BIOMOLECULES AS RECOGNITION ELEMENTS FOR BIOACTIVE POLYPHENOLS IN COFFEE

CHIM/06

Ph.D. Student
Valentina Sinisi

Ph.D. program Coordinator and Thesis Supervisor
Dr. Federico Berti

Thesis Co-supervisor
Dr. Cristina Forzato

ANNO ACCademico 2012/2013
Odi et amo. Quare id faciam, fortasse requiris.
Nescio, sed fieri sentio et excrucior.*

Catullus, carmen 85

* I hate and I love. Wherefore would I do this, perhaps you ask?
I do not know. But I feel that it happens and I am tormented.
# Table of contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table of contents</td>
<td>i</td>
</tr>
<tr>
<td>Abstract</td>
<td>v</td>
</tr>
<tr>
<td>List of abbreviations</td>
<td>xii</td>
</tr>
<tr>
<td>Chapter 1 - Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Coffee polyphenols</td>
<td>3</td>
</tr>
<tr>
<td>1.1.1 Coffee: a successful beverage</td>
<td>3</td>
</tr>
<tr>
<td>1.1.2 The composition of coffee</td>
<td>3</td>
</tr>
<tr>
<td>1.1.3 The aroma of coffee</td>
<td>5</td>
</tr>
<tr>
<td>1.1.4 The taste of coffee</td>
<td>5</td>
</tr>
<tr>
<td>1.1.5 Phenolic compounds</td>
<td>6</td>
</tr>
<tr>
<td>1.1.5.1 Chlorogenic acids (CGAs)</td>
<td>7</td>
</tr>
<tr>
<td>1.1.5.2 Thermal transformation of CGAs</td>
<td>9</td>
</tr>
<tr>
<td>1.1.5.3 Chlorogenic lactones (CGLs)</td>
<td>10</td>
</tr>
<tr>
<td>1.1.5.4 Chlorogenic compounds are absorbed in humans</td>
<td>12</td>
</tr>
<tr>
<td>1.2 Recognition and sensing elements</td>
<td>19</td>
</tr>
<tr>
<td>1.2.1 Sensors: biomolecules and their mimics</td>
<td>19</td>
</tr>
<tr>
<td>1.2.2 Molecular recognition elements</td>
<td>20</td>
</tr>
<tr>
<td>1.2.2.1 Antibodies</td>
<td>20</td>
</tr>
<tr>
<td>1.2.2.2 Receptors</td>
<td>21</td>
</tr>
<tr>
<td>1.2.2.3 Lectins</td>
<td>21</td>
</tr>
<tr>
<td>1.2.2.4 Short peptides</td>
<td>21</td>
</tr>
<tr>
<td>1.2.2.5 Aptamers</td>
<td>24</td>
</tr>
<tr>
<td>1.2.2.6 Peptide Nucleic Acids</td>
<td>25</td>
</tr>
<tr>
<td>1.3 Human Serum Albumin</td>
<td>31</td>
</tr>
<tr>
<td>1.3.1 Human Serum Albumin as a starting point</td>
<td>31</td>
</tr>
<tr>
<td>1.3.2 HSA structure and properties</td>
<td>31</td>
</tr>
<tr>
<td>1.3.3 Binding sites</td>
<td>32</td>
</tr>
<tr>
<td>1.3.3.1 Site I</td>
<td>34</td>
</tr>
</tbody>
</table>
### Table of contents

1.4 HSA100 and mutant libraries .................................................................................. 41
   1.4.1 Directed evolution strategy ........................................................................ 41
   1.4.2 HSA100 ........................................................................................................... 41
     1.4.2.1 HSA100 identification and optimization ............................................... 41
     1.4.2.2 Expression of HSA100 ......................................................................... 42
   1.4.3 HSA100 mutant libraries ............................................................................. 44
     1.4.3.1 Structural analysis ............................................................................... 44
     1.4.3.2 HSA100 site-directed mutagenesis .................................................... 45

Chapter 2 - Aim of research ......................................................................................... 47

2. Aim of research ......................................................................................................... 49

Chapter 3 - Results and discussion .............................................................................. 51

3.1 Quinides synthesis .................................................................................................. 53
   3.1.1 Synthesis of 3,4-O-dicaffeoyl-1,5-γ-quinide (7) ........................................ 53
   3.1.2 Synthesis of 1,3,4-O-tris[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide (10), 3,4-O-bis[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide (11), and 3-O-[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide (12) ........................................... 55
   3.1.3 Synthesis of 3-O-[4-(aminobutanamidoethoxy)feruloyl]-4-O-[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide (28) .......................................................... 59

3.2 HSA, HSA100 and its mutants: interaction studies .................................................. 65
   3.2.1 Human serum albumin (HSA) .................................................................. 65
     3.2.1.1 The choice of fluorescence spectroscopy ........................................... 65
     3.2.1.2 Titration assays ................................................................................. 66
     3.2.1.3 HSA fluorescence quenching ............................................................ 67
     3.2.1.4 Stern-Volmer analysis ..................................................................... 69
     3.2.1.5 Dissociation constants and binding sites ......................................... 71
   3.2.2 HSA100 and its mutants ............................................................................ 77
     3.2.2.1 Binding biomolecules and titration assays ....................................... 77
     3.2.2.2 Fluorescence quenching and dissociation constants ....................... 79

3.3 Surface Plasmon Resonance and Biacore assay ....................................................... 83
   3.3.1 Surface Plasmon Resonance .................................................................... 83
   3.3.2 Evaluation of the association constant ..................................................... 85
   3.3.3 The experiment ......................................................................................... 86
3.4 NMR investigation on the caffeine-quinides interaction ........................................ 91
  3.4.1 Caffeine interacts with polyphenols ................................................................. 91
  3.4.2 Study of the interaction between caffeine and the synthesized quinides ........... 91
    3.4.2.1 Titrations of 7, 10, 11 and 12 with caffeine in DMSO-d6/D2O 75-25% ...... 92
    3.4.2.2 Titrations of 7 with caffeine in D2O/DMSO-d6 75-25% ......................... 96
3.5 Antiviral assays ........................................................................................................ 103
  3.5.1 Further investigations of the bioactivity of the chlorogenic compounds .......... 103
    3.5.2 Influenza virus ................................................................................................. 104
      3.5.2.1 Influenza virus polymerase .................................................................. 104
      3.5.2.2 Influenza endonuclease enzymatic assay ........................................... 105
      3.5.2.3 vRNP reconstitution assay .................................................................. 106
      3.5.2.4 Anti-influenza virus activity ................................................................. 107
  3.5.3 Human immunodeficiency virus (HIV) ............................................................ 108
    3.5.3.1 Anti-HIV activity ...................................................................................... 109
  3.5.4 Other viruses .................................................................................................... 110
    3.5.4.1 CRFK cells: feline coronavirus and feline herpesvirus ......................... 111
    3.5.4.2 Vero cells: Coxsackie B4 virus, Punta Toro virus, sindbis virus, para-influenza 3 virus, reovirus 1 ................................................................. 111
    3.5.4.3 HeLa cells: vesicular stomatitis, Coxsackie B4 virus, respiratory syncytial virus .... 112
    3.5.4.4 HEL cells: herpes simplex virus-1 strain KOS (wild-type and thymidine kinase deficient), herpes simplex virus-2 (G), vaccinia virus, vesicular stomatitis, adenovirus-2 ... 113
  3.5.5 Overview ......................................................................................................... 115

Chapter 4 - Experimental section .................................................................................. 121
4. Experimental section .................................................................................................. 123
  4.1. Instrumentation .................................................................................................... 123
  4.2 Materials and general methods ............................................................................ 123
  4.3 Experimental procedures ...................................................................................... 123
    4.3.1. Synthesis of 3,4-O-dicafeoyl-1,5-γ-quinide 7, improved from the published one ...... 124
    4.3.2 Synthesis of 1,3,4-O-tris[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide 10, 3,4-O-bis[3,4 (dimethoxy)cinnamoyl]-1,5-γ-quinide 11 and 3-O-[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide 12 ............................................................. 130
    4.3.3 Synthesis of (3-O-[4-(aminobutanamidoethoxy)feruloyl]-4-O-[3,4- (dimethoxy)cinnamoyl]-1,5-γ-quinide) 28 ............................................................. 134
  4.4 Fluorescence assay ............................................................................................... 143
  4.5 Biacore assay ....................................................................................................... 143
  4.6 NMR titrations .................................................................................................... 144
Abstract

University of Trieste
PhD School in Science, Technology and Economy of Coffee (Cicle XXVI)

Biomolecules as recognition elements for bioactive polyphenols in coffee
PhD Thesis

PhD student: Valentina Sinisi
Supervisors: Dr. Federico Berti
Dr. Cristina Forzato

Abstract

Coffee is a worldwide diffused beverage and its high impact on the global economy, considering all the steps from its cultivation to the final product selling, makes it a very interesting research topic under many points of view. The success of coffee is due to the perfect combination between the overall involvement of our senses and the well-known stimulating effect.

When we think about a cup of coffee, first of all we remember its typical aroma and its taste. During the years, many studies focused on the volatile fraction, but our knowledge about the taste responsible compounds is still nowadays scarce. Among such compounds present in the coffee beans, chlorogenic acids (CGAs), that belong to the polyphenols family, acquire more and more importance, due to their key-role in determining the coffee flavor, particularly its bitterness, their beneficial health properties, and their possible use as marker to control the industrial process. The roasting of green coffee beans causes a partial loss of CGAs due to the formation of degradation products, such as the corresponding lactones, whose amount depends on the roasting degree; the identification and the quantification of acids and lactones in coffee could be therefore a useful parameter to evaluate the final quality of the beverage.

This PhD project arises from the interest in searching new methodologies for the selective analysis of the lactone fraction, exploiting the recognition properties of biomolecules, proteins or peptides, or of designed sensing elements with high affinity for such lactonic compounds.

The chosen molecules, not commercially available, were first synthesized, namely: 3,4-O-dicaffeoyl-1,5-γ-quinide, 3-O-[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide, 3,4-O-bis[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide, and 1,3,4-O-tris[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide. A direct synthesis of the tri-substituted quinide was also tuned up, starting from D-(−)-quinic acid and 3,4-dimethoxycinnamoyl chloride. The compounds have been fully characterized in order to set up a record of reference data, and considering the potential biological activities of polyphenols, the antiviral properties of these compounds against many viruses have been also evaluated.

A possible approach to develop a selective biosensor is to use natural peptide scaffolds, with stable and highly organized conformations, and reduce its dimensions down to the limit of receptor stability, exploiting even the randomization of the aminoacid within the binding site to improve its ligand properties. In this perspective, the specific binding constants to Human Serum Albumin (HSA), more exactly to its Sudlow site I, of caffeic acid, ferulic acid, 3,4-dimethoxycinnamic acid, 5-O-caffeoyl quinic acid and of the four synthesized quinides were
measured in physiological conditions by fluorescence spectroscopy, reaching promising $K_D$ values, in the range 3-30 $\mu$M; moreover, 3,4- O-dicafeoyl-1,5-$\gamma$-quinide gave a peculiar result, showing a very interesting double binding in the same active site of the protein.

The two diester quinides were also used to test, always by means of fluorescence spectroscopy, the binding ability of a functional 100 aminoacids fragment that replicates the binding site I of the whole protein: this peptide, called HSA100, has been obtained by our research group; even some mutants of HSA100 have been prepared and four of them, randomly chosen, were used in the fluorescence assays. The obtained binding constants with HSA100 and its four mutants remain in the micromolar range, close to those measured with the whole protein, suggesting that proper mutations could lead to a selective biosensor with high affinity for the quinides.

The Surface Plasmon Resonance (SPR) technique could offer a way for the fast screening of the binding properties. In the perspective of setting up a SPR method to quickly select the hits of the mutants libraries, a linker-equipped quinide was synthesized and immobilized on a gold chip, then the binding with HSA was tested in this way, finding a 127 $\mu$M $K_D$ that could be seen as the “zero” point in the SPR relative scale, useful to evaluate the performances of the mutants which will be subsequently tested.

Considering the well-known interaction between polyphenols compounds, such as chlorogenic acid, and caffeine in aqueous solution, NMR titrations were performed to study the behavior of caffeine with the synthesized quinides, in a 75-25% DMSO-$d_6$/D$_2$O environment; with 3,4- O-dicafeoyl-1,5-$\gamma$-quinide the experiment was carried out also in 25-75% DMSO-$d_6$/D$_2$O system, allowing to measure a 69 $\pm$ 5 mmol dissociation constant, that is consistent with the data reported in the literature about similar studies. This kind of binding may be exploited to develop a selective biosensor for quinides by introducing several molecules of caffeine or other xanthines on a suitable scaffold.

Publications:


Further articles will be delivered soon.
Abstract

Università di Trieste
Scuola di dottorato in Scienza, Tecnologia ed Economia nell’Industria del Caffè (Ciclo XXVI)

Biomolecole come elementi di riconoscimento per polifenoli del caffè

Tesi di dottorato

Dottoranda: Valentina Sinisi

Relatori: Dr. Federico Berti
Dr. Cristina Forzato

Abstract

Grazie al suo ben noto effetto stimolante ed alla piacevolezza delle sensazioni di gusto e olfatto che suscita in chiunque lo beva, il caffè è una delle bevande più diffuse al mondo e il suo successo è tale da averlo reso un prodotto di grande impatto sull’economia globale. Le molteplici e delicate fasi necessarie a trasformare le bacche rosse, coltivate e raccolte nei Paesi tropicali, nell’aromatica polvere che tutti conosciamo rendono il caffè oggetto di ricerca costante in diversi ambiti.

Se pensiamo a una tazzina di caffè, le prime cose che ci vengono in mente sono il suo aroma e il suo gusto. Considerando queste due caratteristiche da un punto di vista chimico, molti studi si sono focalizzati sul riconoscimento dei composti volatili, ma molto meno è stato finora scoperto riguardo alle molecole effettivamente responsabili del tipico gusto del caffè. Tra i vari tipi di molecole non volatili, si sta attribuendo un’importanza sempre maggiore agli acidi clorogenici (CGAs), appartenenti alla famiglia dei polifenoli, in quanto, oltre ad avere diverse proprietà benefiche per la salute, sembra abbiano un ruolo determinante nel definire il flavor del caffè in tazza, in particolare la sua nota amara. Tali composti hanno le giuste caratteristiche per essere sfruttati come marker per controllare il processo di produzione: durante la tostatura parte dei CGAs si degrada e compaiono vari tipi di prodotti, in particolare i loro lattoni, la cui quantità finale dipende dalle condizioni e dalla durata della tostatura stessa. Dato che sia i CGAs che i loro lattoni influenzano la qualità finale, poterli riconoscere e quantificare nella bevanda sarebbe un buon modo per controllare e modificare, se necessario, le condizioni di produzione del caffè tostato.

Il progetto affrontato in questo dottorato nasce dall’interesse di trovare uno strumento capace di riconoscere la famiglia dei lattoni, per esempio sfruttando biomolecole, come proteine e peptidi, che già naturalmente si comportano da recettori, o strutture molecolari progettate ad-hoc con le giuste caratteristiche per interagire selettivamente con i lattoni.

Per prima cosa alcuni composti target, non disponibili commercialmente, sono stati scelti e sintetizzati; nello specifico: 3,4-O-dicafeoil-1,5-γ-chinide, 3-O-[3,4-(dimetossi)cinnamoil]-1,5-γ-chinide, 3,4-O-bis[3,4-(dimetossi)cinnamoil]-1,5-γ-chinide, e 1,3,4-O-tris[3,4-(dimetossi)cinnamoil]-1,5-γ-chinide. Il lattone trisostituito inoltre è stato ottenuto tramite sintesi diretta, a partire dall’acido chinico e dal cloruro dell’acido 3,4-dimetossicinnamico. I composti sono stati completamente caratterizzati allo scopo di offrire una serie completa di dati di riferimento; sulla base delle varie attività biologiche riscontrate nei polifenoli, i composti sintetizzati sono stati valutati inoltre come potenziali antivirali.

La progettazione di un biosensore selettivo basato sull’interazione di una biomolecola con un composto target può essere fatta secondo diversi approcci: uno di questi è partire da una
conformazione peptidica naturale, ben organizzata e stabile, e ridurre le sue dimensioni fino al limite della sua stabilità, introducendo anche mutazioni di alcuni residui amminoacidici all’interno del sito attivo nel tentativo di migliorare le proprietà leganti. 

In questa prospettiva è stata studiata tramite spettroscopia di fluorescenza l’interazione tra albumina umana (HSA) e i seguenti composti fenolici: acido caffeico, acido ferulico, acido 3,4-dimetossicinnamico, acido 5-O-caffeoil chinico e i quattro lattoni sintetizzati. Gli esperimenti sono stati fatti in condizioni fisiologiche; i composti testati entrano nel sito dell’HSA denominato Sudlow site I, causando uno smorzamento della fluorescenza e le costanti di dissociazione (K_D) ottenute, tra 3 e 30 μM, sono risultate molto promettenti. Il composto 3,4-O-dicafeoil-1,5-γ-chinide per di più ha dimostrato un comportamento particolare, in quanto sembra che siano due molecole a causare lo spegnimento della fluorescenza, legandosi entrambe nel sito attivo della proteina.

I due chinidi disostituiti sono stati poi usati per testare le proprietà leganti di un frammento funzionale di HSA costituito da 100 amminoacidi, precedentemente sviluppato nel nostro gruppo di ricerca e chiamato HSA100; esso riproduce il sito di legame I della proteina e ne mantiene le caratteristiche. A partire dall’HSA100 sono stati preparati alcuni mutanti; quattro di questi, scelti casualmente senza alcuna selezione, sono stati testati sempre con i due chinidi disostituiti. Le costanti di dissociazione, valutate anche in questo caso tramite spettroscopia di fluorescenza, hanno ancora ordine di grandezza micromolare e il comportamento dei peptidi rispecchia quello della proteina nativa. Inserendo opportune mutazioni probabilmente sarà possibile ottenere con questo approccio un biosensore selettivo e molto affine per la famiglia dei lattoni.

La tecnica di Risonanza Plasmonica Superficiale (SPR) offre la possibilità di eseguire rapidamente screening di libbre peptidiche, aspetto fondamentale per sfruttare davvero la loro produzione. Nell’ottica di mettere a punto un metodo SPR per selezionare velocemente i peptidi più promettenti tra i tanti componenti di una libreria, un chinide con un linker amminico è stato sintetizzato e immobilizzato su un chip d’oro, poi si è misurata l’affinità con l’HSA che in queste condizioni sperimentali presenta una K_D pari a 127 μM; tale valore può essere considerato lo “zero” nella scala relativa dei test SPR, da migliorare con l’introduzione di mutazioni.

Una parte della tesi ha riguardato inoltre lo studio dell’interazione in soluzione acquosa tra composti polifenolici, come ad esempio l’acido 5-O-caffeoil chinico, e la caffeina, effettuando una serie di titolazioni NMR al fine di studiare l’eventuale formazione di complessi tra la caffeina e i chinidi sintetizzati. A causa della scarsa solubilità di questi composti in acqua è stato necessario utilizzare una miscela DMSO-d_6/D_2O 75-25%; con il composto 3,4-O-dicafeoil-1,5-γ-chinide l’esperimento è stato ripetuto in D_2O/DMSO-d_6 75-25% e in questo caso è stato possibile determinare una K_D, risultata pari a 69 ± 5 mmol, valore in linea con quelli riportati in letteratura per studi simili. Questo tipo di interazione può essere anch’esso sfruttato per progettare un biosensore, introducendo per esempio più molecole di caffeina o di altre xantine su un’adatta conformazione molecolare.

Pubblicazioni:

Sinisi V., Forzato C., Cefarin N., Navarini L., Berti F. Interaction of coffee phenolic compounds with human serum albumin. Inviato a Food Chemistry e tuttora sottoposto a revisione.

Ulteriori articoli verranno inviati a breve.
List of abbreviations

Å Ångström
Ala alanine
Arg arginine
Asn asparagine
a. u. arbitrary unit
Boc tert-butoxycarbonyl
br broad signal
°C Celsius degree
CAST combinatorial active-site saturation test
CC₅₀ 50% cytotoxic concentration
CCID₅₀ 50% cell culture infectious dose
CGAs chlorogenic acids
CGLs chlorogenic lactones
CMPF 3-carboxy-4-methyl-5-propyl-2-furanopropanoic acid
CPE cytopathogenic effect
CQAs caffeoylquinic acids
5-CQA 5-O-caffeyl quinic acid
CQLs caffeoyl quinides
3-CQL 3-O-caffeyl-γ-quinide
4-CQL 4-O-caffeyl-γ-quinide
3,4-diCQL 3,4-O-dicafeoyl-γ-quinide
CRFK Crandell-Rees feline kidney
Cys cysteine
d doublet
DABCO diazabicyclo[2.2.2]octane
DCM dichloromethane
dd doublet of doublets
ddd doublet of doublet of doublets
ddddd doublet of doublet of doublet of doublets
DMAP 4-dimethylaminopyridine
DMF dimethylformamide
DMSO dimethyl sulfoxide
DMSO-d₆ deuterated dimethyl sulfoxide
DMQs 3,4-(dimethoxy)cinnamoyl quinides
DPBA 2,4-dioxo-4-phenylbutanoic acid
dsDNA double-strand DNA (virus)
dsRNA double-strand RNA (virus)
ds-10000 dextran sulfate
dt doublet of triplets
EC₅₀ 50% effective concentration
EC₉₀ 90% effective concentration
EC₉₉ 99% effective concentration
EDC 1-ethyl-3-(3-dimethylaminopropyl)-carboodiimide
List of abbreviations

EDC-Cl  \(N\)-(3-Dimethylaminopropyl)-\(N^\prime\) -ethylcarbodiimide hydrochloride
ESI\(^{+/-}\) electrospray ionization (positive/negative mode)
eq. equivalent
FCoV feline coronavirus
FHV feline herpesvirus
FQAs feruloylquinic acids
3-FQA 3-O-feruloylquinic acid
4-FQA 4-O-feruloylquinic acid
FRET Förster resonance energy transfer
GABA \(\gamma\)-aminobutyric acid
Gln glutamine
Glu glutamic acid
Gly glycine
GST glutathione S-transferase
HEK human embryonic kidney
HIV human immunodeficiency virus
His histidine
HOBt 1-hydroxybenzotriazole hydrate
HPLC high-performance liquid chromatography
hPIV human parainfluenza virus
HSA human serum albumin
HSV herpes simplex virus
Hz hertz
IC\(_{50}\) 50% inhibitory concentrations
Ile isoleucine
IR infrared
K Kelvin degree
Leu leucine
Lys lysine
MS mass spectrometry
ND not determined
NMM N-methylmorpholine
m multiplet
MCC minimal cytotoxic concentration
MDCK Madin-Darby canine kidney
Met metionine
MHz megahertz
MIP molecularly imprinted polymer
MOI multiplicity of infection
M.p. melting point
mRNA messenger RNA
MTS 3-(4,5-dimethylthiazole-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium, inner salt
MTT 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
NHS N-hydroxysuccinimide
NMR nuclear magnetic resonance
NNRTI non-nucleoside reverse transcriptase inhibitor
# List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOE</td>
<td>Nuclear Overhauser Effect</td>
</tr>
<tr>
<td>1H-1H NOESY</td>
<td>Nuclear Overhauser Effect Spectroscopy</td>
</tr>
<tr>
<td>NRTI</td>
<td>nucleoside reverse transcriptase inhibitor</td>
</tr>
<tr>
<td>PA-Nter</td>
<td>N-terminal domain of the PA subunit (influenza virus polymerase)</td>
</tr>
<tr>
<td>PFU</td>
<td>plaque forming unit</td>
</tr>
<tr>
<td>Phe</td>
<td>phenylalanine</td>
</tr>
<tr>
<td>PI</td>
<td>protease inhibitor</td>
</tr>
<tr>
<td>PNA</td>
<td>peptide nucleic acid</td>
</tr>
<tr>
<td>ppm</td>
<td>parts per million</td>
</tr>
<tr>
<td>p-TsOH</td>
<td>p-toluenesulfonic acid</td>
</tr>
<tr>
<td>PTV</td>
<td>Punta Toro virus</td>
</tr>
<tr>
<td>q</td>
<td>quartet</td>
</tr>
<tr>
<td>RdRp</td>
<td>RNA-dependent RNA polymerase</td>
</tr>
<tr>
<td>Rf</td>
<td>retardation factor</td>
</tr>
<tr>
<td>RP-HPLC</td>
<td>reverse phase high-performance liquid chromatography</td>
</tr>
<tr>
<td>RSV</td>
<td>respiratory syncytial virus</td>
</tr>
<tr>
<td>r.t.</td>
<td>room temperature</td>
</tr>
<tr>
<td>RU</td>
<td>resonance unit</td>
</tr>
<tr>
<td>s</td>
<td>singlet</td>
</tr>
<tr>
<td>SEM</td>
<td>standard error of the mean</td>
</tr>
<tr>
<td>Ser</td>
<td>serine</td>
</tr>
<tr>
<td>SFS</td>
<td>synchronous fluorescence spectra</td>
</tr>
<tr>
<td>SINV</td>
<td>Sindbis virus</td>
</tr>
<tr>
<td>SPR</td>
<td>surface plasmon resonance</td>
</tr>
<tr>
<td>ssRNA+</td>
<td>positive-sense single-strand RNA (virus)</td>
</tr>
<tr>
<td>ssRNA</td>
<td>negative-sense single-strand RNA (virus)</td>
</tr>
<tr>
<td>t</td>
<td>triplet</td>
</tr>
<tr>
<td>TFA</td>
<td>trifluoroacetic acid</td>
</tr>
<tr>
<td>TMEDA</td>
<td>tetramethylethylenediamine</td>
</tr>
<tr>
<td>THF</td>
<td>tetrahydrofuran</td>
</tr>
<tr>
<td>TK</td>
<td>thymidine kinase</td>
</tr>
<tr>
<td>TLC</td>
<td>thin layer chromatography</td>
</tr>
<tr>
<td>Trp</td>
<td>tryptophan</td>
</tr>
<tr>
<td>Tyr</td>
<td>tyrosine</td>
</tr>
<tr>
<td>UV</td>
<td>ultraviolet</td>
</tr>
<tr>
<td>Val</td>
<td>valine</td>
</tr>
<tr>
<td>vRNA</td>
<td>viral RNA</td>
</tr>
<tr>
<td>vRNP</td>
<td>viral ribonucleoprotein</td>
</tr>
<tr>
<td>VSV</td>
<td>vesicular stomatitis virus</td>
</tr>
<tr>
<td>VV</td>
<td>vaccinia virus</td>
</tr>
<tr>
<td>WT</td>
<td>wild type</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction
1.1 Coffee polyphenols

1.1.1 Coffee: a successful beverage

To drink coffee is a popular habit all over the world. Coffee is one of the world’s most traded products, second in value only to oil: considering the year 2013, about 109 millions of 60 kg-bags have been exported by the growing countries. The coffee industry employs a huge amount of people around the world through its cultivation, processing and trading; in many tropical coffee-producing countries, coffee represents the main source of income. The coffee trees originated in Ethiopia and several legends account its first use, involving birds or flock of goats made unusual vital by eating certain red berries. It is unknown when the plant was domesticated, anyway from the Horn of Africa it began its global spread, to become one of the most consumed beverages in the world: it was cultivated by Arabs during the 14th century; by the 16th century it had reached the rest of the Middle East, Persia, Turkey and northern Africa, then it arrived to Balkans, Italy and to the rest of Europe; during the 17th century it was introduced into the New World and much of the rest of the tropics, where it is still nowadays cultivated.

The genus *Coffea* belongs to the *Rubiaceae* family; about a hundred of different species of coffee plant exists, but only few of them have a commercial importance, in particular *Coffea arabica* and *Coffea canephora*, the latter also known as Robusta. Most of the unsold species are tropical trees and bushes that grow in the shady forests.

Every consumer country has its own recipe of coffee brew, result of the local traditions. However, the reasons of coffee success are always the same: the perfect combination between the overall involvement of our senses, the well-known stimulating effect, and its socializing role.

The big impact of coffee on the global economy makes it a very interesting research topic from many points of view. Chemically, coffee is a boundless subject of study: it is a mixture of thousands of different compounds, belonging to several families; moreover the coffee composition changes according to the coffee species, the geographical origin, the weather conditions, and the farming practices. Each production step influences and transform the chemical content: a deepened knowledge of what happens inside the beans during their ripening, harvesting, drying, roasting, grinding, storage and brewing can allow to really control the key factors and to get in the end a coffee cup of the better possible quality.

1.1.2 The composition of coffee

The chemical mixture inside a green coffee bean includes minerals, carbohydrates, lipids, amino acids, peptides, and carboxylic acids. The relative composition shows variations associated with the coffee species, for example Arabica contains more lipids and trigonelline, while Robusta contains more caffeine and chlorogenic acids (table 1.1.1).

During the roasting process, involving high temperature (usually 170-230 °C for 10-15 minutes) and high pressure inside the beans (up to 25 atm), a large series of chemical reactions is promoted, such as caramelization and degradation of carbohydrates or the Maillard reaction. Water, carbon dioxide and many present volatiles are mostly released,
Chapter 1 – Introduction

while many new compounds are formed, different in molecular weight, volatility and solubility.

<table>
<thead>
<tr>
<th>Class</th>
<th>Arabica</th>
<th>Robusta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caffeine</td>
<td>1.2</td>
<td>2.2</td>
</tr>
<tr>
<td>Lipids</td>
<td>16</td>
<td>10</td>
</tr>
<tr>
<td>Chlorogenic acids</td>
<td>6.5</td>
<td>10</td>
</tr>
<tr>
<td>Aliphatic acids</td>
<td>1.7</td>
<td>1.6</td>
</tr>
<tr>
<td>Total aminoacids</td>
<td>10.3</td>
<td></td>
</tr>
<tr>
<td>- free</td>
<td>0.5</td>
<td>0.8</td>
</tr>
<tr>
<td>Trigonelline</td>
<td>1.0</td>
<td>0.7</td>
</tr>
<tr>
<td>Glycosides</td>
<td>0.2</td>
<td>Traces</td>
</tr>
<tr>
<td>Minerals</td>
<td>4.2</td>
<td>4.4</td>
</tr>
<tr>
<td>- potassium</td>
<td>1.7</td>
<td>1.8</td>
</tr>
<tr>
<td>Carbohydrates (by difference)</td>
<td>58.9</td>
<td>60.8</td>
</tr>
</tbody>
</table>

Table 1.1.1 Composition of green coffee beans (% dry matter).

The composition inside the beans changes according to the roasting degree. It is further different in the cup (figure 1.1.1), because in the brewing the individual solubility drives the extraction of each compound. Obviously, if different extraction methods are used with the same coffee, the composition of the final brews will be different$^{11,12,13}$

Figure 1.1.1 Average chemical composition of green, roasted, and brewed coffee.
1.1.3 The aroma of coffee

Aroma is the first pleasant stimulus that amazes everyone approaching a cup of coffee. Green coffee contains about 300 volatiles, but most of them are lost in the roasting, when on the other side a great number of new volatiles is generated. More than 1000 volatile compounds have been detected in roasted coffee, also very different the one from the others (figure 1.1.2). Gas chromatography (GC), even bidimensional, coupled with mass spectroscopy and/or olfactometry, allowed to identify many of the coffee volatiles, but aroma reconstitution experiments demonstrated that only 25 of them, like 2-furfurylthiol or 3-isobutyl-2-methoxypyrazine, are really needed to mimic the fragrance of a freshly prepared coffee beverage. The roasting and brewing conditions rule the aroma of the coffee cup and can so affect the consumer feelings.

![Molecular structure of some coffee volatiles.](image)

1.1.4 The taste of coffee

The feeling experienced in the mouth while drinking coffee is a combination of tactile and gustatory perceptions. It is not easy to describe the coffee taste, anyway bitterness and sourness are the two predominant notes. Similarly to aroma, taste is due to a large variety of compounds, but the comprehension of the typical coffee flavor molecular basis is only partial and the relative literature is scarce. Starting from the matter of fact that green coffee beans have minimal flavor, the compounds responsible of coffee taste are formed during the bean roasting from tasteless precursors. A series of experiments, including sensory and taste dilution analyses, allowed to identify some bitter molecules (figure 1.1.3), for example the alkaloids caffeine and trigonelline, or several thermally generated heterocycles as furfuryl alcohol, pyrazines, and 2,5-diketopiperazines, but the “recipe” of the coffee bitterness is unclear. Recently the researchers attention has been focused on the family of the chlorogenic acids (CGAs), that belongs to the class of phenolic compounds, the amount and composition of
CGAs and of their roasting degradation products seem to have a key role in defining the final bitterness in the cup.\textsuperscript{27,34,35} Considering the importance attributed to CGAs, this family will be described in detail in the next section.

![Molecular structure of some coffee bitter compounds.](image1)

**Figure 1.1.3** Molecular structure of some coffee bitter compounds.

### 1.1.5 Phenolic compounds

Phenolic compounds are secondary metabolites present in many plants, generally involved in the defense against environmental stresses, like UV radiation or aggression by pathogens and insects;\textsuperscript{33} their biosynthesis and distribution are regulated by the internal defense system of the plants. As a function of the phenol rings number and of the structural elements that bind these rings, the several different phenolic compounds can be generally classified in flavonoids, phenolic acids, lignans and stilbenes (figure 1.1.4),\textsuperscript{36} then into subclasses considering the specific substitutions, their association with carbohydrates or other molecules and their polymerization (table 1.1.2).\textsuperscript{37}

![Molecular structure of some phenolic compounds.](image2)

**Figure 1.1.4** Molecular structure of some phenolic compounds.
Classes and sub-classes | Examples of specific compounds
--- | ---
**Phenolic acids** |  
- Benzoic acids  
gallic acid; protocatechuic acid;  
p-hydroxybenzoic acid  
- Hydroxycinnamic acids  
coumaric acid; caffeic acid;  
ferulic acid; sinapic acid
**Stilbenes** | resveratrol
**Lignans** | Secoisolariciresinol; matairesinol;  
lariciresinol; pinoresinol
**Flavonoids** |  
- Flavonols  
kaempferol; quercetin; myricetin  
- Flavones  
apigenin; luteolin  
- Flavanones  
naringenin; hesperetin  
- Flavanols  
catechins; gallocatechins  
- Anthocyanidins  
pelargonidin; cyanidin; malvidin  
- Condensed tannins or proanthocyanidins  
trimeric procyanidin; prodelphinidins  
- Isoflavones  
daidzein; genistein; glycine

Table 1.1.2 | Main classes of phenolic compounds in higher plants.

The interest about phenolic compounds is ever-increasing: in addition to their well-known antioxidant and free radical scavenging activities, they seem to behave as protective factors against human chronic degenerative disease, cancer and cardiovascular disease and to have anti-histamine, anti-inflammatory, antibacterial, and antiviral properties.

1.1.5.1 Chlorogenic acids (CGAs)

CGAs are esters of quinic acid with different cinnamic acids, such as caffeic, ferulic and p-coumaric acid (figure 1.1.5), acylating the hydroxyl groups at position 3, 4 or 5 of D-(-)-quinic acid, to give origin to a series of mono esters. Also diesters (with the same cinnamoyl moiety or with different cinnamic acids) or even triesters are found, while no acylation at C₁ seems to occur. In literature there is discordance in the CGAs names, anyway in the present thesis the preferred IUPAC numbering system is used.

CGAs are present in several commonly consumed plant foods, like apples, ciders, blueberries, spinach and so on. Green coffee beans are particularly rich of these compounds and their amount, on dry matter basis, ranges from 4 to 8.4% for Coffea arabica and from 7 to 14.4% for Coffea canephora; their amount in the beans is closely related to the quality of coffee brews. The CGAs biosynthesis in coffee plants, that seems to be controlled by the phenylpropanoid pathway and the shikimic acid pathway (figure 1.1.6), produces many different acids, but caffeoylquinic acids (CQAs) represent the main subgroup of the CGAs and 5-O-caffeoylquinic acid (5-CQA, figure 1.1.5, compound 5) is the most abundant one, indeed it is usually called chlorogenic acid.
Figure 1.1.5 Molecular structure of D-(-)-quinic acid (1), caffeic acid (2), ferulic acid (3), p-coumaric acid (4) and 5-O-caffeoylquinic acid (5).

Figure 1.1.6 Possible pathways for 5-caffeoylquinic acid and 5-feruloylquinic acid in coffee plants. Enzyme abbreviations: C3H/C3’H, 4-coumaricacid 3-hydroxylase/4-coumaroylesterase 3’-hydrolase; C4H, cinnamic acid 4-hydroxylase; CCoAMT, caffeoyl-CoA O-methyltransferase; CoAL, CoA ligase; CQT, quinic acid hydroxycinnamoyl transferase; OMT, O-methyltransferase.
Chapter 1 - Introduction

CGAs are mostly concentrated in the seeds, but they are present also in the leaves and in coffee pulp. A qualitative relationship between caffeine and CGAs contents among wild Coffea species has been detected, this feature can be explained by the formation of a 1:1 molar complex between them due to their $\pi$-$\pi$-stacking (figure 1.1.7). 

![Energy-minimized conformation of caffeine-5CQA complex.](Figure 1.1.7)

This kind of complexation is exploited by the plants to control the caffeine compartmentation in the tissues. Caffeine has a toxic effect on insects and some animals due to their poor capacity to metabolize the compound: coffee plant uses it to protect its fruits from predation, indeed caffeine is found in high concentration in leaves and beans which are most exposed to herbivores and insects attack. Due to its amphipilic nature, caffeine permeates most of biological barriers; the formation of a not covalent complex is an efficient strategy to keep caffeine in the cell vacuole, avoiding any possible auto-toxicity.

CGAs have shown several beneficial health properties, such as antioxidant, anti-inflammatory, and antispasmodic activities or as the relative risk reduction of cardiovascular diseases, of type 2 diabetes, and of Alzheimer’s disease. They have also radioprotective and hypocholesterolemic functions, and some dicaffeoylquinic acids can inhibit in vitro the enzyme HIV-Integrase.

1.1.5.2 Thermal transformation of CGAs

Due to their thermal instability, CGAs undergo to several reactions during coffee roasting (figure 1.1.8): they are isomerized, epimerized, dehydrated with the resulting formation of a lactone ring between the carboxyl group of quinic acid and its hydroxyl group at carbon 5, hydrolyzed, degraded into smaller molecules. They also take part in the formation of polymeric compounds like melanoidins. The CGAs loss depends on the roasting conditions, generally about 8-10% of their initial amount is lost every 1% loss of dry matter; a drastic process may cause a decrease of CGAs quantity in coffee up to 95%. In a medium roasted coffee the amount of CGAs is on average 2.4% and 3.1% (dry matter basis) in Arabica and in Robusta, respectively.

At the beginning of the roasting, isomerization reactions take place, accompanied by partial hydrolysis: the amount of 5-CQA decreases, while 3-CQA and 4-CQA are formed; the latter two, having a free hydroxyl group at C5, allow the lactones formation (CGLs), that occurs later in the process.
Chapter 1 – Introduction

The amount of residual CGAs after roasting play a key role in determining the coffee acidity, while both CGAs and CGLs contribute to coffee bitterness, that increases during roasting exactly due to the CGLs release, and astringency. The quantification and characterization of the chlorogenic fraction in coffee is mainly done by means of HPLC, often coupled with mass spectroscopy, but the lack of standards (only 5-CQA is commercially available) complicates their study.

1.1.5.3 Chlorogenic lactones (CGLs)

As well as their acidic precursors, CGLs can be mono- or di-functionalized, due to the possible esterification of the hydroxyl groups on C3 and/or C4 of the quinidic core. Generally, CGLs are absent in green coffee; their rare presence in green beans may be due to heating or dehydration during primary processing. The CGLs amount reaches a maximum in correspondence to a weight loss of about 14%, that means a light medium roasting degree, achieving an average total content about 4 and 4.2 mg/g of coffee (dry matter basis) for Arabica and Robusta, respectively. Then, going on in the roasting, they gradually decrease. Less than 10% of total CGAs in green coffee is converted in lactones.

The caffeoyl lactones (CQLs) are the most abundant group in CGLs; 3-O-caffeoyl-1,5-γ-quinide (3-CQL) is the main lactone with maximum levels of 2.3 and 2.5 mg/g in Arabica and Robusta, respectively; 4-O-caffeoyl-1,5-γ-quinide (4-CQL) is the second major lactone with maximum levels of 1.2 and 1.4 mg/g in Arabica and Robusta, respectively; the di-substituted 3,4-O-dicaffeoyl-1,5-γ-quinide (3,4-diCQL) reaches content about 0.07 and 0.08 mg/g in Arabica and Robusta, respectively, a lower quantity compared to the levels of the mono-substituted feruloyl-1,5-γ-quinides (3-FQL and 4-FQL).

![Outline of some CGAs degradation pathways](image)

**Figure 1.1.8** Outline of some CGAs degradation pathways: R represents the cinnamoyl residue, the dotted arrow the transesterification and the slashed arrow the epimerization of the precursor.

![Molecular structure of the CQLs](image)

**Figure 1.1.9** Molecular structure of the CQLs.
An intramolecular migration of the caffeoyl substituent has been detected in 3-CQL and 4-CQL (Figure 1.1.10): the carbonyl group of the caffeoyl moiety can form an intermediate five-membered orthoformate ring with the two vicinal hydroxyl groups on C3 and C4; depending on which of the bonds with oxygen atoms is brought, one of the two monoester is achieved. This process is reversible, but the equatorial conformation of 3-CQL is preferred, thanks to the lower steric hindrance.

Figure 1.1.10 Intramolecular migration of the caffeoyl substituent in the CQLs.

Even lactones derived from epimerized CGAs has been detected (Figure 1.1.11), in fact the lactone ring can be formed also by the hydroxyl group at C3, giving origin for example to 5-O-caffeoyl-muco-γ-quinide or to 4-O-caffeoyl-muco-γ-quinide, and at C4, leading to molecules like 5-O-caffeoyl-epi-δ-quinide. Some lactones have been isolated by means of preparative chromatography methods or chemically synthesized to study in more details their characteristics.

Figure 1.1.11 Molecular structure of some epimerized CGLs detected in roasted coffee.

In addition to the important role of CGLs in determining the coffee bitterness, as their acidic precursors, they appear potentially bioactive, showing for example inhibition of the human adenosine trasporter, affinity for the human μ opioid receptor and neuroprotective activity. Moreover, due to their neutrality, they would be uptaken in blood and brain even more readily than the parental CGAs.
Chapter 1 – Introduction

1.1.5.4 Chlorogenic compounds are absorbed in humans

Considering the human diet, coffee is one of the main sources of chlorogenic compounds.\textsuperscript{48} The amounts of CGAs and CGLs in a coffee cup depend on the type of roasted coffee used and on the brewing method;\textsuperscript{31,93} in a traditional espresso coffee beverage (30 mL) the CGAs content is on average 70 mg.\textsuperscript{94} The extraction efficiency is higher for CGAs than for CGLs, due to their better water solubility, and in general a lungo (about 120 mL) is more rich in CGAs than an espresso coffee.\textsuperscript{93}

Many studies reported that CGAs are capable to permeate the gastrointestinal barrier and are absorbed in humans, even if with a detectable variability among subjects;\textsuperscript{92,95,96,97,98,99,100} they have been found in plasma both as intact molecules and as their hydrolysis metabolites, particularly as caffeic acid. To our knowledge, no data about CGLs absorption are available in literature yet. CQAs have been detected in plasma even 4 hours after the ingestion but each kind of CGAs shows its own absorption and metabolism mechanisms; for example the average CQA : diCQA molar ratio is lower in plasma compared with the coffee brew used for the experiment, suggesting a favored tissue uptake of CQA or a favored absorption of diCQA thanks to its higher lipophilicity.\textsuperscript{96}

An early absorption of CGAs occurs in the stomach and jejunum,\textsuperscript{101,102} followed by absorption in the small and large intestine.\textsuperscript{103,104} The metabolic fate of CGAs is still not clear; previous studies about 5-CQA have shown that it can be hydrolyzed in the large intestine and in the colon thanks to the microflora esterase activity,\textsuperscript{105,106,107} but not in the stomach or in the small intestine.\textsuperscript{92,108} Ferulic, isofurulic, dihydrocaffeic, dihydroferulic and vanillic acids have been suggested as the main human 5-CQA metabolites (figure 1.1.12);\textsuperscript{96,98,101,109} anyway traces of intact 5-CQA have been detected in urine after coffee consumption,\textsuperscript{92,96} so its hydrolysis in human body seems to be incomplete.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{molecular_structure.png}
\caption{Molecular structure of some human 5-CQA metabolites.}
\end{figure}
1 International Coffee Organization website: www.ico.org
Chapter 1 – Introduction


Moon J. K., Yoo H. S., Shibamoto T. Role of roasting conditions in the level of chlorogenic acid content in coffee beans: correlation with coffee acidity. J. Sci. Food Agric., 2009, 57, 5365-5369.


1.2 Recognition and sensing elements

1.2.1 Sensors: biomolecules and their mimics

The biomolecular recognition is the ability of a biomolecule to selectively interact with another molecule in the presence of structurally similar rival molecules; this kind of events is one of the most significant features of biological and chemical systems.\(^1\) The recognition process is the result of a complex interplay between non-covalent interactions as salt bridges, hydrogen bonds, van der Waals, hydrophobic interactions, and entropic effects.\(^2\) Molecular recognition is the foundation of biosensing, nowadays widespread in many scientific fields.\(^3\) In general, a biosensor is an analytical device consisting of a biological recognition element, responsible of the specific interaction with the target molecule, and a physical transducer, which converts the biorecognition event into a measurable signal (figure 1.2.1).\(^3,4\) Since the first biosensor was developed,\(^6\) several techniques have progressively been associated to provide accurate detection of target analytes and a wide range of possible applications has been proposed, such as environmental,\(^7\) clinical,\(^8,9,10\) and food analysis.\(^11\) High sensitivity and specificity are the main strengths of biosensor technologies.\(^3\)

![Figure 1.2.1](image)

Figure 1.2.1\(^12\) Schematic representation of a single element biosensor containing the biorecognition element, the transducer and the physical output, whose magnitude is related to the concentration of the target molecule.

Biosensors can be classified either by the type of involved biological molecule or by the type of signal transduction they employ (figure 1.2.2). Several bioelement, like enzymes, microbes, antibodies, receptors, plant or animal cells, organelles, or tissues, were first isolated from living systems, then synthesized in the laboratory and sometimes engineered. Recently, artificial materials, such as peptide nucleic acids (PNAs) and molecularly imprinted polymers (MIPs), were added to the available bioelements list. Many methods of signal transduction can be used: optical detection, electrochemical, mass-sensitive, thermal detection, viscosity, and so on.\(^13\)

Actually, one of the main interests in this field is the miniaturization of biosensing chips or devices, using micro- and nano-fabrication techniques,\(^14,15\) to reduce the costs by mass production, to limit the sample volume, to achieve a compact, portable and automated tool.
Figure 1.2.2 Biosensing and transduction classes for in vitro biosensors. Methods of biosensing: antibody/antigen (a); enzyme catalyzed (b); nucleic acid (c); cell-based (d); biomimetic (e). Methods of transduction: optical (f); electrochemical (g); mass-sensitive (h); thermal (i).

useful for simple and real-time analysis. A remarkable example in this sense is the self-monitoring of blood glucose (SMBG) biosensor: the patients themselves routinely measure their glucose level directly introducing a blood sample into the sensor chip, equipped with an enzyme-electrode-based analyser.

Further developments have been offered by several techniques, like immunoreactions, conducting polymers, plasma polymerized films (PPFs), bacterial magnetic particles (BMPs), etc. Even biosensor arrays have been developed: a combination of more recognition elements for the same target allows to gain in specificity, while putting in array biomolecules for different analytes it is possible to get the response about several targets in the same time; some examples are chip arrays for nucleic acids, proteins, toxins or contaminants.

1.2.2 Molecular recognition elements

Molecular recognition elements can be divided into two classes: catalytic and affinity-bases. The former type includes enzymes, microbes, organelles, plant or animal cells, and plant or animal tissues, while antibodies, receptors, nucleic acids, and PNAs belong to the latter. Since the project purpose is the design of a selective biosensor for coffee polyphenols, in the following sections a brief description of just some affinity-bases elements is given.

1.2.2.1 Antibodies

Antibody recognition elements exploit the sensitivity and specificity of antibody–antigen interactions and have been used extensively for detection purposes. The possibility of performing the analysis without any sample purification is the major advantage of antibody biorecognition elements. Large quantities of antibody can be produced using cell clones that specifically produce monoclonal antibodies of choice. Recombinant antibodies, consisting of genetically manipulated fused antigen binding domains of common antibodies, are now
available and their generation turns out to be less expensive and time consuming compared to polyclonal or monoclonal antibodies production.\textsuperscript{36,37} The binding events can be easily detected either through enzyme-labeled fluorescence (ELF) technology\textsuperscript{38} or by surface plasmon resonance (SPR) spectroscopy.\textsuperscript{39}

1.2.2.2 Receptors

Receptors are transmembrane (plasma and intracellular membranes) or soluble proteins. When specific molecules bind to them, a particular cellular response is generated; in addition a variety of toxins and drugs naturally interact with receptors.\textsuperscript{40} High ligand specificity and affinity make receptors attractive biosensor recognition elements, but their use has been significantly limited by their relative instability, their laborious isolation and low yield purification protocols, their transduction difficulties due to absence of signal amplification. However, the advent of recombinant techniques and of a multitude of expression systems allowed to generate large amounts of receptor protein and to reduce their problems.\textsuperscript{41,42} Moreover, surface plasmon resonance spectroscopy permits to directly monitor the receptor–antagonist interaction with high sensitivity.\textsuperscript{43}

1.2.2.3 Lectins

Lectins constitute a broad family of proteins involved in diverse biological processes, occasionally having potent toxic properties; they strongly bind to specific carbohydrate moieties, known as glycans, through multivalent interactions arising from the spatial organization of oligosaccharide ligands.\textsuperscript{44,45} Particular glycans structural profiles and their recognition by lectins have been related to the progression of certain diseases as cancer and autoimmune disease,\textsuperscript{46} making analysis of saccharide–lectin binding processes important as a diagnostic tool.

Lectins are excellent biorecognition elements due to high affinity for saccharide moieties and have been extensively exploited as a basis for biosensor design.\textsuperscript{47,48,49,50}

1.2.2.4 Short peptides

Short, peptide-based, artificial receptors capable of high specificity are a biotechnological strategy to mimic the molecular recognition mechanism occurring in biomolecules such as enzymes, antibodies, drug receptors, and transmembrane proteins,\textsuperscript{51,52} with the aims of retaining their binding properties and of simplifying the receptor production.\textsuperscript{53,54} A growing number of short peptides, composed by 2-50 amino acids, has been reported in the recent literature as active elements in biosensors, thanks to several advantages: an unlimited number of different molecules that can be obtained by combining the 21 natural amino acids, both molecular biology and chemical techniques available for the fast screening of peptide libraries, possibility of automated synthesis and its relatively low cost, and ease of modification to further enhance their binding properties.\textsuperscript{54} Moreover a wide range of target analytes could be detected: whole cells, proteins, small organic molecules (like drugs, hormones, and pollutants), and ions.
Chapter 1 - Introduction

Short peptides can be classified according to the method adopted for their design:

- **designed synthetic peptides**: a stable peptide scaffold is identified, inspired by stable natural peptide domains, and then certain positions in the amino acid chains are randomized to obtain a library of receptors sharing the same constant region; the binding affinity can be iteratively optimized. Many different non-antibody scaffolds, particularly suitable for the recognition of proteins, have been obtained; the zinc-finger motif has been used to build up a zinc Förster resonance energy transfer (FRET) fluorimetric sensor, in which the conformation changes upon zinc binding, and the addition of a fluorophore/quencher couple at the peptide ends allows zinc detection by measuring the fluorescence quenching. The α-helix coiled-coil motif has been also used to detect ions and small organic molecules, as the xanthine-binding peptides prepared in our research group; this motif occurs when a seven amino acid sequence including hydrophobic residues at positions 3, 4, and 5 is repeated along a peptide so that the amphiphilic helices tend to associate on the hydrophobic side, giving rise to parallel or more often antiparallel coils. If the heptad repeat is interrupted by loop sequences, a typical and stable helix-loop-helix motif is obtained, and binding sites can be generated at the loop level (figure 1.2.3).

Unfortunately the modification of a natural structure has some problems: embedding a recognition site in a fixed scaffold leads to a suitable candidate only if the protein scaffold retains its stability in spite of the mutations; stable scaffolds are normally large and difficult to synthesize or express in large quantities; the combinatorial explosion associated with the number of mutation sites (ten sites means possible combinations).

A fully artificial design is also possible, to obtain de-novo peptides with high affinity towards the chosen target through a computational procedure: starting from a poly-alanine of the desired length, an algorithm based on a combination of molecular dynamics, semi-flexible docking, and replica exchange Monte Carlo is able to optimize simultaneously the sequence and conformation of the small peptide until a high affinity degree is reached; the best calculated sequences can be then synthesized and tested. In our research group this approach has been used to design a peptidic-receptor for the Efavirenz, widely used in HIV therapy.

- **artificial, miniaturized receptors**: this is a target-focused approach, in which a natural receptor of the target is identified and reduced to the length of a synthesizable peptide, or miniprotein. Libraries of mutated receptors can then be generated to improve binding and selectivity. Several examples are reported in literature, as a microfluidic accumulation channel/fluorimetric sensor for Botulinum neurotoxin B, that was obtained reducing the toxin receptor synaptogamin II to a 22 amino acids peptide allowing the detection of 16 pg of
target,\textsuperscript{62} or a sensor for the release of nitric oxide from living cells, developed by reducing the fibronectin and laminin binding domains to cell-binding dodecapeptides.\textsuperscript{63} Miniaturized receptors are involved even in the development of an “electronic nose”, that tries to mimic the amazing performances of the mammalian olfactory system integrating a combination of olfactory receptors, odorant-binding protein,\textsuperscript{64} and synthetic olfactory receptor-based polypeptides\textsuperscript{65,66} with a quartz crystal microbalance for gas sensing.

In addition to the synthetic pathway, short peptide phage display can be exploited.\textsuperscript{67,68} Phages are viruses that infect bacteria (typically \textit{Escherichia coli}), and they are used as expression vectors for a foreign DNA strand, which will be expressed by the infected host as a peptide and displayed on the surface of the virion. Phage display libraries are mixtures containing as many as $10^9$ members of different phage clones, each displaying a different peptide generated by randomizing the gene sequence. A binding event between the displayed peptide and an immobilized version of the target is necessary to select and isolate the binder peptides (\textit{figure 1.2.4}); the indispensable introduction of a linker on the target can influence the binding event, and it can be a huge obstacle in particular in case of small molecules.\textsuperscript{69,70} However, phage display allows to generate many binder peptides from fully random libraries, without any kind of design\textsuperscript{71,72,73} and this feature may offer peptides able to discriminate between fine structural differences in the target, and even between different conformations.\textsuperscript{74,75}

![Figure 1.2.4](image.png) The cycle of peptide phage display for the selection of binder peptides.

Finally, short peptides can also be used as active elements for the detection of their own natural receptors; for examples antimicrobial peptides (AMPs) and cell-penetrating peptides (CPPs) are used for sensing bacterial cells,\textsuperscript{76,77,78,79} antigenic peptide sequences for antibody monitoring,\textsuperscript{81} and peptide substrates for enzyme detection.\textsuperscript{82,83,84}

23
1.2.2.5 Aptamers

Aptamers are short, single-stranded nucleic acid ligands (RNA, ssDNA, modified ssDNA, or modified RNA) that are isolated from libraries of oligonucleotides by an in vitro selection process called SELEX (Systematic Evolution of Ligands by EXponential enrichment).\textsuperscript{85,86} Aptamers bind with high affinity and specificity to a broad range of target molecules, and have been produced against a wide range of targets including small molecules, proteins and whole cells, making them very suitable for detection purposes.\textsuperscript{87,88,89} The target recognition event occurs primarily by shape and not by sequence;\textsuperscript{90} the mechanism is still not well understood and this makes often their discovery a trial-and-error procedure in most laboratories.

Aptamers are more flexible than proteins: in solution they are predominantly unstructured, but upon associating with their ligands they fold into molecular architectures, thanks to their self-annealing properties, and the target becomes an intrinsic part of the nucleic acid structure. Their use in biosensors is very promising due to their features: high binding affinity, simple synthesis, easy storage, and wide applicability;\textsuperscript{91} anyway some improvements must be done in biosensing surface covering. Fluorescent detection is widely used, allowing a real-time detection. Aptamer-based molecular beacon (aptabeacons) is a frequently adopted format,\textsuperscript{92,93} derived from traditional molecular beacons.\textsuperscript{94} Many different fluorophores and quenchers are available, moreover the procedure of labeling aptamers with fluorescent dyes is easy. In the case of a hairpin structured aptamer, it can be end-labeled with a fluorophore and a quencher: the binding of the target disrupts the stem, separating the fluorophore from the quencher and leading to fluorescence signals (figure 1.2.3-A). Another possibility is a duplex structure between a fluorophore-labeled aptamer and its complementary DNA sequence bearing a quencher: the presence of targets forces the departure of the complementary strand from the aptamer, accompanied by an increase in fluorescence (figure 1.2.3-B).\textsuperscript{94,95,96}

\textbf{Figure 1.2.3}\textsuperscript{92} Aptabeacons: A an aptamer sequence in a molecular beacon-like hairpin structure, end-labeled with a fluorophore (F) and a quencher (Q); B a fluorophore-labeled aptamer in a duplex structure with a complementary DNA sequence labeled with a quencher. In both the cases fluorophore and quencher are separated after the binding of the target, increasing the fluorescence.
1.2.2.6 Peptide Nucleic Acids

Peptide nucleic acids (PNA) are synthetic DNA analogues in which the negatively charged ribose-phosphate backbone of nucleic acids is replaced by an uncharged polyamidic one, specifically a N-(2-aminoethyl) glycine scaffold, and the nucleobases are attached via a methylene carbonyl linker.\(^97,98\) They are able to form double- and triple-stranded complexes with nucleotides, because the intramolecular distances and configuration of the nucleobases are similar to those of natural DNA or RNA molecules (Figure 1.2.4).\(^3\)

Compared to nucleic acids, PNAs exhibit superior hybridization characteristics and improved chemical and enzymatic stability,\(^99\) moreover the uncharged nature of PNAs makes the PNA–DNA duplexes more thermally and pH stable than DNA–DNA equivalents, enabling even the duplexes formation in low-salt conditions because no positive ions are necessary for counteracting the interstrand repulsion as in the case of two negatively charged DNA filaments. On the contrary single-base mismatches is considerably more destabilizing and the effect depends on the position of the mismatch within the sequence.\(^100\)

Due to their features, PNAs are often used as natural nucleic acids recognition element,\(^99,101,102\) even avoiding the introduction of a label.\(^103\)

\[\text{Fig} \ 1.2.4\] Comparison of the structures of peptide nucleic acid (PNA) and DNA. The deoxyribosephosphate backbone of DNA (brown) is replaced by N-(2-aminoethyl)-glycine units in PNA (blue). The nucleobases are shown in black.
Chapter 1 - Introduction

Chapter 1 - Introduction

1.3 Human Serum Albumin

1.3.1 Human Serum Albumin as a starting point

A first possibility towards the development of a selective biosensor is to start from a natural binding protein showing a relatively high affinity for the ligands of interest and to modify it to improve the binding. The protein sequence can be reduced down to the limit of stability and retention of the binding properties, then peptide libraries can be built on such structural scaffold by random or rational mutation strategies and their screening allows to select the more promising binders; the mutation and selection run could be repeated until the desired degree of affinity improvement is reached.\textsuperscript{1,2,3,4} Some factors that can guide the choice of the starting protein are: a) a broad binding activity in its “wild-type” status, b) an easily manageability, c) the existence of a large panel of information and literature data, useful if eventually problems grow up during the project development, and possibly d) its relatively low cost. On these bases Human Serum Albumin (HSA) represents a valid candidate, thanks to its well-known capability to reversibly bind a large variety of both endogenous and exogenous substances with a micromolar affinity. In the last century many scientists have been attracted by albumin and by its important role in many physiological and pathological processes, so it turns out to be one of the most studied proteins. There are many applications of it, both in clinical medicine and basic research.\textsuperscript{5}

The most essential physiological role of HSA, which is the most abundant plasma protein (600 \textmu M, 42 g/L), is probably the transport of relatively hydrophobic compounds in the bloodstream to target organs such as liver, intestine, kidney, and brain.\textsuperscript{5} It is able to carry nutrients, hormones, fatty acids, amino acids, metals, such as Cu\textsuperscript{2+} and Zn\textsuperscript{2+}, and many pharmaceutical drug compounds, like warfarin, diazepam, and ibuprofen.\textsuperscript{6,7,8} HSA can also sequester potentially toxic molecules.\textsuperscript{9} Furthermore, it can act as an antioxidant, because it can interact or inhibit the formation of various oxidizing agents.\textsuperscript{10} Finally, HSA is significantly involved in the regulation and maintenance of the extracellular fluid levels, of the colloidal osmotic pressure in blood and of the haematic pH.\textsuperscript{5}

HSA also shows a catalytic ability, associated to the binding site located in the albumin subdomain IIA, known as the Sudlow site I,\textsuperscript{11,12} towards certain chemical reactions like \beta-eliminations,\textsuperscript{13} the decomposition of Meisenheimer adducts,\textsuperscript{14} and the Kemp elimination.\textsuperscript{15,16} Moreover, HSA is able to direct the stereoselective reduction of diketones and to catalyze the aldol reaction, how it has been recently reported by our research group.\textsuperscript{17,18,19}

1.3.2 HSA structure and properties

The primary structure of HSA was obtained by classical chemical sequencing methodologies\textsuperscript{20,21} before obtaining the DNA sequence.\textsuperscript{22} The albumin gene contains three polyadenylation signals, and this 3' region presumably arose by triplication of a shorter fragment prior to mammalian radiation,\textsuperscript{23} moreover it exhibits a high degree of DNA polymorphism and appears to have been recently invaded by Alu repetitive sequences,\textsuperscript{22} i.e. short interdispersed elements containing a recognition site for the restriction enzyme AluI.\textsuperscript{24}
HSA is synthesized by the parenchymal cells of the liver and exported as a non-glycosilated protein; it consists of a single polypeptide chain of 585 amino acids with a molecular mass of about 66500 Da. The crystal structure\cite{6,25,26} shows three homologous \( \alpha \)-helical domains, numbered I, II and III from the N-terminal residue, assembled to form a tetrahedral heart-shaped molecule; each domain contains ten \( \alpha \)-helices and can be further divided into subdomains A and B, formed by six and four \( \alpha \)-helices respectively (figure 1.3.1). The binding properties are different in the three domains, even if they are homologous.

**Figure 1.3.1** 3-D structure of HSA. The three domains are color-coded: I, red; II, green; III, blue. The small-molecule binding sites in domains IIA (drug site 1) and IIIA (drug site 2) are shown.

HSA is composed only by amino acids, without prosthetic groups. Unlike the large majority of proteins, HSA has a low content of methionine, glycine and isoleucine, while leucine residues and ionic amino acids such as glutamic acid and lysine are numerous;\textsuperscript{5} this means 185 ions per molecule at pH 7 and explains its high solubility. A typical feature of HSA is the presence of a single tryptophan residue and a high occurrence of cysteine residues, precisely 17 disulphide bridges,\textsuperscript{27} determining high cross-linking degree, and one free cysteine, located in subdomain IA (Cys34). Disulphide bonding is a characteristic of extracellular proteins and contributes to the stability of plasma proteins during the circulation. The disulphide bridges pattern is a structural feature of the albumins. Almost half of the Cys residues are located in contiguous positions along the chain; although adjacent Cys-Cys pairs have been shown to couple each other,\textsuperscript{28} they are generally considered as linked to the nearest residue in the chain before and after the pair. The presence of many disulfide bonds explains the particular resistance of HSA to external conditions.

### 1.3.3 Binding sites

HSA works as a very versatile drug carrier, indeed several bioactive substances reversibly bind the protein in one or very few high-affinity sites of it,\textsuperscript{12,29,30} with typical association
constants in the range of $10^4$ - $10^6$ M$^{-1}$. In addition to the primary sites, several sites of lower or of very low affinity can be found. The binding affinity of any drug to HSA is one of the major factors that determine its pharmacokinetics, i.e. time course of drug absorption, distribution, metabolism and excretion, and its bioavailability: a too low affinity means an unfeasible absorption, while an excessively strong interaction determines a higher dose necessary to achieve the effective concentration in vivo. However interaction studies are primarily focused on HSA high-affinity binding sites, because the plasma concentration of drugs in normal treatments is generally much lower than that of HSA (~ 600 μM), therefore it can be assumed that only high-affinity sites are occupied.

Another important aspect is that the binding of chiral molecules, as are the large majority of sold drugs, to HSA is often enantioselective; it is therefore necessary to investigate how much the stereochemistry of a bioactive molecule influences its delivery and therefore its pharmacokinetics.

The existence of two major binding sites, named Sudlow sites I and II, was demonstrated in a study based on competition with fluorescent probes. A following crystallographic study extended this model, identifying seven binding regions (named FA1-7) for long-chain saturated fatty acids; the site FA7 and the complex FA3 - FA4 corresponds to the already identified Sudlow sites I and II, respectively (figure 1.3.2). Subsequent crystallographic studies have shown that these seven sites bind also intermediate-chain fatty acids, long-chain mono and polyunsaturated fatty acids. Hemin, thyroxine, and several drugs bind to one or more of these seven binding sites (figure 1.3.3). The confirmation of the long-chain fatty acid binding site position in the protein and the determination of their relative affinity, which follows the order FA5 > FA4 > FA2 >> FA1, FA3, FA6, FA7, were achieved by means of NMR spectroscopy.

Figure 1.3.2 HSA structure with various ligands. The colors of the six subdomains of HSA are: blue, IA; cyan, IB; dark green, IIA; light green, IIB; orange, IIIA; red, IIIB. The heme (red) fits the primary cleft in subdomain IB, corresponding to FA1. Sudlow site I (in subdomain IIA, corresponding to FA7) is occupied by warfarin (purple). Sudlow site II (in subdomain IIIA, corresponding to FA3-FA4) and FA6 (in subdomain IIB) are occupied by ibuprofen (magenta). Sites FA2 (at the I-IIA interface) and FA5 (in subdomain IIIB) are occupied by myristate (yellow). Atomic coordinates were taken from Protein Data Bank entries 1O9X, 1H9Z, and 2BXG.
Summary of the ligand binding capacity of HSA as defined by crystallographic studies. Ligands are depicted in space-filling representation; oxygen atoms are coloured in red; all other atoms in fatty acids (myristic acid), other endogenous ligands (hemin, thyroxin) and drugs are coloured in dark-grey, light grey and orange, respectively.

1.3.3.1 Site I

The site I of HSA, also known as warfarin-azapropazone binding site or Sudlow site I, appears in the crystal structure as a pre-formed and conformationally adaptable binding pocket within the core of subdomain IIA and comprise all its six helices and a loop-helix feature of subdomain IB (residues 148-154). The inner wall of the pocket is predominantly non-polar, but two clusters of basic amino acids are located at its bottom (Tyr150, His242, Arg257) and at its entrance (Lys195, Lys199, Arg218, Arg222). The site can be divided into a large central area with three different compartments: two hydrophobic sub-chambers (one on the right and the other on the left) separated by Ile264 and a third compartment which protrudes from the front of the pocket. A model with at least three regions, called Ia, Ib and Ic, has been proposed, updating a previous hypothesis of at least two overlapping subsites (warfarin and azapropazone): basing on this last conjecture the Ia region is located in the middle and plays a connection role between the regions Ib and Ic, which on the contrary seem to be disconnected one from the other and to have independent binding properties, as suggested by equilibrium dialysis experiments. Circular dichroism experiments have also shown that specific markers binding to Ic site, after Ib allosteric change, vary the spatial orientation of specific markers to Ib site, without influencing the affinity.
Dicarboxylic acids and/or bulky heterocyclic compounds with a negative charge located toward the centre of the molecule bind with high affinity to site I of HSA, however these features are not enough to predict whether a molecule can bind to this site or not, because molecules very different in their structure have shown to bind to site I with high affinity (figure 1.3.4). Several examples have been found of independent binding of two different compounds to the site, such as phenylbutazone-indomethacin and azapropazone-indomethacin. Recently the interactive association of three drugs (cinnamic acid, lamivudine and indomethacin) has been proved. This phenomenon supports the presence of multiple subsites overlapped or at least very close to each other and somehow linked together, indeed the binding of a molecule able to change the three-dimensional arrangement of the binding site could influence the structure of the others.

![Chemical structure of some molecules that bind to HSA major site I.](image)

Figure 1.3.4 Chemical structure of some molecules that bind to HSA major site I.

Generally the binding between drugs and protein appears to be dominated by hydrophobic contacts, in the central non-polar portion of the pocket, but even other specific interactions with residues belonging to the inner and outer polar clusters could occur. For example in the absence of fatty acids, warfarin, phenylbutazone, CMPF (3-carboxy-4-methyl-5-propyl-2-furanpropanoic acid) and oxyphenylbutazone are positioned to form a hydrogen bond with Tyr150.

The unique Trp214 residue of HSA is located within this pocket and this makes fluorescence spectroscopy a powerful and widely used technique to monitor what happens in Sudlow site I: molecules that bind to this site, or that alter its structure in an allosteric manner, modify the microenvironment around the residue, thus varying its intrinsic fluorescence signal. The intrinsic fluorescence of HSA appears at 340 nm when excited at 280 nm and it is originating from Trp214, but also from the Tyr and the Phe residues, even if the latter ones give a small contribution due to their low fluorescence quantum efficiency. Synchronous fluorescence spectroscopy, which is performed changing simultaneously the excitation and the monitored emission wavelengths maintaining always the same difference between them, allows to distinguish the two major contribution: a $\Delta \lambda$ set at 15 or 60 nm offers the
characterization of Tyr or Trp residues, respectively.\textsuperscript{58} Another way is to use an excitation wavelength less than 295 nm, which is able to excite both Tyr and Trp residues, and then to subtract the contribution of the only Trp measured at an excitation wavelength in the range 295-305 nm, where the Tyr residue is not excited.\textsuperscript{5}

1.3.3.2 Site II

Site II, also known as indole-benzodiazepine site or Sudlow site II, is located within subdomain IIIA and comprehends all its six helices;\textsuperscript{7,11,26} its overall structure, a largely pre-formed hydrophobic cavity with distinct polar features, is similar to that of site I. However site II is smaller: its principal binding region is similar to the central portion of the site I pocket but it seems to possess just one sub-compartment, corresponding to the right rear of the site I and accessible only after ligand-induced rearrangement of HSA. As a consequence this binding site appears to be more restrictive and highly stereoselective. Moreover a unique cluster of polar amino acids is present (Arg410, Tyr411, Lys414, Ser489), located on one side at the entrance. The clear differences in shape, size and polarity between the two major sites help to account for the different binding characteristics.

Aromatic carboxylic acids, containing a hydrophobic core and the acid moiety peripherally located, like for example non-steroidal anti-inflammatory drugs (FANS), well bind within site II (figure 1.3.5);\textsuperscript{5} also fatty acids interact with a certain selectivity (FA3 and FA4 sites).\textsuperscript{36,37} Bulky molecules, such as bilirubin or hemin, can not fit due to their too large dimensions.

![Figure 1.3.5 Chemical structure of some molecules that bind to HSA major site II.](image)

Replacement of small groups on the ligands may strongly influence the affinity: diazepam, but not its fluorinated analogue flunitrazepam, binds site II with high affinity;\textsuperscript{59} or \(\alpha\)-methylation of an exogenous tryptophan prevents its binding to this site.\textsuperscript{60} Tryptophan offers also a typical example of site II enantioselectivity: the L-enantiomer shows an affinity 100 times higher than the D-one.\textsuperscript{61,62} Site II ligands (diazepam, ibuprofen, diflunisal and indoxyl sulfate) are found, in the X-ray structures, in the central area of the site, with at least one oxygen pointing towards the polar
cluster. Arg410 and Tyr411 residues are generally considered important for the binding of molecules, such as ketoprofen, although testing on recombinant HSA (rHSAs) have shown that Arg410 is not essential for the binding of diazepam.

1.3.4 Influence of endogenous ligands on HSA binding properties

Binding of a molecule to HSA often influences the protein binding properties towards other molecules; this is due to the high conformational flexibility of albumin and to the presence of few high-affinity binding sites. The simultaneous binding of endogenous compounds such as fatty acids, uremic metabolites, bile salts and bilirubin can alter the basic HSA interaction binding properties and so modulate the drug affinities to the protein, changing their pharmacokinetic properties and sometimes causing the onset of toxicity; for example, warfarin binds to HSA with a very high affinity only at low levels of fatty acid, because the conformation of the warfarin binding pocket is significantly altered upon binding of fatty acids and this causes a decrease in interaction.

This risk is especially relevant if two or more drugs compete for the same HSA-high-affinity binding site: 3-carboxy-4-methyl-5-propyl-2-furanpropanoic acid (CMPF), which binds specifically to the site I, competes with drugs that bind to the same site as warfarin, furosemide and bilirubin. Similarly indole-3-acetate toxins, indoxyl sulfate and hippurate, which bind with high affinity to site II, displace drugs that bind to this site. A wide set of site I ligands, such as valproate, ceftriaxone, phenylbutazone, glibenclamide, tolbutamide, warfarin, and furosemide competes with bilirubin, a toxic and insoluble heme catabolism product, which is sequestered and solubilized by HSA. The presence of high levels of free bilirubin can be very dangerous, especially in infants. Therefore, the effect of all drugs that are administered to infants on the binding between HSA and bilirubin should be tested.

Sometimes free drug concentration increases less than expected, because after its displacing it re-binds at another site. This phenomenon was observed with diclofenac in the presence of ibuprofen.

---

Chapter 1 - Introduction


64 Chuang V. T., Otagiri M. Stereoselective binding of human serum albumin. Chirality, 2006, 18, 159-166.
Chapter 1 - Introduction


66 Onishi S., Itoh S., Isobe K., Ochi M., Kunikata T., Imai T. Effect of the binding of bilirubin to either the first class or the second class of binding sites of the human serum albumin molecule on its photochemical reaction. *Biochem. J.*, 1989, 257, 711-714.


1.4 HSA100 and mutant libraries

1.4.1 Directed evolution strategy

In the last years directed evolution\textsuperscript{1,2,3} has become a key method that mimics \textit{in vitro} the natural evolution process to improve essentially any property of a protein, such as thermal robustness,\textsuperscript{4} binding activity,\textsuperscript{5} stereoselectivity.\textsuperscript{6,7,8} Molecular diversity is created by random mutagenesis and/or recombination of a target gene or a set of related genes; the variant libraries are expressed in a suitable bacterial host and screened for an improved property, then the genes encoding the best variants can be used as templates for the next round of direct evolution; all the cycle could be repeated until the desired degree of improvement is reached. Random mutagenesis could generate large libraries of mutant genes ($10^6$-$10^{10}$ members),\textsuperscript{3} and following of protein variants: an efficient screening in these cases becomes hardworking. To address the problem, site-specific saturation mutagenesis can be applied;\textsuperscript{9,10} all possible amino acids are introduced in specific positions in the vicinity or within the site involved in the property of interest; in this way only focused libraries are produced, limiting the following screening. A combinatorial active-site saturation test (CAST)\textsuperscript{11,12,13} is an approach derived from site-specific mutagenesis: considering the three-dimensional structure of the protein, some sets of two or three amino acids in key positions are simultaneously randomized, with the creation of relatively small libraries of mutants; after their screening the gene of each hit is used as a template for a second round of CASTing, modifying other few chosen positions, even combining the beneficial mutations achieved in the first generated libraries.

1.4.2 HSA100

Looking for a selective biosensor for coffee polyphenols, directed evolution seems to be a powerful tool. The choice of HSA as a first binding biomolecule attempt is promising due to its broad ligand properties, but its big size (585 residues) makes its manipulation in mutation cycles hard to handle. The reduction of HSA dimension down to the limit of stability and retention of the binding properties turns out to be a crucial step that can facilitate the mutagenesis strategy. In our research group a peptide corresponding to a 101-residue stretch of the HSA sequence (from Ala194 to Glu294), named HSA100, has been already identified, synthesized and fully characterized;\textsuperscript{14,15} since it retains the binding properties of the parental HSA site I, being in the same time the shortest functional peptide ever derived from the sequence of HSA, it appears as the perfect candidate for the following mutation steps.

1.4.2.1 HSA100 identification and optimization

Structural and functional analyses, performed by Prof. Laio’s group (Sissa, Trieste), led to identify the HSA100 fragment, which corresponds to approximately half of subdomain IIA (figure 1.4.1).
Figure 1.4.1 Structure of the human serum albumin HSA100 domain. The full-length protein and HSA100 domain are shown as white and red ribbons, respectively. The structure was obtained from the Protein Data Bank (1BKE.pdb).

The binding site inside this subdomain is composed of a continuous sequence of amino acids with no contribution from other far residues; all the residues involved in the contact with typical ligands, such as warfarin, are included in this sequence, as well as all the eight cysteine residues forming the four disulphide bridges (Cys200-Cys246, Cys245-Cys253, Cys265-Cys279, Cys278-Cys289), which probably stabilize the polypeptide structure. The lysine residue (Lys199) is also present in the fragment: it is surrounded by a hydrophobic environment and it plays a key role for the enzyme-like catalytic behavior of albumin. Furthermore, the fluorescence properties of the native protein are retained: the tryptophan residue (Trp214) located at the bottom of the main hydrophobic subsite has been preserved and it can be still exploited as a useful internal fluorescent sensor to monitor the binding phenomena with different ligand molecules. Seventy-four of the amino acids are located in six $\alpha$-helical regions ($\alpha_1$: Gln196-Phe206; $\alpha_2$: Arg209-Arg222; $\alpha_3$: Ala229-His247; $\alpha_4$: Leu250-Ser270; $\alpha_5$: Leu284-Gly292; $\alpha_6$: Leu284-Gly292), while the remaining residues form six connecting coil/loop regions.

1.4.2.2 Expression of HSA100

The coding sequence of HSA100 fragment was cloned and expresses in *Escherichia coli*, to simplify the process, using a procariotic expression vector pGEX 4T-1. A preliminary evaluation of protein production was performed by fusing the HSA fragment to either a carrier maltose-binding protein (MBP) or different tag sequences (HIS6 and strep tags): in both cases HSA100 was expressed, purified, and turned out to be active, but the protein yield was extremely low.

A two-step approach has been then used, to improve the result: first the wild type HSA100 DNA sequence was modified using an optimization strategy to maximize both transcription and translation efficiency. After chemical synthesis of the DNA, the fragment coding for
HSA100 was cloned as a fusion protein with a glutathione S-transferase (GST) carrier, generating the GST-HSA100 fusion (Figure 1.4.2).

Figure 1.4.2 Amino acid sequence of GST (black) in fusion with HSA100 (red) and FLAG tag.

In the second step, protein production and purification were optimized by testing a number of different parameters: bacterial strains, temperature of growth, induction condition, and protein extraction protocol. By using the best conditions, 15-20 mg/l of protein was produced, from which up to 4 mg/l of soluble protein was obtained. The soluble fraction was purified by affinity chromatography using a GSH resin, resulting in highly pure, stable, full-length protein (Figure 1.4.3).

Figure 1.4.3 Designed construct bearing the polypeptide GST-HSA100-FLAG tag (a), SDS-PAGE (b), anti-FLAG tag Western blot of purified GST-HSA100 protein (c).

The identity of GST-HSA100 was confirmed by LC-MS/MS and peptide sequencing analyses. The secondary structure composition was defined by means of CD spectroscopy and it is: about 50% α-helix, 30% coil, 16% turn and less than 5% β-sheet. This result is consistent with the GST-HSA100 fusion sequence, which expected composition is 55% helix, 25% coil, 15% turn and 5% sheet, considering the presence of six helices and six coil/loop regions of the HSA100 domain and of two GST domains with a βαβαβαβα folding topology along the first 84 residues followed by an extended α-helical domain of 132 amino acids at the C-terminus.21 GST-HSA100 retains the binding ability of the native protein, how it was demonstrated by fluorescence spectroscopy measurements of its affinity for warfarin and efavirenz, chosen as reference ligand molecules; GST has no role in the interaction and it doesn’t significantly contribute to the fluorescence emission, because its Trp residues are buried in a hydrophobic
environment.\textsuperscript{21} The fragment replicates also the HSA catalytic activity: the chemical control of the diketone reduction and of the aldol reaction is almost the same both in the presence of the whole protein or of its smaller derivative.\textsuperscript{15}

HSA100 is the first albumin fragment successfully expressed in \textit{E. coli}. In the past decade, several larger fragments have been described, all obtained as secreted proteins by expression in \textit{Pichia pastoris}.\textsuperscript{22,23,24,25,26}

### 1.4.3 HSA100 mutant libraries

HSA100, thanks to its reduced dimensions and its preserved binding properties, is the perfect template to be used in directed evolution. In order to improve the affinity of HSA100 fragment for the targets of interest, the site-directed mutagenesis of several amino acids included in the subdomain IIA and directly involved in the interaction between the protein and its ligands was performed.\textsuperscript{14}

#### 1.4.3.1 Structural analysis

Three different regions can be identified in subdomain IIA: a polar entry and two hydrophobic subsites, one larger and one smaller. Seven mutating positions have been identified; bioinformatic analysis of sequence conservation within the set of all mammalian serum albumins was performed to check their non-essential role for correct peptide folding, before any manipulation. Two of the selected positions are placed inside the major hydrophobic pocket (Leu260, Ile264), two inside the minor hydrophobic pocket (Ala215, Leu238) and the remaining three are at the polar site of entry (His242, Arg257, Ser287) (\textit{figure 1.4.4}). The X-ray structure of the warfarin-HSA complex shows that these amino acids are in close contact with the drug.\textsuperscript{27}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{image.png}
\caption{Tridimensional structure of the 100 amino acids fragment with the positions submitted to site-directed mutagenesis.}
\end{figure}
1.4.3.2 HSA100 site-directed mutagenesis

The strategy for cloning the mutant libraries started basing on HSA100 wild-type optimized DNA template: the oligonucleotides were designed to insert separately the first two mutations at Ala215-Leu238, the second pair of mutations at Leu260-Ile264 and the last three mutations at Hys242-Arg257-Ser287. Up to date, the first two of the three planned libraries of HSA100 variants were created by PCR and by molecular assembly. Each position was mutagenized by inserting a NNK triplet into the reverse primer; this triplet encodes for any of the 20 possible amino acids and not for the stop codons. Two PCR rounds were performed for both the produced libraries. Several clones were sequenced, verifying the randomization of selected residues, to check the actual diversity between the members of the same library. The sequences of some analyzed clones of the two libraries are reported in Table 1.4.1 and 1.4.2, confirming that effective randomization occurred.

Table 1.4.1 Sequence analysis of six random clones of mutants library 1 pGex-HSA100 Ala215-Leu238.

<table>
<thead>
<tr>
<th>Mini-library 1</th>
<th>Ala 215 wild type</th>
<th>Leu 238 wild type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clone 1</td>
<td>Threonine (T)</td>
<td>Valine (V)</td>
</tr>
<tr>
<td>Clone 2</td>
<td>Leucine (L)</td>
<td>Lysine (K)</td>
</tr>
<tr>
<td>Clone 3</td>
<td>Asparagine (N)</td>
<td>Lysine (K)</td>
</tr>
<tr>
<td>Clone 4</td>
<td>Lysine (N)</td>
<td>Proline (P)</td>
</tr>
<tr>
<td>Clone 5</td>
<td>Lysine (K)</td>
<td>Valine (V)</td>
</tr>
<tr>
<td>Clone 6</td>
<td>Arginine (R)</td>
<td>Serine (S)</td>
</tr>
</tbody>
</table>

Table 1.4.2 Sequence analysis of six random clones of mutants library 2 pGex-HSA100 Leu260-Ile264.

<table>
<thead>
<tr>
<th>Mini-library 2</th>
<th>Leu 260 wild type</th>
<th>Ile 264 wild type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clone 1</td>
<td>Leucine (L)</td>
<td>Valine (V)</td>
</tr>
<tr>
<td>Clone 2</td>
<td>Valine (V)</td>
<td>Valine (V)</td>
</tr>
<tr>
<td>Clone 3</td>
<td>Serine (S)</td>
<td>Leucine (L)</td>
</tr>
<tr>
<td>Clone 5</td>
<td>Proline (P)</td>
<td>Arginine (R)</td>
</tr>
<tr>
<td>Clone 15</td>
<td>Tryptophan (W)</td>
<td>Histidine (H)</td>
</tr>
<tr>
<td>Clone i</td>
<td>Arginine (R)</td>
<td>Glutamic Acid (E)</td>
</tr>
</tbody>
</table>

No attempts of selection were made at this stage and twenty clones were just randomly chosen without trying to fish out the best binders from the small libraries. An efficient screening of soluble bacterial expression products requires dedicated strategies, and this step comes after the optimization of the peptide libraries production. Each recombinant protein was produced with the optimized protocol previously described for the GST-HSA100. The concentrations of the purified proteins were in a range between 0.2 mg/ml and 4 mg/ml.
Chapter 1 – Introduction

Chapter 2

Aim of research
2. Aim of research

The project of this Ph.D. was developed considering the interest in controlling the coffee quality throughout its production steps, focusing in particular on its roasting. It has been funded by Trieste Coffee Cluster.¹

The roasting process plays a key role on the final chemical composition of coffee, due to the many reactions that occur in the beans subjected to high temperatures and high inner pressures. A mistake in this step can damage aroma and taste of the coffee brew, even starting from green coffee beans of high quality. At present, coffee roasting is monitored basing on the color and the weight loss of the beans, while sensory analysis, gas and liquid chromatography, often coupled with mass spectroscopy, are the most used technique to evaluate the coffee characteristics.

The coffee quality in the cup has been related to the amounts of chlorogenic acids (CGAs) and of their roasting derivatives, in particular of lactones (CGLs); a biosensor or an array of sensors able to discriminate between CGAs and CGLs and to quantify them could be a useful tool in the coffee industry.

In this perspective some chosen target lactones will be first synthesized, then their interaction with Human Serum Albumin, with its fragment HSA100 and with some randomly selected mutants will be studied using fluorescence spectroscopy. To exploit the binding features optimization offered by biomolecules mutagenesis, a linker equipped lactone will be immobilized on a gold chip and a Surface Plasmon Resonance method will be set, to use them in mutants libraries screening.

The interaction of the lactones with caffeine will be also studied by NMR spectroscopy. Moreover, considering that CGLs have shown potential bioactive properties, the possible activity of the synthesized compounds against some viruses, like Influenza and HIV, will be evaluated in biological assays.

Chapter 3

Results and discussion
3.1 Quinides synthesis

3.1.1 Synthesis of 3,4-O-dicafeoyl-1,5-γ-quinide (7)

3,4-O-dicafeoyl-1,5-γ-quinide (3,4-diCQL, figure 3.1.1) is a bitter lactone¹ present in roasted coffee and it is more abundant in C. Canephora than in C. Arabica, because the first is richer in chlorogenic acids (CGAs), which are the direct precursors of the quinides family: during the roasting of the coffee beans, the CGAs undergo different transformations such as lactonisation reactions and 3,4-diCQL is one of the obtained derivatives.²³ This compound, together with 3,4-O-diferuloyl-1,5-γ-quinide and 3,4-O-dicoumaroyl-1,5-γ-quinide, is able to inhibit the human adenosine transporter at low micromolar concentration, potentially counterbalancing the stimulating effect of caffeine.⁴

In the optic of a selective biosensor for the quinides family, 3,4-diCQL was chosen as a first target, to use it for the preliminary studies of the interaction with the possible binding molecules, for example Human Serum Albumin.

The extraction of 3,4-diCQL from the coffee powder is possible but it is very laborious and a preparative reverse phase HPLC is required, moreover its amount is really small, in the best case (Robusta from Uganda, light medium roasting degree) about 30 mg per 100 g of coffee, on dry matter base.³ The direct synthesis is preferable and it is reported in literature;¹⁵ in this thesis its synthesis was achieved following a literature procedure¹ although with some modifications.

The synthesis, illustrated in scheme 3.1.1, starts from D-(−)-quinic acid, which was refluxed in dry acetone in the presence of a catalytic amount of p-toluenesulfonic acid monohydrate to give at the same time the lactone ring formation and the protection of the hydroxyl groups on carbons 3 and 4 as acetonide; to favor the reaction a Soxhleth apparatus was used: molecular sieves, previously activated in the oven, were placed in the extraction thimble, to capture the water present in the refluxing acetone while passing through them. A reaction time of 48 hours gave a better yield compared to the 24 hours reported in literature.

The hydroxyl group on carbon 1 was then protected using 2,2,2-trichloroethyl chloroformate, that was added dropwise to a mixture of lactone 1 and pyridine in dichloromethane. No formation of precipitate was observed and the reaction was quenched after 24 hours with 1M HCl and brine; lactone 2 was achieved after recrystallization from methanol. Trichloroacetic
Chapter 3 – Results and discussion

Acid and water in the same molar amount and in excess compared with compound 2 were used to remove the acetonide and lactone 3 was recrystallized from toluene.

Caffeic acid too needs to be treated before the coupling reaction, to avoid interferences of the phenolic groups, that were protected using methyl chloroformate to yield product 4 which was recrystallized from 50-50 v/v water-ethanol. The one-pot coupling between lactone 3 and acid 4 was also tried in the presence of 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC-Cl) as activating agent, triethylamine as base, and 1-hydroxybenzotriazole (HOBt) as cooperating nucleophile, but no reaction was achieved, moreover the reagents were not recovered, so probably they degraded in these conditions. The protected caffeic acid 4 was therefore activated with thionyl chloride and the product 5 was immediately used to acylate lactone 3, in the presence of a large excess of triethylamine and 4-dimethylaminopyridine (DMAP) as catalyst. The acyl chloride was slowly added in its solid form, to prevent its partial hydrolysis that is easier in solution. The published conditions were previously tried (chloride solved in few mL of dichloromethane, pyridine as base, and no DMAP), but a mixture of monoesters on carbons 3 or 4 of the quinide, a small amount of diester, and other side products was always obtained; their separation by flash chromatography on silica gel was impossible. On the other hand with triethylamine and DMAP the only diester was achieved, always in admixture with other side products: in this case the purification by flash chromatography of the desired compound 6 was possible, even if with some problems because a gradient elution from CH2Cl2/ethyl acetate 98/2 to 92/8 v/v was needed, moreover a quickly run column had to be performed, indeed only a brief contact with the silica allows to avoid the degradation of the product. The purification of this compound was not described in the published method.
The deprotection of all the hydroxyl groups was achieved by warming the diester 6, dissolved in dry pyridine, to 50 °C for seven days in the presence of LiCl in excess. The crude was purified by flash chromatography on polyamide with a gradient elution from ethyl acetate/methanol 80/20 to 50/50 v/v: in this case silica gel is useless because it causes a fast degradation of the quinidic moiety, moreover the fractions must be dried in few time to avoid the degradation that occurs even in solution. Unfortunately with this treatment only fractions rich in the final compound 7 could be achieved. To isolate compound 7 in a purified matter it was necessary to use semi-preparative RP-HPLC with a C18 column and H2O+0.1% TFA and MeOH+0.1% TFA as solvents: the solubility in the starting, more aqueous, eluent was low: the crude had to be splitted into several aliquotes and a 10 mL loop was used for loading the samples. The fractions corresponding to the peak of interest were checked with ESI-MS and then quickly freeze-dried, again to avoid the product degradation. 3,4-diCQL 7 was afforded as a white powder and stored at -18 °C. Also the use of RP-HPLC was omitted in the reference article.

**3.1.2 Synthesis of 1,3,4-O-tris[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide (10), 3,4-O-bis[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide (11), and 3-O-[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide (12)**

In order to broaden the number of possible target molecules and to focus our attention on the quinide core, which is one moiety of the considered molecules that the biosensor would recognize, a new class of quinides has been synthesized (DMQs), bearing 3,4-(dimethoxy)cinnamoyl groups as side chains; the different kind of substituents moreover could be useful to evaluate the specificity of the biosensor among the quinides family. Three of them were purified and fully characterized (figure 3.1.2): 1,3,4-O-tris[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide (10), 3,4-O-bis[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide (11), and 3-O-[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide (12).

![Figure 3.1.2 Molecular structure of 1,3,4-O-tris[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide (10), 3,4-O-bis[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide (11), and 3-O-[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide (12).](image-url)
This kind of quinides has not been detected yet in roasted coffee, but while compound 10 is a “non-natural” molecule, because it is also esterified at position 1, lactones 11 and 12 are probably formed during the roasting process, indeed the corresponding acids, their direct precursors, were recently characterized in green coffee beans by LC-MS\(^6\),\(^7\) and realistically they undergo the same thermal degradation of the other CGAs. These synthesized lactones shall be used as standards in HPLC coffee brew analysis, to detect their possible presence in the mixture, and to study their biological activities and their contribution to the final coffee brew bitterness, defining their bitter threshold level.

The quinide without protective groups was first prepared from D-(-)-quinic acid under acidic catalysis in toluene and DMF, by following a literature procedure.\(^8\) The pure lactone 8 was precipitated from the brown sticky residue obtained after evaporation of the solvents by the addition of methanol. The recrystallization from refluxing ethyl acetate was also tried, as suggested in the literature, but it was not successful.

The esters were synthesized using the same method performed for the synthesis of 3,4-diCQL 7: 3,4-dimethoxycinnamic acid was activated as chloride and immediately used in the coupling with lactone 8, in the presence of triethylamine and DMAP. This reaction was performed at several different molar ratios between the reactants, achieving every time a diverse mixture of compounds.

At a 1:3.5 lactone/acyl chloride molar ratio, the conversion of 1,5-\(\gamma\)-quinide 8 into dimethoxycinnamoyl esters was quantitative, and two products were obtained: the triacylated quinide 10 and the diester 11 in a 1:0.4 molar ratio. The two products could be easily separated by flash chromatography, and the structures of pure 10 and 11 could be assigned on the basis of bidimensional NMR analysis, IR and MS data. In the \(^1\)H NMR spectrum of 10, three 3,4-dimethoxycinnamoyl groups were found, so necessarily all the hydroxyl groups on the quinide reacted.

At a 1:2.2 molar ratio (scheme 3.1.2), three main products could be isolated: triester 10, which was also in this case the most abundant one, the monoester 12, and the diester 11. Three other products could also be recognized, namely 1,3-O-bis[3,4-(dimethoxy)cinnamoyl]-1,5-\(\gamma\)-quinide (13), 1,4-O-bis[3,4-(dimethoxy)cinnamoyl]-1,5-\(\gamma\)-quinide (14), and 4-O-[3,4-(dimethoxy)cinnamoyl]-1,5-\(\gamma\)-quinide (15). The relative abundance of the compounds obtained, determined by NMR analysis, is reported in table 3.1.1 together with the most significant signals related to the protons of the quinidic core. Purification by flash chromatography on silica gel allowed to obtain pure samples of 10, 11 and 12, whereas 13, 14, and 15 were recovered as a mixture.

<table>
<thead>
<tr>
<th>compound</th>
<th>amount (%)</th>
<th>C-1-H (ppm)</th>
<th>C-4-H (ppm)</th>
<th>C-5-H (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>36</td>
<td>5.38</td>
<td>5.73</td>
<td>5.01</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>5.29</td>
<td>5.68</td>
<td>4.94</td>
</tr>
<tr>
<td>12</td>
<td>23</td>
<td>4.92</td>
<td>4.30</td>
<td>4.75</td>
</tr>
<tr>
<td>13</td>
<td>18</td>
<td>5.19</td>
<td>4.94</td>
<td>4.45</td>
</tr>
<tr>
<td>14</td>
<td>7</td>
<td>4.25</td>
<td>5.44</td>
<td>5.04</td>
</tr>
<tr>
<td>15</td>
<td>5</td>
<td>4.17</td>
<td>5.40</td>
<td>4.95</td>
</tr>
</tbody>
</table>

**Table 3.1.1** Composition of the mixture obtained from 8 and 9 at 1:2.2 molar ratio and diagnostic \(^1\)H NMR signals (δ in CDCl\(_3\)).
The results of the MS analysis carried out with all this lactones were peculiar: their molecular ions could be observed only in the ESI$^+$ MS spectra, whereas in ESI$^-$ mode signals for the corresponding acyl quinic acids that result from the opening of the lactone rings were obtained. This behavior should be considered for the analysis of mixtures from natural sources, because often the identities of the components are assigned only by mass spectroscopy analysis.

Scheme 3.1.2 Coupling between lactone 8 and chloride 9 in 1:2.2 molar ratio.
The direct coupling between D-(-)-quinic acid and acyl chloride 9 was also tried in 1:1 and 1:2 molar ratios. The reactions were performed in the presence of triethylamine and DMAP in dichloromethane at room temperature, as in the previous cases. Such reactions led again to the formation of quinide compounds rather than to quinic acid derivatives, and the triacyl derivative 10 was always the most abundant one. At a 1:1 molar ratio, a mixture of 10, 13, 14, and 1-O-[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide (16; diagnostic 1H NMR signals in CDCl3: δ = 4.02 (C3-H), 4.18 (C4-H), and 4.93 (C5-H) ppm) was obtained with percentage abundance 42, 38, 8, and 12%, respectively, whereas at a 1:2 molar ratio, only 10, 13, and 14 (63, 28, and 9%, respectively) were formed; no unreacted D-(-)-quinic acid was found in the crude mixture, but it was realistically washed away during the workup.

The lactonization of quinic acid was unexpected, as similar methodologies were followed by other authors to prepare chlorogenic acids of the same family,79 rather than their lactonization products; inquiring into this surprising result the role of the large excess of base used during the reaction was evaluated and it can be verified that it does not lead to lactone formation with D-(-)-quinic acid alone under the same conditions.

The reaction was therefore studied in more detail with the aim to set up a protocol for the direct and simultaneous acylation and lactonization of quinic acid; by increasing the molar excess of chloride 9 up to 1:7 to force the acylation of all the hydroxyl groups, the triester quinide 10 was obtained as a unique product in 75% yield, calculated after purification by flash chromatography (scheme 3.1.3).

![Scheme 3.1.3 Direct synthesis of triester 10.](image)

The direct coupling between D-(-)-quinic acid and the acyl chloride 9 was also performed, using a 1:0.5 molar ratio in the presence of only 1 equivalent of triethylamine and without DMAP, replicating a literature procedure, to verify if the presence of DMAP was decisive to get the lactonization. On the contrary to what the other authors found, i.e. a mixture of isomers of 3,4-(dimethoxy)cinnamoylquinic acids, the only quinide 16 and the 1-O-[3,4-(dimethoxy)cinnamoyl]quinic acid (17) were obtained in a 1:1.4 ratio (scheme 3.1.4) together with unreacted D-(-)-quinic acid and 3,4-dimethoxycinnamic acid anhydride, as evidenced by 1H NMR spectroscopy of the crude mixture. The 3,4-dimethoxycinnamic acid anhydride was found only in other crude mixtures of attempted coupling without DMAP, but never when the catalyst was added.
On the basis of the result of this uncatalyzed reaction and of the fact that all the coupling products obtained in the direct reaction with D-(-)-quinic acid are functionalized at position 1, a lactonization mechanism may be proposed, involving two molecules of acyl chloride that react sequentially with the carboxylic group of quinic acid to give an anhydride: the first leads to the esterification of C1–OH, and the second activates the five-membered ring formation (Scheme 3.1.5).

The same reaction using protected caffeic acid chloride 5 led again to quinides, so the lactonization occurs also in the presence of other hydroxycinnamic chlorides, independently of their identity. At the moment the only chlorogenic chlorides were used, although it could be interesting to study the reaction utilizing different acyl chlorides in order to understand if the lactonization really occurs in any case or not.

3.1.3 Synthesis of 3-O-[4-(aminobutanamidoethoxy)feruloyl]-4-O-[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide (28)10

A Biacore, Surface Plasmon Resonance (SPR) sensor system, is an instrument that permits to obtain affinity values for the formation of complexes between proteins or peptides and the ligand molecule: one partner is immobilized on a suitable gold chip, usually but not exclusively the biomolecule, while the other one flows on it. This technique has many advantages: automation of the system, fast measurements of the binding constants, short times for the chip regeneration that means the possibility of testing a high number molecules, small amounts (few mg) needed of both the immobilized and the fluxed molecules (Chapter 3, section 3.3).
Chapter 3 – Results and discussion

Libraries of HSA100-derived peptides had been prepared in our research group\textsuperscript{11} (\textit{chapter 1, section 1.4}) and their screening may be a powerful method for the identification of potential selective biomolecule for coffee lactones. To this purpose we have decided to develop a SPR method, immobilizing a suitable quinide on the gold chip and fluxing HSA or the mutant peptides on it. This was the topic of a masterly degree thesis.\textsuperscript{10} A diester quinide, bearing a feruloyl and a 3,4-(dimethoxy)cinnamoyl moieties as side chains, was chosen to be functionalized: a linker equipped with a terminal amino group, necessary to form a covalent amide bond with the carboxylic group placed on the chip surface, was bound to the unique phenolic group present on the feruloyl moiety (\textit{figure 3.1.3}). With this modification, after the immobilization the diester lactone would expose the quinidic core for the interaction; the defined orientation is useful to better select a binding molecule able to recognize selectively the exposed side.

![Figure 3.1.3 Molecular structure of aminic linker equipped diester quinide (compound 28).](image)

In the synthesis design the linker features that must be taken into account are: its length, great enough to allow the interaction between the immobilized molecule and its partner in solution; its amphiphilic behavior, because the synthesis implies the use of organic solvents, while the immobilization on the chip and the binding studies are performed in aqueous buffers, and the essential presence of a terminal amino group, for the immobilization on the chip through the coupling with its carboxylic groups. In our case the final linker was identified as the condensation product of two smaller molecules: \(\gamma\)-aminobutyric acid (GABA) and ethanolamine.

To really control the immobilization step, in order to obtain meaningful data about the interaction with the partner proteins to be tested, it is important that all the molecules have only one linker side chain, always placed in the same position. These are the reasons why a diester with a unique phenol group was chosen: a monoester may give problems of isomerization, with the intramolecular migration of the aromatic chain between the hydroxyl groups on carbon 3 and 4 of the quinidic core, while in a diester both the positions are blocked; a ferulic acid residue, with one phenolic group, at position 3 and a 3,4-(dimethoxy)cinnamic acid, without phenolic groups, at position 4 of the diester were the best choice to have a single precise position available to react with the linker; moreover these two acids don’t require protection and deprotection steps of the aromatic substituents.

The synthesis starts from ethanolamine (\textit{scheme 3.1.6}), that was first Boc-protected, then activated with tosyl chloride in the presence of diazabicyclo[2.2.2]octane (DABCO) as base,\textsuperscript{12} to get compound 19 in 78\% yield, higher than expected. Ferulic acid was protected, to
prevent secondary reactions on its carboxylic group, as methyl ester in refluxing methanol and using p-toluensulfonic acid as a catalyst. The Boc-ethanolamine tosylate 19 and the methyl ester 20 were bounded together by a Williamson etherification:¹³ the phenolate anion was first generated using potassium carbonate as a base, then it was reacted with tosylate 19; 18-crown-6 was added as complexing agent for potassium cation and co-catalyst for the reaction. After flash chromatography on silica gel, pure compound 21 was obtained in a 55% yield. The next step was the Boc deprotection of terminal amino group, that was achieved by stirring compound 21 at room temperature in a 50/50 v/v mixture of trifluoroacetic acid (TFA) and dichloromethane; after several washing in dichloromethane the deprotected compound 22 was obtained in quantitative yield.

To complete the linker chain (scheme 3.1.7), GABA was previously protected at the amino group with di-tert-butil-dicarbonate (Boc₂O) to give compound 23 in 89% yield.

As a first attempt, the synthesis was designed in order to minimize the number of steps: the coupling between the Boc-protected GABA 23 and ethanolamine was performed, in the presence of triethylamine as a base, HOBt and EDC-Cl as coupling reagents; the resulting amide was obtained in 54% yield. The main problem arose in the Williamson etherification between methyl ferulate 20 and the amide formed by GABA and ethanolamine: the activation of the linker alcohol group as tosylate resulted always difficult, leading in each trial to a mixture of compounds, from which it was not possible to extract the desired one. The mixture was used anyway to try the etherification with methyl ferulate 20, but the expected product was not obtained; °H NMR analysis of the result showed that probably the tosylate linker undergoes an intermolecular dimerization before reacting with the feruloyl phenol group. After many attempts without progress, this procedure was left and the here described one was taken.

The Boc-protected GABA 23 was used in the coupling with compound 22, performed by adding 4-methylmorpholine (NMM), EDC-Cl, and HOBt, to afford the linker equipped methyl ferulate 24 in 78% yield.
To continue the synthesis, it was necessary to restore the acid moiety from the methyl ferulate 24: base hydrolysis was exploited at pH 12 with a 3M lithium hydroxide solution. After 48 hours the solution was acidified to pH 3 and acid 25 was obtained in 83% yield.

In the last part of the synthesis (scheme 3.1.8) the esterification of lactone 3 was achieved. The synthesis of compound 26 was a critical step, indeed even if the hydroxyl group on carbon 3 of the quinide is the easier position to esterify, the reaction on the hydroxyl group at position 4 is however possible, also in the presence of less than one equivalent of acid; if the formation of both the isomers occurs, they are not separable by traditional flash chromatography because the silica gel catalyzes the intramolecular transesterification reaction, leading to a change in the starting molar ratio between the two monoesters and to their co-elution.

The coupling was performed with NMM, EDC-Cl and HOBT and the reaction turned out to be very slow, indeed it ended only after ten days, as it was seen by checking every day the reaction mixture by means of ESI-MS spectroscopy. $^1$H-NMR spectra showed that the result was a mixture of the desired monoester in position 3 and a little amount of diester, detectable even as a low intense peak in the mass analysis, while the monoester on the hydroxyl group at carbon 4 was absent. The desired compound was not purified, to prevent the problems due to the contact with the silica gel, and it was directly used for the next step.
As a first attempt, in the hope to avoid the acyl chloride, as the esterification causes hydrochloridric acid release and therefore pH lowering that might be dangerous for the Boc protections, the second esterification at position 4 was tried with the same method of the first one, using 3,4-dimethoxycinnamic acid with NMM, HOBt, and EDC-Cl, but the reaction failed. To force the esterification, we decided to risk: chloride 9, triethylamine and DMAP were added; fortunately the diester 27 was formed without problems and it was finally purified by flash chromatography on silica gel.

The last step was the deprotection of both the terminal amino group of the linker and of the hydroxyl group at position 1 of the quinidic core, that were performed in sequence. The first reaction was carried out with the same method already used for compound 21, with a 50% blend of dichloromethane and TFA, and the outcome was screened by TLC and checked by means of ESI-MS spectroscopy. The trichloroethoxycarbonyl cleavage was then performed following a literature procedure: the Boc-deprotected intermediate was dissolved in THF and zinc powder and acetic acid were added. After the workup with 2% sodium hydrogen carbonate solution, water, and brine the target 28 was afforded in 80% yield.

Scheme 3.1.8 Outline of compound 28 synthesis, third part.
Chapter 3 – Results and discussion

3.2 HSA, HSA100 and its mutants: interaction studies by fluorescence spectroscopy

3.2.1 Human serum albumin (HSA)

3.2.1.1 The choice of fluorescence spectroscopy

Human serum albumin (HSA) has two binding sites, the Sudlow site I in subdomain IIA and the Sudlow site II in subdomain IIIA: they are different in shape, size and polarity, so even in their binding specificity. These two sites have been studied in depth and their interaction preferences were defined, nevertheless it is not always easy to predict in which of them a ligand molecule would go.

Previous studies showed that several polyphenols present in coffee, such as caffeic acid, ferulic acid, and 5-O-caffeoyl quinic acid interact with the protein inside the Sudlow site I, so it might be foreseeable that other similar molecules, like the synthesized quinides, would have the same behavior.

A major characteristic of Sudlow site I is the presence of a unique tryptophan residue (Trp214) within it. Trp is a fluorescent amino acid: if it is excited at 280 nm an emission maximum at 340 nm is often observed; this maximum may vary from 310 nm to 350 nm, depending on its electronic environment. The radiation absorption causes the transition of the indole ring $\pi$-electrons from the ground state to the excited one; their return to the ground state is possible due to the fluorescence deactivation process, but even to other non-radiative processes, like an intersystem crossing, the solvent quenching and the electron or proton transition to an excited state. Fluorescence lifetime and intensity are the result of the competition among all these contributions and depend on the solvent, the pH and the presence of electron donors or acceptors.

The emission spectrum of Trp is strongly influenced by the solvent polarity, indeed the indole emission depends on the formation of hydrogen bonds; as increasing the solvent polarity, the emission maximum shifts towards longer wavelengths. Moreover the fluorescence intensity is highly dependent on the temperature and, of course, on the Trp concentration, but not linearly, because the emitted signal increases upon increasing the fluorophore concentration only until a critical point, after which an emission self-quenching occurs due to self-association.

A molecule able to interact with Trp214 inside the Sudlow site I causes a quenching of its fluorescence; fluorescence spectroscopy therefore may give fast and precious information about the entrance of a possible quencher inside this site of the protein and it is a widely used technique for this purpose. In the presence of a fixed amount of HSA, the monitoring of the fluorescence intensity changes, pursuant to additions in series of known aliquots of ligand, allows to deduce the interaction characteristics: the Stern-Volmer equation shows if the quenching is more static or dynamic, while the Hill linearization gives an evaluation of the process stoichiometry and of the affinity, in terms of a dissociation constant ($K_D$).
3.2.1.2 Titration assays

Fluorescence quenching titrations were performed to study the interaction between essentially fatty acid free HSA and eight polyphenols: four acids (caffeic acid 29, ferulic acid 30, 5-O-caffeoylquinic acid 31, and 3,4-dimethoxycinnamic acid 32) and the four synthesized quinides (3,4-O-dicaffeoyl-1,5-γ-quinide 7, 1,3,4-O-tris[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide 10, 3,4-O-bis[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide 11, and 3-O-[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide 12). The molecular structures of all the ligand compounds are reported in figure 3.2.1.

![Molecular structures of tested ligand compounds](image)

Figure 3.2.1 Molecular structures of the tested ligand compounds.

In all the measurements, the concentration of protein was set to 0.5 μM in 350 μL of solvent, obtained by diluting 135 μL of 10 mM Na2HPO4 and 2 mM KH2PO4 phosphate buffer with 215 μL of mQ water; the pH was 7.4. The ligand concentration was gradually increased during the titration from 1 μM to 500 μM (points at 0, 1, 5, 10, 20, 40, 60, 80, 100, 120, 140, 160, 180, 200, 250, 300, 350, 400, 450, 500 μM), using ligand standard solutions in DMSO; in the case of diester 7 the fluorescence quenching was also measured in a narrower concentration range, from 0.25 to 200 μM. A further titration was carried out with compound 12 in the 0.25-200 μM range, halving the protein concentration to 0.25 μM, to clarify the results. The excitation (λ<sub>exc</sub> range 190-400 nm) and the emission (λ<sub>exc</sub> 280 nm, λ<sub>em</sub> range 300-400 nm) spectra of the protein alone were recorded at the beginning of any experiments. After each addition of the ligand, the emission spectra and the fluorescence intensity (λ<sub>exc</sub> 280 nm, λ<sub>em</sub> 340 nm) were monitored. To distinguish between tryptophan and tyrosine residues the synchronous fluorescence spectra (SFS) were also measured<sup>7,8,9</sup> by setting the excitation wavelength in the 240-320 nm range while the emission was recorded at Δ = 60 nm, in the 300-380 nm range. All the analyses were replicated three times.
3.2.1.3 HSA fluorescence quenching

HSA fluorescence quenching was observed with all the tested ligands and the behavior was similar in all the titrations; a superimposition of the eight sigmoidal trends obtained by plotting the normalized fluorescence intensity against the logarithm of the ligand concentration is shown in figure 3.2.2. Under our experimental conditions, only 3,4-dimethoxycinnamic acid showed interferences due to its intrinsic fluorescence emission, which is greater at lower concentrations (such self-quenching is most likely due to aggregation at higher concentrations).\textsuperscript{10,11,12} To avoid this interference the blank fluorescence of this ligand had been subtracted from the experimental data. It has been also verified that similar corrections were not necessary in all the other measurements.

![Figure 3.2.2](image)

**Figure 3.2.2** Fluorescence quenching comparison of the eight tested ligands (HSA concentration fixed to 0.5 µM, ligand concentration from 0 to 500 µM).

In the hypothesis that the quenching extent is linearly related to the fraction of occupied binding sites, the inflection point of the sigmoidal functions extrapolated from such experimental trends represents the IC\textsubscript{50}, that is the concentration at which half of the available binding sites are occupied by the ligand. The lower is the ligand concentration required to reach the inflection, the greater is the affinity between the ligand and the protein and it is translated in an earlier quenching of fluorescence. The comparison among the eight ligands in figure 3.2.2 shows how the affinities of the tested molecules vary, having however a similar order of magnitude. Qualitatively, the lactones interact better with HSA than the acids, indeed their sigmoids lie in the lower side of the graphic.
The emission spectrum of the Trp residue, excited at 280 nm, changes upon the titration; figure 3.2.3 reports, as an example, the spectra of compounds 29 (a), 31 (b), and 7 (c). The emission intensity decreases while the emission maximum evidently moves to lower wavelengths and the observed blue shifts are at least of 10 nm; this shift is consistent with a change of the environmental polarity surrounding the Trp residue, most likely resulting from replacement of the solvent in the active site by the less polar molecules of ligand.\textsuperscript{13,14,15,16} The use of DMSO for the ligands solutions does not contribute to the blue shift, as it has been verified by adding only DMSO in the same experimental conditions of all the titrations.

![Graphs showing emission spectra of Trp residue](image)

**Figure 3.2.3** Emission spectra of HSA ($\lambda_{ex} = 280$ nm) in the presence of compounds 29 (a), 31 (b), and 7 (c); the concentrations of the ligands increase from top to bottom (0, 1, 5, 20, 60, 120, 180, 250 µM), while the protein concentration is fixed to 0.5 µM.

A further confirmation of this micro-environmental modification near the fluorophore is given by the effects of the ligand addition on the synchronous emission spectra, with a slight blue shift similar in all the ligands (on average the maximum shifts from 280 to 278 nm); figure 3.2.4 shows as examples the results obtained with compounds 31 (a), and 11 (b). As in the emission, also in the synchronous spectra the fluorescence intensity decreases down to zero with increasing ligand concentration and the amount of measured quenching replicates those obtained in the corresponding fluorescence titration, suggesting that the quenching phenomenon is related to tryptophan emission.
Figure 3.2.4 Synchronous spectra of HSA fluorescence emission ($\Delta = 60 \text{ nm}$) in the presence of compounds 31 (a), and 11 (b); the concentrations of the ligands increase from top to bottom (0, 1, 5, 10, 20, 40, 60, 80, 100, 120, 140, 160, 180, 200, 250, 300, 350, 400, 450, 500 $\mu$M), while the protein concentration is fixed to 0.5 $\mu$M.

An increased emission at higher wavelengths, as in figure 3.2.4-b, is observed with several ligands, but only in the case of compound 32 the intensity increases up to over-range on the right side and a correction on the observed quenching data was necessary; figure 3.2.5 shows the SFS spectra of compound 32 alone (a) and in the protein titration (b).

Figure 3.2.5 Synchronous spectra ($\Delta = 60 \text{ nm}$) recorded with compound 32 alone (a) and in the presence of 0.5 $\mu$M HSA (b); the concentrations of the ligand increase from top to bottom (0, 1, 5, 10, 20, 40, 60, 80, 100, 120, 140, 160, 180, 200, 250, 300, 350, 400, 450, 500 $\mu$M).

3.2.1.4 Stern-Volmer analysis

The quenching of Trp214 fluorescence in the presence of another compound can be dynamic, due to the collisional encounters between the fluorophore and the quencher, or static, thanks to the formation of a non-fluorescent ground state complex between the two partners, and this depends on their affinity. In many instances the observed quenching is the result of the combination of both phenomena. From the fluorescence data it is possible to deduce which contribution is the most relevant in the considered case, using the Stern-Volmer equation (equation 3.2.1) that describes the quenching process: \^{6}
Chapter 3 – Results and discussion

\[ \frac{F_0}{F} = 1 + K_q \tau_0 [Q] = 1 + K_{SV} [Q] \]  \hspace{1cm} \text{Eq. 3.2.1}

The variables $F_0$ and $F$ are the emission intensities before and after the addition of the quencher, respectively, $K_q$ is the bimolecular quenching kinetic constant, i.e. a collisional frequency between freely diffusing molecules, $\tau_0$ is the lifetime of the fluorophore (for the tryptophan fluorescence decay $\tau_0$ is about $10^{-8}$ s),$^{17,18}$ $K_{SV}$ is the Stern-Volmer quenching constant and $[Q]$ is the quencher concentration in mol/L; the protein concentration was fixed to 0.5 µM.

The $K_{SV}$ for all the ligands were determined by linear regression of a plot of $F_0/F$ against $[Q]$ (some examples in figure 3.2.6) in the ligand concentration range 0-140 µM, where all the plots were linear: higher concentrations of ligand indeed cause a deviation from the linearity more or less evident in the different cases, probably because large amounts of ligand in solution complicate the quenching mechanism, with the progressive increase of the dynamic quenching contribution; $K_{SV}$ and $K_q$ (calculated using the equivalence $K_q = K_{SV}/\tau_0$) are reported in table 3.2.1.

![Figure 3.2.6](image)

**Figure 3.2.6** Stern-Volmer plots of fluorescence quenching of HSA ($\lambda_{ex} = 280$ nm, $\lambda_{em} = 340$ nm) in the presence of compounds 30 (a), 32 (b), and 10 (c).
Table 3.2.1 Quenching constants according to Stern-Volmer analysis: Stern-Volmer quenching constant ($K_{SV}$) and bimolecular quenching kinetic constant ($K_q$). The reported standard deviations are those of the linear regressions, performed starting from the mean values of fluorescence quenching, which were calculated from the three independent measurements carried out with each compound.

The bimolecular quenching kinetic constants ($K_q$) are at least 3-4 orders of magnitude higher than the higher value for diffusion-limited collisional quenching ($2.0 \times 10^{10}$ L mol$^{-1}$ s$^{-1}$),$^{19, 20}$ thus the static quenching originating from the association of the fluorophore and quenchers in a bimolecular complex is the main contribution to the fluorescence quenching mechanism within the 0–140 µM ligand concentration range. As said before, higher ligand concentrations complicate the quenching mechanism because the dynamic collision contribution becomes more significant, and this causes an upward curvature of the plot.

3.2.1.5 Dissociation constants and binding sites

Once established that the greater contribution to the quenching is static in nature, the fluorescence data can be safely used to deduce the binding parameters. The number of binding sites and the ligand-protein dissociation constants may be obtained by a Hill analysis,$^{21, 22}$ applying equation 3.2.2, that is used in biological and pharmacological fields to describe the relationship between the ligand concentration and the observed effect:

$$\log\left(\frac{Y}{1-Y}\right) = n \log[Q] - \log K_D \quad \text{Eq. 3.2.2}$$

$Y$ is the fraction of occupied binding sites (calculated as 1-F/F$_0$, assuming that the ratio F/F$_0$ gives the fraction of free binding sites), [Q] is the quencher concentration in mol/L, $n$ is the number of binding sites, and $K_D$ is the dissociation constant.

The parameters $n$ and $-\log K_D$ for all the ligands were obtained by linear regression of a plot of $\log \left[ Y/(1-Y) \right]$ against $\log [Q]$ in the ligand concentration range 0-140 µM, because higher concentrations gave some problem on the determination of the slope as an integer number, probably due to a change of the initial stoichiometry; figure 3.2.7 shows the plots of acid 32 (a) and lactone 11 (b) as examples. The dissociation constants $K_D$ were calculated from the intercepts; all the calculated values are reported in table 3.2.2.
Chapter 3 – Results and discussion

Figure 3.2.7 Hill plots in the ligand concentration range 0-140 μM for the binding of HSA (concentration fixed to 0.5 μM) with acid 32 (a) and lactone 11 (b).

<table>
<thead>
<tr>
<th>ligand</th>
<th>n ± SD</th>
<th>K_D ± SD (mol L⁻¹)</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acid 29</td>
<td>1.20 ± 0.03</td>
<td>(2.32 ± 0.06)x10⁻⁶</td>
<td>0.9971</td>
</tr>
<tr>
<td>Acid 30</td>
<td>1.04 ± 0.04</td>
<td>(2.12 ± 0.08)x10⁻⁵</td>
<td>0.9940</td>
</tr>
<tr>
<td>Acid 31</td>
<td>1.12 ± 0.06</td>
<td>(9.15 ± 0.48)x10⁻⁶</td>
<td>0.9895</td>
</tr>
<tr>
<td>Acid 32</td>
<td>0.99 ± 0.03</td>
<td>(3.11 ± 0.11)x10⁻⁵</td>
<td>0.9951</td>
</tr>
<tr>
<td>Lactone 7</td>
<td>0.94 ± 0.07</td>
<td>(5.49 ± 0.33)x10⁻⁶</td>
<td>0.9798</td>
</tr>
<tr>
<td>Lactone 10</td>
<td>1.02 ± 0.02</td>
<td>(1.33 ± 0.02)x10⁻⁵</td>
<td>0.9986</td>
</tr>
<tr>
<td>Lactone 11</td>
<td>1.08 ± 0.02</td>
<td>(4.31 ± 0.06)x10⁻⁶</td>
<td>0.9978</td>
</tr>
<tr>
<td>Lactone 12</td>
<td>1.35 ± 0.03</td>
<td>(3.61 ± 0.08)x10⁻⁷</td>
<td>0.9977</td>
</tr>
</tbody>
</table>

Table 3.2.2 Binding parameters: number of interaction sites (n) and dissociation constants (K_D), calculated according to the Hill analysis in the ligand concentration range 0-140 μM. The reported standard deviations are those of the linear regressions, performed starting from the mean values of fluorescence quenching, which were calculated from the three independent measurements carried out with each compound.

The results show a slope (n) near 1 in most of the cases, which means a 1:1 interaction stoichiometry between ligand and protein, but the values for lactones 7 and 12 are not really acceptable and they need a deeper investigation, that will be discussed further. Focusing on the other six ligands, all the calculated K_D are in the micromolar range, showing a remarkably high affinity of these molecules for the protein. The binding constants of 29, 30 and 31 are comparable to the published ones, thus validating the experimental method. Among the carboxylic acids, compound 10 shows the lowest affinity, probably due to the replacement of both phenolic hydroxyl groups with methoxyl ones: this leaves only the acid moiety capable to form hydrogen bonds, that could strengthen the ligand-protein interaction, with the polypeptide chain within the binding site. As to lactones 10 and 11, the second has a better affinity and this may be explained as a consequence of the free hydroxyl group in position 1 of the quinide, and of its lower dimension compared to lactone 9, that makes easier the entrance of the molecule in the active site.

The monoester 12 presents a good linearity of the data (R = 0.9977), but the slope is too high and the K_D is lower than the HSA concentration (0.5 μM): these results are therefore unrealistic. In an attempt to solve the problem the protein concentration was halved to 0.25
μM, and the ligand was added in the range 0.25-200 μM; figure 3.2.8 shows the decrease of the fluorescence intensity against the logarithm of the monoester molar concentration (a) and the related Hill plot (b).

![Figure 3.2.8](image)

Figure 3.2.8 Effect of the addition of compound 12 on the fluorescence intensity (a) and the related Hill plot (b).

The fluorescence decay follows a sigmoidal trend, as expected, but the Hill plot shows two linear zones with different slope, one in the concentration range between 0.25 and 8 μM, the second between 40 and 200 μM. The parameters obtained for the linear regression on all the data, on the first region and on the second one are reported in table 3.2.3:

<table>
<thead>
<tr>
<th>Range (μM)</th>
<th>n ± SD</th>
<th>-log K_D ± SD</th>
<th>K_D ± SD (mol L⁻¹)</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25-200</td>
<td>1.08 ± 0.06</td>
<td>4.9 ± 0.3</td>
<td>(1.21 ± 0.08)x10⁻⁵</td>
<td>0.9633</td>
</tr>
<tr>
<td>0.25-8</td>
<td>0.51 ± 0.02</td>
<td>1.7 ± 0.1</td>
<td>/</td>
<td>0.9943</td>
</tr>
<tr>
<td>40-200</td>
<td>1.71 ± 0.02</td>
<td>7.5 ± 0.1</td>
<td>/</td>
<td>0.9994</td>
</tr>
</tbody>
</table>

Table 3.2.3 Hill analysis of the data obtained from the titration of HSA 0.25 μM with compound 12.
The reported standard deviations are those of the linear regression, performed starting from the mean values of fluorescence quenching, which were calculated from the three independent measurements carried out with compound 12.

Despite the fact that the average K_D [(1.21 ± 0.08)x10⁻⁵ mol L⁻¹] is now more realistic and close to the K_D of lactone 10, the behavior of this ligand could be seldom explained in the frame of the Hill binding model, which predicts a slope shifting to 2 in the presence of a second binding site, or never exceeding 1 in the presence of cooperativity. A hypothesis for the strange behavior of monoester 12 could be a self-aggregation phenomenon, dependent on the concentration; if the ligand tends to self-aggregate while its concentration in solution increases, it is only partially available to interact with the protein and its apparent affinity is thus affected.
The calculated K_D from the standard method for diester 7 is (5.49 ± 0.33)x10⁻⁶ mol L⁻¹ (table 3.2.2), but the linearity of its Hill plot is too poor (R = 0.9798), as it can be seen in figure 3.2.9.
This compound halves the fluorescence intensity even at concentrations as low as 1 μM; to understand more about this case, the interaction has been studied in the 0-200 μM ligand concentration range, keeping the protein concentration fixed to 0.5 μM but thickening the points during the titration (0, 0.25, 0.5, 0.75, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 40, 60, 80, 100, 120, 140, 160, 180, 200 μM of ligand in the cell). Figure 3.2.10 shows the fluorescence intensity decay due to the addition of ligand (a), and the related Hill plot (b).

Even the first addition, corresponding to 250 nM, causes a significant reduction of the fluorescence intensity, revealing a very high affinity; moreover a double, consecutive sigmoidal behavior can be clearly seen. As in the case of compound 12 (figure 3.2.8-b) the Hill plot shows two linear regions with different slope (figure 3.2.10-b), although the two measurements are not comparable: one is in the 0.25-5 μM ligand concentration range, the other in the 40-200 μM range. Table 3.2.4 reports the parameters obtained for the linear regression on all the data, on the first region and on the second one.
Table 3.2.4 Hill analysis of the data obtained from the titration of HSA 0.5 μM with compound 7. The reported standard deviations are those of the linear regression, performed starting from the mean values of fluorescence quenching, which were calculated from the three independent measurements carried out with compound 7.

<table>
<thead>
<tr>
<th>Range (μM)</th>
<th>n ± SD</th>
<th>-log K_D ± SD</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25-200</td>
<td>0.93 ± 0.05</td>
<td>5.1 ± 0.2</td>
<td>0.9683</td>
</tr>
<tr>
<td>0.25-5</td>
<td>0.80 ± 0.01</td>
<td>4.46 ± 0.08</td>
<td>0.9992</td>
</tr>
<tr>
<td>40-200</td>
<td>2.06 ± 0.09</td>
<td>9.6 ± 0.3</td>
<td>0.9940</td>
</tr>
</tbody>
</table>

The slope in the two regions of the Hill plot is clearly different, it suddenly shifts from about 1 to 2. This may be due to the binding of two molecules of 7 inside the same albumin binding site. Such an unusual event in protein–ligand interaction has been recently reported for albumin, in the simultaneous binding of different drug molecules at Sudlow site I,\textsuperscript{25} therefore the hypothesis of a double interaction with lactone 7 is not so unrealistic.

To obtain the two K_D of the two consequent associations, a different data treatment compared to the usual Hill analysis was needed, and an appropriate equation has been defined (equation 3.2.3). Its derivation started from the usual equilibrium definition, on a series of two consequent and independent association equilibria:

\[
[P] + [L] \rightleftharpoons [PL] \quad \frac{1}{K_1}
\]

\[
[PL] + [L] \rightleftharpoons [PL_2] \quad \frac{1}{K_2}
\]

Where \([P]\), \([L]\), \([PL]\), and \([PL_2]\) are the concentrations in mol/L of the protein, the ligand, the protein-ligand complex with stoichiometry 1:1, and that with stoichiometry 1:2 respectively. The total fluorescence variation (\(\Delta F_{tot}\)) is proportional to the total protein concentration:

\[
\Delta F_{tot} \propto [P]_0
\]

The fluorescence variation at each experimental point (\(\Delta F\)) is given by the sum of two simultaneously quenching events, due to the two association equilibria:

\[
\frac{\Delta F}{\Delta F_{tot}} = \frac{\Delta F_{PL}}{\Delta F_{tot}} + \frac{\Delta F_{PL_2}}{\Delta F_{tot}}
\]

It is important to take in consideration that the quenching efficiency for the first complex is not the same of the second one, so a new coefficient (\(x\)) needs to be introduced and optimized during the fitting, with the constraint 0 ≤ \(x\) ≤ 1. The equation becomes:

\[
\frac{\Delta F}{\Delta F_{tot}} = x \frac{\Delta F_{PL}}{\Delta F_{tot}} + (1 - x) \frac{\Delta F_{PL_2}}{\Delta F_{tot}}
\]

The terms of the previous equation may be correlated with the concentrations of the involved species:
Chapter 3 – Results and discussion

\[
\frac{\Delta F_{PL}}{\Delta F_{tot}} = \frac{[PL]}{[P]_0} \quad \frac{\Delta F_{PL_2}}{\Delta F_{tot}} = \frac{[PL_2]}{[P]_0}
\]

Therefore:

\[
\frac{\Delta F}{\Delta F_{tot}} = x \frac{[PL]}{[P]_0} + (1 - x) \frac{[PL_2]}{[P]_0}
\]

For the complex/total protein ratio, the equation derived from the treatment of two simultaneous and consecutive equilibria can be used:

\[
\frac{[PL_n]}{[P]_0} = \frac{[L]^n}{K_1 \ldots K_n} + \frac{[L]^{n+1}}{K_1 \ldots K_{n+1}} \ldots + \frac{[L]^{n+k}}{K_1 \ldots K_{n+k}}
\]

If \( n = 1 \), this equation becomes equal to the usual hyperbolic binding curve; in the considered case \( n = 2 \) and after the proper substitutions the final equation is obtained (equation 3.2.3):

\[
\frac{\Delta F}{\Delta F_{tot}} = x \frac{[L]}{K_1} + (1 - x) \frac{[L]^2}{K_1 K_2} \quad \text{Eq. 3.2.3}
\]

Where \( \Delta F \) is calculated as \((F_0 - F)\), \( F_0 \) and \( F \) are the fluorescence intensities before and after the addition of the quencher, respectively, \([L]\) is the quencher concentration in mol/L, \( K_1 \) is the first dissociation constant, \( K_2 \) the second dissociation constant, \( x \) is a coefficient that allows to take into account the different quenching efficiency in the two possible complexes. To use equation 3.2.3 the experimental data were transformed as required, plotting the value of \( \Delta F/\Delta F_0 \) against the logarithm of the ligand concentration (figure 3.2.11); in this titration \( \Delta F_0 \) is equal to \( F_0 \), indeed the fluorescence quenching is complete.

![Figure 3.2.11](image)

**Figure 3.2.11** Curve obtained from the quenching fluorescence data of lactone 7, transformed as required by equation 3.2.3.

By a nonlinear regression the optimization of the three parameters (\( K_1, K_2 \) and \( x \)) was achieved; their values are reported in table 3.2.5.

76
Table 3.2.5 Dissociation constants of protein – lactone 7 complex for 1:1 (K₁) and 1:2 stoichiometry (K₂), and quenching efficiency parameter (x). The non-linear regression has been performed starting from the mean values of fluorescence quenching, which were calculated from the three independent measurements carried out with compound 7.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>K₁</td>
<td>9.67 x 10⁻⁷ mol/L</td>
</tr>
<tr>
<td>K₂</td>
<td>2.08 x 10⁻⁵ mol/L</td>
</tr>
<tr>
<td>x</td>
<td>0.65</td>
</tr>
</tbody>
</table>

In addition to the fact that HSA seems to host two molecules of lactone 7 in the same binding site, both the calculated dissociation constants reveal a very high affinity but K₁ is surprisingly low, reaching even the nanomolar range. To our knowledge, this is the lowest affinity one ever reported for albumin, and by far more favorable than the binding of most drugs to this protein.

An overview of the obtained data suggests that HSA binds the chlorogenic compounds with a very high affinity. This result may open a matter: many studies reported that polyphenols are capable to permeate the gastrointestinal barrier and are absorbed in humans, being then found in plasma as both intact molecules and as their hydrolysis metabolites, in particular as caffeic acid;²⁶,²⁷,²⁸,²⁹,³⁰ caffeoylquinic acids have been detected in plasma even 4 hours after the ingestion. It could be interesting to evaluate if the total amount of polyphenols, assumed with a coffee cup, has a competing effect on drug binding by HSA in plasma, and thus on drug pharmacokinetics.

3.2.2 HSA100 and its mutants

3.2.2.1 Binding biomolecules and titration assays

The good results obtained studying the binding between HSA and chlorogenic compounds suggest that this kind of interaction may be really exploited in a biosensor. The selectivity must be improved, but it is difficult to manipulate a big protein as HSA in order to modify its binding properties. To solve this problem in our research group a peptide corresponding to a 101-residue stretch of the HSA sequence (Ala194 to Glu294), named HSA100, has been already identified, synthesized exploiting its expression in E. coli and fully characterized (chapter 1, section 1.4.2).³¹,³² The fragment HSA100, that replicates the Sudlow site I of the parental HSA and retains its binding properties, is a better starting point for the mutation steps: its lower dimension allows to generate mutants libraries with casual variations in known positions within the binding site, influencing thus the binding selectivity.

Seven mutating positions with the right features, i.e. non-essential role for correct peptide folding and close contact with the ligand molecules, have been identified: two of the selected residues are placed inside the major hydrophobic pocket (Leu260, Ile264), two inside the minor hydrophobic pocket (Ala215, Leu238) and the remaining three are at the polar site of entry (His242, Arg257, Ser287) of the binding site (figure 3.2.12).³¹

Once cloned the mutant libraries, they can be screened for the improved affinity: the mutant with the best properties could be then mutated again on other positions, producing a second generation mutants library with the aim to further increase the selectivity. Of course this
process of site-directed mutagenesis could be iterated many times, up to reach the desired characteristics\textsuperscript{33,34,35}

![Figure 3.2.12](image)

Figure 3.2.12\textsuperscript{32} Tridimensional structure of the 100 amino acids fragment with indicated the positions for mutation submitted to site-directed mutagenesis.

In this perspective the interaction properties of wild type HSA100 and of four available clones (table 3.2.6), derived from mini-library 1 (chapter 1, section 1.4.3) which was obtained introducing two mutations in the minor hydrophobic pocket, were studied by fluorescence spectroscopy, following the same method already used with HSA\textsuperscript{36}. The four clones were just randomly chosen without trying to fish out the best binders from the small libraries. Unfortunately the supplied peptides amounts were very low, so the titrations were carried out only one time, making the results interpretation qualitative, and only with the two diester quinides 7 and 11, that were chosen thanks to their good dissociation constants with the whole protein and for their structural similarity. The titrations were performed monitoring the Trp fluorescence intensity in the ligand concentration range 0-200 μM with 0.5 μM peptide concentration in the cell. After each addition of the ligand, the emission spectra ($\lambda_{\text{exc}}$ 280 nm, $\lambda_{\text{em}}$ range 300-400 nm), the fluorescence intensity ($\lambda_{\text{exc}}$ 280 nm, $\lambda_{\text{em}}$ 340 nm), and the SFS (Δ = 60) were recorded, as with HSA.

<table>
<thead>
<tr>
<th>Biomolecule</th>
<th>Ala 215 wild type</th>
<th>Leu 238 wild type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clone TV</td>
<td>Threonine (T)</td>
<td>Valine (V)</td>
</tr>
<tr>
<td>Clone LK</td>
<td>Leucine (L)</td>
<td>Lysine (K)</td>
</tr>
<tr>
<td>Clone NK</td>
<td>Asparagine (N)</td>
<td>Lysine (K)</td>
</tr>
<tr>
<td>Clone KV</td>
<td>Lysine (K)</td>
<td>Valine (V)</td>
</tr>
</tbody>
</table>

Table 3.2.6 The four available mutants from the mini-library 1; for each peptide the mutated amino acids in the selected positions are reported.
3.2.2.2 Fluorescence quenching and dissociation constants

Both lactones cause a complete fluorescence quenching of HSA100 and of the four clones, with slight variations. By overlapping the sigmoidal trends of the titrations with HSA, wild type HSA100 and the four clones (all obtained plotting the normalized fluorescence intensity against the logarithm of the ligand concentration), it is possible to compare the binding properties of the tested biomolecules and to qualitatively evaluate which of them gives the best result: figure 3.2.13 shows the overlapping of the quenching trends with lactones 7 (a) and 11 (b).

Figure 3.2.13 Fluorescence quenching comparison of the tested biomolecules (HSA, HSA100 and clones concentration fixed to 0.5 μM, ligand concentration from 0 to 200 μM): titrations with lactones 7 (a) and 11 (b).

At a glance, in both cases HSA is the best binding molecule and HSA100 is the worst one, while the clones cause fluorescence quenching in the middle with different trend depending on the ligand lactone. The best performance of the native protein is reasonable, due to its more complex geometry and highest folding degree; moreover this result confirms previous data of our research group on binding of drugs such as warfarin and efavirenz to HSA100.\textsuperscript{32} Having established before that the quenching of HSA is mainly due to a static contribution (paragraph 3.2.1.3), the Stern-Volmer analysis was performed also on the data obtained in the HSA100 titrations to control if in this case the quenching mechanism is different.

The $K_{SV}$ for the two ligands were determined using equation 3.2.1 by linear regression of a plot of $F_0/F$ against $|Q|$ in the ligand concentration range 0-10 μM, where the plots were linear: once more, higher concentrations of molecule cause a deviation from the linearity.
Figure 3.2.14 shows the plots obtained with lactone 7 (a) and 11 (b). $K_{SV}$ and $K_q$ (calculated using the equivalence $K_q = K_{SV}/t_0$) are reported in table 3.2.7.

![Figure 3.2.14](image)

**Table 3.2.7** Quenching constants, obtained from the HSA100 titrations data, according to Stern-Volmer analysis: Stern-Volmer quenching constant ($K_{SV}$) and bimolecular quenching kinetic constant ($K_q$). The reported standard deviations are those of the linear regression.

<table>
<thead>
<tr>
<th>ligand</th>
<th>$K_{SV} \pm SD$ ($10^4$ L mol$^{-1}$)</th>
<th>$K_q \pm SD$ ($10^{12}$ L mol$^{-1}$ s$^{-1}$)</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lactone 7</td>
<td>5.26 ± 0.07</td>
<td>5.26 ± 0.07</td>
<td>0.99</td>
</tr>
<tr>
<td>Lactone 11</td>
<td>2.76 ± 0.03</td>
<td>2.76 ± 0.03</td>
<td>0.99</td>
</tr>
</tbody>
</table>

In both cases the plot linearity is very good and the bimolecular quenching kinetic constants ($K_q$) are higher than the higher value for diffusion limited collisional quenching, thus, as before for HSA, the static quenching is the main contribution to the fluorescence quenching mechanism, at least within the 0–10 μM ligand concentration range. About the clones, that reveal a binding behavior in the middle, the quenching may be reputed mainly static by analogy.

The dissociation constants were evaluated in terms of IC$_{50}$ from the sigmoidal trends, determining the inflection point of the curves with a four parametric sigmoidal fitting (**table 3.2.8**); this method was preferred compared to the previous used Hill analysis, due to the availability of only qualitative data.

The affinity of each clone towards lactones 7 and 11 changes slightly within the same order of magnitude, not revealing considerable differences in the binding behavior. Nevertheless, it is important to note that the clones have a higher affinity for compound 7 than for compound 11; probably this is due to the possibility for lactone 7 to stabilize the interaction with the peptides through the formation of hydrogen bonds, thanks to its phenolic groups. Another interesting element to look upon is the opposite interaction tendency of clone LK in the titrations with the two ligand lactones: it is the best binding peptide among the tested inside the mini-library towards compound 7, while it is the worst one towards compound 11.
Chapter 3 – Results and discussion

<table>
<thead>
<tr>
<th>Binding peptides</th>
<th>Lactone 7 $K_D$ (µM)</th>
<th>Lactone 11 $K_D$ (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSA</td>
<td>0.967</td>
<td>4.3</td>
</tr>
<tr>
<td>HSA100</td>
<td>18.5</td>
<td>25.5</td>
</tr>
<tr>
<td>Clone TV</td>
<td>12.6</td>
<td>16.8</td>
</tr>
<tr>
<td>Clone LK</td>
<td>9.9</td>
<td>20.8</td>
</tr>
<tr>
<td>Clone NK</td>
<td>16.0</td>
<td>18.0</td>
</tr>
<tr>
<td>Clone KV</td>
<td>18.5</td>
<td>17.0</td>
</tr>
</tbody>
</table>

Table 3.2.8 Dissociation constants of HSA, HSA100 and clones in the presence of lactones 7 and 11. The values of HSA are the same reported in the previous tables 3.2.5 and 3.2.2 and obtained from the regression according to equation 3.2.3 and from Hill analysis, respectively; on the other hand the $K_D$ of HSA100 and clones have been evaluated in terms of $IC_{50}$ from the sigmoidal trends.

In this mutant, the leucine hydrophobic side chain is replaced by the ionized amino terminal group of lysine, able to establish hydrogen bonds with the phenolic groups of compound 7 but of course this property doesn’t improve the interaction with lactone 11. Clearly, the introduction of the lysine mutation is not sufficient to increase the affinity for the ligand, as clones NK and KV show, indeed the whole peptide influences the binding and not a single amino acid. Anyway it is really interesting that even casual mutations enhance the binding ability of HSA100 wild type towards the tested ligand, and this is a positive first result about the mutant libraries use.

---

Chapter 3 – Results and discussion


3.3 Surface Plasmon Resonance and Biacore assay

3.3.1 Surface Plasmon Resonance

In the last years, Surface Plasmon Resonance (SPR) biosensors have seen a continually growing diffusion. SPR is a tool that allows to get many information, such as protein binding, association/dissociation kinetics, and affinity constants, about biomolecules interactions occurring really close to a transducer surface. This technique presents several advantages: accurate studies in real-time, high reproducibility, label-free reagents, immobilization of only one of the interacting partners on the sensor surface and multiple use of the chip, through a regeneration step of the active surface. Many kinds of molecules may be immobilized (proteins, nucleic acids, carbohydrates, antibodies and so on), even viruses and intact cells; usually this process does not significantly affect the equilibrium constants. Thanks to its versatility, SPR has been applied in many areas, for example clinical diagnosis, virology, drug discovery, and food analysis.

SPR is an anomalous diffraction first described in the early 20th century by Wood. This phenomenon is associated to a charge density oscillation at the interface between a medium as water and a thin metal film, typically gold or silver, having the real part of its complex permittivity of opposite sign; at certain incident wavelength and certain angles a plane polarized light, directed through a glass prism to the metal/media interface, couples to the charge density oscillation and a minimum in reflectivity occurs. The conditions required for the coupling phenomenon are very sensitive to the dielectric properties, in particular to the refractive index, of the medium adjacent to the transducer surface on the side opposite from the reflected light; the resonance phenomenon changes upon variations at the interface (as upon binding of a very small number of biomolecules at the surface), and the variation may be detected in real-time by the system. When the species (biomolecules or cells) immobilized on the surface bind an analyte, the interfacial refractive index changes and this causes a shift in the resonance angle (figure 3.3.1). A sensorgram is the result of the angle shift monitoring over time (figure 3.3.2); the response is expressed in resonance unit (RU), where one RU represents a change of 0.0001° in the SPR angle that for most proteins corresponds to a change in concentration of about 1 pg/mm² on the sensor surface.

![Figure 3.3.1](image.png)

Figure 3.3.1 Variation of the SPR angle due to the interaction between the surface immobilized species and the analyte in solution, monitored in a prism-coupled SPR system.
Chapter 3 – Results and discussion

**Figure 3.3.2** SPR angles before and after the interaction (on the left, $\theta_{\text{SPR}_1}$ and $\theta_{\text{SPR}_2}$ respectively), and the sensogram plotting the response against the time (on the right).

The light used to detect interaction events never enters the sample, avoiding potential absorption phenomena: the wavelength energy dissipates through SPR in the metallic film. Sometimes a background response may occurs, due to a difference in the refractive indices of the running and the sample buffer; in these cases the real binding response is achieved subtracting the background from the data.\(^9\)

SPR data can be improved minimizing the amount of immobilized ligand, to reduce mass transport effects, aggregation and crowding; of course the signal-to-noise ratio must be taken into account. Moreover high flow rates are preferable to avoid concentration gradients in the flow cell, that may occur when the binding rate of the analyte to the ligand is faster than the diffusion of analyte to the surface.\(^{18,19}\) Finally a careful regeneration of the ligand on the chip surface is essential for the data reproducibility.\(^{20}\)

A typical interaction assay is reported in **figure 3.3.3**; information about the affinity of the analyte and the association or dissociation kinetic may be achieved from this plot. Cross-reactivity measurements are also used to evaluate the specificity of the interaction.

**Figure 3.3.3** Typical shape of an SPR-sensogram. It can be divided into four steps: association, steady state or equilibrium, dissociation, and regeneration.
3.3.2 Evaluation of the association constant

The association affinity constant \((K_A)\) is the ratio at equilibrium of free partners and complex concentrations; the dissociation constant \((K_D)\) is simply its inverse. These constants can be directly measured by equilibrium binding analysis injecting a series of analyte concentrations, to exploit the relationship between the binding level at the equilibrium and the analyte amount in solution.

Considering the equilibrium \(A + B \rightleftharpoons AB\), \(K_A\) is expressed by the formula (equation 3.3.1):

\[
K_A = \frac{[AB]}{[A][B]} \quad \text{Eq. 3.3.1}
\]

In the SPR assay, the concentration of the complex, which corresponds to the fraction of occupied receptor, can be measured as the ratio between the steady state response at a given ligand concentration and the maximum response at saturating conditions. Since free analyte is constantly replenished during the sample injection, its concentration may be assumed equal to the bulk analyte concentration. At saturation, the total surface binding capacity is exploited and the response value is the highest one (figure 3.3.4).

\[\text{Figure 3.3.4} \quad \text{Sensorgram plots at different analyte concentrations (C); } R_{\text{max}} \text{ is the highest response value, achieved in saturation conditions.}\]

If the total capacity is known and it is expressed in RU, the amount of free ligand on the surface can be easily obtained simply subtracting the complex concentration, already in RU, and the conversion to absolute surface concentration is not required. In this case the \(K_A\) formula becomes (equation 3.3.2):

\[
K_A = \frac{R_{\text{eq}}}{C \left( R_{\text{max}} - R_{\text{eq}} \right)} \quad \text{Eq. 3.3.2}
\]

where \(C\) is the bulk analyte concentration, \(R_{\text{max}}\) is the total surface binding capacity and \(R_{\text{eq}}\) is the steady state binding level, both in RU. A rearrangement gives equation 3.3.3:

\[
\frac{R_{\text{eq}}}{C} = K_A R_{\text{max}} - K_A R_{\text{eq}} \quad \text{Eq. 3.3.3}
\]
By plotting the steady state response values against the analyte concentrations, a binding hyperbolic isothermal is obtained (figure 3.3.5), from which $K_A$ and $R_{\text{max}}$ could be calculated.

![Figure 3.3.5 Plot of the response obtained at each tested analyte concentration ($R_{eq}$) versus the analyte concentration itself.](image)

The SPR response depends on the number of interacting ligand sites on the surface and on the mass of the immobilized complex, therefore on the molecular weight of the analyte. The analyte response could be expressed by equation 3.3.4:

$$\text{Analyte response} \propto \text{Analyte MW} \times \text{number of interacting analyte molecules} \quad \text{Eq. 3.3.4}$$

while the concentration of ligand sites on the chip is given by the equation 3.3.5:

$$\frac{\text{Ligand sites (pmol/mm}^2\text{)}}{\text{Ligand MW}} \propto \frac{\text{Ligand response}}{\text{Ligand MW}} \times \text{valence} \quad \text{Eq. 3.3.5}$$

where valence is the number of analyte molecules one ligand molecule can bind with, for example if a ligand interacts with two analytes its valence is 2. $R_{\text{max}}$ corresponds to the analyte response when all the available ligand sites have been occupied by the analyte; thus, using equation 3.3.5 to express the number of interacting analyte molecules in equation 3.3.4, $R_{\text{max}}$ is obtained as (equation 3.3.6):

$$R_{\text{max}} = \frac{\text{analyte MW}}{\text{Ligand MW}} \times \text{Ligand response} \times \text{valence} \quad \text{Eq. 3.3.6}$$

### 3.3.3 The experiment

Thanks to its automation and response quickness, SPR could be the choice technique to screen the affinities of the HSA100 mutants (chapter 1, section 1.4) towards quinides. In this perspective a first attempt was performed using HSA as binding molecule, to set the experimental method and to get a first affinity constant, useful as reference value in the relative SPR scale. The experiment was carried out on a Biacore instrument.

The synthesized diester quinide 28 (chapter 3, section 3.1.3) was immobilized on a sensor chip CM5 (figure 3.3.6) by amine coupling chemistry, exploiting the amine linker on the diester and the carboxy groups on the chip surface.
Figure 3.3.6 On the left, schematic illustration of the structure of the sensor chip surface CM5, that carries a matrix of carboxymethylated dextran covalently attached to the gold layer; on the right compound 28 immobilized on the chip surface.

The suggested immobilization procedure\textsuperscript{22} was followed: first the surface carboxy groups were activated, using a mixture of 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC) and N-hydroxysuccinimide (NHS), to give reactive succinimide esters (figure 3.3.7).

Figure 3.3.7 Activation of carboxymethyl dextran with EDC/NHS (the dextran chains are omitted).

The ligand solution was then simply fluxed trough the cell: the reaction between the succinimied esters and the primary amine groups is spontaneous under the buffer conditions (in our case 10 mM Na\textsubscript{2}HPO\textsubscript{4} and 2 mM KH\textsubscript{2}PO\textsubscript{4} phosphate buffer, adjusted to pH 7.4), so the ligand becomes covalently bound to the dextran matrix (figure 3.3.8).

Figure 3.3.8 Amine coupling of ligands to the sensor surface.
Therefore, the potential partner molecule could be carried on the chip surface through a continuous buffer flow to test the interaction. The regeneration of the chip, by dissociating the formed complexes, allows to screen several analytes without damaging the ligand. When small molecules are immobilized, as quinide 28, there is a significant risk of altering their interaction properties, due both to the steric hindrance near the surface, and to the chemical modification of the molecule, necessary for the immobilization itself. A spacer arm between the functional molecule and the chip may be a good strategy.

After the chip preparation, the interaction between compound 28 and HSA was measured. The binding analyses were performed with increasing concentrations of protein (0, 10, 50, 100, 200 and 400 μM); from these data the BIACORE Evaluation Software (version 1.1) gave a dissociation constant of 127 μM. **Figure 3.3.9** shows the plot of the binding response against the protein concentration.

![Figure 3.3.9 Plot of binding response against increasing HSA concentration (0, 10, 50, 100, 200 and 400 μM); the blue dotted line marks the observed K\text{D} value.](image)

This apparent affinity is quite different from that obtained by fluorescence spectroscopy for lactone 11 (**table 3.3.1**), which has the same molecular structure of compound 28, except the linker chain.

<table>
<thead>
<tr>
<th>Assay</th>
<th>K\text{D} (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluorescence (compound 11)</td>
<td>4.3</td>
</tr>
<tr>
<td>SPR (compound 28)</td>
<td>127.0</td>
</tr>
</tbody>
</table>

**Table 3.3.1** Comparison between the affinity values obtained by fluorescence spectroscopy for lactone 11 and by SPR for compound 28.

To explain the disparity in order of magnitude between the two results, the changes in the experimental condition must be taken into account. In the SPR assay, compound 28 is oriented in solution and exposes the lactone core, moreover its immobilization on the chip causes the “loss” of one aromatic ring, that is no more available. These constraints prevent an optimal interaction with HSA, because the best geometrical ligand-protein structure is not
reached, while it can be easily achieved when the ligand is free in solution, as in the fluorescence assay. The linker steric hindrance and the surface may further influence the binding.

A strategy to improve the affinity for HSA can be a change of the linker position; for example if it was bounded to the hydroxyl group at position 1 of the quinic core, after the immobilization of the compound both the aromatic rings would be exposed in solution and so available to interact with the protein, leading probably to a better affinity but preventing the contact between HSA and lactone moiety. Since we are interested to design a biosensor able to discriminate between chlorogenic acids and lactones, that differ only for the quinic core, the current orientation is however better because it allows to expose exactly this side of the molecule, even if at the expense of the affinity.

Anyway the result is satisfactory because the chip functionalized with compound 28 works and it can be used for the fast selection of clones, comparing their affinities to the whole protein K₀ in the SPR relative scale.

---

Chapter 3 – Results and discussion

3.4 NMR investigation on the caffeine-quinides interaction

3.4.1 Caffeine interacts with polyphenols

In aqueous media caffeine tends to self-associate, forming dimers or higher aggregates by “vertical stacking”,\(^1\,^2\,^3\) or in the presence of polyphenols, as in coffee or tea brews, to form heterogeneous stack with them.\(^4\) Several examples of caffeine-polyphenols complexes are reported in the literature: with catechins,\(^5\) theaflavin,\(^6\) quercetin,\(^7\) and chlorogenic acids.\(^8\)

The complex formed by 5-O-caffeoyl quinic acid (5-CQA), a major polyphenolic constituent of green coffee bean, and caffeine was diffusely studied (figure 3.4.1). The potassium chlorogenate was first extracted from green coffee beans one century ago as a 1:1 complex with caffeine:\(^9\) the crystal structure shows stacked alternate layers of 5-CQA aromatic ring and caffeine, and the whole framework is coordinated around the potassium ions.\(^10\) The association constant was determined by spectrophotometric method\(^11\) and \(^1\)H-NMR spectroscopy\(^8,\,^12\) and basing on the last investigation\(^8\) its value is \(30 \pm 4\) M\(^{-1}\), which means a \(K_D\) about 33 mM.

\[\text{Figure 3.4.1 Molecular structure of caffeine (a) and 5-CQA (b) and the energy-minimized conformation of their complex (c).}^{8}\]

This kind of complexation is exploited by the plants to control the caffeine compartmentation in the tissues\(^13,\,^14\) and studying the wild Coffea species a qualitative relationship between caffeine and chlorogenic acid contents has been revealed.\(^15\)

3.4.2 Study of the interaction between caffeine and the synthesized quinides

To go into the caffeine interaction properties, the possible association with the four synthesized quinides (figure 3.4.2) was monitored by \(^1\)H-NMR titrations, following a published method.\(^8\)

Basing on the compounds availabilities, during the experiments the lactones concentrations were kept fixed, while the caffeine content was gradually increased and the addition effects were monitored on the lactone signals.

The best way to study the \(\pi-\pi\)-stacking influence on the \(^1\)H-NMR \(\delta\)-shifts is to use an aqueous medium that allows to enhance the hydrophobic contacts.\(^16,\,^17\) In the current experiments this
was impossible due to the low solubility of the tested quinides in water and an organic cosolvent had to be added to perform the measurements.

Figure 3.4.2 Molecular structure of the synthesized quinides used in the \(^1\)H-NMR titrations with caffeine: 3,4-O-dicaffeoyl-1,5-\(\gamma\)-quinide (7), 1,3,4-O-tris[3,4-(dimethoxy)cinnamoyl]-1,5-\(\gamma\)-quinide (10), 3,4-O-bis[3,4-(dimethoxy)cinnamoyl]-1,5-\(\gamma\)-quinide (11), and 3-O-[3,4-(dimethoxy)cinnamoyl]-1,5-\(\gamma\)-quinide (12).

3.4.2.1 Titrations of 7, 10, 11 and 12 with caffeine in DMSO-\(d_6\)/D\(2\)O 75-25%

A first attempt was done by dissolving the compounds in DMSO-\(d_6\)/D\(2\)O 75-25% to give 1 mM solutions and 1 mL of them were put in the NMR tube; a 0.1 M solution of caffeine, always in DMSO-\(d_6\)/D\(2\)O 75-25%, was used to add one equivalent at a time (10 \(\mu\)L each addition) from 1 to 12 equivalents, then two solid aliquots of 5 mg of caffeine (26 mmol) were added to reach 38 and 64 equivalents in the sample. In all the cases, a slight shift to the upper field of the vinyl and aromatic signals of the quinides was observed and it increased following the addition of caffeine; on the contrary the quinidic core peaks remained unshifted, suggesting that only the aromatic side chains are involved in the interaction with caffeine, as predictable. Since the shift variations are very small, as it could be seen from table 3.4.1, \(K_D\) could not be calculated.

Figure 3.4.3-(a) reports the \(^1\)H-NMR of compound 10, as an example, in the presence of 0 equivalents of caffeine (blue), 12 eq. (red) and 38 eq. (green) in the range 6.5-7.3 ppm, related to the aromatic and vinyl protons. An enhancement in the caffeine concentration corresponds to an up-field shift of the signals. No changes in the chemical shifts of the quinidic protons C3-H, C4-H and C5-H are observed in the \(^1\)H-NMR spectra during caffeine titration (figure 3.4.3-(b)). The plots of the shift variation against caffeine concentration for two aromatic and two vinyl protons are reported in figure 3.4.4.
Table 3.4.1 Mean shift variation and standard deviation, observed on the aromatic and vinyl signals of the four lactones, at 12 and 64 eq. of caffeine during the titrations in DMSO-\textsubscript{d}6/D\textsubscript{2}O 75-25%.

<table>
<thead>
<tr>
<th>Lactone</th>
<th>Caffeine concentration 12 mM</th>
<th>Caffeine concentration 64 mM</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>0.004 ± 0.002</td>
<td>0.040 ± 0.004</td>
</tr>
<tr>
<td>10</td>
<td>0.010 ± 0.002</td>
<td>0.043 ± 0.007</td>
</tr>
<tr>
<td>11</td>
<td>0.007 ± 0.001</td>
<td>0.042 ± 0.006</td>
</tr>
<tr>
<td>12</td>
<td>0.009 ± 0.001</td>
<td>0.047 ± 0.007</td>
</tr>
</tbody>
</table>

Figure 3.4.3 Enlargements from the \textsuperscript{1}H-NMR spectra obtained performing the lactone 10-caffeine titration, as examples; the blue, red and green spectra were recorded with 0, 12 and 38 equivalents of caffeine, respectively. Picture (a): aromatic zone, 6.5-7.3 ppm; picture (b): quinidic zone, 5-5.6 ppm.
Figure 3.4.4 Plot of the shift variation against caffeine concentration in all the titration range (left) and its enlargement between 0 and 12 caffeine equivalents (right); the trends of two aromatic (C<sub>13</sub>-H and C<sub>16</sub>-H) and two vinyl protons (C<sub>9</sub>-H and C<sub>10'</sub>-H) of compound 10 are shown.

The trend of the shift variation seems to reach a first plateau at 10 and 12 mM caffeine (the enlargement on the right side of figure 3.4.4 highlights this sigmoidal behavior), but the further two additions of caffeine cause an even more pronounced shifts. This kind of trend is observed in all the aromatic and vinyl protons of the four lactones, even if with some deviations.

The effect is slightly different on equivalent protons of the four compounds: as an example a comparison of the behavior of vinyl proton C<sub>9</sub>-H in the four titrations is shown in figure 3.4.5; anyway, the observed shifts are really too small to understand if these differences are meaningful or not.

Figure 3.4.5 Shifts of the vinyl proton C<sub>9</sub>-H in DMSO-<i>d</i><sub>6</sub>/D<sub>2</sub>O 75-25%, in all the titration range (left) and between 0 and 12 caffeine equivalents (right).

Even in the case of the monoester 12, it was not possible to understand the stoichiometry and the structure of the complex formed by lactones and caffeine. Several scenarios could be supposed (figure 3.4.6): a 1:1 complex (a), two caffeine molecules interacting with the same aromatic ring forming a “sandwich” (b), the same caffeine in the middle of two aromatic residues giving again a “sandwich” but opposed to that of the previous case (c), a combination of these possibilities depending on the number of the aromatic rings. Moreover, the self-assembling of caffeine that occurs at high concentrations complicates the problem.
Figure 3.4.6 Illustration of the achievable stoichiometries in the titrations: 1:1 (a), 2:1 (b) and 1:2 between aromatic residue (green ovals) and caffeine (orange rectangles), respectively.

In order to get more informations, a 2D $^1$H, $^1$H-NOESY spectrum was recorded on an equimolar mixture of lactone 10 and caffeine (5mM for both molecules, to gain in signal/noise ratio) in DMSO-$d_6$/D$_2$O 75-25%, using a mixing time of 300 msec and a relaxation delay of 1.5 sec. Intermolecular NOEs, revealing a spatial proximity, were found among the lactone aromatic portion and caffeine methyl groups, in particular methyl a (figure 3.4.7).

Figure 3.4.7 2D $^1$H, $^1$H-NOESY spectrum recorded in DMSO-$d_6$/D$_2$O 75-25% on an equimolar mixture of lactone 10 and caffeine; detail of the three methyl groups of caffeine (f2) and of the aromatic region of compound 10 (f1), to highlight the intermolecular NOEs.

The same experiment was repeated on an equimolar mixture of lactone 11 and caffeine (5 mM in DMSO-$d_6$/D$_2$O 75-25%). Intermolecular NOEs were detected between the same regions of the molecules, but the cross peaks in this case were weaker and the noise gave some problems (figure 3.4.8); nevertheless the strongest signal is again in correspondence to the caffeine methyl group a.
Chapter 3 – Results and discussion

Figure 3.4.8 2D $^1$H, $^1$H-NOESY spectrum recorded in DMSO-$d_6$/D$_2$O 75-25% on an equimolar mixture of lactone 11 and caffeine; detail of the three methyl groups of caffeine (f2) and of the aromatic region of compound 11 (f1), to highlight the intermolecular NOEs.

However these results can only confirm the interaction between lactones and caffeine, while no indications about the orientation of caffeine towards the aromatic ring are clearly deducible.

With compounds 7 and 12, the 2D $^1$H, $^1$H-NOESY spectra were not recorded, in the first case due to the too small amount of available lactone, in the second one because the monoester shows instability in solution, due to the tendency to an intramolecular transesterification with the migration of the aromatic residue from the hydroxyl group on carbon 3 to that on carbon 4. This transformation is noticeable even in the titration, indeed in the last two spectra, with 38 and 64 equivalents of caffeine, the set of peaks of the C$_4$ monoester starts to appear. In this case the 2D $^1$H, $^1$H-NOESY spectrum was not performed since it requires much more time compared to the titration (about 14 hours versus 2 hours) and the results could not be interpretable.

3.4.2.2 Titrations of 7 with caffeine in D$_2$O/DMSO-$d_6$ 75-25%

Exploiting the higher solubility in water of compound 7, its titration was repeated in D$_2$O/DMSO-$d_6$ 75-25%, using the same conditions of the previous experiments: 1 mL of a 1 mM solution of 7 was placed in the NMR tube, 1 to 12 equivalents of caffeine were added from a 0.1 M solution in the same solvent (one equivalent at a time, 10 µL each addition). To reach 38 and 64 equivalents in the sample, two 5 mg solid aliquots of caffeine (26 mmol) were then added.
In this case the shift to the upper field of the vinyl and aromatic signals was more evident, even between 0 and 12 equivalents of caffeine (figure 3.4.9), while the quinidic core peaks remained again almost unchanged.

Figure 3.4.9 Enlargement, focused on the aromatic region, of the superimposed spectra recorded during the titration of 7 in D2O/DMSO-d6 75-25% in the presence of 0 (blue), 10 (red) and 38 (green) equivalents of caffeine.

Figure 3.4.10 reports the up-field shift of each proton in presence of 12 equivalents of caffeine (a) and at the end of the titration (b), showing that there are some differences in the shift extents.

Figure 3.4.10 Comparison of the shift variation of the vinyl and aromatic protons, obtained during the titration of lactone 7, in the presence of 12 (a) and 64 (b) equivalents of caffeine.

In figure 3.4.11 are reported the shift variations of the vinyl proton C10-H (a) and of the aromatic one C12'-H (b), as examples, against caffeine concentration. This time, a hyperbolic behavior, similar to that reported in literature, has been observed.
Chapter 3 – Results and discussion

Figure 3.4.11 Plot of the shift variation against caffeine concentration in all the titration range of a vinyl (C<sub>10</sub>-H, a) and an aromatic proton (C<sub>12</sub>'-H, b) of compound 7 during its titration with caffeine in D<sub>2</sub>O/DMSO-d<sub>6</sub> 75-25%.

A regression analysis was performed by fitting the experimental data to a single rectangular, two parameters hyperbola (SigmaPlot™ software), in order to evaluate the limiting value of the shift change (Δ<sub>0</sub>) for each vinyl and aromatic proton. The fitting was done in an iterative cycle, to correct the experimental caffeine concentration using the relation (equation 3.4.1)<sup>18,19</sup>

\[
C_{caf}^{free} = C_{caf} - C_{quin} \frac{\Delta}{\Delta_0}
\]  

Eq. 3.4.1

where \(C_{caf}\) and \(C_{quin}\) are the total concentrations of caffeine and lactone 7, respectively. This correction allows to subtract the amount of caffeine already involved in a complex from its total concentration, therefore to consider only the caffeine effectively available to interact with lactone 7. In a first approximation, the contribution of caffeine self-assembly to the decrease of its amount as isolated molecule in solution was ignored: basing on the self-association K<sub>0</sub> in literature (130 mM<sup>8</sup> in D<sub>2</sub>O at 27 °C and 167 mM<sup>7</sup> in H<sub>2</sub>O/DMSO-d<sub>6</sub> 9:1 v/v at 27 °C and pH 4 ± 0.5) it could be supposed that in D<sub>2</sub>O/DMSO-d<sub>6</sub> 75-25% the amount of dimerized or higher aggregated caffeine is negligible.

After the third calculation run, the \(\Delta_0\) values converged (table 3.4.2).

<table>
<thead>
<tr>
<th>Observed proton</th>
<th>(\Delta_0) (ppm)</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>C&lt;sub&gt;9&lt;/sub&gt;-H</td>
<td>-0.41 ± 0.02</td>
<td>0.99</td>
</tr>
<tr>
<td>C&lt;sub&gt;10&lt;/sub&gt;-H</td>
<td>-0.51 ± 0.02</td>
<td>0.99</td>
</tr>
<tr>
<td>C&lt;sub&gt;12&lt;/sub&gt;-H</td>
<td>-0.56 ± 0.03</td>
<td>0.99</td>
</tr>
<tr>
<td>C&lt;sub&gt;13&lt;/sub&gt;-H</td>
<td>-0.50 ± 0.02</td>
<td>0.99</td>
</tr>
<tr>
<td>C&lt;sub&gt;16&lt;/sub&gt;-H</td>
<td>-0.57 ± 0.03</td>
<td>0.99</td>
</tr>
<tr>
<td>C&lt;sub&gt;9’&lt;/sub&gt;-H</td>
<td>-0.42 ± 0.02</td>
<td>0.99</td>
</tr>
<tr>
<td>C&lt;sub&gt;10’&lt;/sub&gt;-H</td>
<td>-0.40 ± 0.02</td>
<td>0.99</td>
</tr>
<tr>
<td>C&lt;sub&gt;12’&lt;/sub&gt;-H</td>
<td>-0.46 ± 0.02</td>
<td>0.99</td>
</tr>
<tr>
<td>C&lt;sub&gt;13’&lt;/sub&gt;-H</td>
<td>-0.47 ± 0.02</td>
<td>0.99</td>
</tr>
<tr>
<td>C&lt;sub&gt;16’&lt;/sub&gt;-H</td>
<td>-0.45 ± 0.01</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Table 3.4.2 \(\Delta_0\) values obtained for each vinyl and aromatic proton of compound 7 by standard regression analysis, after the third calculation run when the convergence is reached; the reported standard errors are that of the regression.
Figure 3.4.12 shows the results in a more immediate form, using different colors to indicate the relative large (green), medium (orange) and relative small (red) $\Delta_0$ values. The $\Delta_0$ trend can suggest where the caffeine prefers to place itself in the complex: since the shielding effect increases with a smaller distance, the protons that show the greater up-field shifts will be in closer proximity to the caffeine surface than the protons less influenced.\textsuperscript{20,21}

![Diagram of chemical structure with chemical shifts labeled]

Figure 3.4.12 Illustration of the calculated $\Delta_0$; the green, orange and red numbers mean a relative large, medium and relative small value, respectively.

The complex dissociation constant was evaluated on the titration of the vinyl proton C\textsubscript{10}-H, which is an isolated doublet placed in the aromatic region (starting $\delta_H$ 7.56 ppm) and so easily recognizable. The relationship between shift variation and caffeine concentration can be described by equation 3.4.2:\textsuperscript{8}

$$\Delta = \Delta_0 \frac{C_{\text{free}}}{K_D + C_{\text{free}}} \quad \text{Eq. 3.4.2}$$

where $\Delta$ is the actual change of chemical shift of the observed proton, $\Delta_0$ is the limiting value when fully complexed and calculated yet (table 3.4.2), $K_D$ is the dissociation constant and $C_{\text{free}}$ is the experimental caffeine concentration corrected as before by using equation 3.4.1. The $K_D$ obtained from the standard regression of the data, according equation 3.4.2, is 69 ± 5 mM (deviation given by the regression), which is consistent with the data in literature about similar studies.\textsuperscript{6,8,22}

To look into the complex stoichiometry, a Hill analysis\textsuperscript{23,24} was performed, using equation 3.4.3:

$$\log \left( \frac{\Delta}{\Delta_0 \frac{\Delta}{\Delta_0}} \right) = n \log (C_{\text{caf}}) - \log (K_D) \quad \text{Eq. 3.4.3}$$

where the ratio $\Delta/\Delta_0$ is the fraction of occupied sites, $C_{\text{caf}}$ is the caffeine concentration in mol/L, $n$ is the number of binding sites, and $K_D$ is the dissociation constant.
Chapter 3 – Results and discussion

Figure 3.4.13 shows the plot and the equation of the linear regression according to equation 3.4.3 obtained for C_{10}-H, considering the caffeine concentration range 2-64 mM, where the best linearity is reached.

![Linear regression plot](image)

Figure 3.4.13 Linear regression according to equation 3.4.3 for C_{10}-H titration data, obtained on the caffeine concentration range 2-64 mM.

The slope of the calculated straight line is near 1 (1.09 ± 0.07, deviation from the linear regression); the obtained K_{D} value is comparable to the previous one, i.e. 43 ± 2 and 69 ± 5 mM respectively.

The Hill analysis was performed again considering the same proton of the other aromatic ring (C_{10'}-H) and the obtained slope was almost the same (1.10 ± 0.04). Unfortunately the linear regression results are not enough to elucidate the complex stoichiometry, indeed a n value near 1 for all the protons can be explained both by a 2:1 stoichiometry, with a caffeine molecule for each aromatic ring, and 1:1, with the same caffeine molecule in the middle between the two aromatic rings and the formation of a “sandwich” (figure 3.4.14); obviously they can be simultaneously present in solution.

![Stoichiometry illustrations](image)

Figure 3.4.14 Illustration of the possible stoichiometries in agreement with the Hill analysis results, i.e. a 2:1 (a) and a 1:1 complex (b) between compound 7 and caffeine; the green ovals and the orange rectangles are the quinidic aromatic residues and the caffeine molecules, respectively.
From a qualitative point of view, compound 7 and caffeine interact, as already seen in DMSO-\(d_6\)/D\(2\)O 75-25\%, and the higher amount of D\(2\)O strengthens the contacts. This kind of binding could be exploited in the design of a selective biosensor. The millimolar K\(D\) is clearly higher compared to the obtained values with HSA by fluorescence assays (chapter 3, section 3.2), reaching the micromolar range, but the use of \(^1\)H-NMR plays an important role on this difference: the NMR limit of detection requires millimolar concentrations, so, when this technique is used to evaluate binding equilibria, the dissociation constants are just apparent ones. Moreover, the presence of DMSO-\(d_6\) in the current experiment must be taken into account, because it influences the studied interaction.

In order to develop a selective biosensor in this direction, several caffeine molecules could be introduced on a suitable scaffold. Other xanthines could be also exploited, such as paraxanthine, theophylline, theobromine or theacrine, or even a purine like adenine, for example using a peptide nucleic acid (PNA) sequence.

---

Chapter 3 – Results and discussion

3.5 Antiviral assays

3.5.1 Further investigations of the bioactivity of the chlorogenic compounds

Both CGAs and CGLs have revealed beneficial health properties in several studies, as already reported (chapter 1, section 1.1), such as antioxidant,\textsuperscript{1,2} anti-inflammatory,\textsuperscript{3} antispasmodic,\textsuperscript{4} radioprotective\textsuperscript{5} and hypocholesterolemic\textsuperscript{6} functions.

Considering the four synthesized lactones, only 3,4-O-dicaffeoyl-1,5-$\gamma$-quinide (7) was studied to evaluate its potential bioactivity and it turned out to be able, as well as 3,4-O-diferuloyl-1,5-$\gamma$-quinide and 3,4-O-dicoumaroyl-1,5-$\gamma$-quinide, to inhibit the human adenosine transporter at low micromolar concentration, probably counterbalancing the stimulating effect of caffeine.\textsuperscript{7} About the bioactivities of compounds 1,3,4-O-tris[3,4-(dimethoxy)cinnamoyl]-1,5-$\gamma$-quinide (10), 3,4-O-bis[3,4-(dimethoxy)cinnamoyl]-1,5-$\gamma$-quinide (11) and 3-O-[3,4-(dimethoxy)cinnamoyl]-1,5-$\gamma$-quinide (12) still no data are reported in literature.

Since we had the possibility to perform some biological tests, thanks to the kind collaboration of Prof. Dr. Lieve Naesens and her Ph.D. student Annelies Stevaert (Rega Institute for Medical Research, KU Leuven, Belgium), we decided to investigate whether or not the four quinides and the four acids already used in the fluorescence assays (molecular structures reported in figure 3.5.1) show activity against several dangerous viruses for human and/or animal beings.

![Molecular structures of the tested compounds: caffeic acid 29, ferulic acid 30, 5-O-cafeeylquinic acid 31, and 3,4-dimethoxycinnamic acid 32 and the four synthesized quinides 7, 10, 11, and 12.](image-url)
Chapter 3 – Results and discussion

3.5.2 Influenza virus

The seasonal influenza infections, due to human influenza A and B viruses, can be cause of mortality in the more susceptible patients, like children and elderly; moreover unexpected influenza pandemics sometimes come up, as the “Spanish flu” in 1918 or the more recent swine-origin A/H1N1 in 2009. Recently, some influenza viruses have shown resistance to the available anti-influenza virus drugs, which act on the viral M2 ion channel (amantadine and rimantadine) or on neuraminidase (zanamivir and oseltamivir); these drugs are relatively efficient, but their activity would not be sufficient in the presence of a highly virulent strain as that of 1918, which was believed to lead to death within an hour after the contact, in its hemorrhagic form. Influenza vaccination, the most used prophylactic measure, has to be updated every year and it is only partially protective in some target populations, such as the elderly.

In this scenario the necessity of new anti-influenza virus agents is clear, possibly moving towards different modes of action.

3.5.2.1 Influenza virus polymerase

The viral RNA-dependent RNA polymerase (RdRp) represents an attractive target for the development of new anti-influenza virus drugs, because it is highly conserved among influenza A, B and C viruses and its purposes are essential for the viral genome replication. The RdRp is composed of three subunits (PA, PB1 and PB2) and it is involved in the transcription and replication of the eight negative-stranded viral RNA (vRNA), each assembled into a viral ribonucleoprotein (vRNP) complex. The transcription process is initiated by an endonuclease reaction, in which host-cell pre-mRNAs are cleaved to produce 5’-capped oligonucleotides, which serve as primers for transcription of the viral genomic vRNA template into viral messenger RNA (mRNA) strands. This endonucleolytic or “cap-snatching” activity resides in N-terminal domain of the PA subunit (PA-Nter). One or two divalent metal ions (Mg$^{2+}$ or Mn$^{2+}$) are located in the active site; several metal-chelating molecules have been reported to inhibit the influenza virus endonuclease activity, like 2,4-dioxo-4-phenylbutanoic acid (DPBA), N-hydroxamic acid and marchantins (figure 3.5.2).

![Molecular structure of some recognized inhibitors of influenza virus endonuclease activity.](image)

Figure 3.5.2 Molecular structure of some recognized inhibitors of influenza virus endonuclease activity.
3.5.2.2 *Influenza endonuclease enzymatic assay*

An enzymatic assay with influenza PA-Nter endonuclease has been performed to study the possible activity of the eight chlorogenic compounds, according to a published procedure.\(^ {22,28}\) The reaction mixture (25 µl volume) contained: 2 µg recombinant PA-Nter, 1 µg (16.7 nM) single stranded circular DNA plasmid M13mp18 as a substrate, buffer (50 mM Tris-HCl pH 8; 100 mM NaCl; 1 mM MnCl\(_2\); and 10 mM β-mercaptoethanol), and serial dilutions of the test compounds, which were dissolved in DMSO. DPBA was used as the reference compound. After 2 h of incubation at 37 °C, endonucleolytic digestion of the plasmid was visualized by gel electrophoresis with ethidium bromide staining and the amount of remaining intact plasmid was quantified (figure 3.5.3). The 50% inhibitory concentrations (IC\(_{50}\)) were obtained by non-linear least-squares regression analysis and are reported in table 3.5.1. For each active compound, the percentage inhibition of PA endonuclease activity was plotted against the compound concentration on a semi-logarithmic plot (figure 3.5.4). Values were the mean ± SEM (standard error of the mean) of three independent experiments.

![Figure 3.5.3 A typical gel electrophoresis result with acid 29 as an illustration.](image)

<table>
<thead>
<tr>
<th>Compound</th>
<th>IC(_{50}) (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acid 29</td>
<td>16.0</td>
</tr>
<tr>
<td>Acid 30</td>
<td>&gt; 500</td>
</tr>
<tr>
<td>Acid 31</td>
<td>18.2</td>
</tr>
<tr>
<td>Acid 32</td>
<td>&gt; 500</td>
</tr>
<tr>
<td>Lactone 7</td>
<td>33.6</td>
</tr>
<tr>
<td>Lactone 10</td>
<td>&gt; 500</td>
</tr>
<tr>
<td>Lactone 11</td>
<td>&gt; 500</td>
</tr>
<tr>
<td>Lactone 12</td>
<td>&gt; 500</td>
</tr>
<tr>
<td>DPBA</td>
<td>24.2</td>
</tr>
</tbody>
</table>

| Table 3.5.1 Inhibitory activities of the eight tested compounds in the influenza endonuclease enzymatic assay; IC\(_{50}\) values represent the compound concentration causing 50% inhibition of the PA endonuclease activity, calculated by non-linear least-squares regression analysis. |

Acids 29 and 31 and lactone 7 reveal a promising inhibitory activity, reaching IC\(_{50}\) values in the same range as the reference compound DPBA. The other six compounds do not cause any inhibition at 500 µM, the highest tested concentration in this assay.
3.5.2.3 vRNP reconstitution assay

To determine the inhibitory effect of the compounds on reconstituted influenza virus vRNP complexes, human embryonic kidney (HEK) 293T cells were reverse transfected with a transfection mixture containing the four vRNP-reconstituting plasmids and a firefly luciferase reporter plasmid. The cells were incubated with the compounds for 24 h at 37 °C, after which the luciferase activity was determined (the luminescence is a direct measure for the RdRp activity). The influenza virus inhibitor ribavirin 29 (figure 3.5.5-a) was included as a reference compound. The 50% effective concentration (EC$_{50}$) was defined as the compound concentration causing 50% reduction in the vRNP-driven firefly luciferase signal, as compared to cells receiving only medium without compound. These EC$_{50}$ values were calculated by interpolation assuming a semi-log dose-response effect. A more detailed description of the procedure is reported in literature. 30

In parallel, compound cytotoxicity was determined in untransfected HEK293T cells, which were incubated with serial dilutions of the compounds. After 24 h of incubation, the MTS (3-(4,5-dimethylthiazole-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium, inner salt) cell viability assay was performed. 31 In this colorimetric method, MTS is bio-reduced by cells into a colored formazan product that is soluble in tissue culture medium (figure 3.5.5-b) and the absorbance at 490 nm is directly proportional to the number of living cells in culture. The spectrophotometric data were used to calculate the 50% cytotoxic concentration (CC$_{50}$), i.e. the concentration reducing cell viability by 50%, as compared to the wells receiving only medium without compound.
The CC\textsubscript{50} and EC\textsubscript{50} values were obtained by non-linear least-squares regression analysis on data from 2 independent tests and are reported in table 3.5.2.

![Molecular structure of the influenza virus inhibitor ribavirin, used as reference compound (a); the colorless MTS tetrazolium compound is converted into a colored formazan product during the cell viability assay (b).](image)

**Figure 3.5.5** Molecular structure of the influenza virus inhibitor ribavirin, used as reference compound (a); the colorless MTS tetrazolium compound is converted into a colored formazan product during the cell viability assay (b).

<table>
<thead>
<tr>
<th>Compound</th>
<th>CC\textsubscript{50} (\mu M)</th>
<th>EC\textsubscript{50} (\mu M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acid 29</td>
<td>&gt; 200</td>
<td>31.9</td>
</tr>
<tr>
<td>Acid 30</td>
<td>&gt; 200</td>
<td>&gt; 200</td>
</tr>
<tr>
<td>Acid 31</td>
<td>&gt; 200</td>
<td>&gt; 200</td>
</tr>
<tr>
<td>Acid 32</td>
<td>&gt; 200</td>
<td>&gt; 200</td>
</tr>
<tr>
<td>Lactone 7</td>
<td>&gt; 200</td>
<td>134.7</td>
</tr>
<tr>
<td>Lactone 10</td>
<td>&gt; 200</td>
<td>&gt; 200</td>
</tr>
<tr>
<td>Lactone 11</td>
<td>&gt; 200</td>
<td>18.7</td>
</tr>
<tr>
<td>Lactone 12</td>
<td>&gt; 200</td>
<td>&gt; 200</td>
</tr>
<tr>
<td>Ribavirin</td>
<td>&gt; 200</td>
<td>7.1</td>
</tr>
</tbody>
</table>

**Table 3.5.2** Cytotoxicity (CC\textsubscript{50}, determined by MTS viability assay) and antiviral activity (the EC\textsubscript{50} represent the compound concentration producing 50% reduction in vRNP-driven firefly report signal, estimated at 24 h after transfection) of the eight tested compounds in the influenza vRNP reconstitution assay in HEK293T cells; the CC\textsubscript{50} and EC\textsubscript{50} values were obtained by non-linear least-squares regression analysis.

For all the compounds, including the reference compound ribavirin, the cytotoxicity turned out to be higher than the maximum tested concentration (200 \mu M); only the acid 29 and the lactones 7 and 11 showed a reduction in the vRNP-driven firefly luciferase signal, the first two with good EC\textsubscript{50} values, which are slightly higher compared to ribavirin. Since this test involves the entire polymerase complex instead of only N-terminal part of its PA subunit, as in the previous enzymatic assay, and it is performed in cells, the results are already more predictive for activity against the entire virus in cell culture.

### 3.5.2.4 Anti-influenza virus activity

To evaluate the possible activity against the whole influenza virus, Madin-Darby canine kidney (MDCK) cells were infected with A/PR/8/34 influenza virus [multiplicity of infection (MOI): 150 CCID\textsubscript{50} (50% cell culture infectious dose) per well] and treated with the compounds active in the vRNP reconstitution assay (i.e. acid 29 and lactones 7 and 11). The influenza virus inhibitor ribavirin was included as reference compound. After 24 h of incubation at 35 °C, virus yield in the supernatant was estimated by determining the viral genome copy number in a one-step quantitative real-time reverse transcription (qRT)-PCR.
assay. The EC$_{99}$ and EC$_{90}$ values were calculated by interpolation and defined as the compound concentration causing respectively a 2-log$_{10}$ and 1-log$_{10}$ reduction in vRNA copy number, as compared to the virus control receiving no compound. In parallel, the CC$_{50}$ was determined using the MTS cell viability assay in uninfected MDCK cells, which were incubated with serial dilutions of the compounds for 24 h. The results are reported in table 3.5: acid 29 and lactone 7 were not cytotoxic at the maximum tested concentration (200 μM), while lactone 11 showed cytotoxicity in both experiments, with CC$_{50}$ values of 128.7 and 97.4 μM. Due to this observed cytotoxicity, the limiting concentration in the antiviral activity test was lowered to 100 μM for lactone 11. Lactone 7 was the only active compound, with EC$_{90}$ values about ten time higher compared to ribavirin.

<table>
<thead>
<tr>
<th>Compound</th>
<th>experiment</th>
<th>CC$_{50}$ (μM)</th>
<th>EC$_{99}$ (μM)</th>
<th>EC$_{90}$ (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acid 29</td>
<td>1</td>
<td>&gt; 200</td>
<td>&gt; 200</td>
<td>&gt; 200</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>&gt; 200</td>
<td>&gt; 200</td>
<td>&gt; 200</td>
</tr>
<tr>
<td>Lactone 7</td>
<td>1</td>
<td>&gt; 200</td>
<td>100.0</td>
<td>65.2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>&gt; 200</td>
<td>&gt; 100</td>
<td>88.1</td>
</tr>
<tr>
<td>Lactone 11</td>
<td>1</td>
<td>128.7</td>
<td>&gt; 100</td>
<td>&gt; 100</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>97.4</td>
<td>&gt; 100</td>
<td>&gt; 100</td>
</tr>
<tr>
<td>Ribavirin</td>
<td>1</td>
<td>&gt; 200</td>
<td>7.8</td>
<td>5.2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>&gt; 200</td>
<td>9.4</td>
<td>6.1</td>
</tr>
</tbody>
</table>

Table 3.5.3 Cytotoxicity (CC$_{50}$, determined by MTS viability assay) and antiviral activity (EC$_{99}$ and EC$_{90}$ represent the compound concentration causing respectively a 2-log$_{10}$ and 1-log$_{10}$ reduction in vRNA copy number) of the tested compounds in influenza virus yield assay in MDCK cells. All the values were determined at 24 h after infection.

3.5.3 Human immunodeficiency virus (HIV)

The human immunodeficiency virus (HIV) is a lentivirus (slowly replicating retrovirus) that has been recognized to be the cause of the acquired immunodeficiency syndrome (AIDS), two types of HIV have been characterized: HIV-1, worldwide diffused and more virulent, and HIV-2, confined to West Africa. Since the epidemic began, in the early 1980s, more than 75 million people have contracted HIV and nearly 36 million have died of HIV-related causes: HIV/AIDS has become the sixth-largest cause of death worldwide.

AntiRetroviral Therapy (ART) is the recommended treatment of people infected with HIV and it consists in taking daily a combination of three or more anti-HIV drugs from different classes to suppress HIV replication. Traditionally ART uses two nucleoside reverse transcriptase inhibitors (NRTIs) with either a non-nucleoside reverse transcriptase inhibitor (NNRTI), or a protease inhibitor (PI). However, some side effects of the current available drugs, such as liver toxicity, lipid abnormalities, gastrointestinal intolerance or cardiovascular risks, and the always present problem of drug resistance drive to find new therapeutic agents. Previous studies have shown that some dicaffeoylquinic acids can inhibit in vitro the enzyme HIV-integrase, reaching EC$_{50}$ values from 1 to 7 μM in T cell lines.
3.5.3.1 Anti-HIV activity

The anti-HIV assays with the eight chlorogenic compounds were performed according to a published protocol: the method is based on HIV-induced cytopathic effect (CPE) and it measures the degree of cell killing on acute HIV infection. While HIV infection causes the death of almost all untreated cells after five days; inhibition of HIV-replication by active compounds allows to enhance the number of cells that remain alive. Direct inhibitors of the virus or compounds protecting infected cells from HIV-induced apoptosis can be both discovered by means of this assay.

To determine the anti-HIV activity and cytotoxicity of the compounds in human lymphocyte MT-4 cells, supporting the growth of HIV, serial dilutions of the compounds (stock solutions in DMSO) were added to 96-well plates containing the MT-4 cells. To the virus-infected wells, 100–300 CCID₅₀ of HIV-1 [strain IIIB, wild-type (WT) or the NNRTI resistant S0561945 strain, carrying two mutations: Lys103Asn and Tyr181Cys] or HIV-2 (strain ROD) was added. The mock-infected wells received the compounds without virus. After five days of incubation, the spectrophotometric 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay was performed to assess the effect of the compounds on the viability of the mock- and HIV-infected cells: similarly to what happens in the MTS assay, MTT is bio-reduced by cells into a colored formazan product, insoluble in tissue culture medium; the absorbance, measured usually at a wavelength between 500 and 600 nm, is directly proportional to the number of living cells in culture. The CC₅₀ was defined as the compound concentration that reduced the viability of the mock-infected MT-4 cells by 50%. The EC₅₀ was the concentration achieving 50% protection against the viral cytopathic effect in infected cells.

![Figure 3.5.6](image)

**Figure 3.5.6** The colorless MTT tetrazolium compound is converted into a colored formazan product during the cell viability assay.

The assays results are reported in table 3.5.4. Lactone 7 was the only compound that displayed cytotoxicity, and its limiting concentration in the antiviral tests was lowered to 25 μM instead of 125 μM. Lactones 10 and 12 showed activity in both the experiments against HIV-1 IIIB WT strain, while no activity was detected against the NNRTI resistant strain, which indicates that they may possibly be NNRTI-like compounds. Acids 29 and 31 were previously reported in literature as inactive compounds towards HIV-1 replication, but they could have a marginal activity in the same assay against HIV-1 IIIB WT, as detected in one of the two experiments; lactone 11 too has shown a similar behavior. All the other compounds were inactive at the tested concentration range.
Chapter 3 – Results and discussion

The tests on HIV-1 S0561945 strain, a mutant that is already resistant to a certain class of inhibitors, were performed only with the compounds giving the more promising results against HIV-1 IIIB WT strain, i.e. acid 31 and lactones 10 and 12, but no activity was detected. Also, none of the compounds was able to inhibit the virus replication of the HIV-2 ROD strain.

<table>
<thead>
<tr>
<th>Compound</th>
<th>exp.</th>
<th>HIV-1 IIIB WT strain</th>
<th>HIV-1 S0561945 strain</th>
<th>HIV-2 ROD strain</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CC&lt;sub&gt;50&lt;/sub&gt; (µM)</td>
<td>EC&lt;sub&gt;50&lt;/sub&gt; (µM)</td>
<td>CC&lt;sub&gt;90&lt;/sub&gt; (µM)</td>
</tr>
<tr>
<td>Acid 29</td>
<td>1</td>
<td>&gt; 125</td>
<td>&gt; 125</td>
<td>&gt; 125</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>&gt; 125</td>
<td>121</td>
<td>&gt; 125</td>
</tr>
<tr>
<td>Acid 30</td>
<td>1</td>
<td>&gt; 125</td>
<td>&gt; 125</td>
<td>&gt; 125</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>&gt; 125</td>
<td>121</td>
<td>&gt; 125</td>
</tr>
<tr>
<td>Acid 31</td>
<td>1</td>
<td>&gt; 125</td>
<td>55.8</td>
<td>&gt; 125</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>&gt; 125</td>
<td>&gt; 125</td>
<td>&gt; 125</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Acid 32</td>
<td>1</td>
<td>&gt; 125</td>
<td>&gt; 125</td>
<td>&gt; 125</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>&gt; 125</td>
<td>&gt; 125</td>
<td>&gt; 125</td>
</tr>
<tr>
<td>Lactone 7</td>
<td>1</td>
<td>83.5</td>
<td>&gt; 25</td>
<td>&gt; 25</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>36.5</td>
<td>&gt; 25</td>
<td>&gt; 25</td>
</tr>
<tr>
<td>Lactone 10</td>
<td>1</td>
<td>&gt; 125</td>
<td>38.5</td>
<td>68.8</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>&gt; 125</td>
<td>40.2</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Lactone 11</td>
<td>1</td>
<td>&gt; 125</td>
<td>&gt; 125</td>
<td>&gt; 125</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>&gt; 125</td>
<td>89.1</td>
<td>&gt; 125</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Lactone 12</td>
<td>1</td>
<td>&gt; 125</td>
<td>31.1</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>&gt; 125</td>
<td>83.9</td>
<td>&gt; 125</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
</tbody>
</table>

Table 3.5.4 Cytotoxicity (CC<sub>50</sub>, determined by the MTT viability assay) and antiviral activity (EC<sub>50</sub> and EC<sub>90</sub> represent the compound concentration producing respectively 50% and 90% inhibition of virus replication) of the tested compounds in the anti-HIV activity assay in MT-4 cells. All the values were determined at 5 days after infection. ND: not determined.

3.5.4 Other viruses

The same eight chlorogenic compounds were also tested against other viruses, using the same plaque reduction or CPE reduction assays, but on different kinds of cells, depending on the considered viruses. The general procedure is the following: confluent cell cultures (i.e. a cell culture in which all the cells are in contact and thus the entire surface of the culture vessel is covered) in microtiter 96-well plates were inoculated with virus at an MOI of 100 CCID<sub>50</sub> or 20 plaque forming units (PFU). After a 1–2 h virus adsorption period, residual virus was removed, and the cell cultures were incubated in the presence of varying concentrations of the test compounds (stock solutions in DMSO). Viral cytopathicity was recorded as soon as it reached completion in the control virus-infected cell cultures that were
not treated with the test compounds. Antiviral activity was expressed as the EC\textsubscript{50} or compound concentration required to reduce virus-induced cytopathicity or viral plaque formation by 50%. The minimal cytotoxic concentration (MCC) of the compounds was defined as the compound concentration that caused a microscopically visible alteration of cell morphology in uninfected cells. Alternatively, the cytotoxic activity was determined using the MTS cell viability assay and expressed as the CC\textsubscript{50}.

In the following sections the virus descriptions and the results will be classified according to the used kind of cells.

3.5.4.1 CRFK cells: feline coronavirus and feline herpesvirus

Feline coronavirus (FCoV) is a positive-sense single-stranded RNA (ssRNA+) virus that affects cats worldwide, it can cause intestinal infections or feline infectious peritonitis\textsuperscript{50,51}

feline herpesvirus (FHV) is a double stranded DNA (dsDNA) virus, another highly contagious agent for cats, as one of the major causes of their upper respiratory infections or “cat flu”\textsuperscript{52}.

The tests with these two viruses were performed in Crandell-Rees feline kidney (CRFK) cells, which are obtained from the renal cortex of a 12-week-old female cat (Felis catus domesticus) and are extensively used in feline virus research\textsuperscript{53}.

Three reference compounds were added: Hippeastrum hybrid agglutinin (HHA), a mannose-specific lectin derived from Amaryllis\textsuperscript{54};

Urtica dioica agglutinin (UDA), a small-molecular-weight lectin from stinging nettle rhizome\textsuperscript{55} and 9-(1,3-dihydroxy-2-propoxymethyl)guanine (DHPG) or ganciclovir, a synthetic guanine derivative with antiviral activity (figure 3.5.7)\textsuperscript{56}.

HHA and UDA are inhibitors of both FCoV and FHV\textsuperscript{57} while DHPG inhibits only FHV\textsuperscript{58}.

![ganciclovir](image)

**Figure 3.5.7** Molecular structure of ganciclovir.

In the cytotoxicity and antiviral assays in CRFK cell cultures no activity was detected against FCoV or FHV, even though none of the tested chlorogenic compounds displayed cytotoxicity at the highest tested concentration (100 µM).

3.5.4.2 Vero cells: Coxsackie B4 virus, Punta Toro virus, sindbis virus, para-influenza 3 virus, reovirus 1

Coxsackie B4 virus (ssRNA+) affects humans: it can trigger an autoimmune reaction that leads to the destruction of the pancreatic insulin-producing beta cells, one of several different
etiolologies of diabetes mellitus, the family of Coxsackie B viruses seems to have a role in the development of acute onset juvenile (type 1) diabetes. Punta Toro virus (PTV) is a negative-sense single-stranded RNA (ssRNA-) virus belonging to the Phlebovirus genus; it causes a fatal hepatocytic apoptosis in hamsters and mice, with clinical and pathological similarities to the severe haemorrhagic fever caused by the phlebovirus Rift Valley fever virus in humans.

Sindbis virus (SINV, ssRNA+) is transmitted by mosquitoes; it causes sindbis fever in humans and the symptoms include arthralgia, rash and malaise. Human parainfluenza viruses (hPIVs, ssRNA-) are the etiologic agents causing “human parainfluenza”, hPIV-3 is associated with bronchiolitis and pneumonia. Reoviruses (respiratory enteric orphan viruses), members of the family Reoviridae, are a large and diverse group of non-enveloped viruses with segmented double-stranded RNA genomes (dsRNA). Reovirus infection occurs often in humans, but most cases are mild or subclinical, with very rare cases of mild upper respiratory tract illness or enteritis in infants or children; some diseases in animals are also caused by this kind of virus. The tests with these five viruses were performed in Vero cells, which are derived from the kidney of an African green monkey and are one of the most commonly used mammalian continuous cell lines in microbiology, molecular research and cell biology. Ribavirin (figure 3.5.5-a) and dextran sulfate (ds-10000, molecular weight 10 kDa) were used as reference inhibitors for hPIV-3 and PTV; ds-10000 also inhibits SINV. Similarly to the tests in CRFK cell cultures no one of the tested chlorogenic compounds was cytotoxic at the highest tested concentration (100 µM), but they were inactive against all the considered viruses.

3.5.4.3 HeLa cells: vesicular stomatitis, Coxsackie B4 virus, respiratory syncytial virus

Vesicular stomatitis virus (VSV, ssRNA-) belongs to the family Rhabdoviridae; it can cause acute viral vesicular disease in cattle, horses, deer, and pigs, appearing as mucosal vesicles and ulcers generally in the mouth; occasionally it infects also humans leading to a flu-like illness or to headache. It is also a common laboratory virus used to study the properties of viruses in the family Rhabdoviridae, as the well-known rabies virus, and the RNA viruses in general. Respiratory syncytial virus (RSV, ssRNA-) causes respiratory tract infections. It is ubiquitous in all parts of the world and its infection is very common during infancy and childhood, but severe RSV infections have also increasingly been found among elderly patients. The tests with VSV and RSV and a second test with Coxsackie B4 virus were performed in HeLa cells, which are the oldest and most commonly used human cells and are derived from cervical cancer cells; they were already used to test the first polio vaccine in the 1950s. Ribavirin and ds-10000 were used as reference inhibitors. The results are shown in table 3.5.5: all the compounds were not cytotoxic at 100 µM, moreover acid 29 turned out to be active against VSV, while lactone 7 showed a good inhibition of RSV, reaching EC₅₀ values comparable to the ribavirin results. All the other chlorogenic compounds were inactive.
Chapter 3 – Results and discussion

<table>
<thead>
<tr>
<th>Compound</th>
<th>exp.</th>
<th>MCC (µM)</th>
<th>EC₅₀ (µM)</th>
<th>Coxsackie B₄</th>
<th>RSV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acid 29</td>
<td>1</td>
<td>&gt;100</td>
<td>45</td>
<td>&gt;100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>&gt;100</td>
<td>45</td>
<td>&gt;100</td>
<td>&gt;100</td>
</tr>
<tr>
<td>Acid 30</td>
<td>1</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
</tr>
<tr>
<td>Acid 31</td>
<td>1</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
</tr>
<tr>
<td>Acid 32</td>
<td>1</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
</tr>
<tr>
<td>Lactone 7</td>
<td>1</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>20</td>
</tr>
<tr>
<td>Lactone 10</td>
<td>1</td>
<td>≥100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
</tr>
<tr>
<td>Lactone 11</td>
<td>1</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
</tr>
<tr>
<td>Lactone 12</td>
<td>1</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
</tr>
<tr>
<td>ds-10000 (µg/mL)</td>
<td>1</td>
<td>&gt;100</td>
<td>4</td>
<td>100</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>&gt;100</td>
<td>3</td>
<td>58</td>
<td>4</td>
</tr>
<tr>
<td>Ribavirin</td>
<td>1</td>
<td>&gt;250</td>
<td>10</td>
<td>112</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>&gt;250</td>
<td>22</td>
<td>112</td>
<td>22</td>
</tr>
</tbody>
</table>

Table 3.5.5 Cytotoxicity (MCC, determined by microscopy) and antiviral activity (EC₅₀ represent the compound concentration reducing the viral CPE by 50%) of the tested compounds in HeLa cell cultures assays. All the data of ds-10000 are expressed in µg/mL instead of µM.

3.5.4.4 HEL cells: herpes simplex virus-1 strain KOS (wild-type and thymidine kinase deficient), herpes simplex virus-2 (G), vaccinia virus, vesicular stomatitis, adenovirus-2

Herpes simplex virus 1 and 2 (HSV-1 and HSV-2, dsDNA), also known as human herpesvirus 1 and 2, are two members of the Herpesviridae family and they infect humans. Symptoms of herpes simplex virus infection include watery blisters in the skin or mucous membranes of the mouth, lips or genitals. HSV-1 produces mostly blisters on the lips,85,86 while HSV-2 produces most genital herpes;87 they are both ubiquitous and contagious. After the end of the symptoms, HSV-1 and -2 persist in the body by becoming latent and hiding from the immune system in the cell bodies of neurons. After the initial or primary infection, some infected people experience sporadic episodes of viral reactivation or outbreaks.88 Three HSV viruses were used in the assays: HSV-1 WT KOS strain, a mutant thymidine kinase (TK) deficient HSV-189 KOS strain, and HSV-2 G strain.

Vaccinia virus (VV, dsDNA) is a large enveloped virus;90 VV is the active constituent of the vaccine that eradicated smallpox (or variola),91,92 making it the first human disease to be eradicated. Inoculation with VV produces a localized skin infection; in immunocompromised persons, vaccinia may disseminate and cause severe disease.93 Adenovirus (Adv, dsDNA) infects a large variety of vertebrates;94 in humans it causes a wide range of illnesses, from mild respiratory infections in young children to life-threatening multi-organ disease in patients with a weakened immune system.95 The tests with HSV-1 (KOS WT and TK), HSV-2 (G), VV, Adv and a second test with VSV were performed in HEL cells, which are human erythroleukemia cells capable of spontaneous and induced globin synthesis and widely used to study cell biology and differentiation.96,97 Brivudin,98 cidofovir,99 acyclovir100 and ganciclovir were added as reference inhibitors for HSVs and VV, while zalcitabine and alovudine101 were used as...
Chapter 3 – Results and discussion

reference in the tests with Adv-2 (figure 3.5.8, ganciclovir has been shown in figure 3.5.7). Table 3.5.6 reports the results of the tests.

![Molecular structures of the reference inhibitors used in the assays in HEL cells](image)

**Figure 3.5.8** Molecular structures of the reference inhibitors used in the assays in HEL cells.

<table>
<thead>
<tr>
<th>Compound</th>
<th>exp.</th>
<th>MCC (μM)</th>
<th>EC&lt;sub&gt;50&lt;/sub&gt; (μM)</th>
<th>EC&lt;sub&gt;50&lt;/sub&gt; (μM)</th>
<th>EC&lt;sub&gt;50&lt;/sub&gt; (μM)</th>
<th>EC&lt;sub&gt;50&lt;/sub&gt; (μM)</th>
<th>EC&lt;sub&gt;50&lt;/sub&gt; (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acid 29</td>
<td>1</td>
<td>&gt;100</td>
<td>45</td>
<td>100</td>
<td>45</td>
<td>&gt;100</td>
<td>&gt;100</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
</tr>
<tr>
<td>Acid 30</td>
<td>1</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
</tr>
<tr>
<td>Acid 31</td>
<td>1</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
</tr>
<tr>
<td>Acid 32</td>
<td>1</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
</tr>
<tr>
<td>Lactone 7</td>
<td>1</td>
<td>&gt;100</td>
<td>45</td>
<td>45</td>
<td>45</td>
<td>&gt;100</td>
<td>&gt;100</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>100</td>
<td>&gt; 20</td>
<td>&gt;100</td>
<td>&gt;20</td>
<td>&gt; 20</td>
<td>&gt;20</td>
</tr>
<tr>
<td>Lactone 10</td>
<td>1</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
</tr>
<tr>
<td>Lactone 11</td>
<td>1</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
</tr>
<tr>
<td>Lactone 12</td>
<td>1</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
</tr>
<tr>
<td>Brivudin</td>
<td>1</td>
<td>&gt;250</td>
<td>0.02</td>
<td>250</td>
<td>146</td>
<td>17</td>
<td>&gt;250</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>&gt;250</td>
<td>0.04</td>
<td>50</td>
<td>250</td>
<td>17</td>
<td>&gt;250</td>
</tr>
<tr>
<td>Cidofovir</td>
<td>1</td>
<td>&gt;250</td>
<td>4.5</td>
<td>5.8</td>
<td>3.8</td>
<td>29</td>
<td>&gt;250</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>&gt;250</td>
<td>4.5</td>
<td>5.8</td>
<td>5.9</td>
<td>50</td>
<td>&gt;250</td>
</tr>
<tr>
<td>Acyclovir</td>
<td>1</td>
<td>&gt;250</td>
<td>0.2</td>
<td>10</td>
<td>0.2</td>
<td>&gt;250</td>
<td>&gt;250</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>&gt;250</td>
<td>0.3</td>
<td>10</td>
<td>0.2</td>
<td>&gt;250</td>
<td>&gt;250</td>
</tr>
<tr>
<td>Ganciclovir</td>
<td>1</td>
<td>&gt;100</td>
<td>0.03</td>
<td>0.4</td>
<td>0.06</td>
<td>&gt;100</td>
<td>&gt;100</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>&gt;100</td>
<td>0.07</td>
<td>0.8</td>
<td>0.2</td>
<td>&gt;100</td>
<td>&gt;100</td>
</tr>
<tr>
<td>Zalcitabine</td>
<td>1</td>
<td>&gt;250</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>&gt;250</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Alovudine</td>
<td>1</td>
<td>&gt;250</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>&gt;250</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
</tbody>
</table>

**Table 3.5.6** Cytotoxicity (MCC, determined by microscopy) and antiviral activity (EC<sub>50</sub> represent the compound concentration reducing the viral CPE by 50%) of the tested compounds in HEL cell cultures assays. ND: not determined.
None of the tested chlorogenic compounds was cytotoxic in the used concentrations range; acid 29 and lactone 7 show a marginal activity against HSV of both types. Acids 29, 31 and 32 are active against Adv-2 showing EC₅₀ values between 45 and 100 μM, while lactone 7 is particularly promising showing comparable results to those of the references (EC₅₀ between 10 and 20 μM). On the other hand, none of the compounds could inhibit VV and VSV.

3.5.5 Overview

The detected activities in this viruses screening are summarized in table 3.5.7.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Activities</th>
</tr>
</thead>
</table>
| Acid 29  | Marginal activity against herpes simplex virus type 1 and type 2  
|          | Activity against vesicular stomatitis virus  
|          | Marginal activity against adenovirus type 2  
|          | Marginal, if any, activity against HIV 1 IIIB wild-type strain  
|          | Activity in influenza vRNP reconstitution assay  
|          | Activity in influenza endonuclease assay |
| Acid 30  | No activity detected in the performed assays |
| Acid 31  | Activity against adenovirus type 2  
|          | Marginal activity against HIV 1 IIIB wild-type strain  
|          | Activity in influenza endonuclease assay |
| Acid 32  | Marginal activity against adenovirus type 2 |
| Lactone 7| Marginal activity against herpes simplex virus type 1 and type 2  
|          | Activity against adenovirus type 2  
|          | Activity against respiratory syncytial virus  
|          | Activity in influenza vRNP reconstitution assay  
|          | Activity in influenza endonuclease assay  
|          | Activity in influenza virus yield assay |
| Lactone 10| Activity against HIV 1 IIIB wild-type strain, no activity against HIV 1 NNRTI resistant strain or HIV-2 ROD strain |
| Lactone 11| Marginal activity against HIV 1 IIIB wild-type strain  
|          | Activity in influenza vRNP reconstitution assay |
| Lactone 12| Activity against HIV 1 IIIB wild-type strain, no activity against HIV 1 NNRTI resistant strain or HIV-2 ROD strain |

Table 3.5.7 Summary of the detected activities in the performed tests.

Acid 30 is the only compound that did not show any activity in the performed assays; all the others showed a moderate activity against one or more viruses; however it is interesting to note how even very small differences in the molecular structure are enough to completely change the behavior against the viruses.

Considering the experiments related to influenza virus, compounds 29, 31 and 11 turned out to be active in the enzymatic tests (influenza endonuclease assay and/or influenza vRNP reconstitution assay) but were not able to inhibit viral replication in virus infected cells. This behavior can be explained by considering that the influenza virus yield assay is more challenging, since the inhibitors have to have sufficient cell permeability and low cytotoxicity.
to be able to inhibit the viral replication in the cells. It is very common to find more activity in an enzyme assay than in an antiviral assay in cell culture. However, even if the enzyme assays are not very predictive about the behavior against the virus, they give precious information: when some compounds are very active on the enzymatic level, it is always possible to try to figure out why they do not work in the cells, and try to add chemical modifications e.g. to make them enter the cell better or protect them from quick degradation. All the data about activity, inactivity and toxicity can help to determine a structure-activity relationship for the compounds and to plan the synthesis of more effective analogues with a higher selectivity index.

It is also important to consider that when a compound is active in both the enzymatic and cellular assays, that still does not mean for sure that the compound inhibits the influenza endonuclease in the antiviral context. Further experiments are necessary to demonstrate that the compound does not have an unrelated antiviral mode of action in cell culture, resulting in erroneous interpretations when investigating their structure-activity relationship (against PA) in virus-infected cells. The compound might be just an entry inhibitor in cell culture, for example. This could be the case of lactone 7: it has shown a broad antiviral activity against multiple viruses that do not possess an endonucleolytic activity, so it is highly possible that the activity in the antiviral influenza assay is unrelated to inhibition of the influenza endonuclease; the exact inhibition mechanism has to be further investigated.

Chapter 3 – Results and discussion

Chapter 3 – Results and discussion


36 amfAR (the foundation for AIDS research) website: www.amfAR.org

37 from the World Health Organization (WHO) website: www.who.int/hiv/pub/guidelines/arv2013/intro/rag/en


Chapter 3 – Results and discussion


66 Laine M., Luukkainen R., Toivanan A. Sindbis viruses and other alphaviruses as cause of human arthriti.

67 Laine M., Luukkainen R., Toivanan A. Sindbis viruses and other alphaviruses as cause of human arthriti.


Chapter 3 – Results and discussion


Chapter 4

Experimental section
4. Experimental section

4.1. Instrumentation

Thin layer chromatography analyses (TLC) were performed on Merck 60 F_{254} silica gel plates; TLC plates were stained with aqueous permanganate solution, anisaldehyde solution, or examined under ultraviolet light (λ = 254 nm).

Column-chromatographic purifications ("flash chromatography") were carried out with Merck silica gel 60 (0.040-0.063 mm, 230-400 mesh ASTM) or with Polyamide MN-SC-6 0.05-0.16 mm.

Melting points (M. p.) were measured with a Sanyo Gallenkamp apparatus.

Optical activity measurements ([α]) were performed with a Perkin-Elmer 241 polarimeter at the wavelength of sodium D band (λ = 589 nm) by using a quartz cuvette with a length of 10 cm (l = 1 dm).

Nuclear Magnetic Resonance (NMR) 500 MHz 1H-NMR and 125 MHz 13C-NMR spectra were obtained on a Varian 500 spectrometer. Chemical shifts (δ) are reported in ppm using the solvent residual signal as an internal reference (CDCl₃: δH = 7.26 ppm, δC = 77.16 ppm; CD₃OD: δH = 3.31 ppm, δC = 49.00 ppm; DMSO-d₆: δH = 2.50 ppm, δC = 39.52 ppm). Coupling constants (J) are given in Hz. The resonance multiplicity is described as s (singlet), d (doublet), t (triplet), q (quartet), dd (doublet of doublets), dt (double of triplets), ddd (doublet of doublet of doublets), dddd (doublet of doublet of doublet of doublets), m (multiplet), br (broad signal).

Electrospray Ionization mass spectrometry measurements (ESI+/−-MS) were performed with a Esquire 4000 (Bruker-Daltonics) spectrometer.

Infrared spectra (IR) were recorded with a Avatar 320-IR FT-IR (Thermo-Nicolet) spectrometer with thin film of sample on NaCl crystal windows.

Reverse Phase high-performance liquid chromatography analyses (RP-HPLC) were run on Amersham Pharmacia Biotech liquid chromatography equipped with UV Amersham detector, using a Gemini C18 3 μm 2x150 mm column for the analytical runs and a Gemini C18 5 μm 10x250 mm column for the semi-preparative ones.

Elemental analyses (EA) were performed with a Flash 2000 Series CHNS/O Analyzer (Thermo Fisher Scientific).

Fluorescence spectra were recorded on a CARY Eclipse (Varian) spectrometer equipped with a 0.5 cm quartz cuvette.

Surface Plasmon Resonance experiments were run on a BiacoreX100 (GE Healthcare Bio-Sciences) instrument and analyzed with the BiacoreX100 evaluation software.

4.2 Materials and general methods

Chemicals, proteins, solvents of synthetic grade and deuterated solvents were purchased from Sigma-Aldrich Co. (St Louis, MO) and used without further purification. Silica gel 60 and TLC plates were purchased from Merck KGaA (Darmstadt, Germany); Polyamide MN-SC-6 0.05-0.16 mm was purchased from Macherey-Nagel (Düren, Germany).
Chapter 4 – Experimental section

When necessary acetone was dried over CaSO₄ and distilled, methanol was dried over CaH₂ and distilled, dichloromethane was dried over CaCl₂; N,N-dimethylformamide (DMF) and pyridine were bought as anhydrous by Sigma-Aldrich.

Low temperature baths (0 °C) were prepared using a mixture of ice and H₂O.

When anhydrous conditions were required, reaction flasks were flame dried and placed under vacuum (approximately 1 mmHg) using a Schlenk line and purged with Ar. Silicon stoppers were used to fit the reaction flasks necks, and a continuous flux of Ar was used to keep the inert atmosphere. The addition of reagents was done by means of glass syringes.

SPR reagents, buffers and the research-grade Sensor Chip CM5 were obtained from GE Healthcare Bio-Sciences Corp. (Piscataway, NJ).

Gemini C18 3 μm 2x150 mm and 5 μm 10x250 mm columns were bought from Phenomenex Inc. (Torrance, CA).

4.3 Experimental procedures

4.3.1. Synthesis of 3,4-O-dicaffeoyl-1,5-γ-quinide 7, improved from the published one.

3,4-O-isopropyliden-1,5-γ-quinide 1

\[
\text{D-}(-)\text{-quinic acid (3.0 g, 15.6 mmol) and p-toluenesulfonic acid (152 mg, 0.8 mmol) were suspended in distilled acetone (150 mL). The mixture was heated under reflux (56 °C) for 48 hours in a Soxhlet apparatus, which was equipped with an extraction thimble filled with molecular sieves (4 Å, Merck), activated overnight in an oven at 120 °C. The reaction mixture was cooled to 0 °C using an ice-bath; NaHCO₃ (364 mg, 10.3 mmol) was added and the suspension was stirred for 1 hour. The mixture was filtered and the organic phase was evaporated under vacuum to obtain 1 as a white solid (yield 93%).}
\]

\[\text{M.p. 133-136 °C (lit. 140-142 °C)}\]
\[\alpha_d^{25} = -30.3 (c = 1, \text{CH}_3\text{OH})\]

\[\begin{align*}
\text{H NMR (500 MHz, CD}_3\text{OD):} & \quad \delta 1.32 (s, 3H, C\text{CH}_3), 1.49 (s, 3H, C\text{CH}_3), 2.02 (dd, 1H, J_{\text{gem}} = 14.6 \text{ Hz} J_{\text{C}_2\text{H}eq-C\text{H}_3} = 3.0 \text{ Hz}, C_2\text{-H}_{eq}), 2.26 (dddd, 1H, J_{\text{gem}} = 11.7 \text{ Hz} J_{\text{C}_6\text{H}eq-C\text{H}_3} = 6.1 \text{ Hz} J_{\text{C}_6\text{H}eq-C\text{H}_{ax}} = 2.3 \text{ Hz} J_{\text{C}_6\text{H}eq-C\text{H}_{eq}} = 1.4 \text{ Hz}, C_6\text{-H}_{eq}), 2.36 (dddd, 1H, J_{\text{gem}} = 14.6 \text{ Hz} J_{\text{C}_2\text{H}ax-C\text{H}_3} = 7.7 \text{ Hz} J_{\text{C}_2\text{H}ax-C\text{H}_{eq}} = 2.3 \text{ Hz}, C_2\text{-H}_{ax}), 2.53 (d, 1H, J_{\text{gem}} = 11.7 \text{ Hz} J_{\text{C}_6\text{H}_{ax-C\text{H}_{eq}}}, 4.30 (dddd, 1H, J_{\text{C}_4\text{H}_{eq-C\text{H}_3}} = 6.5 \text{ Hz} J_{\text{C}_4\text{H}_{eq-C\text{H}_{ax}} = 2.5 \text{ Hz} J_{\text{C}_4\text{H}_{eq-C\text{H}_{ax}} = 1.4 \text{ Hz}, C_4\text{-H}}, 4.52 (dddd, 1H, J_{\text{C}_6\text{H}_{ax-C\text{H}_{ax}} = 7.7 \text{ Hz} J_{\text{C}_6\text{H}_{ax-C\text{H}_{eq}} = 6.1 \text{ Hz} J_{\text{C}_6\text{H}_{eq-C\text{H}_{eq}} = 2.5 \text{ Hz}, C_6\text{-H}}, 4.67 (dddd, 1H, J_{\text{C}_6\text{H}_{ax-C\text{H}_{ax}} = 6.1 \text{ Hz} J_{\text{C}_6\text{H}_{ax-C\text{H}_{eq}} = 2.5 \text{ Hz}, C_6\text{-H}};
\end{align*}\]

\[\begin{align*}
\text{C NMR (125.4 MHz, CD}_3\text{OD):} & \quad \delta 24.54 (q, \text{CH}_3), 27.32 (q, \text{CH}_3), 35.55 (t, \text{CH}_3), 38.98 (t, \text{CH}_3), 72.28 (d, \text{C}_5), 72.90 (d, \text{C}_5), 73.62 (s, \text{C}_1), 76.62 (d, \text{C}_3), 110.74 (s, \text{C}_3), 180.03 (s, \text{C}_7);
\end{align*}\]
IR (cm⁻¹): 3583, 2932, 1777, 1315, 1161, 1075, 980;
ESI-MS: 215 m/z (50, [M+H]⁺), 237 m/z (100, [M+Na]⁺).

1-O-(2,2,2-trichloroethoxycarbonyl)-3,4-O-isopropyliden-1,5-γ-quinide 2

Compound 1 (592 mg, 2.77 mmol) and pyridine (0.55 mL, 6.80 mmol) were dissolved in CH₂Cl₂ (6 mL). A solution of 2,2,2-trichloroethyl chloroformate (0.42 mL, 3.09 mmol) in CH₂Cl₂ (2 mL) was added dropwise to the reaction mixture at 0 °C. The solution was stirred at 0 °C for 1 h and then at room temperature for 24 h. The mixture was sequentially washed with 1 M HCl (two times, 5 mL at a time) and with brine (two times, 8 mL at a time) and the organic layer was dried on anhydrous Na₂SO₄. The solvent was eliminated under reduced pressure to obtain an orange oil. The crude was dissolved in refluxing MeOH (6 mL) and cooled, the solution was stored overnight at 4 °C and precipitate as a white powder (yield 57%).

M.p. 153-156 °C (lit. 164-166 °C);
Rᵢ = 0.80 (dichloromethane – ethyl acetate 7:3, TLC stained with permanganate solution);
[α]D²⁵ = -7.8 (c = 1, CH₂Cl₂);
¹H NMR (500 MHz, CDCl₃): δ 1.33 (s, 3H, CCH₃), 1.52 (s, 3H, CCH₃), 2.40 (dd, 1H, J_gem = 14.7 Hz J₂eq·CH = 2.4 Hz, C₂-H eq), 2.56 (ddd, 1H, J_gem = 14.7 Hz J₂Hax·CH = 7.7 Hz, J₂Hax·CHeq = 2.3 Hz, C₂-Hax), 2.65 (d, 1H, J_gem = 11.4 Hz, C₆-Hax), 3.06 (dddd, 1H, J_gem = 11.4 Hz J₆Heq·CH = 6.5 Hz J₆Heq·CH₂Hax = 2.3 Hz J₆Hax·CHeq = 1.1 Hz, C₆-H eq), 4.34 (ddd, 1H, J₄H·CH = 5.3 Hz J₄H·CH₂ = 2.3 Hz J₄H·CHeq = 1.1 Hz, C₄-H), 4.55 (ddd, 1H, J₃H·C₂Hax = 7.7 Hz J₃H·CH = 5.3 Hz J₃H·CH₂Hax = 2.4 Hz, C₃-H), 4.72 (d, 1H, J_gem = 11.8 Hz, C₁₀-H), 4.80 (dd, 1H, J₃H·CH₆Heq = 6.5 Hz J₃H·CH₆H = 2.3 Hz, C₃-H), 4.82 (d, 1H, J_gem = 11.8 Hz, C₁₀-H);
¹³C NMR (125.4 MHz, CDCl₃): δ 24.41 (q, CH₃), 27.08 (q, CH₃), 30.34 (t, C₆), 35.42 (t, C₅), 71.15 (d, C₁), 72.46 (d, C₆), 75.47 (d, C₅), 77.09 (t, C₁₀), 78.90 (s, C₁), 94.03 (s, CCl₃), 110.21 (s, C₁), 151.50 (s, C₁₀), 172.61 (s, C₇);
IR (cm⁻¹): 2995, 2939, 1809, 1765, 1380, 1241, 1075, 736;
ESI-MS: 411 m/z (97, [M+Cl⁺+Na⁺]), 413 m/z (100, [M+Cl⁺+Cl⁺+Na⁺]), 415 m/z (25, [M+Cl⁺+Cl⁺+Cl⁺+Na⁺]).
Chapter 4 – Experimental section

1-O-(2,2,2-trichloroethoxycarbonyl)-1,5-γ-quinide 3

Trichloroacetic acid (373 mg, 2.28 mmol) was added to water (42 μL) and the mixture was heated until a clear solution was obtained. When the acidic aqueous solution reached room temperature, 2 (252 mg, 0.65 mmol) was added and reaction was stirred for 4 hours. Ice-cooled water (5.7 mL), ethyl acetate (11.4 mL) and 40% aqueous NaHCO₃ solution (11.4 mL) were added in this order. The organic layer was separated from the aqueous one, that was further extracted with ethyl acetate (15 mL); the collected organic phases were washed with 2% NaHCO₃ (15 mL) and water (12 mL). The organic phase was then dried over anhydrous Na₂SO₄, filtered and the solvent was removed under vacuum to obtain a powder. The residence was crystallized from toluene to afford 13 as white crystals (yield 61%).

M.p. 132-133 °C (lit. 130-131 °C);^2 Rₛ = 0.26 (dichloromethane – ethyl acetate 7:3, TLC stained with permanganate solution);

[α]D^25 = -4.6 (c = 1, CH₂Cl₂);

¹H NMR (500 MHz, CDCl₃): δ 1.67 (br, 1H, OH), 2.17 (dd, 1H, J₆₋₇ = 11.2 Hz, C₆-H₇), 2.39 (ddd, 1H, J₆₋₇ = 11.2 Hz, J₆₋₈ = 11.6 Hz, C₆-H₈), 2.84 (br, 1H, OH), 3.06 (ddd, 1H, J₆₋₇ = 11.2 Hz, J₆₋₈ = 11.6 Hz, C₆-H₈), 4.03 (ddd, 1H, J₆₋₇ = 11.2 Hz, J₆₋₈ = 11.6 Hz, C₆-H₈), 4.18 (dd, 1H, J₆₋₇ = 11.2 Hz, J₆₋₈ = 11.6 Hz, C₆-H₈), 4.74 (d, 1H, J₆₋₇ = 11.2 Hz, C₆-H₈), 4.81 (d, 1H, J₆₋₇ = 11.2 Hz, C₆-H₈), 4.93 (dd, 1H, J₆₋₇ = 11.2 Hz, J₆₋₇ = 11.2 Hz, C₆-H₈);

¹³C NMR (125.4 MHz, CDCl₃): δ 32.65 (t, C₆), 36.53 (t, C₆), 65.86 (d, C₉), 65.95 (d, C₉), 76.20 (d, C₅), 77.11 (s, C₅), 79.04 (t, C₅), 94.03 (s, C₆), 151.58 (s, C₅), 171.44 (s, C₅);

IR (cm⁻¹): 3412 (br), 1782, 1764, 1379, 1244, 1037, 665;

ESI-MS: 347 m/z (20, [M+3CI⁻H]), 349 m/z (20, [M+3CI⁻H]), 351 m/z (5, [M+3CI⁻H]), 383 m/z (80, [M+3CI⁻]), 385 m/z (100, [M+3CI⁻]), 386 m/z (50, [M+3CI⁻]), 389 (20, [M+3CI⁻]).

3,4-O-dimethoxycarbonyl caffeic acid 4
Caffeic acid (2.01 g, 11 mmol) was dissolved in 1 M aqueous NaOH (40 mL) and cooled to 0 °C. Methyl chloroformate (2.04 mL, 26.4 mmol) was added dropwise and the mixture was stirred for 1 h at 0 °C and for 1 h at room temperature: in a few time a yellow powder began to precipitate. The reaction mixture was acidified with 2 M aqueous HCl to pH 1, the solid was collected by filtration and washed with water. Recrystallization from 50-50 v/v water-ethanol (28 mL) gave 4 as a yellow-earth powder (yield 90%).

M.p.: 140-141°C;
$R_f$ = 0.43 (dichloromethane – methanol 9:1, UV lamp at 254 nm)
$^1$H NMR (500 MHz, CDCl$_3$): $\delta$ 3.92 (s, 3H, OCH$_3$), 3.93 (s, 3H, OCH$_3$), 6.41 (d, 1H, $J_{C_1H-C_2H} = 15.9$ Hz, C$_1$-H), 7.34 (d, 1H, $J_{ortho} = 8.4$ Hz, C$_5$-H), 7.46 (dd, 1H, $J_{ortho} = 8.4$ Hz $J_{meta} = 1.9$ Hz, C$_4$-H), 7.49 (d, 1H, $J_{meta} = 1.9$ Hz, C$_8$-H), 7.72 (d, 1H, $J_{C_2H-C_1H} = 15.9$ Hz, C$_2$-H);
$^{13}$C NMR (125.4 MHz, CDCl$_3$): $\delta$ 56.10 (q, 2C, OCH$_3$), 118.92 (d, C$_1$), 122.71 (d, C$_8$), 123.72 (d, C$_5$), 127.03 (d, C$_4$), 133.26 (s, C$_3$), 142.86 (s, C$_6$), 144.07 (s, C$_8$), 144.91 (s, C$_4$), 153.03 (s, OCOO), 153.20 (s, OCOO), 171.58 (s, COOH);
IR (cm$^{-1}$): 2919, 1761, 1693, 1633, 1438, 1265, 932;
ESI$^-$-MS: 295 m/z (100, [M-H$^-$]).

3,4-O-dimethoxycarbonyl caffeic acid chloride 5

Thionyl chloride (2.4 mL, 33 mmol) was added dropwise to compound 4 (1.41 g, 4.76 mmol) and the mixture was heated to 90 °C until a homogeneous brown solution formed without gas development (ca. 2 h). Before stopping the reaction, the mixture was checked by $^1$H-NMR in CDCl$_3$. The unreacted thionyl chloride was removed under vacuum and the brown solid residue was recrystallized from toluene (10 mL) to obtain 5 as a yellow powder (yield 50%), which was used immediately after its preparation.

$^1$H NMR (500 MHz, CDCl$_3$): $\delta$ 3.92 (s, 3H, OCH$_3$), 3.93 (s, 3H, OCH$_3$), 6.61 (d, 1H, $J_{C_1H-C_2H} = 15.6$ Hz, C$_1$-H), 7.38 (d, 1H, $J_{ortho} = 8.5$ Hz, C$_5$-H), 7.48 (dd, 1H, $J_{ortho} = 8.5$ Hz $J_{meta} = 2.1$ Hz, C$_4$-H), 7.52 (d, 1H, $J_{meta} = 2.1$ Hz, C$_8$-H), 7.76 (d, 1H, $J_{C_2H-C_1H} = 15.6$ Hz, C$_2$-H);
$^{13}$C NMR (125.4 MHz, CDCl$_3$): $\delta$ 56.16 (q, 2C, OCH$_3$), 123.36 (d, C$_5$), 123.89 (d, C$_8$), 127.03 (d, C$_4$), 132.66 (s, C$_3$), 142.86 (s, C$_6$), 144.07 (s, C$_8$), 144.91 (s, C$_4$), 153.03 (s, OCOO), 153.20 (s, OCOO), 166.88 (s, COCl);
ESI$^+$-MS: 334 m/z (100, [MOCH$_3$+Na$^+$]).
1-O-(2,2,2-trichloroethoxycarbonyl)-3,4-O-bis[3,4-O-(dimethoxycarbonyl)caffeoyl]-1,5-γ-quinide 6

4-dimethylaminopyridine (DMAP, 10 mg, 0.08 mmol) and Et₃N (0.35 mL, 2.5 mmol) were added to a solution of 3 (136 mg, 0.39 mmol) in anhydrous CH₂Cl₂ (10 mL). After cooling the solution to 0 °C, chloride 5 (600 mg, 1.9 mmol) was slowly added and the yellow solution was stirred for 1 h at 0 °C and then for 24 h at room temperature. The reaction mixture was sequentially washed with 1 M HCl (twice, 15 mL at a time), 2% NaHCO₃ (15 mL) and brine (10 mL); the organic layer was dried over Na₂SO₄ and the solvent was removed by vacuum evaporation. The crude was purified by flash chromatography on silica gel (glass column 2.5 x 35 cm, gradient elution from CH₂Cl₂/ethyl acetate 98/2 to 92/8 v/v) to obtain 6 as a pearly powder (yield 40%).

Rₗ = 0.35 (dichloromethane – ethyl acetate 9.5:0.5, UV lamp at 254 nm);
¹H NMR (500 MHz, CDCl₃): δ 2.46 (dd, 1H, J₆Cᵣ,₁H = 11.8 Hz, J₆C₂H,₁H = 11.6 Hz, C₂H₆), 2.56 (dd, 1H, J₆Cᵣ,₁H = 11.8 Hz, J₆C₂H,₁H = 6.8 Hz, J₆C₂H,₁H = 2.7 Hz, C₂H₆), 2.71 (d, 1H, J₆Cᵣ,₁H = 11.6 Hz, C₂H₆), 3.23 (dd, 1H, J₆Cᵣ,₁H = 11.6 Hz, J₆C₂H,₁H = 5.9 Hz, J₆C₂H,₁H = 2.7 Hz, C₂H₆), 3.89 (s, 3H, OCH₃), 3.90 (s, 3H, OCH₃), 3.92 (s, 3H, OCH₃), 3.93 (s, 3H, OCH₃), 4.75 (d, 1H, J₆Cᵣ,₁H = 11.8 Hz, C₂H₆), 4.85 (d, 1H, J₆Cᵣ,₁H = 11.8 Hz, C₂H₆), 5.02 (dd, 1H, J₆Cᵣ,₁H = 11.8 Hz, C₂H₆), 5.36 (ddd, 1H, J₆Cᵣ,₁H = 11.6 Hz, J₆C₂H,₁H = 6.8 Hz, J₆C₂H,₁H = 4.7 Hz, C₃H), 5.71 (dd, 1H, J₆C₂H,₁H = 4.9 Hz, J₆C₂H,₁H = 4.7 Hz, C₃H), 6.28 (d, 1H, J₆C₁₁H,₁₂H = 16 Hz, C₁₁-H), 6.45 (d, 1H, J₆C₁₁H,₁₂H = 15.9 Hz, C₁₁-H), 7.27 (d, 1H, J₆ortho = 8.5 Hz, C₁₅-H), 7.34 (m, 2H, C₁₄-H and C₁₅'H), 7.42 (m, 2H, C₁₄-H and C₁₄'H), 7.52 (d, 1H, J₆meta = 1.9 Hz, C₁₈-H), 7.59 (d, 1H, J₆meta = 16 Hz, C₁₂-H), 7.67 (d, 1H, J₆meta = 15.9 Hz, C₁₂-H);
¹³C NMR (125.4 MHz, CDCl₃): δ 33.78 (t, C₆), 33.87 (t, C₆), 56.07 (q, 2C, OCH₃), 56.11 (q, 2C, OCH₃), 65.05 (d, C₆), 66.12 (d, C₆), 73.74 (d, C₆), 78.78 (t, C₆), 94.04 (s, CCl₃), 114.2 (s, C₆), 117.89 (d, C₁₁), 118.11 (d, C₁₁'), 122.45 (d, C₁₈), 122.51 (d, C₁₈'), 123.69 (d, C₁₃), 123.81 (d, C₁₃'), 127.05 (d, C₁₄), 127.23 (d, C₁₄'), 132.94 (s, C₁₃), 133.16 (s, C₁₃'), 142.83 (s, C₁₇), 142.93 (s, C₁₇'), 144.00 (s, C₁₆), 144.21 (s, C₁₆'), 144.35 (d, C₁₁), 144.94 (d, C₁₂), 151.54 (s, C₆), 152.97 (s, C₁₉), 152.99 (s, C₁₉'), 153.15 (s, C₂₀), 153.21 (s, C₂₀'), 164.61 (s, C₁₈), 164.78 (s, C₁₈'), 170.10 (s, C₇);
IR (cm⁻¹): 358, 2918, 1771, 1722, 1441, 1260, 1146, 728;
ESI⁺-MS: 929 m/z (100, [M⁺Cl⁺Na⁺]).
LiCl (229 mg, 5.4 mmol) was added to a suspension of compound 6 (438 mg, 0.48 mmol) in dry pyridine (stored on molecular sieves, 4.4 mL) and the mixture was stirred for 7 days at 50 °C. During the reaction time the suspension turned to a brown solution. The solvent was removed under vacuum and the residue was dissolved in ethyl acetate (20 mL), then sequentially washed with 2 M HCl (two times, 12 mL each one), 2% NaHCO₃ (two times, 10 mL each one) and brine (8 mL); the organic phase was dried on Na₂SO₄ and the vacuum removal of the solvent gave an orange residue. The crude was previously treated by flash chromatography on polyamide MN-SC-6 (glass column 2 x 30 cm, gradient elution from ethyl acetate/methanol 80/20 to 50/50 v/v) and the fractions rich in the target molecule were subsequently purified by semi-preparative RP-HPLC on a Phenomenex Gemini C18 5 μm 10x250 mm column (15 mg of crude for each run, loop 10 mL), using a gradient of H₂O+0,1% TFA (A) and MeOH+0,1% TFA (B) (20 min A 80% B 20%, from 20 to 90 min increase of B until A 40% B 60%, from 90 to 110 min A 5% B 95%, from 110 to 125 min A 5% B 5%) at a flow rate of 2 mL/min. The elution was monitored with a UV/vis detector at λ 214, 288 and 325 nm; the fractions corresponding to the peak of interest were checked with ESI+-MS (molecular ion [M-H]+ 499 m/z) and then freeze-dried: 6 was obtained as a white powder (yield 20%).

M.p.: 134-136 °C;
Rf = 0.55 (ethyl acetate, UV lamp at 254 nm);

¹H NMR (500 MHz, CD₃OD): δ 2.16 (dd, 1H, J_gem = 11.8 Hz J_C2Hax-C3H = 11.6 Hz, C₂-Hax), 2.28 (ddd, 1H, J_gem = 11.8 Hz J_C2Heq-C3H = 6.8 Hz J_C2Heq-C6Heq = 2.4 Hz, C₂-Heq), 2.47 (ddd, 1H, J_gem = 11.9 Hz J_C6Heq-CSH = 5.7 Hz J_C6Heq-C2Heq = 2.4 Hz, C₆-Heq), 2.59 (d, 1H, J_gem = 11.9 Hz, C₆-Heq), 2.63 (d, 1H, J_SH-C6Hax = 5.7 Hz J_SH-C4H = 5.1 Hz, C₅-H), 5.18 (ddd, 1H, J_SH-C2Hax = 11.6 Hz J_SH-C2Heq = 6.8 Hz J_SH-C4H = 4.7 Hz, C₅-H), 6.51 (dd, 1H, J_SH-CSH = 5.1 Hz J_SH-CH = 4.7 Hz, C₅-H), 6.14 (d, 1H, J_SH-CH = 15.9 Hz, C₅-H), 6.37 (d, 1H, J_SH-C10H = 15.8 Hz, C₁₀-H), 6.68 (d, 1H, J_oortho = 8.2 Hz, C₁₃-H), 6.80 (m, 2H, C₁₂-H and C₁₂'-H), 6.98 (m, 2H, C₁₆-H and C₁₆'-H), 7.08 (d, 1H, J_meta = 2.0 Hz, C₁₆'-H), 7.18 (d, 1H, J_oortho = 15.9 Hz, C₁₆'-H), 7.63 (d, 1H, J_oortho = 15.8 Hz, C₁₀'-H);

¹³C NMR (125.4 MHz, CD₃OD): δ 37.36 (t, C₁), 38.70 (t, C₂), 65.89 (d, C₄), 67.79 (d, C₃), 72.89 (s, C₁), 75.07 (d, C₅), 113.95 (d, C₆), 114.07 (d, C₈), 114.81 (d, C₁₀), 115.34 (d, C₁₆'), 116.42 (d, C₁₆), 116.55 (d, C₁₃), 123.41 (d, C₁₂), 123.49 (d, C₁₂'), 127.45 (s, C₁₁), 127.50 (s, C₁₁'), 146.81 (s, C₁₅), 146.89 (s, C₁₃), 147.87 (d, C₁₀), 148.55 (d, C₁₆), 149.81 (s, C₁₂), 150.00 (s, C₁₄), 167.37 (s, C₈), 167.54 (s, C₈'), 178.22 (s, C₇);

IR (cm⁻¹): 3406, 2951, 1791, 1633, 1269, 1021, 645;
ESI+-MS: 499 m/z (100, [M+H]+).
4.3.2 Synthesis of 1,3,4-O-tris[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide 10, 3,4-O-bis[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide 11 and 3-O-[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide 12

1,5-γ-quinide 8 (starting from a published procedure)

D(-)-quinic acid (1.50 g, 7.83 mmol) and p-toluenesulfonic acid monohydrate (165 mg, 0.37 mmol) were suspended in a mixture of toluene (20 mL) and dry DMF (stored over molecular sieves, 3.75 mL), and the mixture was heated under reflux with a Dean–Stark apparatus for 20 h. The solvent was removed under vacuum, and the brown sticky residue was recrystallized from methanol (10 mL) to afford 8 as an ivory solid (yield 50%).

M.p.: 170-172 °C (lit. 182-183 °C);
Rf = 0.25 (dichloromethane – methanol 9:1, TLC stained with permanganate solution);
[α]D25 = -29.7 (c = 1, CH3OH);
1H NMR (500 MHz, CD3OD): δ 1.87 (dd, 1H, Jgem = 11.4 Hz Jc2Hax-C3H = 10.9 Hz, C2-Hax), 2.02 (ddd, 1H, Jgem = 11.4 Hz Jc2Heq-C3H = 6.6 Hz Jc2Heq-C6Heq = 2.4 Hz, C2-Heq), 2.23 (ddd, 1H, Jgem = 11.2 Hz Jc6Heq-C6H = 5.5 Hz Jc6Heq-C6Heq = 2.4 Hz, C6-Heq), 2.47 (d, 1H, Jgem = 11.2 Hz, C6-Hax), 3.70 (ddd, 1H, Jc3H-C2Hax = 10.9 Hz Jc3H-C3Heq = 6.6 Hz Jc3H-C4H = 3.9 Hz, C3-H), 3.98 (dd, 1H, Jc4H-C5H = 4.1 Hz Jc4H-C6H = 3.9 Hz, C4-H), 4.70 (dd, 1H, Jc5H-C6H = 5.5 Hz Jc5H-C4H = 4.1 Hz, C5-H);
13C NMR (125.4 MHz, CD3OD): δ 37.97 (t, C6), 40.26 (t, C2), 66.97 (d, C3), 67.47 (d, C4), 73.24 (s, C1), 78.01 (d, C5), 179.69 (s, C7);
IR (cm⁻¹): 3583, 3281, 2919, 1795, 1142, 1038, 979;
ESI-MS: 197 m/z (100, [M+Na]+);
EA: C7H10O5 (174.15 g/mol) calculated C 48.28, H 5.79; found C 48.26, H 5.90.

3,4-dimethoxycinnamic acid chloride 9

Thionyl chloride (2.4 mL, 33 mmol) was added dropwise to 3,4-dimethoxycinnamic acid (1.41 g, 4.76 mmol) and the suspension was heated to 90 °C until a homogeneous grey-green solution formed without gas development (ca. 2 h). The unreacted thionyl chloride was removed under vacuum, and the olive-brown solid residue was recrystallized from toluene (10 mL) to afford 9 as yellow solid (yield 50%), which was used immediately.
**Chapter 4 – Experimental section**

$^1$H NMR (500 MHz, CDCl$_3$): $\delta$ 3.91 (s, 3H, OCH$_3$), 3.92 (s, 3H, OCH$_3$), 6.49 (d, 1H, J$_{C1H-C2H}$ = 15.4 Hz, C$_1$-H), 6.89 (d, 1H, J$_{ortho}$ = 8.3 Hz, C$_5$-H), 7.02 (d, 1H, J$_{meta}$ = 1.9 Hz, C$_8$-H), 7.07 (dd, 1H, J$_{ortho}$ = 8.3 Hz, J$_{meta}$ = 1.9 Hz, C$_4$-H), 7.75 (d, 1H, J$_{C2H-C1H}$ = 15.4 Hz, C$_2$-H);

$^{13}$C NMR (125.4 MHz, CDCl$_3$): $\delta$ 55.88 (q, OCH$_3$), 55.97 (q, OCH$_3$), 110.11 (d, C$_8$), 111.11 (d, C$_5$), 119.57 (d, C$_1$), 124.57 (d, C$_4$), 125.90 (s, C$_3$), 149.37 (s, C$_7$), 150.79 (d, C$_2$), 152.66 (s, C$_6$), 165.84 (s, COCl);

IR (cm$^{-1}$): 2938, 2838, 2596, 2360, 1681, 1624, 1595, 1517, 1268, 1141, 1023.

**Coupling of 8 and 9**

DMAP (27 mg, 0.22 mmol) and Et$_3$N (0.9 mL, 6.4 mmol) were added to a suspension of 8 (181 mg, 1.04 mmol) in anhydrous CH$_2$Cl$_2$ (20 mL) and the mixture was cooled to 0 °C. Chloride 9 (796 mg, 3.5 mmol) was slowly added, and the orange solution was stirred for 1 h at 0 °C and then for 24 h at room temperature. The reaction mixture was sequentially extracted with 1 M HCl (two times, 10 mL at a time), 2% NaHCO$_3$ (12 mL), and brine (10 mL); the organic layer was dried with Na$_2$SO$_4$, and the solvent was removed by vacuum evaporation. The crude was analyzed by $^1$H NMR spectroscopy and it was a mixture of esters 1,3,4-O-tris[3,4-(dimethoxy)cinnamoyl]-1,5-$\gamma$-quinide 10 and 3,4-O-bis[3,4-(dimethoxy)cinnamoyl]-1,5-$\gamma$-quinide 11 in a molar ratio of 1:0.4 (conversion of 1,5-$\gamma$-quinide in esters 100%).

The same coupling was performed with a starting molar ratio between 8 and 9 of 1:2.2 (8 172.5 mg, 1.0 mmol; 9 497 mg, 2.2 mmol; DMAP 24.5 mg, 0.2 mmol; Et$_3$N 0.85 mL, 6 mmol; 20 mL of CH$_2$Cl$_2$). In this case, the crude product was a mixture of esters mainly composed of 10, 3-O-[3,4-(dimethoxy)cinnamoyl]-1,5-$\gamma$-quinide 12, 1,3-O-bis[3,4-(dimethoxy)cinnamoyl]-1,5-$\gamma$-quinide 13 (36, 23, and 18%, respectively) and minor amounts of 11, 1,4-O-bis[3,4-(dimethoxy)cinnamoyl]-1,5-$\gamma$-quinide 14, and 4-O-[3,4-(dimethoxy)cinnamoyl]-1,5-$\gamma$-quinide 15 (<10% each; conversion of 1,5-$\gamma$-quinide in esters 100%).

Both the crude mixtures were purified by flash chromatography on silica gel (glass column 2.5x35 cm, elution in gradient from dichloromethane/ethyl acetate 7:3 to 6:4 v/v); 10, 11, and 12 were obtained as pure products.
Chapter 4 – Experimental section

1,3,4-O-tris[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide 10

\[
\begin{align*}
\text{HCO}_3 & \quad \text{OCH}_3 \\
\text{HCO}_3 & \quad \text{OCH}_3
\end{align*}
\]

This compound was recovered from both crudes as an ochre powder.

Compound 10 was also obtained in the following conditions: DMAP (11.7 mg, 0.096 mmol) and Et₃N (2.35 mL, 16.7 mmol) were added to a suspension of D-(-)-quinic acid (133 mg, 0.69 mmol) was suspended in anhydrous CH₂Cl₂ (10 mL) and the mixture was cooled to 0 °C. Chloride 9 (796 mg, 3.5 mmol) was slowly added and the solution was stirred for 1 h at 0 °C and then for 24 h at room temperature. The reaction mixture was sequentially extracted with 1 M HCl (twice, 10 mL at a time), 2% NaHCO₃ (12 mL), and brine (10 mL); the organic layer was dried with Na₂SO₄, and the solvent was removed by vacuum evaporation. The product was purified by flash chromatography on silica gel (glass column 2.5×35 cm, gradient elution from dichloromethane/ethyl acetate 7:3 to 6:4 v/v); compound 10 was obtained as a yellowish powder (yield 75%).

M.p.: 103-105 °C;

\[R'= 0.64 \text{ (dichloromethane – ethyl acetate 7:3, UV lamp at 254 nm);} \]

[α]D²⁵ = +266.7 (c = 1, CH₂Cl₂), +260.8 (c = 0.5, CH₂Cl₂);

¹H NMR (500 MHz, CDCl₃): δ 2.51 (ddd, 1H, J₂-heq,3-heq = 11.7 Hz), J₂-heq,4-heq = 6.8 Hz), J₂-heq,6-heq = 2.5 Hz, C₂-Heq), 2.64 (dd, 1H, J₂-heq,3-heq = 11.7 Hz, J₂-heq,4-heq = 11.6 Hz, C₂-Hax), 2.89 (d, 1H, J₂-heq,3-heq = 11.7 Hz, C₂-Hax), 3.13 (ddd, 1H, J₂-heq,3-heq = 11.7 Hz), J₂-heq,4-heq = 6.1 Hz, J₂-heq,2-heq = 2.5 Hz, C₂-Heq), 3.80 (s, 3H, OCH₃), 3.88 (s, 3H, OCH₃), 3.92 (s, 6H, 2 OCH₃), 3.93 (s, 6H, 2 OCH₃), 5.01 (dd, 1H, J₂-heq,3-heq = 6.1 Hz, J₂-heq,4-heq = 5.1 Hz, C₁₀-H), 5.38 (ddd, 1H, J₂-heq,3-heq = 11.6 Hz, J₂-heq,4-heq = 6.8 Hz, J₂-heq,6-heq = 4.7 Hz, C₂-H), 5.73 (dd, 1H, J₂-heq,3-heq = 5.1 Hz, J₂-heq,6-heq = 4.7 Hz, C₁₀-H), 6.20 (d, 1H, J₂-heq,3-heq = 15.9 Hz, C₁₀-H), 6.34 (d, 1H, J₂-heq,3-heq = 15.9 Hz, C₁₀-H), 6.40 (d, 1H, J₂-heq,3-heq = 15.9 Hz, C₁₀-H), 6.77 (d, 1H, J₂-heq,3-heq = 8.3 Hz, C₁₀-H), 6.87 (d+d, 2H, J₂-heq,3-heq = 8.3 Hz, C₁₀-H and C₁₁-H), 6.95 (d, 1H, J₂-heq,6-heq = 1.8 Hz, C₁₁-H), 7.01 (dd, 1H, J₂-heq,3-heq = 8.3 Hz, J₂-heq,4-heq = 1.8 Hz, C₁₁-H), 7.07 (d+d, 2H, J₂-heq,3-heq = 1.8 Hz, C₁₁-H and C₁₂-H), 7.12 (dd+d, 2H, J₂-heq,3-heq = 8.3 Hz, J₂-heq,4-heq = 1.8 Hz, C₁₁-H and C₁₂-H), 7.57 (d, 1H, J₂-heq,3-heq = 15.9 Hz, C₁₀-H), 7.67 (d, 2H, J₂-heq,3-heq = 15.9 Hz, C₁₀-H and C₁₁-H);

¹³C NMR (125.4 MHz, CDCl₃): δ 33.87 (t, C₂), 34.71 (t, C₂), 55.78 (q, OCH₃), 55.92 (q, OCH₃), 55.93 (q, OCH₃), 55.94 (q, OCH₃), 55.98 (q, OCH₃), 56.01 (q, OCH₃), 64.92 (d, C₂), 65.99 (d, C₂), 73.97 (d, C₅), 76.57 (s, C₅), 109.54 (d, C₁₀), 109.67 (d, C₁₀), 109.77 (d, C₁₀), 110.93 (d, C₁₁), 111.02 (d, 2C, C₁₀ and C₁₁), 114.05 (d, C₂), 114.24 (d, C₂), 114.37 (d, C₂), 122.81 (d, C₁₀), 123.08 (d, C₁₀), 123.45 (d, C₁₀), 126.85 (s, C₁₁), 126.89 (s, C₁₁), 127.01 (s, C₁₁), 146.01 (d, C₁₁), 146.66 (d, C₁₁), 146.86 (d, C₁₀), 149.13 (s, C₁₁), 149.27 (s, C₁₁), 149.32 (s, C₁₂), 151.32 (s, C₁₂), 151.63 (s, C₁₂), 151.65 (s, C₁₂), 165.15 (s, C₂), 165.27 (s, C₂), 165.51 (s, C₂), 171.45 (s, C₂);

IR (cm⁻¹): 2938, 1804, 1716, 1630, 1598, 1514, 1464, 1422, 1263, 1138, 1023, 981, 734;
3,4-O-bis[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide 11

This compound was recovered as a yellow powder from the crude mixture of the first coupling run (1:3.5 molar ratio between quinide 8 and chloride 9).

M.p.: 104-106 °C;
Rf = 0.35 (dichloromethane – ethyl acetate 7:3, UV lamp at 254 nm);
[α]D25°: +219.4 (c = 0.5, CH2Cl2), +213.6 (c = 0.25, CH2Cl2);
1H NMR (500 MHz, CDCl3): δ 2.26 (dd, 1H, Jgem = 11.9 Hz, J2Hax-C2H = 11.7 Hz, C2-Hax), 2.37 (ddd, 1H, Jgem = 11.9 Hz, J2Heq-C3H = 6.8 Hz, J2Heq-C4H = 2.3 Hz, C2-Heq), 2.50 (ddd, 1H, Jgem = 11.9 Hz, J2Heq-C3H = 5.6 Hz, J2Heq-C4H = 2.3 Hz, C2-Heq), 2.64 (d, 1H, Jgem = 11.9 Hz, C2-Hax), 3.26 (br, 1H, C1-OH), 3.79 (s, 3H, OCH3), 3.87 (s, 3H, OCH3), 3.91 (s, 3H, OCH3), 3.92 (s, 3H, OCH3), 4.94 (dd, 1H, JCH1-COeq = 5.6 Hz, JCH1-C4H = 5.1 Hz, C5-H), 5.29 (ddd, 1H, JCH1-C2Hax = 11.7 Hz, JCH1-C2Heq = 6.8 Hz, JCH1-C4H = 4.8 Hz, C3-H), 5.68 (dd, 1H, JCH1-C3H = 5.1 Hz, JCH1-C4H = 4.8 Hz, C2-H), 6.20 (d, 1H, JCH1-C1H = 15.9 Hz, C6-H), 6.38 (d, 1H, JαH1-C1H = 15.8 Hz, Cγ-H), 6.77 (d, 1H, Jortho = 8.3 Hz, C13-H), 6.87 (d, 1H, Jortho = 8.3 Hz, C13-H), 6.94 (d, 1H, Jmeta = 1.8 Hz, C16-H), 6.96 (dd, 1H, Jortho = 8.3 Hz, Jmeta = 1.8 Hz, C12-H), 7.05 (d, 1H, Jmeta = 1.8 Hz, C16-H), 7.11 (dd, 1H, Jortho = 8.3 Hz, Jmeta = 1.8 Hz, C12-H), 7.56 (d, 1H, J10H-C9H = 15.9 Hz, C10-H), 7.67 (d, 1H, J10H-C9H = 15.9 Hz, C10-H);
13C NMR (125.4 MHz, CDCl3): δ 36.90 (t, C3), 37.42 (t, C5), 55.76 (q, OCH3), 55.90 (q, OCH3), 64.56 (d, C6), 66.15 (d, C3), 72.01 (s, C1), 73.95 (d, C5), 109.70 (d, C16), 109.74 (d, C10), 110.90 (d, C13), 111.03 (d, C13), 114.02 (d, C9), 114.39 (d, C6), 122.81 (d, C2), 123.08 (d, C12), 126.79 (s, C1), 126.99 (s, C11), 145.99 (d, C10), 146.61 (d, C10), 149.10 (s, C13), 149.35 (s, C18), 151.29 (s, C4), 151.62 (s, C14), 165.39 (s, C3), 165.54 (s, C3), 176.78 (s, C5);
IR (cm⁻¹): 3462, 2938, 1797, 1714, 1630, 1598, 1514, 1464, 1261, 1138, 1065, 1022, 981, 734;
ESI⁺-MS: 577 m/z (100, [M+Na]+);
EA: C30H28O11·1/2H2O (563.54 g/mol) calculated C 61.81, H 5.54; found C 61.94, H 5.48.
Chapter 4 – Experimental section

3-O-bis[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide 12

This compound was recovered from the crude mixture of the coupling in 1:2.2 molar ratio between quinide 8 and chloride 9; the last fractions collected from the column were a blend of 12 and 15; recrystallization from CHCl₃ gave 12 as a white solid.

M.p.: 188-189 °C; 
Rᵣ = 0.16 (dichloromethane – ethyl acetate 7:3, UV lamp at 254 nm); 
[α]D²⁵ = +3.2 (c = 0.5, CH₃OH)

1H NMR (500 MHz, CD₃OD): δ 2.09 (dd, 1H, J_gem = 11.7 Hz J₂Hax-C₃H = 11.6 Hz, C₂-Hax), 2.17 (ddd, 1H, J_gem = 11.7 Hz J₂Heq-C₃H = 6.2 Hz J₂Heq-C₆Heq = 2.4 Hz, C₂-Heq), 2.30 (ddd, 1H, J_gem = 11.6 Hz J₂Heq-C₃H = 5.5 Hz J₂Heq-C₆Heq = 2.4 Hz, C₆-Heq), 2.57 (d, 1H, J_gem = 11.6 Hz C₆-Hax), 3.87 (s, 6H, 2 OCH₃), 4.30 (dd, 1H, J_C₄H-C₃H = 4.8 Hz J_C₄H-C₆Heq = 4.9 Hz, C₅-H), 4.75 (dd, 1H, J_C₅H-C₆Heq = 5.5 Hz J_C₅H-C₆Hax = 4.8 Hz, C₃-H), 4.92 (ddd, 1H, J_C₆H-C₆Hax = 11.6 Hz J_C₆H-C₅Heq = 6.2 Hz J_C₆H-C₅Hax = 4.9 Hz, C₆-H), 6.44 (d, 1H, J_C₁₀H-C₉H = 15.9 Hz, C₁₂-H), 6.98 (d, 1H, J-meta = 8.2 Hz, C₁₃-H), 7.18 (dd, 1H, J-meta = 1.9 Hz, C₁₂-H), 7.22 (d, 1H, J-meta = 1.9 Hz, C₁₆-H), 7.71 (d, 1H, J₁₁H-C₁₀H = 15.9 Hz, C₁₀-H)

13C NMR (125.4 MHz, CD₃OD): δ 36.89 (t, C₂), 37.77 (t, C₆), 56.43 (q, OCH₃), 56.51 (q, OCH₃), 64.83 (d, C₄), 70.26 (d, C₃), 73.01 (s, C₁), 77.64 (d, C₅), 111.58 (d, C₁₆), 112.65 (d, C₁₃), 116.05 (d, C₉), 124.09 (d, C₁₂), 128.78 (s, C₁₁), 146.99 (d, C₁₀), 150.81 (s, C₁₅), 152.98 (s, C₁₄), 167.74 (s, C₈), 178.87 (s, C₇)

IR (cm⁻¹): 3583, 2360, 1788, 1708, 1636, 1597, 1517, 1264, 1135, 1021, 812, 666

ESI⁻-MS: 387 m/z (100, [M+Na]⁻);

EA: C_{18}H_{20}O₈·½H₂O (373.35 g/mol) calculated C 57.91, H 5.67; found C 58.08, H 5.46.

4.3.3 Synthesis of (3-O-[4-(aminobutanamidoethoxy)feruloyl]-4-O-[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide) 28

Boc-ethanolamine 18

Et₃N (2.02 g, 2.80 mL, 20 mmol) was added to a solution of ethanolamine (1.22 g, 1.20 mL, 20 mmol) in dry dichloromethane (20 mL) and di-tert-butyl-dicarbonate (3.94 g, 16 mmol) in dry CH₂Cl₂ was added dropwise to the reaction mixture at 0 °C, under mechanical stirring and Argon atmosphere. The mixture was kept under stirring for 7 hours at room temperature,
then washed with 0.1 M Na$_2$SO$_4$ (three times, 10 mL at a time) and brine (twice, 10 mL at a time). The organic layer was dried with Na$_2$SO$_4$ and the solvent was eliminated by vacuum evaporation to afford 18 as a colourless oil (yield 62%).

$R_f = 0.55$ (dichloromethane – methanol 9:1, TLC stained with anisaldehyde solution);
$^1$H NMR (500 MHz, CDCl$_3$): $\delta$ 1.43 (s, 9H, 3CH$_3$), 2.69 (br, 1H, OH), 3.26 (dt, 2H, $J_{CH2-CH2} = 5.4$ Hz $J_{CH2-NH} = 1.5$ Hz, C$_2$-H$_2$), 3.67 (dt, 2H, $J_{CH2-CH2} = 5.4$ Hz $J_{CH2-OH} = 2.2$ Hz, C$_1$-H$_2$), 5.01 (br, 1H, NH);
$^{13}$C NMR (125.4 MHz, CDCl$_3$): $\delta$ 28.35 (q, 3C, 3CH$_3$), 43.23 (t, C$_2$), 62.40 (t, C$_1$), 79.69 (s, C$_4$), 156.79 (s, C$_3$);
IR (cm$^{-1}$): 3349, 2977, 2934, 1689, 1171, 1069;
ESI$^+$-MS: 184 m/z (100, [M+Na]$^+$, 200 m/z (30, [M+K]$^+$).

Boc-ethanolaminetosylate 19

1,4-diazabicyclo[2.2.2]octane (DABCO, 0.70 g, 6.20 mmol) was added to a solution of compound 18 (0.50 g, 3.10 mmol) in anhydrous CH$_2$Cl$_2$ (6 mL) under argon atmosphere. The solution was cooled to 0 °C with an ice-bath, then p-toluenesulfonyl chloride (0.89 g, 4.65 mmol) was slowly added to the reaction mixture over 5 minutes, and a precipitate was immediately formed. The reaction was kept at 0 °C for 1 hour and then at room temperature for further 2 h. The precipitate was filtered off and the mother solution was extracted with tert-butylmethylether (four times, 6 mL at a time). The combined organic layers were collected and washed with 2M HCl, 5% NaHCO$_3$, and brine, then dried over Na$_2$SO$_4$. The solvent was eliminated under reduced pressure to obtain 19 as an oil (yield 78%).

$R_f = 0.39$ (petroleum ether - ethyl acetate 2:1, UV lamp at 254 nm);
$^1$H NMR (500 MHz, CDCl$_3$): $\delta$ 1.40 (s, 9H, 3CH$_3$), 2.44 (s, 3H, C$_8$-CH$_3$), 4.06 (t, 2H, $J_{CH2-CH2} = 5.0$ Hz, C$_2$-H$_2$), 7.26 (d, 2H, $J_{ortho} = 8.2$ Hz, C$_7$-H and C$_9$-H), 7.78 (d, 2H, $J_{ortho} = 8.2$ Hz, C$_6$-H and C$_10$-H);
$^{13}$C NMR (125.4 MHz, CDCl$_3$): $\delta$ 21.63 (q, C$_8$-CH$_3$), 28.27 (q, 3C, 3C$_2$-CH$_3$)$\times$3C), 39.76 (t, C$_2$), 69.45 (t, C$_1$), 79.78 (s, C$_4$), 127.92 (d, 2C, C$_6$ and C$_9$), 129.93 (d, 2C, C$_7$ and C$_9$), 140.75 (s, C$_3$), 145.00 (s, C$_3$), 155.88 (s, C$_3$);
IR (cm$^{-1}$): 3401, 2978, 2932, 1701, 1519, 1364, 1175, 664;
ESI$^+$-MS: 338 m/z (100, [M+Na]$^+$), 354 m/z (20, [M+K]$^+$).
Methyl ferulate 20

To a solution of ferulic acid (1.00 g, 5.15 mmol) in MeOH (30 mL) p-toluenesulfonic acid (0.10 g, 0.515 mmol) was added. The mixture was heated under reflux overnight and then cooled to room temperature. The solvent was eliminated under reduced pressure and the product thus obtained was dissolved in dichloromethane (30 mL). The organic layer was washed with saturated NaHCO₃ (2 x 15 mL) and dried over Na₂SO₄. The solution was then concentrated in vacuum to afford a transparent yellow oil, which was crystallized from ethyl ether to give 20 as a white solid (yield 92%).

M. p.: 60-64 °C;
Rᵣ = 0.65 (dichloromethane - methanol 9:1, UV lamp at 254 nm);
¹H NMR (500 MHz, CDCl₃): δ 3.79 (s, 3H, C₁-OCH₃), 3.92 (s, 3H, C₈-OCH₃), 6.29 (d, 1H, J_C2H=C₃H = 15.9 Hz, C₂-H), 6.91 (d, 1H, Jortho = 8.2 Hz, C₆-H), 7.02 (d, 1H, Jmeta = 1.8 Hz, C₄-H), 7.07 (dd, 1H, Jortho = 8.2 Hz Jmeta = 1.8 Hz, C₅-H), 7.62 (d, 1H, J_C3H=C2H = 15.9 Hz, C₃-H);
¹³C NMR (125.4 MHz, CDCl₃): δ 51.74 (q, C₁-OCH₃), 56.07 (q, C₈-OCH₃), 109.51 (d, C₆), 114.87 (d, C₅), 115.29 (d, C₃), 123.16 (d, C₉), 127.08 (s, C₄), 145.09 (d, C₇), 146.91 (s, C₈), 148.12 (s, C₉), 167.86 (s, C₁)
IR (cm⁻¹): 3407 (br), 2951, 1701, 1634, 1515, 1434, 1270, 1172; ESI⁻-MS: 209 m/z (40, [M+H]⁻), 231 m/z (100, [M+Na]⁻).

4-(Boc-aminoethoxy)methyl ferulate 21

To a solution of 20 (325 mg, 1.56 mmol) in anhydrous acetone (25 mL), potassium carbonate (1.10 g, 7.96 mmol), 18-crown-6 (21 mg, 7.96x10⁻³ mol), and compound 19 (738 mg, 2.34 mmol) were added in order and the mixture was heated under reflux (56 °C) for 72 h. The mixture was then cooled to room temperature, filtered, and evaporated under reduced pressure. A brown-orange oil was obtained and the crude was purified by silica gel chromatography (dichloromethane-ethyl acetate 4:1 v/v), to afford 20 as a white solid (yield 55%).

Rᵣ = 0.44 (dichloromethane - ethyl acetate 4:1, UV lamp at 254 nm);
Chapter 4 – Experimental section

^1^H NMR (500 MHz, CDCl3): δ 1.44 (s, 9H, 3 C13-CH3), 3.56 (m, 2H, JCH2-CH2 = 5.0 Hz, C11-H2), 3.79 (s, 3H, C1-OCH3), 3.89 (s, 3H, C8-OCH3), 4.10 (t, 2H, JCH2-CH2 = 5.0 Hz, C10-CH2), 5.12 (br, 1H, NH), 6.31 (d, 1H, J2H-CH3 = 15.9 Hz, C2-H), 6.88 (d, 1H, Jortho = 8.3 Hz, C6-H), 7.04 (d, 1H, Jmeta = 1.9 Hz, C5-H), 7.08 (dd, 1H, Jmeta = 1.9 Hz, C3-H), 7.62 (d, 1H, JCH2-CH2 = 15.9 Hz, C3-H);
^13^C NMR (125.4 MHz, CDCl3): δ 28.37 (q, 3C, 3 C13-CH3), 39.98 (t, C11), 51.61 (q, C1-OCH3), 55.84 (q, C8-OCH3), 68.68 (t, C10), 79.52 (s, C13), 110.23 (d, C6), 113.42 (d, C5), 115.82 (d, C2), 122.40 (d, C3), 128.04 (s, C1), 144.60 (d, C12), 149.64 (s, C7), 150.15 (s, C7), 155.88 (s, C12), 167.57 (s, C1);
IR (cm⁻¹): 3369, 2976, 1756, 1709, 1513, 1257, 1167, 1033; ESI^+^-MS: 374 m/z (100, [M+Na]^+), 390 m/z (30, [M+K]^+).

4-(aminoethoxy)methyl ferulate 22

![Diagram of 4-(aminoethoxy)methyl ferulate 22](image)

Compound 21 (300 mg, 0.85 mmol) was stirred in 50% TFA/CH2Cl2 (15 mL) for 90 minutes at room temperature. The solvents were removed by vacuum evaporation; the solid residue was dissolved in CH2Cl2 and dried under vacuum five times. Product 22 was obtained as a white solid (yield 100%).

^1^H NMR (500 MHz, CDCl3): δ 3.33 (t, 2H, JCH2-CH2 = 4.6 Hz, C11-H2), 3.79 (s, 6H, C1-OCH3 and C8-OCH3), 4.22 (t, 2H, JCH2-CH2 = 4.6 Hz, C10-H2), 6.28 (d, 1H, J2H-CH3 = 16.1 Hz, C2-H), 6.87 (d, 1H, Jortho = 8.3 Hz, C6-H), 6.98 (d, 1H, Jmeta = 1.6 Hz, C5-H), 7.04 (dd, 1H, Jortho = 8.3 Hz Jmeta = 1.6 Hz, C3-H), 7.56 (d, 1H, JCH2-CH2 = 16.1 Hz, C3-H), 8.37 (br, 2H, NH2);
^13^C NMR (125.4 MHz, CDCl3): δ 39.33 (t, C11), 51.70 (q, C1-OCH3), 55.73 (q, C8-OCH3), 66.30 (t, C10), 110.71 (d, C6), 116.09 (d, C5), 116.88 (d, C2), 122.22 (d, C12), 129.85 (s, C7), 144.05 (d, C12), 148.54 (s, C13), 150.05 (s, C7), 167.34 (s, C1);
IR (cm⁻¹): 3541, 2920, 2360, 2340, 1693, 1514, 1203, 1136, 666; ESI^+^-MS: 252 m/z (100, [M+H]^+), 274 m/z (70, [M+Na]^+).
A solution of 4-aminobutanoic acid (1.00 g, 9.7 mmol) in triethylamine/distilled MeOH (1:7 v/v, 22.5 mL) was stirred at 0 °C for 10 minutes. A solution of di-tert-butyl dicarbonate (2.24 g, 10.67 mmol) in MeOH (20 mL) was added dropwise in 10 minutes. The mixture was stirred for 1 h under Argon atmosphere. The temperature was gradually risen to room temperature, and then the mixture was heated under reflux (60 °C) overnight. The solution was cooled and evaporated under reduced pressure, the residue was dissolved in CH₂Cl₂ and washed with 0.2 M HCl. The organic layer was separated, dried over Na₂SO₄, filtered, and concentrated to give 23 as a brown oil (yield 89%) that was used in the next step without further purification.

R_f = 0.69 (dichloromethane - methanol 9:1, TLC stained with permanganate solution);

^1H NMR (500 MHz, CDCl₃): δ 1.42 (s, 9H, 3 C₆H₃), 1.80 (m, 2H, J_C3H2-C2H2 = 7.2 Hz, C₃-H₂), 2.38 (t, 2H, J_C2H2-C3H2 = 7.2 Hz, C₂-H₂), 3.17 (br, 2H, C₄-H₂), 4.17 (br, 1H, NH);

^13C NMR (125.4 MHz, CDCl₃): δ 25.14 (t, C₃), 28.35 (q, 3C, 3 C₆H₃-C), 31.25 (t, C₂), 39.74 (t, C₄), 79.54 (s, C₆), 156.17 (s, C₅), 178.21 (s, C₁);

IR (cm⁻¹): 3348, 2978, 2936, 1709, 1526, 1368, 1253, 1170;

ESI⁻-MS: m/z 226 (100, [M+Na]⁺).

4-(Boc-aminobutanamidoethoxy)methyl ferulate 24

Compound 23 (180 mg, 0.89 mmol)) was dissolved under Argon atmosphere in dry dichloromethane (stored over CaCl₂, 16 mL) and kept at 0 °C for 10 minutes. 1-hydroxybenzotriazole (HOBt, 156 mg, 1.15 mmol) and 4-methylmorpholine (NMM, 0.8 mL, 7.12 mmol, reached pH 8.5-9) were added and the mixture was stirred for further 10 minutes. 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC-Cl, 221 mg, 1.15 mmol) and compound 22 (290 mg, 1.15 mmol) were added. The mixture was stirred for 1 h at 0 °C and then for 72 h at room temperature, then it was washed with 10% citric acid (twice, 5 mL at a time), saturated NaHCO₃ (5mL), and brine (5 mL) and dried with Na₂SO₄. The solvent was removed under reduced pressure to afford 24 as a white solid (yield 78%).

R_f = 0.25 (dichloromethane - ethyl acetate 4:1, UV lamp at 254 nm);
4-(Boc-aminobutylamidoethoxy)ferulic acid 25

Compound 24 (304 mg, 0.70 mmol) was dissolved in tetrahydrofuran (6 mL). A 3 M solution of LiOH (1.51 g, 35.99 mmol) in H₂O (6 mL) was added dropwise to the reaction mixture to reach a pH close to 12. After 48 h the reaction was stopped and acidified to pH 3 with 3 M HCl. The reaction mixture was extracted with ethyl acetate (six times, 12 mL at a time) and the organic layer was dried over Na₂SO₄. The solution was filtered and the solvent was evaporated under reduced pressure to afford compound 25 as a white powder (yield 83%).

¹H NMR (270 MHz, CDCl₃): δ 1.41 (s, 9H, 3 C₁₇-CH₃), 1.82 (m, 2H, J₁₁₁₂−C₁₄₁₂ = 6.7 Hz, C₁₄-H₂), 2.26 (t, 2H, J₁₁₁₂−C₁₄₁₂ = 6.7 Hz, C₁₃-H₂), 3.15 (m, 2H, C₁₅-H₂), 3.70 (m, 2H, J₁₁₁₂−C₁₄₁₂ = 5.1 Hz C₁₁-H₂), 3.88 (s, 3H, C₆-OCH₃), 4.14 (t, 2H, J₁₁₁₂−C₁₄₁₂ = 5.1 Hz, C₁₀-H₂), 4.80 (br, 1H, C₁₆-NH), 6.29 (d, 1H, J₁₂₁₃−C₁₄₁₅ = 15.8 Hz, C₂-H), 6.64 (br, 1H, C₁₂-NH), 6.89 (d, 1H, J_ortho = 8.0 Hz, C₆-H), 7.04 (s, 1H, C₈-H), 7.08 (d, 1H, J_ortho = 8.0 Hz, C₈-H), 7.67 (d, 1H, J₁₂₁₃−C₁₄₁₅ = 15.8 Hz, C₆-H);

¹³C NMR (125.4 MHz, CDCl₃): δ 26.26 (t, C₁₄), 28.38 (q, 3C, 3 C₁₇-CH₃), 33.62 (t, C₁₅), 38.88 (t, C₁₁), 39.77 (t, C₁₅), 55.86 (q, C₆-OCH₃), 68.03 (t, C₁₀), 79.42 (s, C₁₇), 110.38 (d, C₆), 113.43 (d, C₆), 115.55 (d, C₆), 122.82 (d, C₆), 127.88 (s, C₆), 146.27 (d, C₆), 149.63 (s, C₆), 150.29 (s, C₆), 156.45 (s, C₆), 171.21 (s, C₆), 173.16 (s, C₁₂);

IR (cm⁻¹): 3582, 3231, 3065, 2933, 1680, 1627, 1529, 1136, 635;

ESI-MS: 421 m/z (100, [M - H]).
1-O-(2,2,2-trichloroethoxycarbonyl)-3-O-(4-(Boc-aminobutanamidoethoxy)feruloyl)-1,5-γ-quinine) 26

Compound 25 (125 mg, 0.30 mmol) was dissolved under Argon atmosphere in dry dichloromethane (stored over CaCl₂, 15 mL) and kept at 0 °C for 10 minutes. HOBT (52 mg, 0.39 mmol) and NMM (5 μL, 0.68 mmol, reached pH 7-7.5) were added, then the mixture was stirred for further 10 minutes. EDC-Cl (74 mg, 0.39 mmol) and compound 3 (103 mg, 0.30 mmol) were added. The mixture was stirred for 1h at 0 °C and then for 10 days at room temperature. The reaction mixture was was washed with 10% citric acid (5 mL), saturated NaHCO₃ (5 mL), and brine (twice, 5 mL at a time) and dried over Na₂SO₄. The solvent was removed under reduced pressure to obtain a crude compound (yield 87%) that was used for next step without purifications.

¹H NMR (500 MHz, CDCl₃): δ 1.41 (s, 9H, C₃₂₀-CH₃), 1.80 (m, 2H, C₂₁-H₂), 2.15 (m, 1H, C₂₂-H₁), 2.23 (m, 2H, C₂₂-H₂), 2.37 (m, 1H, C₂-H₆), 2.78 (t, 1H, J₆-H = 11.3 Hz, C₂-H₆), 3.04 (m, 1H, J₆-H = 11.3 Hz, C₂-H₆), 3.14 (m, 2H, C₂₄-H₂), 3.68 (m, 2H, C₂₀-H₂), 3.88 (s, 3H, C₁₇-OCH₃), 4.13 (m, 2H, C₂₁-H₂), 4.45 (m, 1H, C-H), 4.76 (m, 2H, C₂-H₂), 4.91 (m, 1H, C-H), 5.14 (m, 1H, C₃-H), 6.36 (d, 1H, J₁₁₁H-C₁₁₂H = 15.9 Hz, C₁₁-H), 6.53 (br, 1H, C₂₁-NH), 6.88 (d, 1H, Jortho = 8.4 Hz, C₁₈-H), 7.03 (d, 1H, Jmeta = 1.9 Hz, C₁₈-H), 7.08 (dd, 1H, Jortho = 8.4 Hz Jmeta = 1.9 Hz, C₁₁-H), 7.65 (d, 1H, J₁₁₂H-C₁₁₁H = 15.9 Hz, C₁₂-H);

¹³C NMR (125.4 MHz, CDCl₃): δ 26.25 (t, C₂₁), 28.39 (q, 3C, C₂₀-CH₃), 32.57 (t, C₃), 33.61 (t, C₂₂), 36.55 (t, C₂), 38.85 (t, C₁₀), 39.73 (t, C₁), 55.89 (q, C₁₇-OCH₃), 64.22 (d, C₄), 68.03 (d, C₁₀), 68.11 (t, C₁₉), 76.28 (d, C₁₂), 76.93 (t, C₁₀), 79.84 (s, C₁), 79.06 (s, C₂₁), 93.92 (s, C₁₇-CCL₃), 110.38 (d, C₁₆), 113.44 (d, C₁₉), 114.51 (d, C₁₁), 122.45 (d, C₁₄), 128.95 (s, C₁₁), 146.33 (d, C₁₂), 149.66 (s, C₁₇), 150.30 (s, C₁₆), 151.45 (s, C₁₉), 156.41 (s, C₂₁), 165.24 (s, C₁₀), 171.40 (s, C₂₁), 172.92 (s, C₁);

IR (cm⁻¹): 3343, 2936, 1800, 1699,1513, 1245, 1036, 732;

ESI-MS: 775 m/z (95, [M+Cl⁺+Na⁺]), 777 m/z (100, [M⁺Cl⁻+Cl⁺+Na⁺]), 779 m/z (25, [M⁺Cl⁻+Cl⁺+Na⁺]).
1-O-(2,2,2-trichloroethoxycarbonyl)-3-O-[4-(Boc-aminobutanamidoethoxy)feruloyl]-4-O-[3,4-(dimethoxy)cinnamoyl]-1,5-γ-quinide 27

Compound 26 (0.199 g, 0.264 mmol) was dissolved in CH₂Cl₂ (12 mL) and the solution was cooled to 0 °C under stirring. DMAP (6.5 mg, 0.05 mmol), triethylamine (0.2 mL, 1.58 mmol) and chloride 9 (0.090 g, 0.40 mmol) were added. The solution was kept at 0 °C for 1 hour and then it reached gradually the room temperature. After 24 hours the mixture was washed with 1M HCl (twice, 3 mL at a time), 2% NaHCO₃ (4 mL) and brine (twice, 4 mL at a time). The organic layer was dried over Na₂SO₄ and the solvent was removed by vacuum evaporation to afford a yellow pale oil. The crude was purified by silica gel chromatography (dichloromethane - methanol 9:1 v/v). Product 27 was obtained as a yellowish solid (yield 12%).

R₇ = 0.47 (dichloromethane - methanol 9:5:0.5, UV lamp at 254 nm);
[α]D₀⁺ = -79.5 (c =1.0, CHCl₃);
¹H NMR (500 MHz, CDCl₃): δ 1.41 (s, 9H, 3 Cₛ-C₃H₃), 1.80 (m, 2H, C₃H₂), 2.24 (m, 2H, C₂H₂), 2.48 (m, 1H, C₂H₂), 2.57 (m, 1H, C₂H₂), 2.74 (d, 1H, J₀ = 11.5 Hz, C₀H₂), 3.14 (m, 2H, C₂H₂), 3.20 (m, 1H, J₀ = 11.5 Hz, C₀H₂), 3.68 (m, 2H, J₀ = 5.2 Hz, C₂H₂), 3.79 (s, 2H, OCH₃), 3.88 (s, 3H, OCH₃), 3.91 (s, 3H, OCH₃), 4.10 (t, 2H, J₀ = 5.2 Hz, C₂H₂), 4.80 (m, 2H, C₂H₂), 5.03 (m, 1H, C₂H₂), 5.36 (m, 1H, C₂H₂), 5.71 (m, 1H, C₂H₂), 6.21 (d, 1H, J₀ = 15.9 Hz, C₃H₂), 6.37 (d, 1H, J₀ = 15.9 Hz, C₃H₂), 6.52 (br, 1H, C₂H₂), 6.80 (d, 1H, J₀ = 8.3 Hz, C₂H₂), 6.88 (d, 1H, J₀ = 8.3 Hz, C₂H₂), 6.95 (d, 1H, J₀ = 1.7 Hz, C₂H₂), 7.00 (dd, 1H, J₀ = 1.7 Hz, C₂H₂), 7.05 (d, 1H, J₀ = 1.9 Hz, C₂H₂), 7.12 (dd, 1H, J₀ = 8.3 Hz, C₂H₂), 7.57 (d, 1H, J₀ = 15.9 Hz, C₂H₂), 7.68 (d, 1H, J₀ = 15.9 Hz, C₂H₂);
¹³C NMR (125.4 MHz, CDCl₃): δ 26.20 (t, C₂), 28.36 (q, 3C, 3 Cₛ-C₃H₃), 33.60 (t, C₃), 33.72 (t, C₃), 33.81 (t, C₃), 38.76 (t, C₂), 39.74 (t, C₂), 55.77 (q, OCH₃), 55.95 (q, OCH₃), 56.04 (q, OCH₃), 64.69 (d, C₃), 65.77 (d, C₃), 68.04 (t, C₂), 73.75 (d, C₃), 76.83 (t, C₂), 78.77 (s, C₂), 79.31 (s, C₁), 93.89 (s, Cₛ-C₃Cl), 109.64 (d, C₁), 109.64 (d, C₂), 111.07 (d, C₂), 113.25 (d, C₂), 113.78 (d, C₂), 114.55 (d, C₁), 122.69 (d, C₁), 123.22 (d, C₁), 126.74 (s, C₃), 127.68 (s, C₃), 146.07 (d, C₂), 146.87 (s, C₂), 149.34 (s, C₂), 149.50 (s, C₂), 150.28 (s, C₂), 150.62 (s, C₂), 151.42 (s, C₂), 156.57 (s, C₂), 165.06 (s, C₂), 165.39 (s, C₂), 170.07 (s, C₂), 172.80 (s, C₂);
IR (cm⁻¹): 3583, 3379, 2973, 1808, 1714, 1513, 1242, 1139, 1034, 731;
ESI⁻-MS: 975 m/z (92, [M⁺Cl⁺Na⁺]), 977 m/z (100, [M⁺Cl⁺Na⁺]), 979 m/z (15, [M⁺Cl⁺Cl⁺Na⁺]).
Compound 27 (23 mg, 0.024 mmol) was stirred in 50% TFA/CH₂Cl₂ (10 mL) for 2 h at room temperature. The solvents were removed by vacuum evaporation; the solid residue was dissolved in CH₂Cl₂ and dried under vacuum five times, to obtain a yellowish solid (yield 100%).

ESI-MS: 843 m/z (100, [M⁺Cl⁺Na⁺]), 845 m/z (97, [M⁺Cl⁺Na⁺]), 847 m/z (28, [M⁺Cl⁺Cl⁺Na⁺]).

To complete the deprotection of the trichloroethoxycarbonyl group, the Boc-deprotected intermediate (20 mg, 0.025 mmol) was dissolved in THF (200 μL), then zinc powder (4.9 mg, 0.075 mmol) was added under stirring to the solution. Acetic acid (200 μL) was added and the mixture was left under stirring at room temperature for 3 hours. Afterwards the reaction was diluted with THF and filtered to remove the zinc powder aggregates. The solvents were evaporated under reduced pressure, then the amorphous solid was dissolved in ethyl acetate (10 mL) and the organic layer was washed with 2% NaHCO₃ solution (2 mL), water (twice, 2 mL at a time) and brine (2 mL). The organic phase was dried over Na₂SO₄ and the solvent was removed by vacuum evaporation. The final compound 28 was obtained as a pale yellowish layer (yield 80%).

¹H NMR (500 MHz, CDCl₃): δ 1.88 (m, 2H, C₁₀-H₂), 2.29 (m, 2H, C₁₀-H₂), 2.26 (m, 1H, C₇-H₅), 2.40 (m, 1H, C₇-H₅), 2.51 (m, 1H, C₆-H₆), 2.64 (m, 1H, C₆-H₆), 2.72 (m, 2H, C₁₂-H), 3.65 (m, 2H, C₁₀-H₂), 3.80 (s, 3H, OCH₃), 3.90 (s, 6H, 2 OCH₃), 4.09 (m, 2H, C₁₁-H₂), 5.01 (m, 1H, C₁₂-H), 5.34 (m, 1H, C₁₂-H), 5.65 (m, 1H, C₁-H), 6.21 (m, 1H, C₁₀-H), 6.34 (m, 1H, C₁₀-H), 6.46 (br, 1H, C₁₀-CH₃), 6.79 (m, 1H, C₁₀-H), 6.87 (m, 1H, C₁₀-H), 6.94 (m, 1H, C₁₀-H), 7.01 (m, 1H, C₁₂-H), 7.05 (m, 1H, C₁₁-H), 7.11 (m, 1H, C₁₂-H), 7.56 (m, 1H, C₁₀-H), 7.66 (m, 1H, C₁₂-H);

¹³C NMR (125.4 MHz, CDCl₃): δ 29.02 (t, C₁₀), 33.60 (t, C₁₀), 36.92 (t, C₁₀), 37.49 (t, C₁₀), 39.46 (t, C₁₀), 41.56 (t, C₁₀), 55.76 (q, OCH₃), 55.90 (q, OCH₃), 55.96 (q, OCH₃), 64.56 (d, C₁₀), 66.15 (d, C₁₀), 68.50 (t, C₁₁), 73.95 (s, C₁), 73.95 (s, C₁), 109.70 (d, C₁₀), 109.74 (d, C₁₀), 110.90 (d, C₁₁), 111.03 (d, C₁₀), 114.02 (d, C₁₀), 114.39 (d, C₁₁), 122.81 (d, C₁₂), 123.08 (d, C₁₂), 126.79 (s, C₁₁), 126.99 (s, C₁₁), 145.99 (d, C₁₀), 146.61 (s, C₁₀), 149.10 (s, C₁₁), 149.35 (s, C₁₁), 151.29 (s, C₁₁), 151.62 (s, C₁₁), 165.39 (s, C₁₁), 165.54 (s, C₁₁), 172.19 (s, C₁₁), 176.78 (s, C₁₁);

IR (cm⁻¹): 3462, 2938, 1796, 1714, 1630, 1597, 1513, 1464, 1261, 1138, 1065, 1022, 981, 734;

ESI⁺-MS: 669 m/z (100, [M+H⁺]).
4.4 Fluorescence assay

Caffeic acid, ferulic acid, 3,4-dimethoxycinnamic acid, chlorogenic acid hemihydrate, compounds 7, 10, 11, and 12 stock solutions (7 mM, 1.4 mM, 350 μM and 87.5 μM for 7) were prepared in DMSO.

Steady state fluorescence spectra were recorded at 25 °C (λ_{exc} 280 nm, λ_{em} 340 nm). The emission corresponding to λ_{exc} 280 nm was recorded in the λ_{em} range 300-400 nm. Synchronous fluorescence spectra (SFS) were measured by setting the excitation wavelength in the 240-320 nm range, and the emission was recorded at Δ = 60 nm in the 300-380 nm range. The slit width at the excitation was set to 10 nm, at the emission to 20 nm. The concentration of essentially fatty acid free HSA (A3782) solutions was 0.5 μM in 350 μL of solvent (135 μL of 10 mM Na₂HPO₄ and 2 mM KH₂PO₄ phosphate buffer diluted in 215 μL of mQ water, pH 7.4) for all measurements; the ligand concentration was gradually increased during the titration from 1 μM to 500 μM (points 0, 1, 5, 10, 20, 40, 60, 80, 100, 120, 140, 160, 180, 200, 250, 300, 350, 400, 450, 500 μM); for compound 7 the fluorescence quenching with a titration in a narrower concentration range (from 0.25 to 200 μM, points 0.25, 0.5, 0.75, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 40, 60, 80, 100, 120, 140, 160, 180, 200 μM) was also measured; with compound 12 a further titration from 0.25 to 200 μM halving the protein concentration to 0.25 μM, in the same solvent condition, was performed and in this case the slit widths at the excitation and at the emission were set to 20 nm. For HSA100 and clones, the titration was performed from 0.25 to 200 μM with 0.5 μM peptide concentration. After each addition of the ligand, the emission spectra, the fluorescence intensity, and the SFS were recorded. All the analyses were replicated three times, except with HSA100 and its clones, for which the analysis was single.

The data were treated using Excel™ (Microsoft Office™) and SigmaPlot™.

4.5 Biacore assay

For covalent ligand immobilization a CM5 chip was used. The entire protocol may be divided in three main steps: immobilization, binding and evaluation.

The immobilization involves an amide bond formation: the carboxyl groups on the surface of the sensor chip were first activated with a mixture of EDC and NHS to give reactive succinimide esters, which could react with the free amine group present on the side chain of the ligand compound 28, that was prepared as a 2 mM solution dissolving it in 10mM HCl with 5% in DMSO to improve the solubility. The running buffer used was a 10 mM Na₂HPO₄ and 2 mM KH₂PO₄ phosphate buffer, adjusted to pH 7.4 (the same used for fluorescence assay).

The binding protocol consists of 100 seconds for contact time and 300 seconds for dissociation time. To complete the dissociation a regeneration step was achieved with 30% CH₃CN in 1 mM NaOH solution for 15 seconds. The stabilization time was equal to 180 seconds. Biding events were studied at different essentially fatty acid free HSA (A3782) concentrations: 0 μM, 50 μM, 100 μM, 200 μM, and 400 μM. The HSA solutions buffer and the sunning buffer were the same used as running buffer.
Chapter 4 – Experimental section

Binding analysis was performed using BIAevaluation software, the affinity constant (K_D) was determined using numerical integration of steady state affinity model of the binding sensorgrams, that showed a good curve fitting. All the analysis were carried out at 25 °C.

4.6 NMR titrations*

All NMR measurements were performed at 25 °C on a Varian 500 MHz spectrometer. The data were treated using Excel™ (Microsoft Office™) and SigmaPlot™.

4.6.1 Titrations of 7, 10, 11 and 12 with caffeine in DMSO-d₆/D₂O 75-25%

Compounds 7, 10, 11 and 12 were dissolved in DMSO-d₆/D₂O 75-25% to give 1 mM solutions and 1 mL of them were put in the NMR tube. A 0.1 M solution of caffeine in DMSO-d₆/D₂O 75-25% was used to add 1 equivalent at a time (10 μL each addition) from 1 to 12 equivalents, then two solid aliquots of 5 mg of caffeine (26 mmol) were added to reach 38 and 64 equivalents in the sample. A proton spectrum with the quinides alone and after each caffeine addition was recorded using the same settings: lock on DMSO-d₆; range -1/11 ppm; 64 scan; reference on DMSO-d₆ (2.50 ppm).

4.6.2 2D ¹H, ¹H-NOESY spectrum recorded in DMSO-d₆/D₂O 75-25%

Two samples were prepared in DMSO-d₆/D₂O 75-25%, one equimolar in compound 10 and caffeine (5 mM for both molecules) and the second equimolar in compound 11 and caffeine (always 5 mM for both molecules). The two 2D ¹H, ¹H-NOESY spectra were recorded using the same settings: lock on DMSO-d₆; range -1/11 ppm; reference on DMSO-d₆ (2.50 ppm); 32 scans for t₁ increments; 400 t₁ increments, a relaxation delay of 1.5 s, a mixing time of 300 ms.

4.6.3 Titration of 7 with caffeine in D₂O /DMSO-d₆ 75-25%

Compounds 7 was solved in D₂O/DMSO-d₆ 75-25% to give a 1 mM solution and 1 mL of it was put in the NMR tube. A 0.1 M solution of caffeine in D₂O /DMSO-d₆ 75-25% was used to add 1 equivalent at a time (10 μL each addition) from 1 to 12 equivalents, then two solid aliquots of 5 mg of caffeine (26 mmol) were added to reach 38 and 64 equivalents in the sample. A proton spectrum with 7 alone and after each caffeine addition was recorded using the same settings: lock on D₂O; range -1/11 ppm; 64 scan; reference on D₂O (4.79 ppm).
Chapter 5

Conclusions
5. Conclusions

In this thesis work new data about the phenolic compounds present in coffee, particularly the lactones, have been reached.

A method for the synthesis of quinides, with caffeoyl and 3,4-dimethoxycinnamoyl side chains, has been defined and the full characterization of some of them allowed to set up a record of reference data, useful in coffee characterization, and to evaluate their biological activities against many viruses.

In the perspective of designing a selective biosensor for the quinides family, firstly the interaction between HSA and eight phenolic compounds (four acids and four synthesized quinides) has been studied by means of fluorescence spectroscopy: a good affinity has been revealed for all the tested ligands, reaching $K_D$ in the range 3-30 $\mu$M. The two diester quinides have been then used to measure, always by fluorescence spectroscopy, the binding ability of the HSA-derived fragment HSA100 and of four HSA100-mutants, randomly chosen: again, the measured $K_D$ values were in the micromolar range.

In order to exploit the generation of HSA-100 mutants libraries to obtain a selective biosensor for the quinides, a first SPR method for the quickly screening of the mutants has been set up, thanks to the synthesis and the immobilization on a gold chip of a proper linker-equipped quinide. The method has been tested measuring the binding with HSA, finding a 127 $\mu$M $K_D$.

Moreover, the behavior of caffeine in the presence of the four synthesized quinides has been studied by NMR titrations, showing that the interaction occurs even in 75-25% DMSO-$_d_6$/D$_2$O environment and this kind of binding could be too exploited to develop a selective biosensor.

These results suggest that following the attempted approaches it will be possible to develop the desired biosensor able to selectively recognize the quinides family.
Aknowledgments

Many thanks to:

Trieste Coffee Cluster for this great opportunity

Dr. Luciano Navarini, Dr. Silvia Colomban and all the members of illy-Aromalab for their kind hospitality and their teachings

Prof. Dr. Lieve Naesens and Annelies Stevaert (Rega Institute for Medical Research, KU Leuven, Belgium) for all the biological assays, the remarkable patience and the countless detailed answers

Prof. Daniele Sblattero and Dr. Immacolata Luisi (Department of Life Sciences, University of Trieste) for the samples of HSA100 and its mutants

Dr. Filomena Guida (Department of Life Sciences, University of Trieste) for her essential help with the HPLC and the kind availability to perform the SPR assays

Dr. Katarína Boronová and Nicola Cefarin for their substantial contribution to this thesis

Prof. Fabio Benedetti for his suggestions during the group meetings

Prof. Paolo Tecilla for the attempts about the stoichiometry of the complexes

Dr. Federico Berti and Dr. Cristina Forzato for all the advices, the corrections and the encouragements

All the people who have shared with me the life in the laboratories 330-333, particularly Valeria Laboragine, Elena Guercia, Elena Pellizzoni, Dr. Silvia Pavan, Dr. Adrian Oštríč and Dr. Giampaolo Fontanive

And last but not least my family, for the continuous support, and Alessio for his lovable obstinacy

Grazie!