diluted or contained an addition of a concentrated activity buffer, this in an effort to simulate conditions that might be obtained the online approach. These test mixtures were stirred for 60 min with the 41-mer DNA ON and afterwards analyzed by IPC.

Figure VIII-2 depicts the chromatograms in which the presence or loss in the enzymatic activity can be visually appreciated. A dilution 1/10 of the IPC mobile phase gradient at the end of the gradient (82% TEA 100 mM pH 5.5, 18% ACN) in the activity buffer (50 mM Tris buffer pH 7.5, 5 mM CaCl$_2$ and 5 mM MgCl$_2$) results in approximately 1.8% of ACN in the final mixture, while the concentration of the cations Ca$^{2+}$ and Mg$^{2+}$ is not significantly changed to diminish the enzymatic activity. The chromatogram corresponding to this latter mixture is depicted in Figure VIII-2A, in which it can be noticed that the enzymatic activity has not been lost, as the resulting chromatogram (solid-red) presents the same degradation peaks as the chromatogram (dashed-green) corresponding to a sample in which the DNase I silica particles were only incubated in the activity buffer. An overlay with a blank chromatogram (dashed-blue) corresponding to the ON incubated with deactivated silica particles without enzyme is also displayed, in which the peak eluting at 32 minutes corresponds to the unreacted ON. The pH of the resulting IPC mobile phase-activity buffer was measured as 7.3, which proves to be adequate for the enzymatic activity. In a similar way, a small addition of a concentrated activity buffer solution to the IPC mobile phase (section VIII.2.4.1) also allowed obtaining adequate, yet somewhat less performant conditions for the enzymatic activity (Figure VIII-2B). It can therefore be visually appreciated that the ON main peak has not completely reacted as was the case in the chromatogram presented in Figure VIII-2A. This effect can be mainly attributed to the higher presence of acetonitrile and due to the increased ionic strength of the resulting mixture when compared to a dilution process of the IPC mobile phase.
Figure VIII-2. IPC chromatograms of an R41 ON representing the effect of mixing IPC and IEC LC mobile phases with the activity buffer over the immobilized DNase I activity. The blue small dotted line chromatogram corresponds to the R41 ON incubated with blank silica particles. The green dotted line chromatogram corresponds to the R41 ON incubated with DNase I silica particles in the activity buffer. The solid red line chromatogram corresponds to the ONs incubated with DNase I silica particles in the respective mixture with the mobile phase activity buffer. The encircled region points out the degradation degree of the ON which can be visually estimated as a decrease of its peak height compared to the blank. A) 1/10 dilution of an IPC mobile phase (MP) in activity buffer (resulting in ~1.8% of B in activity buffer). B) Addition of a concentrated activity buffer solution to the IPC MP (resulting in ~18% of B in activity buffer). C) 1/10 dilution IEC mobile phase (MP) in activity buffer (resulting in ~8.0% of B in activity buffer). (For details refer to sections VIII.2.4 and VIII.3).
In Figure VIII-2C, the effect on the enzymatic activity of the 1/10 dilution of an IEC mobile phase at the end of the gradient (80%: 20/80 ACN/water 1.25 M NaCl (pH 11.5); 20%: 20/80 ACN/water (pH 11.5)) [13] with the activity buffer is presented. It can be noticed that the increased ionic strength of the medium also does not allow for a complete degradation of the ON. Moreover, when performing an addition of activity buffer to the IEC mobile phase, no enzymatic activity could be appreciated (chromatogram not shown), making this approach unsuitable. An accurate estimation of the enzymatic activity at other dilution factors and mobile phase compositions is possible. Nevertheless, its calculation falls out of the main scope of this study at this stage, where the main message is to point out the possibility of coupling DNase I based reactors between both IPC dimensions of an off-line comprehensive set-up.

From a practical point of view, the addition of activity buffer to the LC mobile phase presents several advantages. First and most importantly, there is minimal dilution of the LC effluent which could result in poor detectability in the second dimension. Second, in order to allow degradation of the ON in the IMER, long residence times need to be established, meaning that the cumulative flow rate from combining the LC mobile phase in the first dimension plus the activity buffer needs to be as low as possible. In the described setup (section VIII.2.4.1), a flow rate of 20 µL/min inevitable requires the usage of a column with a narrow I.D. of 1 mm, this in order to operate the column close (yet somewhat below) to the optimal linear velocity to obtain satisfactory chromatographic separation performance. However, under these conditions the separation of a mixture of 6 ONs requires 240 min (Figure VIII-3A). Therefore, the implementation of even lower flow rates cannot be considered convenient. On the other hand a 1/10 dilution of the complete effluent of this mobile phase would result in total flow rate through the IMER of 200 µL/min, for which we encountered insufficient reaction of the ONs using this IMER [14]. This latter approach may be improved by the usage of more active IMERs. Another strategy for implementing the dilution would consist in splitting the LC mobile phase before mixture with the buffer. However, that approach is the less preferable as it complicates the experimental setup in addition to drastically reducing the detectability of the ONs in the second dimension.
Figure VIII-3. A) IPC chromatogram of a mixture of 6 ONs. B) IPC-IMER x IPC contour plot. C) IPC chromatogram of a fraction corresponding to the elution of the R41 ON. For conditions refer to section VIII.2.4.

Figure VIII-3B presents the contour plot of an IPC-IMER x IPC comprehensive analysis of a mixture of 6 ONs. It also confirms that the IMER activity is not lost after performing the online LC-IMER setup with addition of activity buffer to the IPC mobile phase. It can also be observed in Figure VIII-3C that the degradation profile of the 41-mer DNA ON is similar to the one obtained in batch analysis (Figure VIII-2B). Moreover, it can be noticed that the degradation profiles are specific for each ON, proving the applicability of this approach for an improved characterization of mixtures of ONs. In Figure VIII-4,
an overlay of the IPC-IMER x IPC plot with a blank plot generated using a deactivated IMER without enzyme is presented, whereby it can be observed that certain ONs degrade more than others. These differences are attributed to the nucleotide composition, purine pyrimidine ratio and to the ability of an ON to hybridize between itself. DNase I cleaves preferentially double strands [22], this effect can be particularly visualized for the R41 ON, which due to its length and random nucleotide composition is more likely to form double strands between itself when compared to the shorter ONs (DNA12, AC15, and DNA20).

Figure VIII-4. 3D plot overlay of an IPC-IMER x IPC and IPC-(blank IMER) x IPC analysis of a mixture of 6 ONs. For conditions refer to section VIII.2.4.
Additionally, the flexibility of the DNA strand also influences the degree of cleavage imposed by the DNase I [23]. DNA regions rich in A and T nucleotides are less flexible [24], explaining the lower cleavage observed for the T30 ON. The selectivity of the enzyme is also illustrated through its degradation of the DNA molecules while imposing no degradation on the RNA bases ONs. The molecules used here offer a limited number of examples to illustrate the high selectivity provided by IMERs in such a comprehensive platform.

VIII.4 Conclusion

A comprehensive IPC-IMER x IPC platform for an improved characterization of mixtures of ONs has been developed. A DNase I IMER was successfully coupled online with IPC in the post column configuration. This proves the possibility for the implementation of other orthogonal LC techniques such as IEC in the first dimension and/or second dimension for instance. The hyphenation of MS detection with a suitable LC technique in the second dimension is realistic and would provide additional valuable information on the cleaved ON fragments. Moreover, the usage of an IMER combined with LC provides reproducible reactions over longer periods of time, as demonstrated before [14]. The possibility to couple IMERs with nucleases which are specific or not to the sugar moiety in addition to their cleavage mechanism (endonuclease, exonuclease) may also be explored. This last approach may be performed either by the simultaneous immobilization of various nucleases to a support or by the interconnection of IMER subunits as many nucleases show similar activation conditions. Finally, the introduction of an enzymatic reaction in a comprehensive type of platform expands its selective power, while providing unique data regarding the stability of mixtures of ONs towards nucleases. The approach offer potential applications for the mapping, characterization and stability assessments of ONs.

VIII.5 References

Offline comprehensive liquid chromatography in combination with a Deoxyribonuclease I immobilized enzymatic reactor for selective screening of oligonucleotide mixtures

General conclusion

One dimensional chromatographic separation techniques are limited in terms of the capability they offer for the separation of oligonucleotides (ONs). The slower diffusivity of such solutes compared to smaller molecules (< 1000 Da) typically leads to diminished chromatographic separation efficiencies under the conventional flow rates used in HPLC (e.g. 0.3-1 mL/min) due the slower mass transfer. Additionally, the synthetic nature of ONs in addition to the various chemical modification strategies used to improve their resistance towards nucleases lead to increased sample complexities which are rapidly exceeding the separation capabilities of the commonly employed one dimensional HPLC approaches such as Ion-pair chromatography (IPC), Ion-exchange chromatography (IEC) and Hydrophilic interaction liquid chromatography (HILIC). Electrodriven techniques such as capillary gel electrophoresis on the other can provide somewhat improved, but still insufficient peak capacities to unambiguously separate the broad variety of impurities which can theoretically be formed during manufacturing of ONs.

One of the strategies to improve the peak capacity in separation techniques comprises the combination of one dimensional analytical approaches into two dimensional separation platforms. Under this approach, the comprehensive approach subjects all the fractions of one dimension to a second dimension resulting in an improved peak capacity that in theory is the product of the individual single dimension peak capacities. In the first research component of this work a successful increase in the separation capability of complex mixtures of oligonucleotides has been attained by the combination of liquid chromatography with capillary electrophoresis under such a comprehensive off-line approach, as described in chapter IV. This approach offers the possibility to use IPC or IEC in the first dimension in combination with CE or multiplexed CE in entangled polymer
General conclusion

solutions in the second dimension, delivering an effective tool for the separation of ONs based on their size and nucleotide composition. The technique has been developed for UV detection as the vast majority of routine drug purity measurements are performed with this robust detector.

Subsequently the performance of an in-house developed deconvolution algorithm was tested to improve the separation of partially co-eluting ON signals of various types as described in chapter V.

Although the strategy was successful for the deconvolution of fairly dissimilar ONs the approach offered limited benefits for ONs differing less in composition, as is typically the case between active pharmaceutical ingredients and the associated impurities.

As mentioned, many chemical modifications have been implemented to enhance the stability of ONs in vivo. The current strategies for evaluating their stability include their incubation in solution with nucleases followed by their analysis. This approach presents limitations regarding reproducibility and throughput. The hyphenation of an immobilized enzyme reactor (IMER) with LC for the stability evaluation of ONs towards nucleases was investigated in the second research section of this work. The endonuclease DNase I was successfully immobilized on silica particles which were packed afterwards into a column to manufacture the IMERs. Afterwards, this IMER was coupled to LC in an on-line pre-column configuration as described in chapter VII. A successful hyphenation with IPC and IEC was achieved, generating a novel tool allowing on-line evaluation of the nuclease resistance of various ONs. The repeatability and reproducibility of the developed methodology proved satisfactory as analyses of the degraded ONs demonstrated very comparable profiles over extended periods of time. The developed on-line IMER-IPC approach might allow for possible coupling to MS, offering additional information on the cleaved ONs fragments. Another foreseeable improvement is the immobilization of other nucleases that cleave DNA or RNA in different ways. A drawback of the approach described in chapter VI, is that it requires introduction of pre-purified ON’s in the system for the proper evaluation of the corresponding nuclease resistance.

As a solution for testing nuclease resistance and to characterize mixtures of ONs, the hyphenation of a DNase I IMER in an off-line comprehensive LC x LC platform was investigated. This approach described in chapter VIII, illustrates the principle of this approach for the characterization of mixtures of ONs. Further developments on comprehensive platforms integrating IMERs with different enzymes or interconnections of IMER subunits may allow for high throughput selective analysis of ONs in
various complex matrixes. This approaches contribute to support for the future the incursion of biomolecules to enhance the current analytical techniques.
Since information on the human genome become available, an increased rise in the antisense technology has been observed. Under this approach, a single stranded DNA or RNA fragment is used to inhibit the formation of proteins. Those DNA or RNA based single stranded short fragments, better known as oligonucleotides (ONs), have been implemented as therapeutic agents amongst a variety of other biochemical applications. Concomitantly there has been a growing need for the development of additional and improved analytical tools for the analysis of the quality and purity of these solutes. Overall the conventional analytical techniques typically employed for the analysis of such molecules provide individually too limited separation capabilities to address complex mixtures of oligonucleotides. Therefore, a first research component of this work comprised the development of analytical separation methodologies offering improved characterization capabilities for oligonucleotide analysis.

In Chapter II the principles of the separation techniques typically used for the analysis of oligonucleotides such as liquid chromatography and capillary electrophoresis are described. The various separation mechanisms, the differing experimental conditions and technical implementation are thereby described. This to allow for a better comprehension of the hyphenated and multidimensional methodologies described further. Additionally, a short explanation on the approaches used to assess the performance of uni- and multidimensional separations is provided. An overview covering the most relevant properties of oligonucleotides and of the current analytical techniques used for ON analysis is provided in Chapter III.

In Chapter IV, two offline two-dimensional liquid chromatography (LC) × capillary gel electrophoresis (CGE) and LC × (24) multiplexed-CGE methodologies for the separation of oligonucleotides of therapeutic size are described. Both ion-pair chromatography (IPC) and ion-exchange chromatography (IEC) were thereby studied as methods in the first dimension HPLC separations. Single and multiplexed capillary electrophoresis methods in entangled polymer solutions were used in the second dimension separations. The performance of electrokinetic and of pressure injection were compared for the analysis of the collected LC fractions and the comprehensive separation was
optimized with standard mixtures of poly adenosine, thymidine, cytosine and uracil homodeoxyoligonucleotides up to 35 bases long. Highly orthogonal methodologies and overall peak capacities of 6435 and 6993 for IPC × CGE and IEC × CGE, respectively, were obtained within a few hours analysis time.

In Chapter V, a chemometrical deconvolution method is discussed as it provides a fast alternative *in silico* tool for the separation of partially overlapping ON signals that present certain UV spectral differences, and it might be easily implemented as an additional feature in LC-PDA, CGE-PDA and LCxCGE-PDA ON analyses. The in-house-developed MATLAB script differentiates between the ONs based on their dissimilarity in the UV-spectrum and retention time, using the entire 3D retention time UV-spectrum. Consequently, the script generates individual signal chromatograms and UV-spectra for each of the compounds. The possibilities and limitations of the deconvolution script in the separation of partially overlapping ONs were explored. Homo-oligonucleotides and ONs with random base composition were used to evaluate this computational tool.

In Chapter VI, an overview of the main aspects related to the manufacturing and characterization of immobilized enzyme reactors (IMERs) for hyphenation with LC is covered to allow the reader to better understand the new proposed methodologies described in chapters V and VI. Emphasis is thereby set on typically used manufacturing procedures relevant to this work and on providing an introduction in kinetic studies required for characterization of IMER performance.

In Chapter VII, the development of an on-line approach based on an immobilized enzyme reactor (IMER) coupled to liquid chromatography (LC) is described as an alternative methodology for improved ON stability testing. More in detail, Deoxyribonuclease I (DNase I) was immobilized on epoxy-silica particles of different pore sizes and packed into a column for the construction of an IMER. Subsequently, the hyphenation of the IMER with ion-pair chromatography (IPC) and ion-exchange chromatography (IEC) was evaluated, leading to the successful development of two on-line methodologies: IMER-IPC and IMER-IEC. More specifically, natural and modified DNA and RNA oligonucleotides were used for testing the performance of the methodologies. Both methodologies proved to be simple, automatable, fast and highly reproducible for the quantitative and qualitative evaluation of ON degradation. In addition, the extended IMER life time in combination with a more
straightforward control of the reaction kinetics substantiate the applicability of the IMER-LC platform for ON stability tests and its implementation in routine and research laboratories.

Finally, in Chapter VIII, the development of a comprehensive ion-pair chromatography-immobilized enzyme reactor x ion-pair chromatography (IPC-IMER x IPC) methodology for the advanced characterization of DNA/RNA oligonucleotides (ON) mixtures is described. DNase I IMER was thereby coupled post-column to IPC, followed by the collection of the eluting fractions and by offline reanalysis by IPC. The influence of the mobile phase over the IMER activity was qualitatively evaluated. The methodology proved to generate relevant ON degradation profiles that might be correlated with the ON stability towards nuclease. The developed platform shows potential for further implementation in selective analysis of ON mixtures and in mapping studies.
Samenvatting

Nu er steeds meer informatie over het menselijk genoom beschikbaar is, wordt er ook een verhoging van de anti-sense technologie waargenomen. Bij deze benadering wordt een enkelstreng DNA of RNA-fragment daarbij gebruikt om de vorming van eiwitten te remmen. Deze DNA of RNA gebaseerde korte enkelstreng fragmenten, beter bekend als oligonucleotiden (ON), worden geïmplementeerd als therapeutische middelen en ook voor andere verschillende biochemische toepassingen. Daarnaast, er is een groeiende behoefte aan de ontwikkeling van bijkomende en verbeterde analytische methoden voor de analyse van de kwaliteit en zuiverheid van deze moleculen. De conventionele analytische technieken die typisch gebruikt worden voor de analyse van oligonucleotiden bieden te beperkte scheidingsmogelijkheden om complexe mengsels van oligonucleotiden aan te pakken. Een eerste onderzoekscomponente van dit werk omvatte daarom de ontwikkeling van analytische scheidingsmethodologieën die verbeterde mogelijkheden vertonen voor de karakterisatie van ONs.

De beginselen van de scheidingstechnieken die typisch gebruikt worden voor de analyse van oligonucleotiden zoals vloeistofchromatografie en capillaire elektroforese worden beschreven in hoofdstuk II. De verschillende scheidingsmechanismen, de verschillende experimentele condities en technische uitvoering worden hier beschreven. Dit laat een beter begrip toe van de gecombineerde en multidimensionale methoden die in de verdere secties worden behandeld. Bovendien wordt een korte uitleg beschreven van de benaderingen die gebruikt zijn om de prestaties van uni- en multidimensionele scheidingen te evalueren. Een overzicht met de meest relevante eigenschappen van oligonucleotiden en van de huidige analytische technieken die worden gebruikt voor de analyse van ONs worden in hoofdstuk III beschreven.

In hoofdstuk IV worden twee off-line tweedimensionale vloeistofchromatografie (LC) x capillaire gelelectroforese (CGE) en LC x (24) multiparallele-CGE methoden voor de scheidings van oligonucleotiden met therapeutische grootte beschreven. Ion-paar chromatografie (IPC) en ionenuitwisselingschromatografie (IEC) werden daarbij onderzocht als eerste dimensie scheidingsmethoden. Enkelvoudige en multiparallele capillaire elektroforese methoden met behulp van verknoopte polymeeroplossingen werden in de tweede scheidingsdimensie onderzocht. De
Samenvatting

prestaties van elektrokinetische en drukinjectie werden vergeleken voor de analyse van de verzamelde LC fracties, en de volledige comprehensieve scheiding werd geoptimaliseerd met standaard mengsels van poly-adenosine, thymidine, cytosine en uracil homodeoxyoligonucleotides tot 35 basen lang. Sterk orthogonale methoden en algehele piekcapaciteiten van 6435 en 6993 voor de IPC × CGE en IEC × CGE, respectievelijk, werden binnen een paar uur tijd analyse verkregen.


In hoofdstuk VI wordt een overzicht gegeven van de belangrijkste aspecten van de productie en karakterisering van geïmmobiliseerde enzym-reactoren (IMERs) voor de combinatie met LC, zodat de lezer een beter inzicht heeft in de voorgestelde nieuwe methodologieën die in de hoofdstukken V en VI beschreven zijn. De nadruk ligt op de conventionele gebruikte productie procedures die in dit werk worden gebruikt. Daarnaast wordt ook een introductie uiteengezet over kinetische studies die nodig zijn voor de karakterisering van IMERs.

In hoofdstuk VII wordt de ontwikkeling beschreven van een online benadering op basis van geïmmobiliseerde enzym-reactoren (IMER) gekoppeld met vloeistofchromatografie (LC) als een alternatieve methode voor verbeterde stabilitesten. Meer in detail, Deoxyribonuclease I (DNase I) werd geïmmobiliseerd op epoxy-silicadeeltjes met verschillende poriegroottes en gepakt in een kolom voor de bouw van een IMER. Vervolgens werd de koppeling van de IMER met ionenpaar chromatografie (IPC) en ionenuitwisselingschromatografie (IEC) geëvalueerd, wat leidde tot de succesvolle ontwikkeling van twee on-line methodes: IMER-IPC en IMER-IEC. Natuurlijke en
Samenvatting
gemodificeerde DNA en RNA oligonucleotiden werden gebruikt voor het testen van de prestaties van de methoden. Beide methoden bleken eenvoudig, automatiseerbaar, snel en zeer reproduceerbaar voor de kwantitatieve en kwalitatieve evaluatie van ON degradatie. Bovendien onderbouwt de verlengde IMER levensduur, in combinatie met een eenvoudigere regeling van de reactiekinetiek, de toepasselijkheid van het IMER-LC platform voor ON stabiliteitstesten en voor gebruik in routine- en onderzoekslaboratoria.

Tenslotte wordt in hoofdstuk VIII de ontwikkeling beschreven van een uitgebreid ioniennpaarchromatografie-geïmmobiliseerde enzymreactor x ioniennpaar chromatografie (IPC-IMER x IPC) methode voor de geavanceerde karakterisatie van DNA / RNA oligonucleotiden (ON) mengsels. DNase I IMER was hierbij in post-kolom configuratie gekoppeld met IPC. Vervolgens werden alle eluerende fracties dan verzameld voor offline analyse door IPC. De invloed van de mobiele fase over de IMER activiteit werd kwalitatief geëvalueerd. De methode bleek relevant om ON degradatieprofielen te genereren die kunnen worden gecorreleerd met de ON stabiliteit met betrekking tot nucleasen. Het ontwikkelde platform toont potentieel voor verdere implementatie in de selectieve analyse van ON mengsels en in mapping studies.
List of abbreviations

$\mu_e$  Electrophoretic mobility

[P]  Product concentration

[S]  Substrate concentration

$<\beta>$  First dimension broadening factor

$\mu_a$  Apparent solute mobility

$\mu_{eo}$  Electroosmotic mobility

1D  One dimensional

2D  Two dimensional

ACN  Acetonitrile

Ads  Adsorption

AIDS  Acquired immune deficiency syndrome

AMP  Adenosine monophosphate

ARGO 2  Argonaute 2 (enzyme)

$A_{sp}$  Specific activity

BDMA  Butyldimethylammonium

C  Polymer concentration

C  Chromatographic profile

CDI  1,1-carbonyldiimidazole

CE  Capillary electrophoresis

CGE  Capillary gel electrophoresis

CIM  Convective Interaction Media

$C_M$  Concentration of the analyte in the mobile phase

cm  centimeter

$C_m$  Constant of the C term of the Van Deemter equation

CMV  Cytomegalovirus

CPG  Controlled pore glass

$C_s$  Concentration of the analyte in the stationary phase

CZE  Capillary zone electrophoresis

D  Convoluted data matrix

Det  Detection

DICER  Endoribonuclease Dicer (enzyme)

$D_i$  Diffusion

$D_m$  Diffusion coefficient

DNA  Deoxyribonucleic acid

Dnase I  Deoxyribonuclease I

$d_p$  Particle diameter

DSS  Disuccinimidylsuberate

E  Electric field strength
### List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EOF</td>
<td>Electroosmotic flow</td>
</tr>
<tr>
<td>ESI</td>
<td>Electrospray ionization</td>
</tr>
<tr>
<td>Ε</td>
<td>Permittivity or dielectric constant of the medium</td>
</tr>
<tr>
<td>F</td>
<td>Flow rate</td>
</tr>
<tr>
<td>$F_E$</td>
<td>Electric force</td>
</tr>
<tr>
<td>$F_F$</td>
<td>Frictional force</td>
</tr>
<tr>
<td>GPTMS</td>
<td>(3-Glycidoxypropyl) trimethoxysilane</td>
</tr>
<tr>
<td>H</td>
<td>Column plate height</td>
</tr>
<tr>
<td>h</td>
<td>Peak height</td>
</tr>
<tr>
<td>HFIP</td>
<td>Hexafluoroisopropanol</td>
</tr>
<tr>
<td>HILIC</td>
<td>Hydrophilic interaction liquid chromatography</td>
</tr>
<tr>
<td>HPLC</td>
<td>High performance liquid chromatography</td>
</tr>
<tr>
<td>HPMC</td>
<td>(Hydroxypropyl)methyl cellulose</td>
</tr>
<tr>
<td>IAM</td>
<td>Immobilized artificial membrane</td>
</tr>
<tr>
<td>IEC</td>
<td>Ion-exchange chromatography</td>
</tr>
<tr>
<td>IMER</td>
<td>Immobilized enzyme reactor</td>
</tr>
<tr>
<td>lnj</td>
<td>Injection</td>
</tr>
<tr>
<td>IP</td>
<td>Ion paring</td>
</tr>
<tr>
<td>IPC</td>
<td>Ion-pair chromatography</td>
</tr>
<tr>
<td>iRNA</td>
<td>Interference RNA</td>
</tr>
<tr>
<td>$k'$</td>
<td>Retention factor</td>
</tr>
<tr>
<td>$k_{cat}$</td>
<td>Catalytic constant</td>
</tr>
<tr>
<td>$K_M$</td>
<td>Michaelis–Menten constant</td>
</tr>
<tr>
<td>$K_V$</td>
<td>Kilovolt</td>
</tr>
<tr>
<td>$K_{VO}$</td>
<td>Permeability</td>
</tr>
<tr>
<td>L</td>
<td>Column length or capillary length</td>
</tr>
<tr>
<td>l</td>
<td>Effective capillary length</td>
</tr>
<tr>
<td>LC</td>
<td>Liquid chromatography</td>
</tr>
<tr>
<td>LNA</td>
<td>Locked nucleic acid</td>
</tr>
<tr>
<td>LOD</td>
<td>Limit of detection</td>
</tr>
<tr>
<td>LOQ</td>
<td>Limit of quantification</td>
</tr>
<tr>
<td>m</td>
<td>mg of enzyme in solution or immobilized</td>
</tr>
<tr>
<td>MALDI</td>
<td>Matrix-assisted laser desorption/ionization</td>
</tr>
<tr>
<td>MCR-ALS</td>
<td>Multivariate curve resolution-alternating least squares algorithm</td>
</tr>
<tr>
<td>$m_M$</td>
<td>Quantity of an analyte in the mobile phase</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>MS</td>
<td>Mass spectrometry</td>
</tr>
<tr>
<td>$m_s$</td>
<td>Quantity of an analyte in the stationary phase</td>
</tr>
<tr>
<td>$M_W$</td>
<td>Molecular weight</td>
</tr>
<tr>
<td>N</td>
<td>Column plate number (Separation efficiency)</td>
</tr>
<tr>
<td>n</td>
<td>Peak capacity (multidimensional separations)</td>
</tr>
<tr>
<td>η</td>
<td>Viscosity</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>$N_A$</td>
<td>Avogadro number</td>
</tr>
<tr>
<td>NARP</td>
<td>Non-aqueous reversed phase chromatography</td>
</tr>
<tr>
<td>nm</td>
<td>nanometers</td>
</tr>
<tr>
<td>NPC</td>
<td>Normal phase chromatography</td>
</tr>
<tr>
<td>ON</td>
<td>Oligonucleotide</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>$P_c$</td>
<td>Peak capacity</td>
</tr>
<tr>
<td>PCA</td>
<td>Principle component analysis</td>
</tr>
<tr>
<td>PDA</td>
<td>Photodiode array detector</td>
</tr>
<tr>
<td>PEI</td>
<td>Polyethyleneimine</td>
</tr>
<tr>
<td>PNA</td>
<td>Peptide nucleic acid</td>
</tr>
<tr>
<td>PPO</td>
<td>Poly-propylene oxide</td>
</tr>
<tr>
<td>PS-DVB</td>
<td>Poly(styrene-divinylbenzene)</td>
</tr>
<tr>
<td>$q$</td>
<td>Ion charge</td>
</tr>
<tr>
<td>$r$</td>
<td>Column radius</td>
</tr>
<tr>
<td>$R_g$</td>
<td>Radius of gyration</td>
</tr>
<tr>
<td>RISC</td>
<td>RNA-induced silencing complex</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RNase H</td>
<td>Ribonuclease H</td>
</tr>
<tr>
<td>RPC</td>
<td>Reversed phase chromatography</td>
</tr>
<tr>
<td>$R_s$</td>
<td>Resolution</td>
</tr>
<tr>
<td>RSD</td>
<td>Residual standard deviation</td>
</tr>
<tr>
<td>$S$</td>
<td>Spectral profile</td>
</tr>
<tr>
<td>SEC</td>
<td>Size exclusion chromatography</td>
</tr>
<tr>
<td>$S_i$</td>
<td>Conductivity</td>
</tr>
<tr>
<td>SVD</td>
<td>Single value decomposition</td>
</tr>
<tr>
<td>$t$</td>
<td>Migration time</td>
</tr>
<tr>
<td>$t_0$</td>
<td>Void time</td>
</tr>
<tr>
<td>$t_1-t_2$</td>
<td>Elution time window</td>
</tr>
<tr>
<td>TBA</td>
<td>Tributylamine</td>
</tr>
<tr>
<td>TEA</td>
<td>Triethylamine</td>
</tr>
<tr>
<td>TEAB</td>
<td>Triethylammonium bicarbonate</td>
</tr>
<tr>
<td>TEEA</td>
<td>Triethylammonium acetate</td>
</tr>
<tr>
<td>temp</td>
<td>Temperature gradients</td>
</tr>
<tr>
<td>$t_f$</td>
<td>Retention time of the final peak</td>
</tr>
<tr>
<td>$t_G$</td>
<td>Gradient run time</td>
</tr>
<tr>
<td>$t_i$</td>
<td>Retention time of the first peak</td>
</tr>
<tr>
<td>TLC</td>
<td>Thin layer chromatography</td>
</tr>
<tr>
<td>$T_R$</td>
<td>Retention time</td>
</tr>
<tr>
<td>$t_s$</td>
<td>Sampling time</td>
</tr>
<tr>
<td>$u$</td>
<td>Linear velocity</td>
</tr>
<tr>
<td>$U$</td>
<td>Enzymatic activity</td>
</tr>
<tr>
<td>UHPLC</td>
<td>Ultrahigh pressure liquid chromatography</td>
</tr>
</tbody>
</table>
### List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>V</td>
<td>Voltage</td>
</tr>
<tr>
<td>V</td>
<td>Enzymatic reaction rate</td>
</tr>
<tr>
<td>VBA</td>
<td>Visual Basic for Applications</td>
</tr>
<tr>
<td>$V_M$</td>
<td>Volume of the mobile phase in a column</td>
</tr>
<tr>
<td>$V_S$</td>
<td>Volume of the stationary phase in a column</td>
</tr>
<tr>
<td>$W_b$</td>
<td>Peak width at the base</td>
</tr>
<tr>
<td>$W_h$</td>
<td>Peak width at half height</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Selectivity factor</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Phase ratio</td>
</tr>
<tr>
<td>$\beta_0$</td>
<td>Spreading angle (degree of similarity between two dimensions)</td>
</tr>
<tr>
<td>$\Delta\mu_a$</td>
<td>Difference between the apparent mobilities of two peaks</td>
</tr>
<tr>
<td>$\Delta P$</td>
<td>Column pressure drop</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Term dependent on the column packing</td>
</tr>
<tr>
<td>$\mu_0$</td>
<td>Average free mobility of oligonucleotides</td>
</tr>
<tr>
<td>$\mu_m$</td>
<td>micrometer</td>
</tr>
<tr>
<td>$\nu_e$</td>
<td>Ion velocity</td>
</tr>
<tr>
<td>$\xi$</td>
<td>Zeta potential</td>
</tr>
<tr>
<td>$\xi_m$</td>
<td>Average mesh size</td>
</tr>
<tr>
<td>$\pi$</td>
<td>Pi number</td>
</tr>
<tr>
<td>$\rho_p$</td>
<td>Polymer density</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>$\tau$</td>
<td>base line resolution</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Obstruction factor of the packing bed</td>
</tr>
<tr>
<td>$\phi$</td>
<td>Volume fraction</td>
</tr>
<tr>
<td>$\phi_*$</td>
<td>Entanglement threshold</td>
</tr>
<tr>
<td>$\overline{\mu}_a$</td>
<td>Average apparent mobility</td>
</tr>
<tr>
<td>$\overline{W}$</td>
<td>Average peak width</td>
</tr>
<tr>
<td>$\overline{W_h}$</td>
<td>Average peak width at half height</td>
</tr>
</tbody>
</table>
LIST OF PUBLICATIONS


- Alvarez-Porebski, P., Hernandez-Garcia-Diego, L., Gomez-Ruiz, H., “Comparison of Soxhlet extraction, ultrasonic bath and focused microwave extraction techniques for the simultaneous extraction of PAH’s and pesticides from sediment samples”, Scientia Chromatographica, 6(2), (2014), 124-138.

WORKS PRESENTED IN CONGRESSES AND SYMPOSIA


- 2014. Combining Multiplexed Gel Capillary Electrophoresis with Liquid Chromatography for Offline Comprehensive Analysis of Complex Oligonucleotide Samples. 13th International
Curriculum vitae

Symposium on Hyphenated Techniques in Chromatography and Separation Technology. Bruges, Belgium.

- **2013.** Off-line comprehensive IP-RPLC x CE and IEX-LC x CE for oligonucleotide impurity mapping. HPLC 2013. Amsterdam, Netherlands.


- **2011.** Study on the water quality in Lacantún river basin. XXIV National Congress on Analytical Chemistry, Mexico.

- **2010.** Basal studies for the evaluation of water quality in the Lacantún river basin. Symposium of Environmental Monitoring of Aquatic Systems in Tropical Zones, Chiapas, Mexico.


- **2009.** Evaluation of extraction techniques for Polycyclic Aromatic Hydrocarbons and Pesticides from sediment samples of the Lacantún river basin. 44th Mexican Congress on Chemistry, Mexico.

- **2004.** Spectrophotometric determination of caffeine in coffee, tea and carbonated beverages samples. XVII National Congress on Analytical Chemistry, Mexico.

SCIENTIFIC AWARDS


- Second Poster Award of HTC-13 for the most innovative poster contribution (from ± 180 posters). “Combining Multiplexed Gel Capillary Electrophoresis with Liquid Chromatography for Offline Comprehensive Analysis of Complex Oligonucleotide Samples”. 13th international symposium on Hyphenated Techniques in Chromatography and separation technology, Bruges, Belgium. January 2014.
In first place I would like to thank my promotor Professor Frederic Lynen for giving me the opportunity to be part of his research group and for all the support during my PhD studies, whose valuable knowledge and feedback allowed me to successfully complete this doctoral research.

To Professor Pat Sandra whom introduced me to the research group and also gave me the opportunity to form part of it.

To Professor Annemieke Madder for proving a number of oligonucleotide samples and orientation during part of this research.

To all my colleagues from the separation science group: Pieter, Maarten, Seppe, Sander, Mike, Kevin, Tim, Jente, Engda, Kai, Olivier, Tomas and all other, whose appreciated company, help and advice made an enjoyable unique experience during my PhD studies.

To all the administrative personnel of the S4 building whose valuable work allows the adequate and smooth conditions for performing a doctoral research.

To the Separation Science Group of Ghent University, the National Council of Science and Technology of Mexico (CONACyT) and the Special Research Fund (BOF) of Ghent University for financially supporting this project.

To all the members of the examination committee for the effort of evaluating the thesis and feedback.

To all the friends that I met during this doctoral research period, with all whom unforgettable experiences were lived and made possible to enjoy the long research pathway.
Acknowledgements

I also would like to thank professors in Mexico: Humberto Gomez, Lucia Hernandez, Maria del Pilar Canizares and Carlos Ramirez for motivating me to pursue a doctoral degree. The knowledge and experiences I acquired while working with them form a valuable part of my academic formation.

My sincere gratitude to Sibel, for being a very valuable part of the doctoral studies experience and for her moral support and advises especially during the last stage of my doctorate.

Finally and most important, I would like to thank my family. In particular my parents, my mother Barbara, my father Francisco and my sister Anabel. Without you all this would not be possible. Thanks for all your help in the different aspects, your moral support and advises which allowed me to successfully accomplish this doctoral studies. Despite the distance, I never felt your absence.