STUDY OF INTERACTION BETWEEN POLYPHENOLIC COMPOUNDS AND PROTEIN USING COMPUTAIONAL AND CAPILLARY ELECTROPHORESIS TECHNIQUES

By

Myalowenkosi Innocent Sabela

Submitted in fulfilment of the requirements of the degree of Master of Technology: Chemistry in the Faculty of Applied Sciences at the Durban University of Technology

August 2012

Declaration

I Myalowenkosi Innocent Sabela declare that the thesis submitted for the degree of Master of Technology (MTech): Chemistry at the Durban University of Technology has not been submitted to any other university and no portion of this or any other closely related work is under consideration for publication elsewhere in any medium. All the work was done by the Author

Signature of the student

31/08/12

Date

Signature of the promoter

3/8/12

Date

Acknowledgements

I would like to extend my sincere gratitude to my promoter, Prof K Bisetty for his benevolent support and direction throughout my research. I would also like to thank our collaborational research team Lucia Asensi-Bernardia, Laura Escuder-Gilabert, Yolanda Martín-Biosca and particularly Prof Salvado Sagrado-Vives and Prof Maria J. Hernendaz from University of Valencia, Spain, for their remarkable input in making this research a success.

I am grateful of the support received from Dr Parvesh Singh for sharing his experience and remarkable support in the field of computational chemistry to make this project a success. I would also like to extend my acknowledgements to Njabulo J. Gumede who has been very compassionate and informative from the preliminary studies of this project up to the last day.

Finally, my deepest gratitude goes to my sisters, brothers and parents, for their benevolent support throughout my studies. This thesis is dedicated to them.

Abstract

The present work involves the interaction studies of chiral compounds with the Human Serum Albumin (HSA) protein using computational and experimental methods. The HSA protein has multiple binding sites that forms the basis for its exceptional ability to interact with many organic and inorganic molecules, which makes this protein an important regulator of intercellular fluxes and the pharmacokinetic behaviour of many drugs. This study was undertaken to evaluate the related pharmacokinetic and enantioselective binding parameters of the racemic catechin enantiomers with the HSA. Accordingly, this work involved a method development for the chiral separation of a racemic compound, by capillary electrophoresis-electrokinetic chromatography (CE-EKC) with a highly sulphated beta-cyclodextrin (HS-β-CD) as a chiral selector. The experimental work was supported by two molecular docking studies. The first included the mimicking of the host-guest interactions between a chiral selector and an enantiomeric compound. The second study included the estimation of the pseudo enantioselective (ES) binding of catechin to HSA.

Overall, it was found that CE-EKC is the preferred method for the(±)-catechin binding to HSA protein evaluation. Moreover, the technique used in this work is not restricted to HSA or polyphenols, but can also be applied to other proteins and ligands that possess chirality. Furthermore, the molecular docking approaches also proved to be very useful for the evaluation of chiral recognition systems and for elucidation of the ligand-protein interactions.

List of Contents

List of Tables

List of Figures

Acronyms and Symbols

ACEC	affinity capillary electrochomatography
AEKC	affinity electrokinetic chromatography
AA	Amino Acid
R	Avogadro's constant
BGE	background electrolyte
k'	capacity factor or retention factor
CE	capillary electrophoresis
CE-FA	capillary electrophoresis in frontal analysis mode
CGE	capillary gel electrophoresis
CZE	capillary zone electrophoresis
d	concentration of free drug
cmc	critical micelle concentration
CD	Cyclodextrin
E	electric field strength
EOF	electroosmotic flow
μ_{EOF}	electroosmotic flow mobility
v_{EOF}	electrophoretic flow velocity (electro-osmotic flow)
ee	Enantiomer excess/Enantiomeric excess
ES	Enantioselectivity
FF	Force field
GS	Grubbs' test statistic
HPAC	High Performance Affinity Chromatography
HPLC	High Performance Liquid Chromatography
HPLC-FA	High Performance Liquid Chromatography in the Frontal Analysis mode
HSA	Human Serum Albumin
HS-β-CD	highly sulphated β-cyclodextrin
ID	Identity of the analyte or D/P mixture
pI	isoelectric point

*L*d	length of capillary to the detector
LOD	limits of detection
LOQ	limits of quantification
MeOH	methanol
MEKC	micellar electrokinetic chromatography
MECC	micellar electrokinetic chromatography
*t*m	migration time
m	multiple binding
NSB	non-specific binding
NaN	not available number
PBS	Phosphate Buffer Saline
PLP	Piecewise Linear Potential
PMF	Potentials of Mean Force
rpm	revolutions per minute
RMSD	Root Mean Square Deviation
SDS	Sodium Dodecyl Sulphate
*L*t	total capillary length
D	total concentration of free drug
P	total concentration of protein
UV	ultra violet
v	velocity
η	viscosity
V	voltage

Aim

The main focus of this work was to study the interactions between a ligand and the HSA protein using computational and experimental approaches.

Objectives

- To develop a rapid analytical method to provide a baseline resolution for the two enantiomers of catechin.
- To use molecular modelling to elucidate the mechanism of the interaction between enantiomers and chiral selector.
- To develop a novel strategy relating the concentrations of protein, drug and protein-drug complex, to allow for the direct and indirect estimation of the pharmacokinetic parameters.
- To evaluate the number of primary binding sites in a protein with 'verification plots'.
- To perform robust statistics as part of the experimental design to avoid challenges due to the conversion of the protein-binding data analysis.
- To utilize molecular modeling approaches for elucidation of the ligand–protein interactions and stereoselective binding mechanism.

Poster Presentations

Diep in Die Berg Conference and Function Centre, Pretoria
ASSAf-DST-NRF Second Annual South African Young Scientists' Conference
26-27 September **2011**
Challenges for the robust assessment of stoichiometry and enantioselective binding of xenobiotics to human serum albumin mimicking physiological conditions. (±)-Catechin as example

Oral Presentations

Durban University of Technology
Institutional Research Day
21 November **2011**
Challenges for the robust assessment of stoichiometry and enantioselective binding of xenobiotics to human serum albumin mimicking physiological conditions. (±)-Catechin as example

Publications

Myalowenkosi I. Sabela, Njabulo J. Gumede, Laura Escuder-Gilabert, Yolanda Martín-Biosca, Krishna Bisetty, María-Jose Medina-Hernández, and Salvador Sagrado, Connecting simulated, bioanalytical and molecular docking data on the stereoselective binding of (±)-Catechin to human serum albumin, *Anal. Bioanal. Chem.*, 402 (**2012**)1899–1909

Myalowenkosi Innocent Sabela, Parvesh Singh, Njabulo Joyful Gumede, Krishna Bisetty and Sagrado Sagrado, Evaluation of Enantioresolution of (±)-Catechin using Electrokinetic Chromatography and Molecular Docking, *Journal of Scientific Research in Pharmacy,* 1(2), **(2012)** 1-4

CHAPTER 1

INTRODUCTION

Proteins play an important role in the binding of drugs. The characterization of the structure and energetics of molecular complexes is thus a key factor for the understanding of biological functions. However, accurate energy estimates for the ligand-protein binding presents a challenging problem and therefore most methods use various descriptors, such as interface planarity, shape, surface complementarity, propensity and residue compositions in an attempt to find combination of properties that would be reliable for identification of significant interfaces [1, 2]. Consequently, the prediction of the strength of non-covalent associations and structures of molecular complexes, has therefore becoming a primary objective in pharmacokinetic studies of computational and bioanalytical chemistry.

Computational chemistry offers biological simulation approaches that uses principles of computer science to assist in solving chemical problems related to pharmacokinetics. It is a well-known fact that apart from relatively recent results concerning the hydrogen molecular ion, the quantum n-body problem cannot be solved analytically especially in a closed form. However its results in most cases support the findings obtained by bioanalytical experiments, sometimes it is useful in prediction of unobserved chemical phenomena. The Capillary Electrophoresis (CE), on other hand, is an adequate analytical separation technique that has remarkable potential in a wide range of applications, ranging from micromolecules such inorganic ions, organic acids, amino acids, peptides, drugs, nucleosides, nucleotides, vitamins, steroids, and carbohydrates, to macromolecules, such as nucleic acids, hormones, proteins, etc. The use of CE to determine the physicochemical and biochemical properties of a variety of molecules makes this technique very attractive from other analytical techniques [3-6]. Separation in CE is based upon differences in mobility, which mainly depends on the charge to mass ratio of the molecule. The alternative ways of changing mobilities includes varying the pH of the electrolyte/buffer or adding a suitable complexing agent to the buffer. Usually, the

CE analyses are relatively fast, requires little samples and reagents, and available at a cheaper cost than chromatography or conventional electrophoresis.

1.1 Pharmacokinetics

Pharmacokinetics is the study of the progression of a drug in relation to time within the body and includes the biological processes of Absorption, Distribution, Metabolism and Excretion (ADME) as shown below.

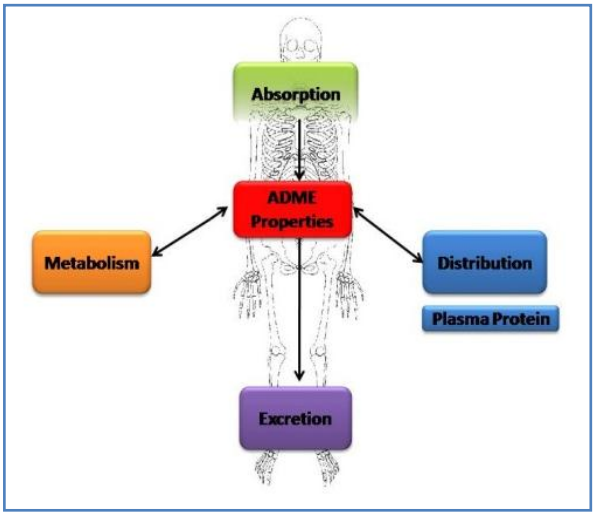


Figure 1.1: Influence of drug levels using ADME criteria

The measurement of drug concentrations in blood or plasma is used to derive the pharmacokinetic parameters. These parameters are very important for the dose regimen and dose size [7]. A number of drugs are administered orally for convenience and compliance reasons. It is well understood that a certain percentage of the dose know as bioavailability reaches the blood circulation system, reason being after the drug is administered it dissolves in the gastro-intestinal tract; part of it is get absorbed through the gut walls as it passes into the liver to enter the blood circulation [8]. From there, the drug will be distributed to various tissues and organs in the body to a certain degree depending on the structural and physicochemical properties of the compound. Finally, the drug will bind to its molecular target, where it will perform desired action [8, 9].

1.2 Enantiomers

Enantiomers are two chemically identical molecular species which differ from each other as non-superimposable mirror images. The most simple and vivid model for enantiomeric structures are the left and right hands. Enantiomers, in addition to diastereomers are thus a special case of stereoisomers [10]. The stereochemistry of such compounds cannot be ignored because of their significant effect of the biological activity of the compounds. Thus the separation of chiral compounds is of key importance in different areas of research, especially in the pharmaceutical field. After separation optical purity testing is performed to assure their safety and effectiveness. Enantiomers do not differ in their physicochemical properties (except for a few chiroptical properties) in an achiral medium. This means that both enantiomers of any chiral compound have exactly the same mobility in free solution. Enantioselective chromatography and capillary electrophoresis are widely used in the determination of enantiomeric compositions such as enantiomeric excess or optical purity of chiral compounds [10, 11]. The tremendous successful development of the CE arise from it common advantages such as impressive peak efficiency, rapidity of analysis, low consumption of sample and reagents and the ease of operation especially when changing chiral selectors. At present separations which are attributed to capillary zone electrophoresis (CZE) are conceptually based on different mobilities of analytes not in free solution, but in free buffers containing some additive. Enantioseparations using neutral non-micellar chiral selectors belongs to this type. Currently, cyclodextrins (CDs) remains the most popular resolving agents (chiral selector) used in CE analysis [12, 13]. The applications of computational chemistry in the area of cyclodextrins have not been fully explored thus far, despite its extensive use in the separation chemistry [14]. In terms of computational studies where, for example, geometry optimizations have been carried out, structure and energy are inextricably linked. In all cases of chiral resolution transient complexes are formed between one or both of the enantiomers and the chiral additive, and differences in the free energy of formation of these complexes leads to enantiomeric separation [15]. Ideally, the enantiomers have different affinities to form these associates but if they migrate with the same mobility, the enantioseparations will be on the difference in their binding constants.

In this study a molecular docking technique was used to gain insight into the selector-enantiomer interaction energy and to support information for the enantiomeric separations observed experimentally using CE.

1.3 Enantioselective Protein Binding

Stereoselectivity in protein binding can have a significant effect on the bioactivity of chiral molecules interacting with the human body [3]. In particular, human serum albumin (HSA), the most abundant protein in the circulatory system exhibits the highest potential enantioselectivity among the plasmatic proteins, thus is a key piece for the pharmacokinetic characterization of chiral xenobiotics. Protein-binding studies cannot be undertaken without the use of analytical techniques. The most important analytical approaches to characterize drug–protein binding have been recently reviewed, including a critical evaluation of their strengths and weaknesses [3, 16]. In general, plasma protein binding is a reversible interaction process. Most studies devoted to the evaluation of stereoselective binding of chiral molecules to plasma proteins are conducted in three steps: (i) incubation process (equilibration of xenobiotic and protein mixtures according to a given experimental design); (ii) analytical process which can involve two phases, the separation of free xenobiotic (unbound fraction) from xenobiotic–protein complex (bound fraction) and the determination of the enantiomer concentration in one of the fractions (unbound or bound) by means of an achiral or chiral separation technique, depending on the use of individual enantiomers or racemic compounds, respectively; (iii) parameter estimation (e.g. protein binding (percentage), *PB*, number of binding sites or stoichiometry, affinity constants and enantioselectivity to the protein) following a given mathematical strategy [3, 16]. However, the impact of the experimental design and the mathematical approaches on the results have been underestimated in the past, entailing a risk in the lack of quality results [4, 17]. For instance, stoichiometry values lower or higher than unity can be considered abnormal for the main active site under physiological conditions [4]. Such results are limited, incomplete and obtained at non-physiological conditions therefore provides no evidence of enantioselective binding to HSA. It is unfortunate that such results are reported in the literature

without any verification, justification or commentary [18-23]. Accordingly, the main focus of this work is on the enantioselectivity of (±)-catechin to HSA. Other important goals are the verification of the stoichiometry of the complexes, and their affinity constants in mimicking physiological conditions, using a consistent experimental design and mathematical approaches linked to a simulation study [4]. The univariate statistical approach used for estimating affinity constants (and enantioselectivity), instead of the traditional regression strategy, introduces clear statistical advantages (i.e. outlier checking, hypothesis testing and robust statistics and realistic uncertainty). A challenge of this work to improve the general quality scheme of these studies, thus, all drawbacks identified in previous estimations have been detected and avoided. The present work is mainly intended to influence traditional schemes in order to improve the experimental and mathematical processes used in protein binding studies, in view of the accurate future work in this area of research.

1.4 Chiral Selectors

Since the enantiomers of an analyte exhibit the same charge density, their electrophoretic mobilities are the same and they cannot be separated with a simple CZE technique. Therefore, it is important to add a chiral selector which is able to interact with the analyte in the presence of the background electrolyte. Hence making the choice of a suitable chiral selector a crucial step for a complete chiral resolution in CE. There are two different separation mechanisms that are possible [24]. The enantioselective interaction results in chiral selector either transforming the enantiomers at a different rate into new chemical entities (kinetic enantioselectivity) or forming labile complexes of different stability with each enantiomer (thermodynamic enantioselectivity) [10]. Since enantiomers can interact differently with a chiral additive this results in different effective mobilities of enantiomers. Thus, in the case of neutral chiral buffer additives, the resolution is controlled not by the different mobilities of enantiomers in free solution, but by their different interactions with a chiral buffer additive or by different mobilities of the transient diasteriomeric complexes of each enantiomer with the chiral additive [25]. The mechanism described depends on the experimental conditions such as the concentration of the chiral selector and concentration of the buffer. However, related conditions often overlay each other. Chiral selectors can either be an suitable chiral molecule or a chiral surface.

Cyclodextrins and their derivatives are the most commonly used chiral selectors that have been used in capillary electrophoresis for the enantioseparation of different classes of compounds. However, in the past decade, researchers have shown interest in performing protein based chiral separations. The method is called affinity electrokinetic chromatography (AEKC), or affinity capillary electrophoresis (ACE). The principal is mainly about adding adequate amount (concentration or volume) of the protein the running buffer and optimizing the conditions for optimum resolution [26, 27]. This method can be used for the separation of chiral therapeutic substances provided that both enantiomers bind with the protein to different extents. In this technique, the mechanism of separation is based on the interaction of the analytes with the protein (within the running buffer) resulting in a change in the net mobility

of the analytes due to analyte and analyte-protein complex having significantly different mobilities. Another approach to enantioseparations by capillary electrophoresis is to use immobilized protein as chiral selectors. The sufficient applied voltage applied allows the buffer and analyte migration through the capillary, while the enantiomers interact with the immobilized protein. This technique is slight similar to that used in liquid chromatography, therefore it is generally termed affinity capillary electrochromatography (ACEC) [28].

1.4.1 Cyclodextrins

Cyclodextrins (CD's) are by far the most popular chiral selectors used in CE and will therefore be discussed in details in this section. Cyclodextrins have many uses, one of which is a medium for chemical separations, especially chiral separations [29, 30]. For many years separation scientists have been aware of the ability of cyclodextrins to selectively bind guest molecules, and many uses of cyclodextrins for separations now exist. Because these molecules are inexpensive, easy to modify, and come in varying sizes, they have been adopted for use in planar, gas, liquid, and super- and subcritical fluid chromatography, as well as in membrane separators. The CD's shown in Figure 1.2 are torus shaped cyclic D-glucooligosacharides produced starch by enzymatic degradation [31]

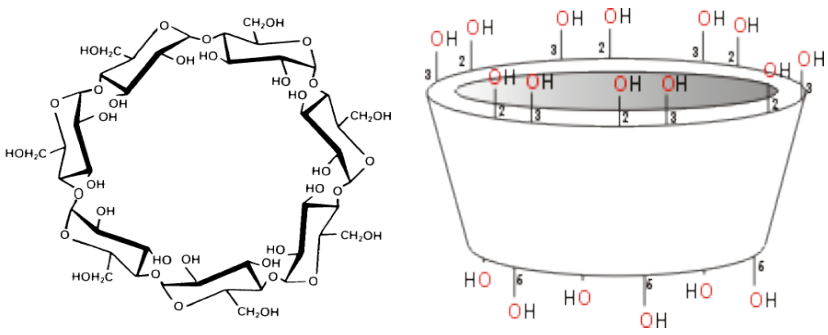


Figure 1.2: Structure of β cyclodextrin with seven glucose units

Special mention should be made of the CD's cyclic oligosaccharides having an external hydrophilic surface and a hydrophobic cavity. This cavity plays a crucial

role is separation because it include other compounds by hydrophobic interactions. This mechanism is also regarded as sterically selective because analyte/s are required to fit the size of the cavity, of which the diameter depends on the number of glucose units in the CD structure (6, 7, 8, forα, β and γ-CDs, respectively) [32]. The CD's contains chiral hydroxyls groups in the glucose molecules, making the inclusion complex formation chirally selective [13]. The hydroxyl groups present on the rim of the CDs can easily be modified by substitution reactions to obtain CD derivatives of interests. The performance of the CD or it derivative depends on several parameters such as separation medium, nature of the enantiomers, pH etc. the most commonly used CD derivatives in CE for chiral analysis includes, the uncharged methylated-, hydroxyl-ethylated-, hydroxyl-propylated-, acetylated-CDs and the charged ones such as methylamino-, carboxymethylated-, sulphated-, phosphated-CDs, etc [11]. In this work, Highly Sulphated- β-Cyclodextrin (HS-β-CD) was used as chiral selector, as it is more widely used in the enantiomeric separation of a variety of analytes.

1.4.1.1 Structural Features of HS-β-CD

Structural features of cyclodextrins have been addressed using data derived mostly from X-ray crystallographic analysis of solid-state structures and from NMR studies of cyclodextrins and their complexes in solution [14]. In the structure of HS-β-CD in Figure 1.3, the hydroxyl groups present on the rim of the CDs are substituted with sulphate groups. Most computational studies of cyclodextrins involve host-guest complexes and are classified based on the energetics of binding and structural analyses of inclusion complexes [14].

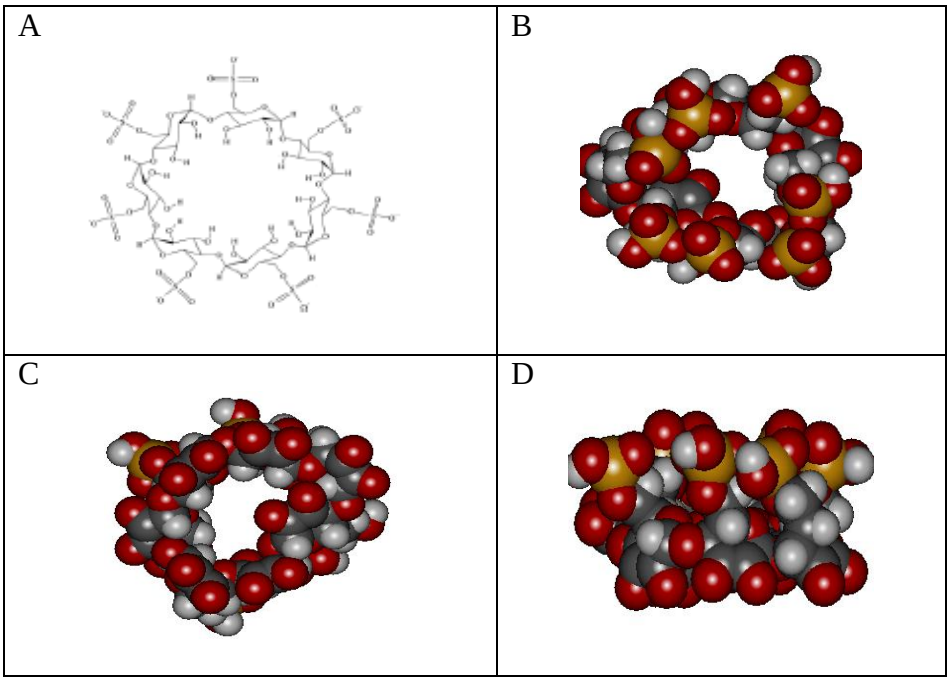


Figure 1.3: Structures of highly sulphated-β-cyclodextrin: A: 2D image, B: top, C: bottom and D: side view.

1.4.2 Proteins as chiral selectors

Proteins can be used as additives to the running buffers (affinity electrokinetic chromatography (AEKC)) or as immobilized selectors (affinity capillary electrochromatography (ACEC)).Proteins have been found to be very useful in the separation of chiral compounds due to the broad range of their enantioselectivity which arise from multiple interactions with chiral compounds. In addition, the enantiomeric drug separation with protein can provide useful information on the pharmacokinetics differentiation of the drug enantiomers [32, 33]. The most commonly used protein is human serum albumin (HSA), the most abundant protein in the circulatory system (i.e. with the largest complexation potential), exhibiting the highest potential enantioselectivity among the plasmatic proteins, thus a key piece for pharmacokinetic evaluation of chiral xenobiotics. The major limitation of using proteins as selectors is that they are very expensive, easily degradable and they possess strong absorptivity at the UV region, hence, detection limits of the analyte becomes higher.

1.5 Macromolecular Target/Proteins

Proteins are macromolecules consisting of different numbers and sequences of amino acids, which allow them to adopt different 3D structures and possesses unique biological functions. An important convention needed for understanding much of the information available for a particular protein is the designation of the individual atoms and structural elements of a protein. The backbone of a protein consists of a repeated sequence of three atoms, belonging to one amino acid residue, generally represented as N_i, C_i^{α}, and $C_i^{'}$, where *i* the number of the residue starting from the amino end of the chain, as shown in Figure 1.4 below.

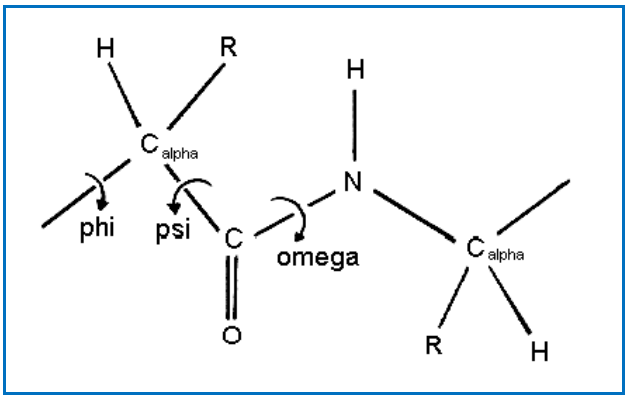


Figure 1.4: Typical backbone of a protein

Generally, in proteins and polypeptides monomers of the amino acids are joined by peptide bonds, (-CO-NH). The peptide bond is almost planar and, in most cases, assumes a *trans* conformation, with the exception of proline, which has a high probability of assuming a *cis* conformation [34, 35]. The 3D construction of proteins is stabilized considerably by hydrogen bonds. Due to their weakness, however, they can undergo prominent changes during complexation, conformational change, and folding [36]. "Proteins can be divided into three main classes, which correlate with typical tertiary structures: globular proteins, fibrous proteins, and membrane proteins. Almost all globular proteins are soluble and many are enzymes. Fibrous proteins are often structural, such as collagen, the major component of connective tissue, or keratin, the protein component of hair and nails. Membrane proteins often serve as receptors or provide channels for polar or charged molecules to pass through the cell membrane" [37]. The significance of protein can be compared to other biological

macromolecules such as polysaccharides and nucleic acids, they are critical parts of the organisms and play an effective role in every process within cells. "The protein is characterized by its remarkable ability to bind a broad range of hydrophobic small molecule ligands including fatty acids, bilirubin, thyroxine, bile acids and steroids; it serves as a solubilizer and a transporter for these compounds and, in some cases, provides important buffering of the free concentration" [38] The most important biological functions of protein includes ligand transport, cellular signal generation, biochemical reaction catalysis, and structural support. Many proteins are classified as enzymes that catalyzes biochemical reactions and play a crucial role during metabolism. Proteins need to be present in animals' diets, a source of amino acids that cannot be synthesized biologically by the animals [39, 40]. Since the main focus of the work is based on HSA, hence a detailed discussion of this protein is essential.

1.5.1 Human Serum Albumin (HSA)

"Albumin is synthesized in the liver as prepro-albumin which has an N-terminal peptide that is removed before the nascent protein is released from the rough endoplasmic reticulum. The product, pro-albumin, is in turn cleaved in the Golgi vesicles to produce the secreted albumin" [41].

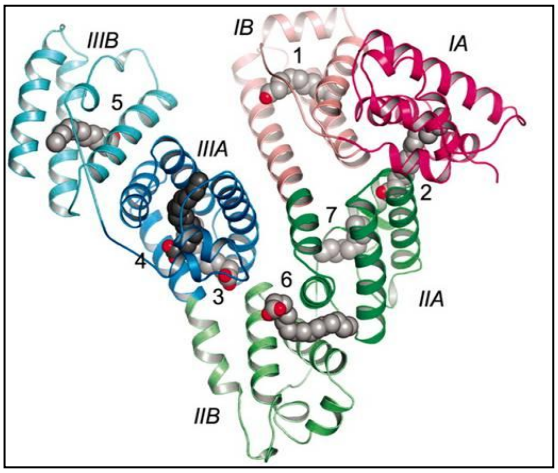


Figure 1.5: Crystallographic structure of Human Serum Albumin (HSA) complexed with fatty acids at different subdomains [41].

Human Serum Albumin (HSA) is a principal extracellular protein with a high concentrationin blood plasma (HSA; 35–50 mg/ml in plasma; MW: 66 500). "HSA is a globular protein composed of three structurally similar domains (I, II and III), each containing two subdomains (A and B) and stabilized by 17 disulphide bridges. Despite the internal structural symmetry, the three domains have different capacities for binding fatty acids, thyroxine and drugs" [38, 42]. The availability of multiple binding sites on HSA is the main cause of it remarkable ability to interact with many organic and inorganic molecules and make this protein an important regulator of intercellular fluxes, as well as the pharmacokinetic behaviour of many drugs [42, 43]. Hence the evaluation of micromolecules with respect to albumin binding is impressive and of fundamental importance especially in the field of pharmaceutical [44]. "The most important sites of HSA are I and II, also known as warfarin and benzodiazepine binding sites respectively" [43, 45]. HSA also binds covalently or reversibly with a wide variety of drugs in two primary sites which overlaps with the binding locations of endogenous and exogenous ligands [38]. HSA interacts reversibly with broad spectrum of therapeutic agents and they are often present at the primary sites, several sites of lower or of very low affinity. The *in-vivo* concentration of administered drugs/ligands is almost always much lower than that of HSA (about 0.6 mM).

1.5.1.1 Binding Sites and Residues of Human Serum Albumin HSA

Generally, proteins are made up of a larger number of residues bonded in a distinct conformation. Because they are arranged non-uniformly, pockets are found within them. "The X-ray diffraction studies on HSA have shown that this protein has three homologous domains (I–III), and that each of them is composed of two subdomains (A and B). The subdomain A and B contains six and four α-helices, respectively, which are connected by flexible loops" [45]. There are several binding sites (formed as pockets) which can be found within these subdomains and they are characterized as follows:

- **Site I**: Is formed as a pocket in subdomain IIA and involves the lone tryptophan residue of the protein (Trp214), (Figure 1.5) shows the crystal view of the residues). "Site I has been described as a large hydrophobic cavity present in subdomain IIA. The inside wall of the pocket is formed by hydrophobic side chains, whereas the entrance to the pocket is surrounded by positively charged residues. Lys 199, His 242, and Arg 257 function sterically and electrostatically in ligand binding" [45]. A large population of drugs bind in this site includes Non-Steroidal Anti-Inflammatory Drugs (NSAIDs), phenylbutazone, indomethacin, many sulfonamides and even many anti-epileptic drugs like phenytoin etc. This site is also called Warfarin binding site or as Azapropazone binding site. It is a relatively large binding site but also flexible, thus a large number of individual ligand binds independently of each other but sometimes compete for the site [45].

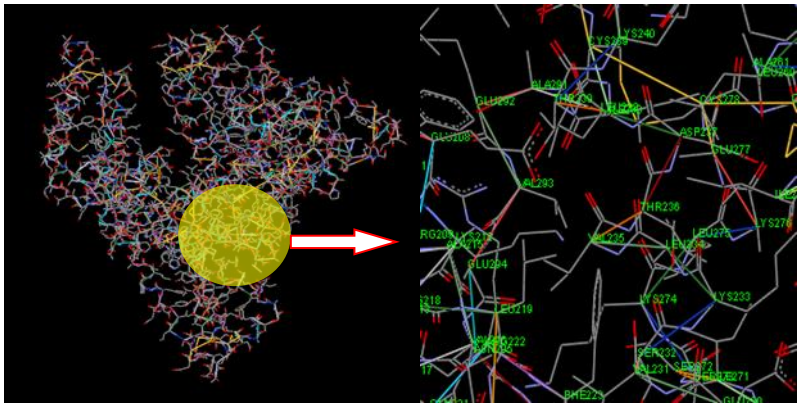


Figure 1.6: Close view of residues available in subdomain IIA. These are the residues that participate in binding of the ligand to this particular site.

- **Site II**: Has been known to be located in subdomain IIIA which contains a hydrophobic, branched, T-shaped cavity. The cavity has two hydrophillic entrances that are separated by about 10 Å, each of which contains residues that interacts with the carboxylate group of the bound ligand. "The binding pocket has two spots of hydrogen bond partners. "The first spot is located across the II–III domain interface provided by Ser 342 and Arg 348 from IIB and by Arg 485 from within IIIA (refer to Figure 1.5). The other spot is

provided by Arg 410, Tyr 411, and Ser 489, all from subdomain IIIA" [46]. Site II is smaller or narrower compare to site I, because no large ligands apparently bind to it. It flexibility is also less, so the binding mode that takes place in this site is often affected by stereoselectivity of both residues and ligand [38, 45, 46]. It is commonly known as Diazepam binding site. Benzodiazepines, medium chain fatty acids, ibuprofen, ketoprofen, etc. bind extensively at the very site. This is due to structural changes of these drugs having high and specific affinity for the sites. This site appears to be smaller, or more narrow, than site I, because no large ligands apparently bind to it [45]. At both the sites I & II many drugs are known to bind. "Aromatic and heterocyclic ligands normally bind within two hydrophobic pockets in subdomains IIA and IIIA" [42].

- **Site III**: This very protein site is called as Digitoxin binding site [42, 45].

- **Site IV:** This is referred as Tamoxifen binding site. At the sites III & IV very few drugs are known to bind. For more information refer to ref. 37.

1.5.2 α_1-Acid Glycoprotein (α_1AGP)

The acid glycoprotein exists with a 44,000 molecular weight and comprises of 0.04 to 0.1 g/g % of the total plasma concentration of proteins called the Oromucoid, as it mainly binds to basic drugs like Imipramine, Desipramine, Lidocaine, Quinidine, etc

1.5.3 Lipoproteins

Lipoproteins are molecular aggregates consisting of various types of lipids and apolipoproteins. They are divided into several subclasses according to density. High-density lipoprotein and low density lipoprotein are considered to be the most important due to their higher plasma concentrations [47].

1.5.4 Plasma

Plasma protein binding plays an important role in the whole-body disposition of drugs/ligands. Pharmacokinetic properties, such as hepatic metabolism rate, renal excretion rate, membrane transport rate and distribution volume, are functions of the ratio of free fraction. From a pharmacodynamic point of view, knowing the ratio of free fraction of drugs in plasma is extremely important because unbound drugs in plasma can easily reach the target organ, whereas bound drugs pass the blood capillary wall with great difficulty. Therefore precise information on the free drug based on the plasma fraction is essential for drug development and for determining the safety of a drug in clinical trials [48].

1.6 Ligands

1.6.1 Polyphenols

Flavonoids are groups of more than 4000 polyphenolic compounds that occur naturally in foods of plant origin. They are benzo-γ-pyronederivatives consisting of rings that are phenolic (ring A and B) and pyrane (ring C). They are categorized into a variety of subclasses according to the saturation level and opening of the central pyrene ring, mainly into flavones, flavanols, isoflavones, flavonols, flavanones, and flavanonols. The basic structure of flavonoids with the flavanol nucleus is shown in Figure 1.7 below:

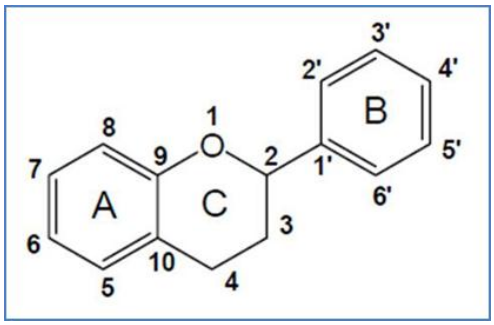


Figure 1.7: Basic structure of flavonoids containing 15 carbon atoms derived from a C6-C3-C6 skeleton

They are present in dietary plants, fruits and vegetables. Therefore, they are consumed in considerable amounts and are also stable at physiological temperature [49]. They form part of normal diet, and bind to biomolecules and biomembranes

[50]. Additionally, flavonoids are found in several medicinal plants, and herbal remedies containing flavonoids have been used in traditional medicines around the world [49, 51, 52]. Oxidation reactions can produce free radicals, which start chain reactions that damage biological cells and results in human diseases such as cancer, and cardiovascular diseases, and some pathological disorders of gastric and duodenal ulcers, allergies, vascular fragility, and viral and bacterial infections [53-57]. Flavanols are capable of terminating these chain reactions by reacting with the free radicals being oxidized themselves, thus protecting the biological cells [53]. Other important biological functions include antiallergic, anti-inflammatory, antimutagenic, anticarcinogenic, and modulation of enzymatic activities [49, 55, 56]. However, their *in vivo* activity still remains uncertain, and questions concerning their absorption, metabolism and bioavailability remains unanswered. Based on the current understanding of ADM, it is suggested that factors such as protein binding may ruin polyphenol absorption and bioavailability and even lower their ability to scavenge free radicals [57-59]. As the interest in flavonoids multiplies in the pharmaceutical and biomedical sciences it certainly leads to a need for the development of analytical methods for their enantioseparations [60]. Moreover, the interactions between polyphenols and HSA have a major biological importance and can be used as a model for polyphenol–protein interaction. For this purpose, we investigated the interaction of catechin with human serum albumin using computational and experimental approaches.

1.6.2 Catechin

Catechin has two chiral centres (carbons 2 and 3) and has therefore four diastereomers; those isomers in *trans* configuration are called (+)-catechin (2R-3S) and (-)-catechin (2S-3R), abbreviated in this work as (+)-Catechin and (-)-Catechin, respectively. Structures are shown in Figure 1.8 below.

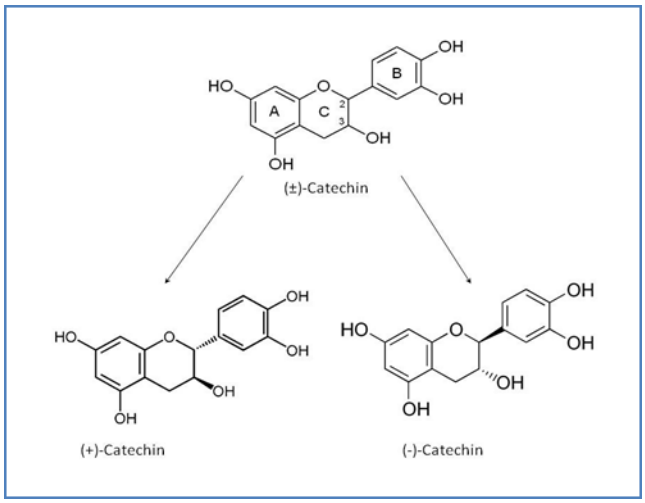


Figure 1.8: Structure of catechin and it two *trans* diastereomers

Some differential biological activity between enantiomers of (±)-catechin has been reported. (+)-catechin and (-)-catechin seem to have stereospecific opposite effects on glycogen metabolism in isolated rat hepatocytes [61]. Bais et al. reported that (-)-catechin, but not (+)-catechin, had allelochemical activity and that only (+)-catechin showed antibacterial activity [62]. Several inhibitory effects of (+)-catechin have been observed, for instance, on intestinal tumor formation in mice [63] oxidation of low density lipoproteins [64], histidine decarboxylase (i.e. inhibiting the conversion of histidine to histamine) [65] or monoamine oxidase of the type MAO-B, which could be used as part of the treatment of Parkinson's and Alzheimer's patients [66]. Incubation experiments with (+)-catechin show a prevention of human plasma oxidation [67]. On the other hand, it has been reported that (-)-catechin suppresses the expression of Kruppel-like factor 7 and increases the expression and secretion of adiponectin protein in 3T3-L1 cells [68]. All these observations reveal the interest in expanding the studies of bio-differentiation of these enantiomers.

1.7 Ligand-Protein Interactions

Protein-ligand interactions are thought to be specific, which means that a given ligand is likely to interact only with particular protein types and in particular regions of the protein surface. The data for the interacting protein and conformational

details of it residues is very useful in understanding biochemical processes. This information is very crucial in drug discovery and design. The features such as protein-protein binding and ligand-protein binding in particular manner are the main reasons for importance of biological processes and the energetics often provides the most important and useful link between structure and function of biomolecular systems [44, 69]. Drug–protein interactions are determining factors in the therapeutic, pharmacodynamic and toxicological drug properties. Plasma proteins, such as HSA or α_1-Acid Glycoprotein, act as carrier proteins for numerous endogenous and exogenous compounds. Furthermore, the proteins have an ability to replace the bioavailable or free drug that is removed by metabolic or elimination processes by dissociation of drug–protein complexes [70, 71]. This process is highly dependent on the strength of binding. Strong binding can also have a significant impact on the pharmacokinetics of drugs such as delaying *in vivo* half-life of the therapeutic agent. On the other hand, it can prevents drug/molecule release in tissues, resulting in HSA binding information being a key feature of its ADME properties [72]. The other important possible interaction is the co-binding of two drugs or displacement of one drug by another which may alter the therapeutic drug level and can lead to serious health conditions. Drug distribution within the body is determined mainly by the bioavailability or free (unbound) concentration of drug in circulating plasma. The unbound fraction, in turn, depends on the kinetics of the equilibrium between the bound and unbound form and mostly on drug absorption by plasma proteins [73].

Most drugs/ligands bind to plasma proteins; to a varying degree, these bonds are reversible. Biological macromolecules have a potential to differentiate between enantiomeric substrates in several approaches. There is a probability that the binding favors one form of enantiomer over the other to a varying degree and depending on the nature of the protein; stereoisomers may also have opposite effects such as one form being an agonist while it isomer is an antagonist at the same receptor. A bound drug always remains confined to a specific tissue or to a particular protein site with which it possess a greater affinity. In general, the drug-protein complex has weak chemical bonds such as ionic bonds, hydrophobic bonds, hydrogen bonds or van der Waals forces, therefore the binding process are defined as a reversible. For a better understanding of pharmacokinetics and pharmacodynamics, it is essential to use a

sensitive and simple method to quantify the concentration of unbound ligands in human plasma, and then determine its plasma protein binding [48]. For methods that use the change in peak area or plateau height for the determination of the concentrations of unbound and bound ligand requires special care because the concentration of bound ligand is calculated on the bases of difference between the total and unbound concentrations of the ligand. This method assumes that there is no loss of the ligand on the capillary or at stage during of the analysis [74].

1.7.1 Types of Binding

1.7.1.1 Competitive in Binding

This type of binding is very familiar in the interactions of racemic compounds with proteins where the two ligands are capable of binding to the same sites. It is important to consider the behaviour of each individual form of the ligand since the binding to protein can be stereoselective in nature, depending on the chemical properties of the protein [75]. Displacement interactions are predominant ones among these reactions. If we consider ligand A bound to a specific site on the target molecule or protein and if a second compound, ligand B, is administered having a greater affinity towards the bound site it would effectively displace ligand A. This phenomenon is said to be displacement reaction. In a case where two or more drugs have same or identical affinity for a same site then the phenomenon is said to be competitive because they compete with one another to bind at the same site [76, 77]. The ligand/drug which has been removed from its binding site is said to be displaced ligand while the one that does the displacement is called as displacer [59]. Protein-polyphenol association is a well-known phenomenon. However, considerable information has recently been obtained in the area of how the structure of either the protein or the polyphenol may affect the interaction [78].

1.7.1.2 Non-Specific Binding (NSB)

The most common problem to deal with in receptor-ligand interactions is non-specific binding (NSB). In general, the NSB model requires no assumption and can

account for complicated systems. Binding to the receptor of interest is called specific binding, the part of a ligand used in a competitive-binding assay that is found in the bound fraction, independent of the binding reaction has a potential to bind in the other sites, the phenomenon is called nonspecific binding [79].

.

CHAPTER 2

MOLECULAR MODELLING AND DOCKING STUDIES

Molecular modelling includes approaches such as protein modeling, which uses quantum mechanical methods to assess the potential for interaction between the micro and macromolecules that participate in ADME processes. The latest improvements in the calculation of free energies have smoothed the curve for the understanding of ligand-protein and host-guest interactions [80]. Most commonly used methodologies for evaluation of protein-ligand are molecular docking approaches which are based on molecular dynamics, robotics, or rotamer libraries [81]. There is still a need for a better understanding of chiral mechanism in chromatographic, electrophoretic and other separations techniques, thus adding demand for computational modelling of enantioseparations. In this work, molecular docking was used to study the bound conformations of catechin to HSA proteins receptor and to elucidate the enantioresolution mechanism in capillary electrophoresis [8].

2.1 Docking

The applications of computational chemistry techniques to study structural and functional properties of biological macromolecules has increased dramatically due to rapid advances in computer power, improvements in force fields, and development of numerical algorithms [82]. An important approach in structure-based rational drug design is the virtual fitting of a ligand into a macromolecular target at the binding site and estimating its binding constant, a process known as molecular docking. Several software packages have been developed, such as FRED, GOLD [83] GLIDE [82, 83, 84] LigandFit [83, 85] Surflex [83, 86] AutoDock [87, 88] and DOCK [86] that addresses the task in a similar fashion by fitting a flexible or rigid molecule in a rigid protein binding site. However, there is still a room for improvement regarding estimations of binding affinities [89] especially on the robustness of the methods with regard to both structural and energetic predictions. It is understood from all of the

above programs that both false negatives (poor docking of active compounds) and false positives (over-predicted binding affinity) can be obtained [90].

The theoretical complexity and accuracy vary greatly, ranging from simple statistical multivariate equations to computationally intensive free energy perturbation methods [82]. Docking algorithms consist of a search algorithm and a scoring function. The search algorithm searches through the rotational and translational degrees of freedom of the ligand in the protein binding site. Dihedral angles of the ligand are also searched, while the receptor is generally considered to be rigid. The scoring function assesses the complementarity of the different ligand poses that are generated and rank ordered. Validation studies of docking algorithms mostly focus on binding mode predictions or on virtual screening enrichment. In binding mode validation studies, known protein–ligand complex structures are used to assess the ability of the algorithm to reproduce the experimentally observed binding mode of the ligand. The accuracy of the docked binding poses is assessed by calculating the Root Mean Square Deviation (RMSD) of the pose with respect to the crystallographically observed binding mode [91]. The scoring functions that attempt such computational predictions are essential for analyzing the outputs of molecular docking, therefore important for drug discovery, chemical biology and structural biology. Each scoring function assumes a predetermined theory-inspired functional form for the relationship between the variables that characterize the complex, which also include parameters fitted to experimental or simulation data and its predicted binding affinity. The inherent problem of this rigid approach is that it leads to poor predictions for those complexes that do not conform to the modeling assumptions [80]. However, the main reason for keeping the protein rigid is to reduce the time required for the computations. An alternative approach of assessing the ligand-protein interaction is to use Pharmacophore models, when there are no structural details on the protein or target molecule [8].

2.2 Scoring Functions

A ligand scoring is a method used to for quick estimation of the binding affinity of a ligand, it is highly dependent on the conformation of the docked ligand into a target receptor structure. The scoring function is a key element of a protein-ligand docking algorithm, because it directly determines the accuracy of the algorithm. Speed and accuracy are the two important aspects of a scoring function. In general, a perfect scoring function is defined as the one that would be both computationally efficient and reliable. Numerous scoring functions have been developed in the past decades and can be grouped into three basic categories according to their methods of derivation: force field, empirical, and knowledge-based scoring functions [82, 92, 93].

2.3 Molecular Mechanics Method

2.3.1 Force Field Scoring Function

Force field (FF) scoring functions are developed based on physical atomic interactions, including van der Waals (VDW) interactions, electrostatic interactions, and bond stretching/bending/torsional forces [82, 92].

2.3.2 Empirical Scoring Function

The second kind of scoring functions are empirical scoring functions, which estimate the binding affinity of a complex on the basis of a set of weighted energy terms

$$\Delta G = \sum_i W_i \Delta G_i \qquad (2.1)$$

Where ΔG_i represents different energy and W_i is the corresponding coefficients W_i determined by fitting the binding affinity data of a training set of protein–ligand complexes with known three-dimensional structures. Compared to the force field scoring functions, the empirical scoring functions are much faster in binding score calculations due to their simple energy terms [92]. Empirical does not have a correct setting for all instances but instead is set to give optimal results for a certain type of

structure hence there is still a need for a simple empirical function which yields as accurate an estimate of the binding affinity as possible. Their major limitation is the dependence on the size, composition and generality of the training set used to derive the weights [93].

2.3.3 Knowledge-Based Scoring Functions

A third kind of scoring functions are knowledge-based scoring functions (also referred to as statistical-potential based scoring functions), which employ energy potentials that are derived from the structural information embedded in experimentally determined atomic structures. The idea is that sufficiently large experimental data of the property of hundreds of similar compounds can serve to derive rules and general principles inherently stored in the knowledge base [93]. The principle behind knowledge-based scoring functions is simple: Pair-wise potentials are directly obtained from the occurrence frequency of atom pairs in a database using the inverse Boltzmann relation. For protein–ligand studies, the potentials are calculated by

$$w(r) = -k_B T \ln[g(r)], g(r) = \frac{p(r)}{p} * (r) \quad (2.2)$$

where k_B is the Boltzmann constant, T is the absolute temperature of the system, p(r) is the number density of the protein–ligand atom pair at distance r, and p*(r) is the pair density in a reference state where the interatomic interactions are zero. Compared to the force field and empirical scoring functions, the knowledge-based scoring functions offer a good balance between accuracy and speed [92]. This approach has been successfully applied in the field of protein-fold predictions.

2.3.4 Physics-Based Energy Functions

This approach is challenging because the physics of ligand binding is complicated. Factors such as electronic polarizability, entropic losses in the ligand and receptor, and solvation effects can all impact absolute and relative binding

affinities. Two important strengths of physics-based scoring methods are: (i) they do not require parameterization using ligand binding affinity data and crystal structural information, and hence are not subject to concerns about over-fitting, and (ii) it is possible to systematically pursue improvements to the scoring function by using more sophisticated energy models and sampling schemes [82].

2.3.5 Consensus Scoring

Consensus scoring is not really a specific type of scoring function but a technique in protein-ligand docking. It improves the probability of finding a correct solution by combining the scoring information from multiple scoring functions in hopes of balancing out the errors of the individual scoring functions. Therefore, the main issue in consensus scoring is how to make the combination rule for individual scores so that the true binders can be discriminated from others according to the consensus rule. MultiScore and X-Cscore are two examples of consensus scoring methods [85]. Despite considerable progress, current scoring functions are still far from being universally accurate, considering the test set-dependency of their performance and the fact that many of the scoring functions failed on one or two of the three widely-used criteria.

2.3.6 Molecular Mechanics/Poisson-Boltzmann Surface Area (MM-PB/SA)

This is a hybrid method, MM-PB/SA, combining molecular mechanics and continuum solvent calculations. MM-PB/SA and similar methods instead employ the free-energy cycle shown in Figure 2.1.

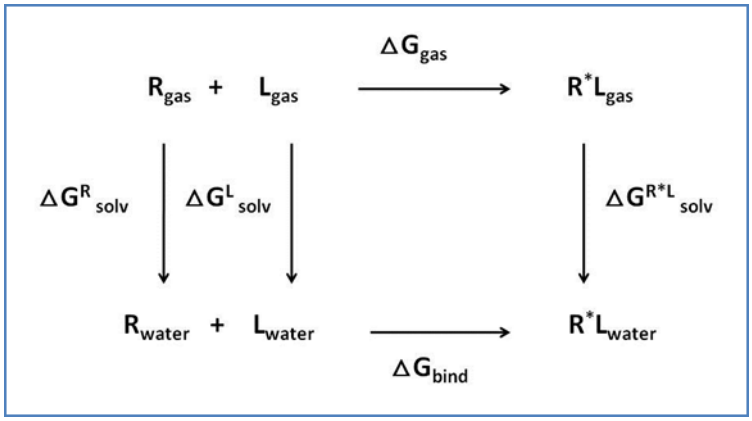


Figure 2.1: A hybrid method of molecular mechanics and continuum solvent calculations (MM/PBSA)

Free-energy cycles used to calculate the binding free energy of ligand (L) to the receptor (R) in a solution. R_{gas} is the free receptor in vacuum, R_{water} is the free receptor in solution, L_{gas} is the free ligand in vacuum, L_{water} is the free ligand in solution, ΔG^{R}_{solv} is the free energy to solvate R, and $\Delta G^{R^*L}_{solv}$ is the free energy to solvate complex R*L. This approximation allows the absolute binding free energy of ligand L to the receptor to be estimated by decomposing the binding energy into a gas-phase free energy and a salvation free energy of transferring the free ligand, free receptor and ligand–receptor complex from the gas phase to aqueous solution (Eqn. 2.3) where the free energies for each species are evaluated individually (Eqn. 2.4):

$$\Delta G_{bind} = \Delta G^{R*L}_{water} - \Delta G^{R}_{water} - \Delta G^{L}_{water} \tag{2.3}$$

$$G_{water} = G_{gas} + G_{solv} = \left(H_{gas} - TS\right) + G_{solv} \tag{2.4}$$

The MM-PB/SA method, as originally formulated, averages the gas phase enthalpy and solvation free energy over multiple configurations sampled from molecular dynamics (MD) simulations with explicit solvent [94]. In some cases, estimation of the entropy losses upon binding are also included. The estimation of the individual contributions is a challenging task due to limited knowledge of the relative relevance of the interacting forces that are themselves strongly dependent on the character of individual complexes. Moreover, individual term can be large whereas

the overall result is usually small, thus increasing the level of uncertainty on that result. The MM-PB/SA method has widely been applied to predict ligand–receptor binding geometries and to calculate absolute or relative binding affinities in good agreement with experimental data providing a mean absolute error of ±1 to 2 kcal/mol. However, a fundamental question with the MM-PB/SA approach normally arises from it accuracy on the contribution from the entropy change upon binding. If the absolute binding free energy is to be estimated, then the entropic contribution is to be determined in a manner that is consistent and also provide significant results. Generally, this is a challenging task more especially if the conformational changes are significantly high, and even relative entropies are difficult to determine with good accuracy. "In calculations of relative binding free energies for a series of ligands to a common protein receptor, the entropic contribution is often assumed to cancel when the ligands are of similar size" [69, 95]. Moreover the explicit consideration of water molecules in protein binding sites is then an interesting alternative to improve the quality of scoring methods that are based on the implicit description of solvent effects. The ability to predict the structural and thermodynamic properties of water molecules in protein binding sites has been investigated by numerous theoretical methods [96]. Additionally machine learning-based scoring functions constitute an effective way to assimilate the fast growing volume of high-quality structural and interaction data in the public domain and are expected to lead to more accurate and general predictions of binding affinity. This will enhance the interpretability of features in terms of the intermolecular interactions [80].

2.4 LigandFit

LigandFit is an algorithmic tool for protein modeling available in the Discovery Studio software for computational chemistry and biology designed for applications in drug development. LigandFit was originally developed by Venkatachalam and co-workers [85]. Docking is based on an initial shape match to the binding site**.** This approach consists of two major parts. In the first step, the potential binding sites are identified using detection algorithms and ranked in an order of importance. The

second step is when the actual docking of a ligand to a specified binding site through a Monte Carlo conformational sampling procedure. This procedure generates binding poses that match the binding site on the bases of shape filtering. Hereafter the rigid body energy minimization is performed on candidate binding poses using the DockScore energy function. In LigandFit, the ligand is accurately docked into protein active sites. The method employs a cavity detection algorithm for detecting invaginations in the protein as candidate active site regions. Candidate poses are minimized in the context of the active site using a grid-based method for evaluating protein-ligand interaction energies. Ligand binding poses can be further evaluated by external scoring functions, including LigScore1, LigScore2, PLP1, PLP2, Jain, PMF, PMF04, and the LUDI scoring function [85, 97-99]. The MM/MD and QM/MM/MD methods are normally used subsequently to docking in order to get a better understanding of the host-guest interactions.

CHAPTER 3

THEORETICAL PRINCIPLES OF EXPERIMENTAL DESIGN

Since the notion of the experimental work is to generated data can be analyzed, it is always good practice to organize the experiment approach properly to ensure that the experiment is not bias and correct type of data is available to answer the questions of interest as clearly and efficiently as possible. This preparation process is known as experimental design**.** In the ligand-protein interaction studies, it is common to begin with a process model with several discrete or continuous input factors that can be controlled that is, varied experimentally and one or more measured output responses obtained. The experimental data are used to derive an empirical (approximation) model linking the outputs and inputs. Since the reversible interactions play a key role in many biochemical and physiological processes, the assessment of these interactions in terms of affinity constants (K_a) and stoichiometry is an important part in describing and understanding such systems. In this section we look at the corresponding approaches that can be used to acquire such information.

3.1 Ligand and Protein Concentrations

The reciprocal of the association constant (K_a), i.e., the dissociation constant (K_d), is the ligand concentration that occupies half of the maximum number of binding sites on the protein. This value allows for the evaluation of a range of ligand concentrations that should be used to build the desired part of the binding curve. In practice, a range of total drug (D) concentrations extending from 0.1 to 10 K_d (where the ratio of total drug (D)/ total protein (P) << 1) seems to be a reasonable drug range concentration for practical use [4, 100]. Unbound drug concentrations show better correlation to the pharmacological activity than the sum drug concentration.

3.2 Mimicking Physiological Conditions:

The pH of the buffer can have significant impact on drug–protein interactions whereby there are conformation changes in the protein receptors or there could be changes on the net charge of either the drug or the protein. These charges are deemed to alter coulombic interactions. Overall the ionic strength of the buffer can also have an effect on the interaction of the drug with the protein [16]. Therefore, the experiments were conducted at as close as possible to the physiological conditions (i.e., with a phosphate buffer at pH 7.4, 67 mM). Sodium dihydrogen phosphate (NaH_2PO_4) buffer was chosen because the electrophoretic current generated with Na^+ ions was lower at a given voltage.

3.3 Randomization of Experimental Measurements

Due to the complexity of the experiments, therefore it could be extremely difficult to eliminate bias using only expert judgments or data screening however an additional approach of randomization in experiments is common practice. In a randomized experimental design, objects, sampling points or individuals are randomly assigned (by chance) to an experimental group. Using randomization is the most reliable method of creating homogeneous treatment groups, without involving any potential biases or judgments [101]. Moreover there is no consistent or systematic effect in the assignment that is tied in with the method. There are several variations of randomized experimental designs, two of which are complete randomization and randomized block design.

3.4 Limit of Detection (LOD)

The limit of detection is the smallest concentration of the analyte in the sample that can be reported with a certain level of confidence [102]. The 1975 International Union of Pure and Applied Chemistry (IUPAC) definition for LOD can be stated as, "*A number expressed in units of concentration (or amount) that describes the lowest concentration level (or amount) of the element that an analyst can determine to be statistically different from an analytical blank*" [103]. LOD's may be calculated

based on the standard deviation of the response (SD) and the slope of the calibration curve (S) at levels approximating the LOD according to the formula: LOD = 3.3(SD/S) [104, 105]. The standard deviation of the response can be determined based on the standard deviation of y-intercepts of regression lines. The limit of detection is occasionally used indiscriminately or confused with the sensitivity of the method. The analytical sensitivity is the ratio of the calibration curve slope to the standard deviation of the analytical signal at a given analyte concentration or mass of the test analyte. Spectrophotometric detectors used in CE may be an order of magnitude less sensitive compared to those available for HPLC because the former detectors use shorter path lengths. However, the reduced sensitivity is partially compensated by the high separation efficiency of CE, which allows for more precise integration of peak areas. Limits of detection can also be improved by lowering detection wavelength down to ± 185 nm and by using sample stacking or a bubble capillary for on-column concentration of samples. The normal detection limits that can be obtained with CE coupled with a UV detector is in low µM (or µg/mL) range, which can also be reduced to as low as few nM (or ng/mL) if the sample is pre-concentrated.

3.5 Limit of Quantification (LOQ)

According to the IUPAC, Limit of Detection is defined as, "*The minimum detectable value of the net signal (or concentration) is that value for which the false negative error is β, given α. "α"is defined as the probability for a false positive ("analyte present" result when that is wrong) and "β" is defined as the probability of a false negative ("analyte absent" result when that is wrong)*" [106]. It is important that this term be defined in a manner such that it is easy to compare values reported by different analysts and laboratories. For most modern analytical methods, the LOQ may be divided into two components, instrumental detection limit (IDL) and method detection limit (MDL) [102, 107]. The calculation method is also based on the standard deviation of the response (SD) and the slope of the calibration curve (S) according to the formula: LOQ = 10(SD/S) [104]. The standard deviation of the

response can also be determined based on the standard deviation of y-intercepts of regression lines.

3.6 Statistical Analysis of Binding Data

In order to assess and improve the quality of the data, different statistical tools and approaches are used. Since outlier tests are mostly based on probabilities, their interpretation is naturally subject to possible errors. Outlier tests are carried out at some predefined significance level, α, that defines the probability of rejecting a null hypothesis that is true.

3.6.1 Z-Score

The z-scores assign how many standard deviations, *s*, the value of individual laboratory, *x*, is above (to the right of) or below (to the left of) the consensus value, x. Values of *x* that are larger than the mean have positive z-scores and values of *x* that are smaller than the mean have negative z-scores. If *x* equals the consensus value, then *x* has a z-score of 0.

$$z = \frac{x - X}{s} \tag{3.1}$$

In this equation the *x* value is the result obtained by a single laboratory for a given analysis; *X* is the assigned consensus value for the level of the analyte, and *s*is the target value for the standard deviation of the results (often the standard deviation of participants used instead) [108].

3.6.2 Grubbs` test

Grubbs` test is the most commonly used and recommended by ISO test for outliers is the Grubbs' test [109]. This test compares the deviation of the suspect value from the sample mean with the standard deviation of the sample. Grubbs' test is

also known as the maximum normed residual test. The Grubbs' test finds the change of standard deviation expressed as percent:

$$GS = 100\frac{|s - s_H|}{s} \tag{3.2.1}$$

or

$$GS = 100\frac{|s - s_L|}{s} \tag{3.2.2}$$

where GS is Grubbs' test statistic, s is standard deviation of whole data set, sL is standard deviation of data set, where the lowest result is left out, sH is standard deviation of data set, where the highest result is left out. Bigger GS is compared with a tabulated critical value from the t-distribution found in the table. Grubbs' single test detects one outlier at a time [110]. This outlier is eliminated from the dataset and the test is iterated until no outliers are detected. However, multiple iterations change the probabilities of detection, and the test should not be used for sample sizes of six or less since it frequently tags most of the points as outliers. In order to achieve more confidence Grubbs' double test is also used [111]. It determines whether the two largest or two smallest values at a time might be outliers. It is very similar to Grubbs` single test, only difference is that outliers are left out in pairs. Grubbs` double test is usually used in connection with Grubbs` single test.

3.6.3 Scatchard Plot

Over the last thirty years, the so-called Scatchard plot has been the traditional method for analysis of binding data until the introduction of sophisticated non-linear curve-fitting software. Scatchard plot is a very powerful tool for identifying deviations from simple models, which without deviations, is represented by a straight line on the Scatchard plot [112].

3.6.4 Klotz Plot

The graphical representation of the Klotz plot is used in two ways: determining the suitability of chosen ligand concentrations and analyzing the data for co-operativity [113].

3.6.5 Non-Linear Curve Fitting

Careful inspection of the data, using the methods linear regression, will allow to most suitable data fitting to a model that accounts for the physical reality. If the protein has one receptor site or two, it is advisable to reanalyze data with several models and compare them to each other with statistical robustness [75].

3.7 Factors Affecting Protein Drug Binding.

3.7.1 Drug Related Factors

A drug in blood exists in the bound and unbound forms. Depending on a specific drug's affinity for plasma protein, a proportion of the drug may become bound to plasma proteins, with the remainder being unbound. If the protein binding is reversible, then a chemical equilibrium will exist between the bound and unbound states, such that:

$$\text{Protein} + \text{Drug} \rightleftharpoons \text{Protein} - \text{Drug Complex}$$

(a) Physiochemical characteristics of the drug.

Lipophillicity is the most desirable physiochemical parameter that is perquisite for protein binding to occur. Also an increase in the lipid content of drug moiety eventually enhances the rate as well as extends of protein binding process [114].

(b) *Concentration of drug in the body.*

Alteration in the concentration of drug substance as well as the protein molecules or surfaces subsequently brings alteration in the protein binding process [115].

(c) *Drug's affinity towards protein.*

This factor entirely depends upon the degree of attraction or affinity that the protein molecules have towards drug moieties. For Digitoxin has more affinity for cardiac muscles proteins as compared to that of proteins of skeletal muscles or those in the plasma like HSA.

3.7.2 Protein Related Factors

(a) *Concentration of protein/binding component.*

This is the most important protein related parameter to be given priority. As the human serum plasma proteins constitute the major portion of the plasma proteins, a large number of molecules undergo an extensive binding with them as compared to the concentration of other protein molecule.

(b) Number of binding sites on the protein

In association to the concentration of proteins molecules available the number of binding sites available in the protein molecules is also significant. Albumin not only possesses large number of binding sites but also has greater potential of carrying out binding process. Numerous drug exhibit multiple site binding with albumin molecules in plasma like fluocloxacillin, ketoprofen, indomethacin etc [41].

3.8 Estimation of Pharmacokinetic Parameters

The different constants do not describe exactly the same situation; however their values are generally used equivalently. In the following only the abbreviation K_i will be used. The binding constant describes the strength of the interaction between protein and ligand [116]. It is a thermodynamic state function and reflects how much

of the ligand is, on the average, bound to the protein. The smaller the value of K_i, the stronger does the ligand bind to the protein. If the concentration of the ligand is significantly less than K_i, then only a small percentage of ligand molecules will be bound to protein molecules [117]. Often, instead of the K_i value, a so-called IC_{50} value is determined. This quantity indicates at which ligand concentration the activity of the protein (usually enzyme) has been decreased by 50%. In contrast to the K_i value, the IC_{50} value depends on the concentration of the enzyme. Experience has shown that both values run approximately parallel so that IC_{50} values, which are easier to determine, are very well suited for the characterization of a ligand in comparison to other structures.

The dissociation constant is commonly used to describe the affinity between a ligand (L) (such as a drug (D)) and a protein (P) i.e. how tightly a ligand binds to a particular protein [116]. The formation of a ligand-protein complex (PL) or (DP) can be described by a two-state process. When we assume that the stoichiometry of the binding between *D* and *P* is 1:1, the reversible reaction between *D* and *P* is expressed by Eqn. 3.3.1. The equilibrium association constant, K_1, is used to characterize this reaction, which is governed by the law of mass action:

$$K_1 = K_a = \frac{[PL]}{[P][L]} \tag{3.3.1}$$

the corresponding dissociation constant is defined as

$$K_{-1} = K_d = \frac{[P][L]}{[PL]} \tag{3.3.2}$$

where [P], [L]=[D] and [PL]=[DP] represent molar concentrations of the protein, ligand or drug and complex, respectively [118].

The number of total ligand bound per protein, *r*, is expressed as shown in Eqn.3.4.1.

$$r = \frac{b}{P} = \frac{D - d}{P} = n_i \frac{K_i d}{1 + K_i d} \tag{3.4.1}$$

where *r* is the fraction of bound enantiomer per molecule of protein, *P* and *D* are the total protein and enantiomer concentrations, respectively, *b* and *d* represent the bound and free (unbound) concentrations of the enantiomer after the equilibrium, respectively, *m* corresponds to the number of classes of independent active sites in the protein, each class *i* accounting for n_i binding sites (the apparent stoichiometry) and K_i, the affinity constant of the interaction (in M^{-1} units; all concentrations must be in M units) [4]. The 1:1 binding is often an oversimplification of the reality. Therefore, a more complex model, including multiple interaction types and binding sites, is frequently required and it is given by the following equation:

$$r = \frac{b}{P} = \frac{D - d}{P} = \sum_{n=1}^{m} n_i \frac{K_i d}{1 + K_i d} \tag{3.4.2}$$

where *m* is the total number of different classes of binding sites and n_i is the number of binding sites possessing the same affinity constant for a ligand/drug. The binding constants and stoichiometries are estimated by nonlinear least-square fitting of the experimental data to the binding isotherm of the mathematical program such as GraphPad Prism, Statgraphics, Unscrambler, Matlab etc [4, 100]. The most import prerequisite of the study is that there should be a significant difference in electrophoretic mobilities of the drug, protein, and protein-drug complex. The dissociation constant has molar units (M), which correspond to the concentration of ligand/drug D at which the binding site on a particular protein is half occupied, i.e. the concentration of ligand, at which the concentration of protein with ligand bound, b, equals the concentration of protein with no ligand bound P. Generally, non-linear regression is more accurate and precise for the estimation of binding constants than linear regressions following algebraic manipulation (e.g., inverse plots) [74]. Plots of the number of bound ligand molecules per molecule of protein (*r*) as a function of the free ligand concentration yield a binding curve. Direct fitting to the experimental data using nonlinear regression analysis for obtaining the binding parameters is recommended. However statistical evaluation has to be implemented in results obtained experimental prior to estimation of any parameters.

CHAPTER 4

CAPILLARY ELECTROPHORESIS

4.1 Principles of Capillary Electrophoresis (CE)

Capillary electrophoresis is a highly efficient analytical separation technique. Separation is based upon differences in mobility, which mainly depends on the charge to mass ratio of the molecule. The most common ways to influence mobilities are changing the pH of the buffer or by adding an appropriate complexing agent to the buffer. CE analyses are usually very fast, use little sample and reagents, and cost much less than chromatography or conventional electrophoresis. Although modern CE is still in its teenage years, it has demonstrated tremendous potential for a wider range of applications, from small molecules that include inorganic ions, organic acids, amino acids, peptides, drugs, nucleosides, nucleotides, vitamins, steroids, and carbohydrates to larger molecules, such as hormones, proteins, nucleic acids, and even living cells. The use of CE to determine the physicochemical and biochemical properties of a variety of molecules make this technique very attractive from other disciplines. Such measurements are carried out on a very small amount of sample (about 5 to 10 µl) and can provide valuable information suitable for the design of biologically active molecules.

4.2 Electrophoretic Mobility (μ_e)

CE differentiates charged species on the basis of electrophoretic mobility (μ_e), under the influence of an applied electric field. The value of μ_e is directly related to the charge of a molecule and inversely related to its hydrodynamic drag [119]. It is important to note that electrophoretic mobility (μ_e) is proportional to electric charge on the species and inversely proportional to certain frictional forces. In short, molecules with a positive charge will tend to flow toward the negative terminal, negative species toward the positive terminal, with neutral species having no electrophoretic mobility. This phenomenon can be described according to the Eqn. 4.1 shown below.

$$\mu_e = \frac{q}{6\pi\eta r} \tag{4.1}$$

Where q is the ion charge, η is the viscosity of the solution and r is the ion radius. The second method of movement of an individual species is based on the bulk flow of the buffer solution. This phenomenon is called electro-osmotic flow (μ_{EOF})denoted by Eqn. 4.2.

4.3 Electroosmotic Flow (μ_{EOF})

The movement of the buffer within the capillary under the influence of the electric field is referred to as electroosmosis. The inner surface of a fused silica capillary is covered with silanol groups (Si-OH), which are ionized to SiO^- at pH > 2. At high pH, where the silanol groups are predominantly deprotonated, the EOF is significantly greater than at low pH where they become protonated [120]. The negatively charged surface is counterbalanced by positive ions from the buffer, forming the so-called electric double layer create a potential very close to the wall (as shown in Figure 4.1). Under the influence of the electric field, the positive ions in the diffuse part of the double layer migrate towards the cathode; in doing so they entrain the waters of hydration, which results in electroosmotic flow most pronounced property of the EOF is the flat flow profile in Figure 4.1B compared with the parabolic flow (drop like) profile of hydrodynamic flows, common in chromatography [31].

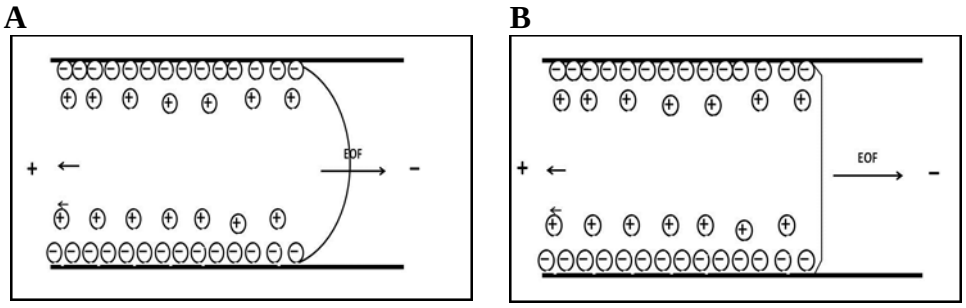


Figure 4.1: Flow velocity profiles of pressure and electrodriven flow.

In many experiments, the electroosmotic mobility, μ_{EOF}, is greater than the electrophoretic mobility μ_e, and consequently, all species will be transported from the sample end to the collection chamber with separation being determined by charge and frictional forces. The mobility of EOF is denoted by the following equations:

$$v_{EOF} = \left(\frac{\varepsilon\zeta}{\eta}\right) E \tag{4.2.1}$$

or

$$\mu_{EOF} = \frac{\varepsilon\zeta}{\eta} \tag{4.2.2}$$

where μ_{EOF} is the electroosmotic flow mobility, v_{EOF} is the velocity, ε is the dielectric constant of the solvent [$C^2.J^{-1}.m^{-1}$], ζ is the potential of the capillary surface andη is the viscosity of the solution. EOF can be effectively controlled by changing several experimental conditions, among which buffer pH (affecting the dissociation of the wall silanols), ionic strength (affecting the ζ potential), organic solvents (acting on both ζ potential and buffer viscosity) and buffer additives (e.g. surfactants) should be mentioned [30]. The time it takes for a species to reach the detector is the *migration time* and is given by:

$$t_m = L_d L_t \tag{4.3}$$

where L_d is the length of the capillary to the detector, L_t is the total length of the capillary, and V is the applied voltage across the capillary.

4.4 Instrumentation

Capillary electrophoresis is made of a few major components, which includes a high voltage power supply, a capillary, capillary cassette, two buffer reservoirs, two electrodes, and a detector. The applied potential can either be negative or positive if a positive of 10 to 30 kV is applied; all sample components migrate towards the negative electrode due to electro-osmotic flow (EOF). Capillaries are typically of 50 µm inner diameter and 0.5 to 1 m in length. A small volume of sample (+/-10 nl) is

injected at the positive end of the capillary and the separated components are detected near the negative end of the capillary. CE detection is similar to detectors in High Performance Liquid Chromatography (HPLC), and includes absorbance, fluorescence, electrochemical, and mass spectrometry. The capillary can also be filled with a gel, which eliminates the electro-osmotic flow. Separation is accomplished as in conventional gel electrophoresis but the capillary allows higher resolution, greater sensitivity, and on-line detection. A schematic representation of a basic capillary electrophoresis instrument is shown in Figure 4.2.

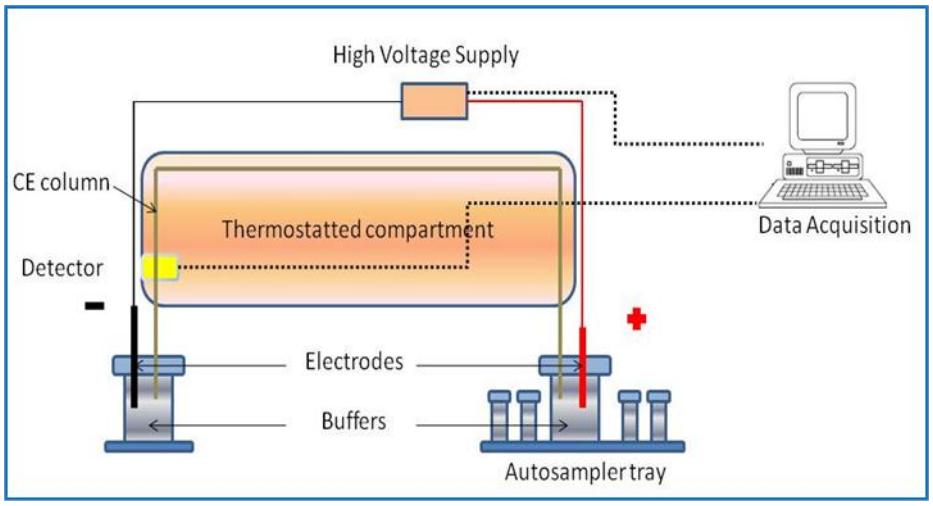


Figure 4.2: The basic schematic instrumental setup of CE, which consists of a high voltage power supply (0 to 30 kV), a fused silica (SiO_2) capillary, two buffer reservoirs, two electrodes, and an on-column Photo Diode Array UV detector

4.4.1 High Voltage Power Supply

The high voltages up to 30kV can be supplied to the capillary to facilitate the separations by generating the electroosmotic and electrophoretic flow of buffer solutions and ionic species, respectively, within the capillary. When a voltage is applied across the capillary, cations in the diffuse portion of the double layer migrate in the direction of the cathode, carrying water with them. The result is a net flow of buffer solution in the direction of the negative electrode [25]. In order to sustain a good reproducibility of the migration time a constant voltage supply throughout the analysis is highly recommended. Voltage is expected to reach the desired voltage within the first few seconds (in approximately 4 s) it is applied. Figure 4.3 below

show the ramp of the voltage within 0.2 s thereafter it remains constant throughout the run.

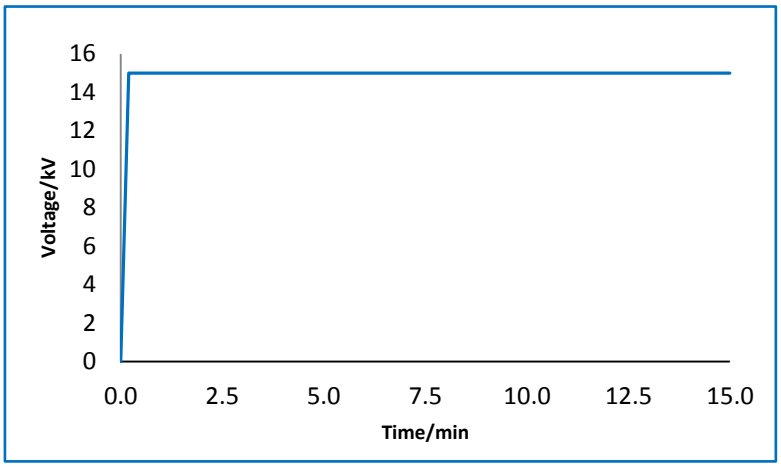


Figure 4.3: Typical output of the voltage applied through the capillary.

4.4.2 Capillaries

CE columns are packed with stationary phases. Ideally, the capillary should be: chemically and physically resistant, accurately and precisely produced with narrow internal diameters (I.D.) of 20–100 µm and lengths from 20–100 cm. They should also be transparent to UV radiation. The internal surface may be simply ‘‘uncoated’’ silica, directly in contact with buffers and solutes, or ‘‘coated’’ with polymers, shielding silica from interactions with solutes. Fused-silica capillaries meet almost all these requirements, but tend to adsorb analytes, particularly proteins [30]. Absorbance measurements on CE are often made directly on the capillary as shown in schematic diagram Figure 4.2, where the detection window of the capillary is aligned to the detector. A new capillary requires extensive conditioning before it can be used for analysis. The capillary is flushed for 10 min with 1.0 M sodium hydroxide, 5 min with deionized water, and 20min with an electrolyte (run buffer). This conditioning procedure is important to ensure that the surface of the capillary is fully and uniformly charged. For some methods it is necessary to regenerate this surface between runs with 0.1 M sodium hydroxide, and in extreme cases, 1.0 M sodium hydroxide is used. The regeneration procedure is frequently necessary if

migration times change on a run-to-run basis. This is most common when using buffers in the pH 2-6. Also in drug-protein studies because protein adsorption to the bare silica wall in capillary can cause problems in separation and loss of protein and drug. Hence the capillary has to be reconditioned between each run to remove any drug or protein adsorbed on the stationery surface of the capillary, preferably in the following sequence: 2 min with deionized water, 2 min with 0.1 M sodium hydroxide, 2 min with deionized water and 15 min phosphate buffer (separation buffer). Stationary phases made up by immobilized serum albumins from human and from a variety of mammalian species have also been used for the enantioresolution of a number of neutral and acidic drugs. (See details in chapter 1)

4.4.3 Pre-Treatment of the Capillary

Reproducibility of the migration time in CE is strongly dependent on the silanol concentration at the residual surface and to some extent also on amount of impurities in the silica, so rigorous pretreatment and preconditioning procedures should be applied in order to render the capillary in a well-defined reproducible state of surface hydroxylation. Details of the methodology used for pre-treatment are given in section 5.1.5.

4.4.4 Temperature Control

CE separations are generally carried out at constant temperature of the thermostated compartment. Nonetheless, efficient control of temperature of the separation compartment is fundamental for analytical performances: increases in capillary internal temperature, due to inadequately dissipated Joule heat, cause changes in the buffer viscosity and consequently in the migration velocity of analytes [30]. The two approaches generally used to control the temperature are to bath the capillary in a high velocity air stream or in a liquid. A benefit of the air thermostated system is the instrument simplicity and ease of handling the separation capillary.

4.5 Sample Introduction

In order to avoid excessive band broadening only a small fraction of the capillary is loaded with the measuring solution. Different modes of injection can be used to introduce the sample into the capillary: these include hydrostatic, electrostatic, hydrodynamic (pressure or vacuum), and split flow [120]. The type of sample introduction that is commonly used is hydrodynamic mode.

4.5.1 Hydrodynamic Injection

There are different ways in which the analytes, samples or buffers can be introduced dynamically into a capillary, namely vacuum, pressure, or split flow. In all cases amount injected depend on the dimensions of the column, pressure applied for a set time and on the viscosity of the solution. This volume (V_{inj}) can be calculated using Hagen-Poiseuille Eqn. 4.4.

$$V_{inj} = \frac{\Delta P d^4 \pi t_{inj}}{128 \eta L} \tag{4.4}$$

Where ΔP is the pressure difference across the capillary, d is the capillary inside diameter, t_{inj} is the injection time, η buffer viscosity and L is the total length of the capillary [30, 120, 121]. The advantage in using vacuum over pressure injection is that only the vial at the detector end (outlet side) of the capillary that requires a gas-tight seal and not sample vials (inlet side). However this approach is difficult to implement when the detection system has inherent physical constrains such as for a mass spectrometer. The pressure injection technique is particularly attractive and convenient when used in combination with MS detection [122]. During injection it is important that the sample plug length be minimized. If the length is longer than the dispersion caused by diffusion, efficiency and resolution will be sacrificed. The current instrumentation can also be used to load small volumes reproducibly, under typical conditions but detection limit constraints often necessitate longer injection lengths.

4.6 Modes of Capillary Electrophoreses Operation

A number of CE modes are available for separation of proteins, anions, cations, drugs, polypeptides etc. These modes include capillary zone electrophoresis (CZE), affinity capillary electrophoresis (ACE), micellar electrokinetic capillary chromatography (MECC) or micellar electrokinetic chromatography (MEKC), capillary sieving electrophoresis (CSE), capillary isoelectric focusing (CIEF),capillary isotachophoresis (CITP), capillary electro-chromatography (CEC), affinity electrokinetic chromatography (AEKC), affinity capillary electro-chomatography (ACEC), vacancy affinity capillary electrophoresis (VACE) and capillary electrophoresis frontal analysis (CEFA) [120].

4.6.1 Capillary Zone Electrophoresis CZE

Capillary zone electrophoresis (CZE), also known as free solution capillary electrophoresis, is the simplest form of CE. In CZE, the capillary is filled with a running buffer, sample is introduced at the inlet and the voltage is across the capillary to generate electric fields. The separation mechanism is based on differences in the charge-to-mass ratio. Small, highly charged ions move fast, large, lowly charged ions move slow [30, 120, 121]. The electrophoretic migration velocity (v) of a charged particle depends on its electrophoretic mobility (μ_{EOF}) and on the applied electric field (E):

$$v = \mu_{EOF} E \tag{4.5}$$

where μ_{EOF} is described by Eqn.4.2.2

Zone electrophoresis is particularly useful for the separation of peptides, such as those prepared from tryptic digests. There has been interest in the development and incorporation of solid phase immobilized reactors for on-line digestion before electrophoresis analysis [29].

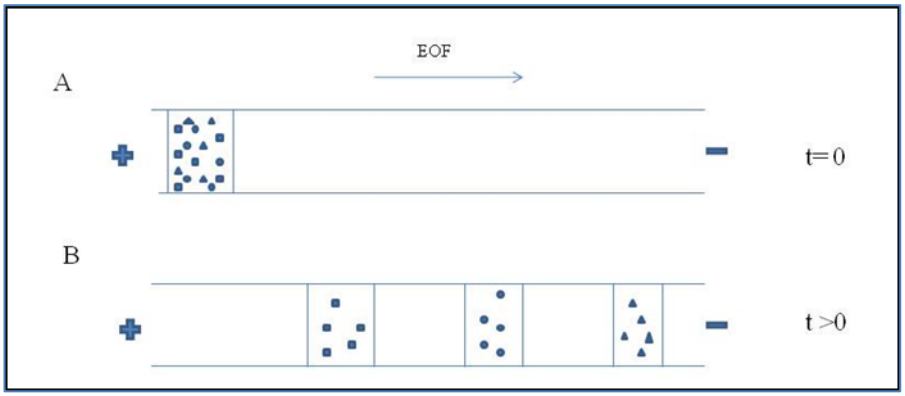


Figure 4.4: Separation profile of molecules with different charges under the influence of an applied electric field. ►positively charged molecules, ● neutral molecules ■ negatively charged molecule

4.6.2 Affinity Capillary Electrophoresis (ACE)

ACE refers to the separation by CE of substances that participate in specific or non-specific affinity interactions during separation [123]. In the past two decades, ACE has been one of the most rapidly growing analytical techniques for studying a variety of non-covalent interactions and determining binding constants and stoichiometries [74]. In this technique; free and bound forms of a receptor can be distinguished as a function of the concentration of free ligand. In a traditional ACE study increasing concentrations of ligand in the running buffer causes a shift in the migration time of the receptor. Analysis of this change in migration time relative to a standard(s) yields a value for the binding constant [119].

4.6.3 Capillary Electrochromatography (CEC)

CEC is a hybrid technique combining electrokinetic pumping and stationary-phase retention capacity of high-performance liquid chromatography. Columns for CEC typically have larger inner diameter than those used for normal CE methods, thus facilitates MS detection schemes. The CEC column can be packed with an alkyl silica stationary phase that may be modified for a specific or selective retention. In other cases, the capillary surface or packing material is chemically modified to reduce absorption effects and stabilize EOF [29]. At the interface between silica (of the capillary walls and of the packing particles) and the buffer, an EOF is generated

as shown in Figure 4.1. This improves greatly the kinetics of partitioning equilibria and consequently the separation efficiency. Currently, great attention is being paid to CEC, particularly for the separation of hydrophobic compounds and for its coupling with MS, but the technique is still in a stage of development and its applications are scarce [30].

4.6.4 Capillary Isoelectric Focusing (CIEF)

In CIEF a pH gradient under the influence of an electric field is set up using ampholytes. Ampholytes are zwitter ionic molecules that contain both an acidic and a basic moiety and can have a isoelectric point (pI) value that span the desired pH range of the CIEF experiment (pH 3 to 9 for example). After filling the capillary with the mixture of solute and ampholyte, the pH gradient is formed. With a basic solution at the cathode and an acidic solution at the anode, upon application of the electric field the charged ampholyte and proteins migrate through the medium until they reach a region where they become uncharged (at their pI). The status of the focusing process is indicated by the current. Once complete, a steady-state is reached and the current no longer flows. After focusing, the solute and the ampholytes are mobilized by the hydraulic pressure and the zone passes through the detector [120, 121]. One important application of the CIEF is determination of the isoelectric point (pI) of a proteins [122].

4.6.5 Capillary Isotachophoresis (CITP)

Capillary Isotachophoresis (CITP) is a “moving boundary” electrophoretic technique. In CITP, a combination of two buffer systems is used to create a state in which the solutes all move as connected, separate bands and at the same velocity. The zones remain sandwiched throughout the capillary between a so-called leading and terminating electrolytes. Capillary Isotachophoresis separation a mechanism is shown in Figure 4.5.

Figure 4.5: Illustration of CITP mechanism. T terminating buffer, L leading buffer, •, Δ and ■ represents solute mechanism

4.6.6 Micellar Electrokinetic Chromatography (MEKC) or Micellar Electrokinetic Capillary Chromatography (MECC)

MEKC is a mode that is closest to the reverse-phase HPLC, and the retention of the analyte is similarly described in terms of the capacity factor, k'. Surfactants are used to modify the mobility of proteins and other analytes to manipulate adsorption of compounds to capillary walls. If the concentration of the micelle exceeds the critical micelle concentration (cmc), the micelles can act as a pseudo stationary phase, in which case the separation is called micellar electrokinetic capillary chromatography (MECC). Typically, MECC is performed on unmodified or uncoated fused silica [29, 30]. The most commonly used surfactant is sodium dodecyl sulphate (SDS). The anionic SDS micelles are electrostatically attracted towards the anode, but, because of the prevalent velocity of the EOF, they slowly migrate towards the cathode, i.e., in the direction of the detector. Consequently, micelles decrease selectively the migration of the nonionic solutes they interact with (by partitioning mechanisms), which otherwise would migrate at the same velocity as the EOF. Depending on the individual partitioning equilibria of the different analytes between the hydrophobic core of the micelles and the aqueous buffer, a different retarding effect on the neutral analytes will be observed, determining their separation in the capillary [30]. Using MEKC it is possible to evaluate surfactant critical micelle concentration (*cmc*). The retention factor (k') of an analyte in MEKC at low micelle concentration is directly proportional to the micelle concentration:

$$k' = P_{mv}\mathrm{V}([\mathrm{S}] - \mathrm{cmc}) \tag{4.6}$$

Where P_{mv} is the partition coefficient of the solution between the micelle and aqueous phase, V is the molar volume of the surfactant, and [S] is the total concentration of the surfactant. A plot of *k'* vs. [S] for several analytes gives a value of the *cmc* extrapolated to the point where *k'* is zero. The *cmc* of a surfactant can also be obtained by measuring the change in current with increasing concentration of the surfactant break in a linear relationship between current and concentration occurs at the *cmc* of the surfactant due to the difference in the conducting properties of it monomeric and micellar forms. The *cmc* values obtained by researchers using CE are in agreement with values determined by viscosity measurements [122]. The development of micellar electrokinetic capillary chromatography (MECC) has extended CE applications to the separation of both neutral and charged molecules through the use of micelles in the separation buffer.

4.6.7 Affinity Capillary Electrochomatography (ACEC)

In this technique protein is immobilized on the wall of the capillary acting as chiral selectors (in enantioresolution) or receptor (in ligand-protein interaction studies). Stationary phases made up by immobilized serum albumins from human (eg human serum albumin (HSA)) and from a variety of mammalian species have been used for the enantioresolution of a number of neutral and acidic drugs [28].

4.6.8 Vacancy Affinity Capillary Electrophoresis (VACE)

VACE is a relatively new method in CE for the estimation of binding constants [124]. The capillary is filled with a running buffer containing both drug (D) and protein (P). The concentration of either D or P is fixed and that of the other component is varied. The binding constants are estimated from the migration behavior of D as well as ACE. An advantage of the VACE is the fact that the absolute number of the binding sites, *n*, can be calculated from the change in μ and free drug (d) as compared to ACE. In addition, VACE is particularly convenient for the estimation of binding constant of a weakly soluble D in water because the solubility should be increased in the running buffer containing P [124, 125].

4.6.9 Capillary Electrophoresis Frontal Analysis (CE-FA)

In frontal analysis (FA), the equilibrium between D and P is obtained in the sample vial. Thereafter the equilibrated sample is introduced hydrodynamically into the capillary as a large plug with buffer/electrolytes, the free and bound components are separated due to the difference in mobility. Bound drug remains within the protein zone. Injection of the large sample plugs give rise to the appearance of plateau peaks. The height of the free drug plateau peak is proportional to the free drug concentration in the original sample and the degree of binding can be determined by aid of a calibration curve. The key advantages of CE-FA are the near-physiological conditions and rapidity of analysis, ease of automation, small sample consumption, robustness and accuracy of results, which are all extremely useful for the accurate and rapid characterization of binding processes [9, 47]

The primary prerequisite in CE-FA is the requirement of sufficiently different migration times between the free and bound form (complex) of one of the species studied. FA is more robust because the height and thus concentration is not affected by changes in migration times, EOF, length of the capillary, and applied voltage to the same extent in a plateau peak as in a zonal peak because of the non-dispersed zone constituting the plateau region [47].

4.6.10 Affinity Electrokinetic Chromatography (AEKC)

Proteins can be used as additives to the running buffer (Affinity Electrokinetic Chromatography, AEKC) or as immobilized selectors (Affinity Capillary Electrochromatography, ACEC). AEKC has indeed several advantages over ACEC. Methodological developments in AEKC are easily carried out since this technique does not require protein immobilization. In addition, the use of soluble proteins eliminates the possibility of altering both the structure and the binding properties of the protein. Furthermore, higher peak efficiency is obtained with AEKC [28]. In this technique, if the analyte and the analyte–protein complex have different mobility, the interaction of the compound with the protein results in a change in the net mobility of the analyte [125, 126]. The critical problem in AEKC is the background UV absorption of a protein solution. The convenient way to overcome low detection

sensitivity is the partial filling technique. In this technique, the capillary is not completely filled hydrodynamically with the macromolecules such as protein or any other biological receptor prior to the injection of the analyte therefore the method becomes partial filling-Affinity Electrokinetic Chromatography. In addition the experimental conditions which includes column temperature, pH of running buffer, voltage applied, and plug length are optimized such that the analyte is detected in the absence of selector plug. This experimental approach has shown to be useful in the separation of enantiomers [71, 125]. Accordingly, this method was chosen as a preferred mode of separation in the present work.

4.7 Detection Modes

Detection in CE faces challenging problems due to the miniaturization of the separation compartment: small volumes (ηl) and amounts (ηg–pg) of analytes injected, the electrophoretic zones have very low volumes and often are closely adjacent. UV-visible adsorption is the most widely used detection method primarily due to its nearly universal detection nature. The high transparency of the fused-silica capillaries wall allows the use of low UV wavelengths (down to 190 nm) and the use of the on-capillary detection in the capillary electrophoresis improves the efficiency of detection. On the other hand, due to the limits of the capillary diameter (20–100 μm), the path-length is very short and, according to the Beer's law ($A=bC\varepsilon$, where A = absorbance; b = optical path length; C = concentration; ε = molar absorptivity), this limits enormously the absorbance to be measured. Besides, because of its round section, the capillary has a poor design as an optical ''cell''. Consequently, ''in column'' UV detection displays moderate concentration sensitivity (about 10 M) and linear range (up to about 0.5 mAU). A way to gain in sensitivity by increasing the optical path length is to use axial (instead of orthogonal) illumination, with the light beam passing through a Z-shaped capillary cell. Thus, the sensitivity can be increased substantially (i.e. ten times or even more), but the resolution is compromised [30]. An alternative to increase the optical path-length is the ''bubble cell'' design, which has recently been introduced commercially. Other less common detection systems are electrochemical, fluorescence, Laser Induced Fluorescence

(LIF) and mass spectrometry (MS), which can provide greater sensitivity and selectivity. The use of MS and fluorescence detectors is restricted due to the expensive instrumentation and the need for pre-treatment of compounds, which is time consuming and increases the cost of analysis.

4.8 Separation Efficiency

The efficiency of a separation of two components (1 and 2) is well described by its resolution (R_s), defined as:

$$R_s = \frac{1.18(t_2 - t_1)}{w_1 + w_2} \quad (4.7)$$

where t_2 and t_1 is the migration time of the second and first eluted component respectively, w_1 and w_2 is the peak width of first and second eluted peaks. If the value of R_s is greater than 1.5 peaks are regarded as being well resolved.

CHAPTER 5

METHODOLOGY

5.1 Analytical Experiments:

5.1.1 Instrumentation

A ^{3D}CE capillary electrophoresis system (Hewlett-Packard, Waldbronn, Germany) equipped with a diode array detector (DAD) and ^{3D}CE Chemstation software was used. A 50 µm inner diameter (id) and 363 µm outer diameter (od) fused-silica capillary with total and effective lengths of 56 and 47.5 cm, respectively was employed (Agilent Technologies, Germany). Electrophoretic solutions and samples were filtered through 0.45 µm pore size nylon membranes (Micron Separation, Weestboro, MA, USA) and degassed in an ultrasonic bath (JP Selecta, Barcelona, Spain) prior to use. A Crison Micro-pH 2000 pH meter (Crison Instruments, Barcelona, Spain) was employed to adjust the pH of buffer solutions. A thermostated bath (JP Selecta) was used for samples incubation. For ultrafiltration, Amicon® Ultra-0.5 mL centrifugal filter devices with Ultracel®-10K regenerated cellulose membranes of a nominal molecular weight limit of 10 KDa (Millipore Corporation, Bedford, MA, USA) and a centrifuge (Alvarez-Redondo, S.A., Madrid, Spain) were used.

5.1.2 Data Processing

Experimental data was treated with routines made/adopted in MATLAB$^{(R)}$ 4.2 and EXCEL $^{(R)}$ 2007 was used for verification. The nonlinear fitting was performed with SIPMLEX $^{(R)}$ algorithm in MATLAB $^{(R)}$ [4]. The Script created for calculations is shown in Appendix 5.

5.1.3 Reagents and Solutions

All reagents were of analytical grade. HSA Fraction V was purchased from Sigma (St. Louis, MO, USA); sodium dihydrogen phosphate dihydrate, sodium

chloride and sodium hydroxide were from Scharlab (Barcelona, Spain); (+)-catechin hydrate was from Cayman Chemical Company (Ann Arbor, MI, USA); (±)-catechin hydrate was from Sigma and 20 %w/v highly sulphated-β-cyclodextrin (HS-β-CD) aqueous solution was purchased from Beckman Coulter (Fullterton, CA, USA). Ultra Clear TWF UV ultra pure water (Siemens Water Technologies, Barsbüttel, Germany) was used to prepare solutions. In order to approximate the physiological ionic strength, a phosphate buffer saline (PBS) was used in the incubation stage, consisting in a 67 mM phosphate buffer at pH 7.4 and 9 g/L of sodium chloride. For the separation/determination in CD-EKC a 30 mM phosphate buffer pH 7 solution was used as background electrolyte. 2000 μM of (±)-catechin and 1000 μM HSA stock solutions were daily prepared by weighing the corresponding amount of compound powder and dissolving it with PBS.

5.1.4 Preparation of Stock Solutions

An approximate amount of (±)-catechin and HSA were weighed and dissolved in phosphate buffer at pH 7.4 to prepare stock solutions of 2 mM and 1 mM respectively. These solutions were used to prepare series of standards that were to be used for experimental design. Standards with varying concentrations were placed in an ultrasonic bath for 4min to remove air bubbles and to ensure that the reaction between drug and protein goes to completion.

5.1.5 Capillary Conditioning

New capillaries were conditioned by flushing for 15 min with 1 M NaOH at 60 °C. Then, they were rinsed for 5 min with deionized water and 15 min with phosphate buffer at 37 °C. The capillary was cleaned and conditioned prior to each injection as follows: (i) 2 min rinse with deionized water, (ii) 2 min rinse with 0.1 or 1 M NaOH, (iii) 2 min rinse with deionized water and (iv) and 2 min with phosphate buffer at 1000 mbar.

5.1.6 Procedure for the Enantioseparation of (±)-Catechin by CD-EKC

A 0.25% w/v HS-β-CD (prepared by dilution of 20% w/v HS-β-CD) in phosphate buffer was used as chiral selector. This solution was flushed through the capillary by applying approximately 1000 mbar for 120 s, thus the capillary was completely filled. The inlet and outlet vials of the separation system are free of chiral selector assuring an extremely low consumption of the CD. Ultrafiltrated and standard solutions were injected hydrodynamically at 50 mbar for 5 s. Separation was performed in normal polarity by applying 15 kV voltage. The capillary was thermostated at 37 ºC and the UV-detection wavelength was set at 220 nm. Peak areas corrected by migration times (in order to compensate for velocity discrepancies between peaks) were used as analytical signal. Since the resolution at the maximum calibration concentration (250 µM) was adequate (*Rs* ~ 2.5) no further optimization (out of the scope of this work) was performed. Enantiomers were identified by spiking (+)-catechin to the racemic mixture. HSA was also used as a chiral selector.

5.1.7 Procedure for the Interaction of (±)-Catechin–HSA by CE–FA

Approximate amount of Sodium dihydrogen phosphate dihydrates was dissolved in deionized water to make a final concentration of 50 mM adjusted with 0.1 M NaOH to pH 7.4. Stock solutions of 2 mM (±)-catechin and 1 mM HSA were prepared respectively by dilution of an adequate amount with 50 mM phosphate buffer. Two sets of mixtures were prepared: (i) concentration of HSA was kept constant at 525 µM while that of catechin was varied from 20-200 µM. (ii) concentration of catechin was kept constant at 40 µM while that of HSA was varied from 300-675 µM HSA. All the mixtures were placed in a water bath for 30 min at 36 ^{0}C followed by ultrasonication to remove dissolved gases prior to injection. Samples were injected hydrodynamically at 50 mbar into a buffer filled capillary for 60 s while the 15 kV was kept in the normal polarity during electrophoretic separation.

5.1.8 Pharmacokinetic Evaluation by Incubation, Ultracentrifugation and EKC

Racemic (±)-catechin (standards or mixtures with HSA) solutions were prepared by dilution of their corresponding stock solutions in PBS. All these mixtures were allowed to reach equilibrium for 30 min in the thermostated bath at 36.5 ºC and were filtered through the regenerated cellulose filters by centrifugation at 9000 rpm for 30 min. The ultrafiltrate was ultrasonicated during 4 min and injected into the EKC system.

5.2 Computational: Molecular Modeling and Docking

5.2.1 Computer Programs

Schrödinger's Maestro 9.1 was used as the primary graphical user interface (GUI) [127]. The Ligands were prepared using Ligprep [128]. The HSA protein was prepared using the protein preparation wizard [129-132]. All docking calculations were performed by using Glide 5.0 [133-135]. The docked poses were further used for the calculations of relative binding free energies using Prime MMGBSA [132].

5.2.2 Enantioresolution

In this study, the inclusion complexes between highly sulphated-β-cyclodextrin (HS-β-CD) and both catechin enantiomers were modelled by molecular docking methods to investigate the binding mode and to predict the elution order for chiral separation. The molecular structure of β-cyclodextrin (CD) was obtained from the crystal structure and the HS-β-CD was built by adding seven sulphite (SO_3^{2-}) functionalities to each primary hydroxyl group of β-CD. The initial structures of (+)-catechin and (-)-catechin were geometrically optimized considering quantum mechanical and molecular mechanics (QM-MM) force fields using the DMol3 as the quantum level in Discovery Studio (DS). Docking calculations were performed using the CDOCKER module of DS. The CDOCKER module is a grid-based molecular docking method where the receptor is held rigid while the ligands are allowed to flex during the refinement. Due to lack of information about the binding site, the whole

CD molecule was selected as a target and a binding sphere of dimensions; 0.587 (X), 0.421 (Y) and 0.485 (Z), was generated. The CHARMm FF was used as an energy grid force field for docking and scoring function calculations. Random ligand conformations were generated from the initial structure through high temperature molecular dynamics, followed by random rotations which were further refined by grid-based (GRID 1) simulated annealing and a final grid-based minimization. Of the 10 best poses, one (conformation) having highest docking score was used for the binding energy calculations and further analysis.

5.2.3 Enantioselective Binding of (±)-Catechin to HSA

5.2.3.1 Protein Structure Preparation

The typical structure file from the PDB is not suitable for immediate use in molecular modeling calculations. A typical PDB structure file consists only of heavy atoms and may include a co-crystallized ligand, water molecules, metal ions, and co-factors. Some structures are multimeric, and may need to be reduced to a single unit [129]. Moreover, since we are working under physiological conditions it is important to prepare the protein at pH 7.4 and to assign a protonation state for a co-crystallized ligand and the lowest energy state of the ligand will be used for further analysis. All the protein structures were prepared by using Maestro 9.1 protein preparation wizard [129]. The HSA protein complex structures used were 2BXD and 2BXF. The structures were pre-processed in the following order: (i) assigning bond orders, (ii) addition of hydrogens, (iii) creating zero-order bonds to metals, (iv) convertion selenomethionines to methionines, (v) orientation of hydroxyl groups and amide groups of Asn and Gln residues, (vi) optimization of His residues charge state (vii) removal of water molecules beyond 5 Å of hetero-atomic groups. A restrained minimization of the protein was performed using the default constraint of 0.30 Å RMSD and the OPLS 2005 force field.

5.2.3.2 Ligand Structure Preparation

The ligands were drawn by ChemDraw from Chem Office, Software Inc, 2001. Hydrogens were then added to the ligands and the geometry of the ligands was

cleaned before selected for further use. Maestro Gui. Ligprep [128] was used to prepare 3D coordinates for all ligands under study. Protonation states were generated using Epik at pH 7.4 in order to mimic physiological conditions [130]. All possible chiral conformations were generated. One low energy conformation of the ligands under study was generated.

5.2.3.3 Glide Grid Generation

Glide searches for favorable interactions between one or more ligand molecules and a receptor molecule, usually a protein. The shape and properties of the receptor are represented on a grid by several different sets of fields that provide progressively more accurate scoring of the ligand poses. Grids for each conformer of the receptors that adopt more than one conformation on binding can be prepared, to ensure that possible active conformations are not missed.

Glide can, however, handle different hydroxyl conformations with a single grid generation [135]. The Grid file for 2BXD containing R-Warfarin site 1 and 2BXF which contains Diazepam, a marker for site 2, were prepared. Ligand showing markers within the receptor were selected to ensure that the ligand is excluded from the receptor. The centroid of selected ligand was selected and the enclosing box occupying the active site that contains the ligand was visible in the Maestro Gui. The default size of the enclosing box was used since it was assumed that ligands of similar or less in size dock to the site markers, if for example we were going to dock ligands larger than those in the workspace the size of the enclosing box was going to be increased.

The Receptor Grid Generation panel was used to define Glide constraints for the receptor grids to be generated. Glide constraints are receptor-ligand interactions that are believed to be important to the binding mode, based on structural or biochemical data. Setting constraints enables Glide to screen out ligands, conformations, or poses that do not meet these criteria early on in their evaluation for docking suitability [133]. Positional constraint was selected by looking at a functional group in the ligand where a hydrogen bond can likely to occur. Accordingly a crystal structure of a prepared receptor ligand complex like 2BXD

where we selected the carbonyl group of Warfarin since it is responsible for a hydrogen bond with Tyr 150 and His 242 as shown in Figure 1.5. Ghuman and co-worker/s used Crystallography to study the binding of drugs to HSA and witnessed a similar binding pattern as the one observed in this work [43]. The hydrogen bond constraints were done by selecting a receptor residue that shows a potential to form hydrogen bonding with the ligand. Tyr 150 was chosen as the residue that is thought to exhibit a central role in site 1 drug interactions [43]. To further prepare a grid a hydrophobic constraint was generated by running a job to form a hydrophobic map of the receptor site indicating the hydrophobic region. After generating hydrophobic maps the rotatable residues in the receptor were selected. The hydroxyl groups in residues such as Ser, Thr, and Tyr can adopt different orientations with different ligands. Glide can therefore allow such groups to adopt different orientations when ligands are docked, to produce the most favorable interaction. For Ser and Thr, the hydroxyls can be oriented in any of the three local minima; for Tyr, they can be in either of the two local minima. The Rotatable Groups tab allows one to choose which of these hydroxyl groups should be treated flexibly. However, all the residues containing ¯OH groups were selected except for Tyr150 in Site 1 and Tyr411 for site 2 because they were chosen for hydrogen bonding. The grids were then generated by running a job that took a few minutes.

5.2.3.4 Glide Docking Methodology

Glide (Grid-based Ligand Docking with Energetics) is described as docking funnel that uses a series of filters to sample the protein binding site and search for acceptable poses [134]. Glide docking can be performed in two modes i.e. flexible mode and rigid mode. In the flexible mode the residues containing ¯OH groups in the active site are allowed to rotate or to move freely during docking and the rest of the protein is held rigid where there are no backbone or side-chain movements. In rigid docking mode the protein residue and side chains are held rigid during docking. In the flexible docking mode, Glide generates a set of conformers for each input ligand and then performs an exhaustive search for possible positions and orientations of ligand over the active site [136].

A ligand possesses a combination of position and orientation of a ligand relative to the receptor in flexible docking. The ligand poses that Glide generates pass through a series of hierarchical filters that evaluate the ligand's interaction with the receptor. The initial filters test the spatial fit of the ligand to the defined active site, and examine the complementarities of ligand-receptor interactions using a grid-based method pattern after the empirical Chem Score function [137]. Glide offers its greatest value when conformations are generated internally. For each core conformation, an exhaustive search of possible locations and orientations is performed over the active site of the protein. The search begins with the selection of "site points" on an equally spaced 2 Å grid that covers the active-site region. This selection is made by evaluation of distances from the site point to the receptor surface at a series of pre-specified directions and sorted into distance ranges ("bins") of width 1 Å. Poses that pass these initial screens enter the final stage of the algorithm, which involves evaluation and minimization of a grid approximation to the OPLS-AA non-bonded ligand-receptor interaction energy. Final scoring is then carried out on the energy-minimized poses. By default, Schrödinger's proprietary Glide Score multi-ligand scoring function is used to score the poses. The Glide Score is based on Chem Score, but includes a steric-clash term, adds buried polar terms devised by Schrödinger to penalize electrostatic mismatches, and has modifications to other terms:

$$\text{Gscore} = 0.065*\text{vdW} + 0.130*\text{Coul} + \text{Lipo} + \text{Hbond} + \text{Metal} + \text{BuryP} + \text{RotB} + \text{Site} \qquad (5.1)$$

The choice of best-docked structure for each ligand is made using a model energy score (*Emodel*) that combines the energy grid score, the binding affinity predicted by Glide Score, and (for flexible docking) the internal strain energy for the model potential used to direct the conformational-search algorithm.

In this work Glide SP (standard precision) mode was used as well as Glide XP (extra precision) mode as it has been observed that for Glide SP, scores of -10 or lower represents good binding. For some targets (e.g., with shallow active sites), scores of -8 or -9 might be very good. Metalloproteins and Glide XP both tend to produce lower docking scores (-12 to -15 or even less). Therefore, it is difficult to

compare the two methods since they are built on different paradigms. We opted to first dock our ligands with SP and re-dock the poses with Glide XP. Moreover, Glide Score is an empirical scoring function with many terms, including force field (electrostatic, van der Waals) contributions and terms rewarding or penalizing interactions known to influence ligand binding. It has been optimized for docking accuracy, database enrichment, and binding affinity prediction. Glide Score should be used to rank poses of different ligands, for example in virtual screening. While the XP Glide Score shares many terms with the SP and HTVS (High-throughput Virtual Screening) Glide Score, they do have significant differences and have been optimized separately; as a result, these two Glide Scores cannot be compared directly [134, 136].

For Glide XP, the pose-selection procedure is a bit more complicated, though it still involves Emodel and Glide Score. Poses produced by Glide XP have an "XP Pose Rank" property that shows how Glide XP has ranked the poses of a given ligand [133]. Glide sorts its results by Glide Score. This means that if you save multiple poses per ligand, the apparent ranking of poses for a given ligand (by Glide Score) will not reflect the actual ranking Glide used for pose selection. Therefore, looking at the value of Emodel or XP Pose Rank properties to determine the "best" pose for a ligand is the best option.

5.2.3.4.1 MM-GBSA Methodology

Molecular mechanics generalized born surface area (MM-GBSA) approaches can be applied as a way of including implicit solvation into the estimation of the free energy of ligand binding [133]. In this study, MM-GBSA calculations were carried out using the PRIME-MM-GBSA utility. As a post-docking step, docked ligand poses generated with Glide XP were used to calculate ligand binding energies and ligand strain energies for both (-)-catechin and (+)-catechin and a single receptor, using the MM-GBSA technology available with Prime. The flexible region of the protein was chosen and defined as any residue within 12 Å of the ligand in the active site. The Prime MMGBSA binding energy was then calculated for the previously generated poses by Glide XP.

CHAPTER 6

RESULTS AND DISCUSSION

This chapter focuses on the results obtained for the experimental and computational part of this work. The experimental work was undertaken in three parts which included the separation of the enantiomers using a chiral selector, followed by the interaction of the ligand with the protein using CE-FA and finally the enantioselective binding of ligands to protein using EKC.

The purpose of the computational work was two folds. Firstly, to get a better understanding of the conformational profile of the inclusion complexes between HS-β-CD and both (+) and (-) enantiomers, and to validate the order of elution obtained experimentally. Secondly, to better understand the binding modes of the enantiomers to the binding sites of HSA protein.

6.1 Separation of (±)-Catechin Enantiomers with CE-EKC

Figure 6.1 shows the electropherogram of the chiral separation of (±)-catechin at pH 7.4 in the absence of a chiral selector. The migration time of 10.6 min observed for (±)-catechin showed no separation.

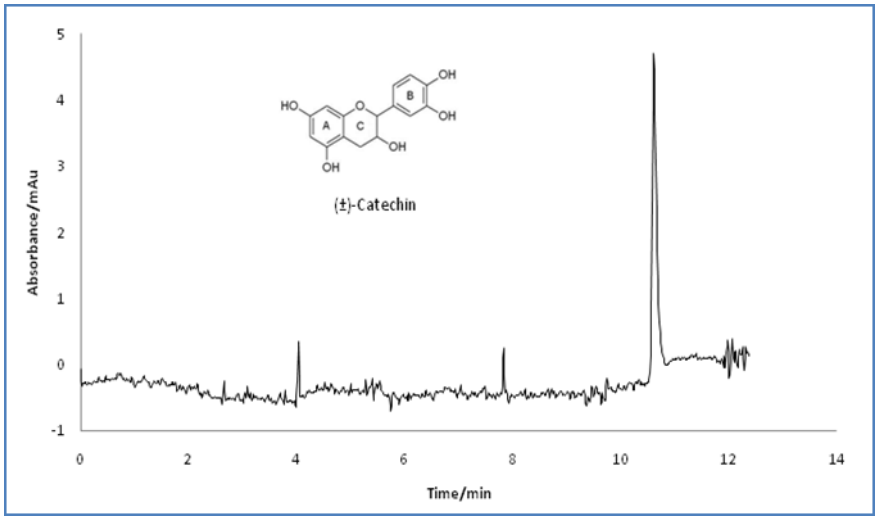


Figure 6.1: Shows the unresolved peak of (±)-catechin obtained in the absence of the chiral selector. Capillary thermostated at 37 ºC. BGE 30 mM phosphate buffer at pH 7.4. Normal polarity of 15 kV. The UV-detection wavelength was set at 220 nm.

The high migration time at pH 7.4 was due mainly to the neutral charge of (±)-catechin and the negatively charged molecules that have a reverse direction of electrophoretic mobility to the electroosmotic flow (EOF). A baseline resolution of (±)-catechin was achieved after the addition of the chiral selector HS-β-CD to the background electrolyte (BGE). The corresponding resolution profile is shown in Figure 6.2. The (-)-catechin enantiomer eluted at 13.4 min while the second (+)-catechin enantiomer eluted after 13.8 min. Enantiomers were identified by spiking the racemic standard with (+)-catechin because of its commercial availability. The electropherogram of the spiked standard is shown by dotted line in Figure 6.2. It is observed that the absorbance signal of the first peak is tripled, confirming that (+)-catechin eluted first.

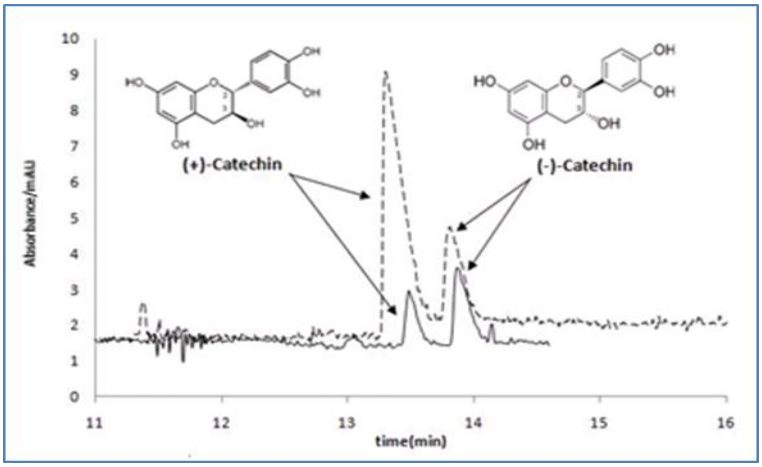


Figure 6.2: EKC electropherogram of (—) 100 µM (±)-catechin baseline resolved in the presence of 0.1w/v HS-β-CD and (---) superimposed electropherogram for a 100uM (+)-Catechin spiked solution. Capillary thermostated at 37 ºC filled with HS-β-CD by applying 1000 mbar for 120 s. BGE 30 mM Sodium dihydrogen phosphate buffer at pH 7.4. Normal polarity of 15 kV. The UV-detection wavelength was set at 220 nm.

These results clearly suggests that the mechanism of the two complexes (adduct) formed are different, and therefore results in the separation of the enantiomers as depicted in Figure 6.2. However in Figure 6.3 it can be observed that the (±)-catechin ligand migrates faster than the chiral selector HS-β-CD thus illustrating the complex formation along the migration towards the outlet or cathode end of the capillary. The direction and migration speeds of the molecules within the

capillary are dependent on the charge and size of the molecules. Hence the order of injection is to be varied in such a way that there will be interaction between the racemic compound and the chiral selector. Accordingly in this work the chiral selector was injected prior to the injection of (±)-catechin, as depicted in Figure 6.3.

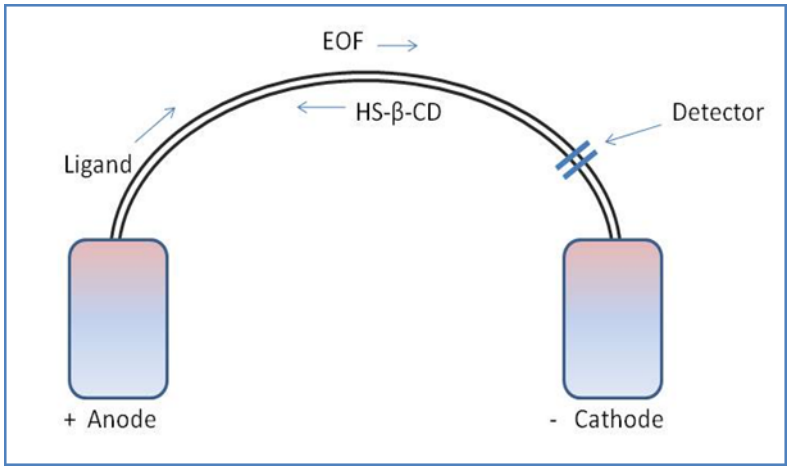


Figure 6.3: Electrokinetic chromatographic mechanism of separation when HS-β-CD is used a chiral selector.

In case the chiral selector is either bound to the capillary surface or incorporated or bound to a gel matrix, the net velocity of the complex will be zero. In case the chiral selector is added to the BGE, the net velocity of the complex will not necessarily be zero, but (in most cases) differ from the velocity of the free analyte as can be seen in Figure 6.4. The complex formation observed has a different average velocity of the analyte than that of the free analyte. Consequently we can postulate that the difference in complex stability of the two enantiomers leads to difference in the average velocity of these compounds. Therefore conditions are maximized such that there is significant difference in average velocity between the two complexes, which favors better enantioseparation.

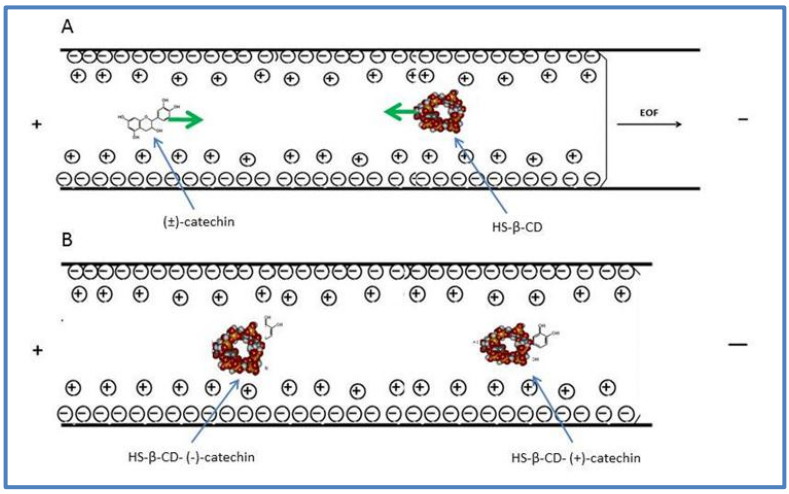


Figure 6.4: EKC mechanism electropherogram of (—) 100 µM (±)-catechin baseline resolved in the presence of 0.1w/v HS-β-CD and (---) superimposed electropherogram for a 100uM (+)-Catechin spiked solution. Capillary thermostated at 37 ºC filled with HS-β-CD by applying 1000 mbar for 120 s. BGE 30 mM phosphate buffer at pH 7.4. Normal polarity of 15 kV. The UV-detection wavelength was set at 220 nm.

The resolution of the two peaks observed in Figure 6.2 are affected by the change in pH of the BGE, the concentration of the chiral selector present in the BGE and the capillary temperature. Hence the efficiency of separation of two compounds assessed by their resolution is given by equation 4.7 (see chapter 4). In this work, the resolution was found to be 1.8, a characteristic feature of separated compounds. Since the area of a peak is proportional to the analyte concentration (peak height) and peak residence time in the detector (peak width), a direct estimation of the peak area of components resolved within a CE separation requires peak area normalization. If the data are not corrected then an overestimation of the latter migrating enantiomer will occur in terms of area count [138, 139]. In this work very small amounts of HS-β-CD chiral selector were used to achieve a baseline resolution of (±)-catechin, but the separation is dependent on the concentration of the chiral selector used. Furthermore, it was observed that the fluctuation of the electro-osmotic flow (EOF) adversely influenced the electrophoretic mobility, thus careful control or correction of EOF was required. Changes in the EOF were mainly due to changes in the zeta potential. However, the coating of the inner wall capillary used to control the EOF could be used to avoid solute adsorption which may lead to peak broadening and

inaccurate mobility measurements [75]. For this purpose, the surface of the inner wall was continuously renewed with sodium hydroxide throughout this work.

In addition to separation with HS-β-CD, another attempt for the separation of (±)-catechin was performed using HSA protein as a chiral selector. The HSA protein is known to exhibit the highest potential enantioselectivity amongst the plasmatic proteins as discussed in chapter 1. This makes enantioselectivity a key parameter for enantioresolution when using protein as a chiral selector. Thus, the concentration of HSA was varied from 50 to 200 µM to achieve maximum enantioselectivity. As the concentration increased the corresponding peak also increased, but there was a negative impact on the (±)-catechin signal as it was suppressed resulting in an improved separation on the enantiomers as can be seen in the figure below.

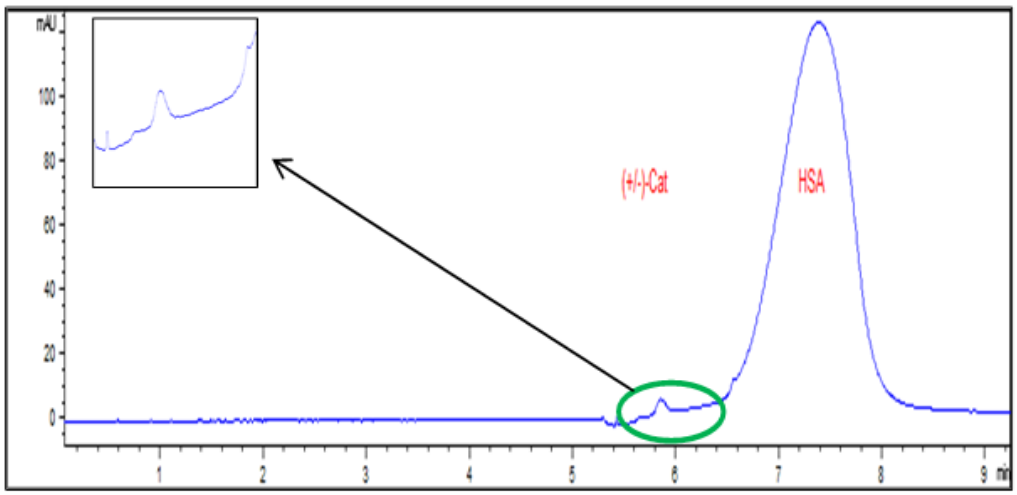


Figure 6.5: Shows the unresolved peak of (±)-catechin obtained when HSA is used as chiral selector. Concentration of HSA. Capillary thermostated at 37 ºC filled with HSA by applying 1000 mbar for 5 s. BGE 30 mM Sodium dihydrogen phosphate buffer at pH 7.4. Normal polarity of 15 kV. The UV-detection wavelength was set at 220 nm.

In Figure 6.5, the peaks for the ligand and protein were easily identified because of the strong absorbance of HSA in the UV region as compared to the racemic catechin or any other ligand. Furthermore, a significant difference in migration times of 5.92 and 7.38 min for the ligand and protein were observed respectively. This difference depicts that (±)-catechin was loosely bound to HSA within the capillary, hence it was able to migrate faster. It is clear that the chiral

resolution of (±)-catechin with HSA as a chiral selector was impossible resulting in a single peak observed in the electropherogram.

6.1.1 Enantiomer Excess/Enantiomeric Excess (ee)

The enantiopurity of a substance is usually reported in terms of percent enantiomeric excess (ee %). It represents the percent by which one enantiomer E_1 is in excess in a mixture of the two compounds (E_1, E_2). It can be used to measure the purity of substances especially in the synthesis of chiral compounds. The enantioseparation results obtained in Figure 6.2 showed that the HS-β-CDs possess good chiral recognition abilities to (±)-catechin, and therefore enantiomeric excess was determined based on the absorbance signals of the corresponding resolved peaks. Enantiomeric excess was found to be 54 % for (-)-catechin and its reproducibility was also studied over a four day period with a view to establishing possible conformational changes with time. The results shown in Appendix 2A confirm that there were no changes in the concentration of the enantiomers, hence the stock solutions were deemed to be stable for the duration of four days. It is vital that a reproducibility study is performed prior to interaction studies so that the experimental design can be prepared accordingly, especially when the analysis is to be done on a sequential basis over a 24 hour period.

6.1.2 Factors affecting Enantioresolutions

6.1.2.1 pH

The migration times influenced by the charge of both the analyte and the anionic chiral selector as well as by the EOF is depicted in Figure 6.6 [12]. The poor resolution observed at low pH is due to high mobility within the capillary arising from the positively charged ions which are attracted towards the negatively charged electrode (see Figure 6.6 below). As the pH increased, the system shifted in an opposite direction, thus resulting in an improved resolution between the two enantiomers.

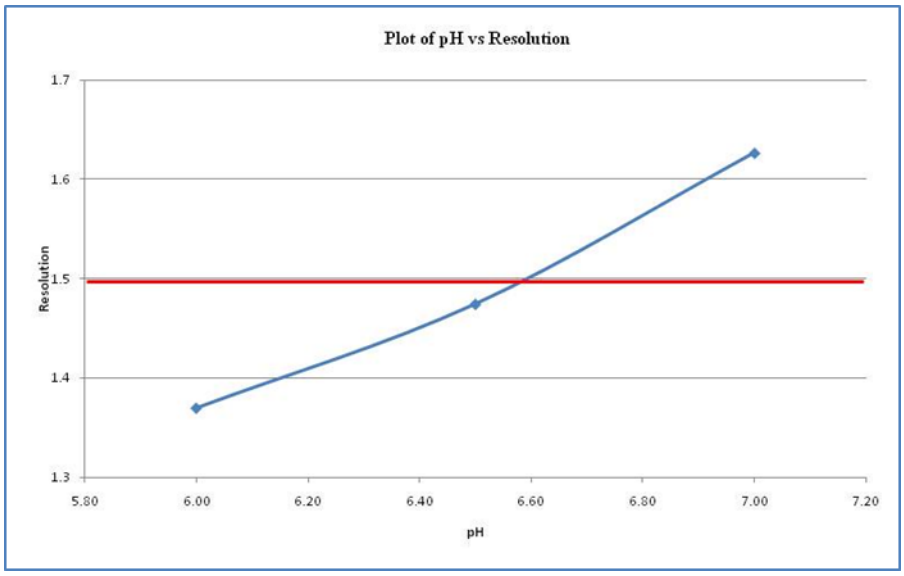


Figure 6.6: Effect of pH on the resolution of 100 µM (±)-catechin in the presence of the HS-β-CD. 15 kV, 30 mM Sodium dihydrogen phosphate buffer at pH 6, 6.5 and 7.

Apart from the high-resolution capability of HS-β- CDs in CE, it is worth mentioning that peak tailing were very seldom observed within the evaluated pH range. The rest of the electropherograms are shown in Appendix 2B.

6.1.2.2 Ionic Strength of the Buffer

The selection of the appropriate BGE (type of buffer and concentration, ionic strength, pH) is of paramount importance in order to achieve the optimum experimental conditions for a successful separation of enantiomers. In this work, studies on the effects of the buffer concentration on the enantiomeric resolution were carried out with racemic catechin. The resolution of catechin enantiomers were calculated using Eqn. 4.7 (see chapter 4). The efficiency of the separation increased, when the concentration increased from 20 to 50 mM, while keeping the chiral selector at 0.25% w/v as can be seen in Figure 6.7. In general, varying the concentration of the buffer affected both the migration time and the EOF.

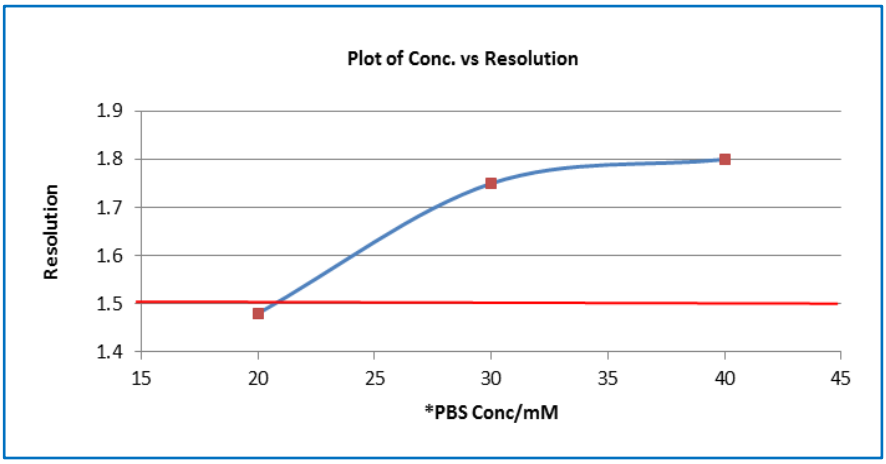


Figure 6.7: Effect of Sodium dihydrogen phosphate buffer (PBS) with pH 7 on the resolution of 100 µM (±)-catechin in the presence of the HS-β-CD. Polarity 15 kV. *PBS conc: concentration of phosphate buffer

In this case the migration time increased as the concentration of the buffer was increased because of the labile diastereomeric complex formation with HS-β-CD during the electrophoretic run, thus decrease the mobility of the analytes towards the detection window. This decrease is also dependant mainly on the concentration of the chiral selector and on the rate constant of the complex formed. Moreover, the effect of concentration with a high ionic strength BGE is counter-balanced by an increased Joule-heating as a consequence of an increased current. The addition of a relatively low concentrated chiral selector, HS-β-CD, allowed for the achievement of an efficient separation with a shorter migration time.

6.1.2.3 Column and Cassette Temperature

The capillary temperature is another important parameter to be controlled in chiral separation analysis using electromigration methods. The total analysis (or migration) time are affected by temperature because both the thermodynamics and kinetics of adsorption processes are functions of temperature. In order to study the effects of the capillary temperature on the optical resolution of the analytes using HS-β-CD, experiments were carried out by varying the temperature from 15 to 25 °C as shown in Figure 6.8.

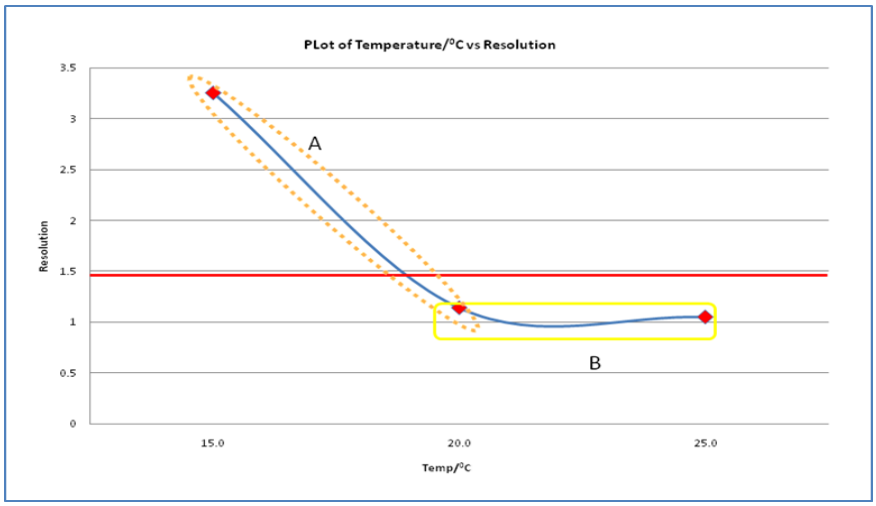


Figure 6.8: Shows a relationship between temperature of the capillary and resolution. Resolution above the red line (Rs =1.5) is regarded a based separation between the two peaks. Normal poarity 15 kV, BGE 30 mM Sodium dihydrogen phosphate buffer at pH 7.

In part A, it is observed that the resolution decreases drastically due to the minimal time for interaction of the enantiomer with the chiral selector. However in part B the high temperature tends to decrease the stability of the labile diastereomeric complexes formed between the analytes and the chiral selector. In addition, there is minimum interaction with the chiral selector hence the constant resolution observed, due mainly to differences in the binding constant of the enantiomers to the chiral selector rather than to the mobilities of the complexes [140]. Related electropherograms are shown in Appendix 2C.

6.2 Simulation of (±)-Catechin Chiral resolution

Molecular docking was further performed to assess the interactions between HS-β-CD and each enantiomer of catechin and to explain their differential separations in CE. The docked complexes of (+)-catechin and (-)-catechin with the HS-β-CD suggested important interactions and are diagrammatically shown in Figure 6.9, while the different docking results are summarized in Table 6.1.

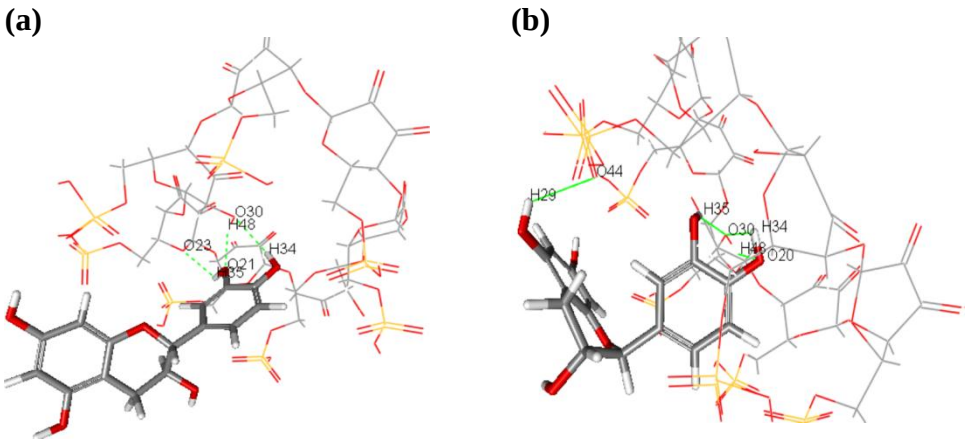


Figure 6.9: Docked complexes of (+)-catechin (a) and (-)-catechin (b) with HS-β-CD. Ligand is presented in stick form, while HS-β-CD is shown as Lines. All interacting atoms are labeled. Hydrogen bonds are shown as dotted green lines.

Clearly, it can be observed that the evolution of the enantiomers near the ends of the cyclodextrin (CD)are similar as they tend to locate themselves on the cavity axis while introducing a single ring (B-ring) inside the CD.A closer inspection of Figure 6.7 revealed that the geometric orientations of both enantiomers were similar and both complexes are stabilized by electrostatic (O---H) and hydrophobic interactions between the phenyl ring (Ring B, Figure 6.10) of ligands and cavity of the HS-β-CD. Of the two possible binding modes (A, B, Figure 6.10), only one mode (A) was observed for both enantiomers in their complexes. Both (+)-catechin and (⁻)-catechin penetrate partially into the cavity of HS-β-CD, where two hydroxyl (-OH) functionalities in the B-ring of both enantiomers form hydrogen bonds with the hydroxyl (-OH) groups of the HS-β-CD.

Table 6.1: Docking results obtained for complexes of (+)-catechin and (-)-catechin with HS-β-CD.

Ligand (Lig)	Quantum Energy (QM)	No. of H-bonds	Donor-acceptor	Bond Distance (Å)	CDocker Interaction energy ($kcalmol^{-1}$)	Binding energy* ($kcalmol^{-1}$)
(+)-catechin	-3996.1	3	Lig (A-RING)34H---O30(CD)	2.43	-21.6	-38.02
			Lig (A-RING)21O---H48(CD)	2.16		
			Lig (A-RING)35H---O23(CD)	2.37		
(-)-catechin	-3995.0	4	Lig (B-RING)29H---O44(CD)	2.40	-23.3	-39.74
			Lig (A-RING)35H---O30(CD)	1.86		
			Lig (A-RING)34H---O30(CD)	2.36		
			Lig (A-RING)20O---H48(CD)	1.88		

*Binding energy = Energy of Complex - Energy of Ligand - Energy of Receptor

However, the extent of hydrogen bonding was more predominantin the case of (-)-catechin (4 bonds, Table 6.1) suggesting the additional participation of an A-ring hydroxyl group in hydrogen bonding with the sulphate moiety of HS-β-CD. Moreover, the measured hydrogen bond distances were comparatively shorter in the complex of (-)-catechin with the HS-β-CD. The lower CDocker interaction energy of -23.3 kcal/mol for (-)-catechin further suggested that the energy required for its proper interaction with the HS-β-CD is lower compared to the (+)-catechin. These observations clearly show that numerous C–H---O hydrogen bonds are formed. The resulting host–guest interactions therefore comprise a third significant component besides the O–H---O hydrogen bonds and van der Waals interactions.

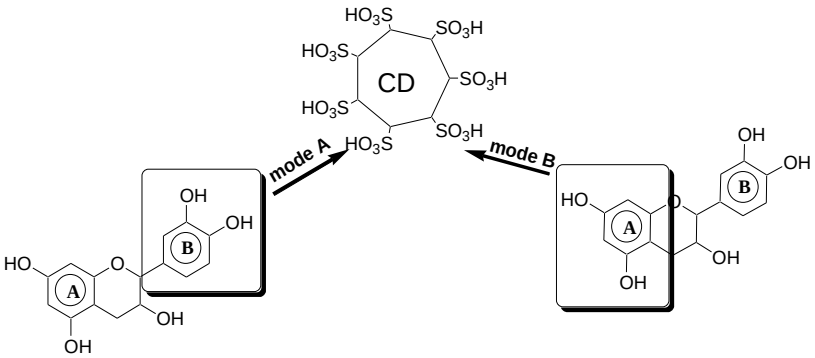


Figure 6.10: Two different binding modes of catechin towards the HS-β-CD.

It is reported that the electrostatic host-guest interactions plays a significant role in the energies of binding of a number of cyclodextrin inclusion complexes. Moreover, the greater strength of the hydrogen bond energies over van der Waals interactions are well documented in literature [141, 142]. From the present docking results the computed free energy of binding (Table 6.1) suggested that the (-)-enantiomer has a stronger propensity to interact with HS-β-CD relative to the (+)-enantiomer. Furthermore, the most stable (Figure 6.9) position for the enantiomers is not located at the cavity centre, in both enantiomers of catechin. Our results suggest that, the Coulombic forces significantly contribute to the enantiodifferentiation; while the van der Waals term is the main contributor to interaction energy, and therefore determines the orientation of the guest molecule, it is the electrostatic term that results in the difference in energies between the most stable position of (+) and (-

)-catechin. The position and orientation of the guest molecule for which the lowest potential energy was obtained, is postulated to be the geometry of the inclusion complex for both enantiomers. The binding energies of -39.74 kcal/mol and -38.02 kcal/mol, for the complexes of HS-β-CD with the (-)-enantiomer and (+)-enantiomer respectively, were found to be consistent with the elution order observed in the CE experiments, and thus explains the stronger (-)-enantiomer interaction with HS-β-CD. The difference in interaction energies of the two enantiomers (E = -1.68 kcal/mol), represents the energetic contribution to the enantioselectivity, and thus plays a significant role in the enantioseparation. The magnitude of the change in energy signifies of the driving force towards complexation. The more negative the binding energy, the more thermodynamically favorable is the inclusion complexing nature [141]. The binding geometries in the two complexes are different, but they are affected by numerous host–guest C–H--O interactions, resulting from induced fits of the hosts toward the guests. It is believed that the C–H---O bonds can make energetically favourable contributions to the stability as well as to the strength of host–guest interactions [142]. In both (+)-catechin or (-)-catechin inclusion complexes, several C-H--O hydrogen bonds were observed. In general, hydrogen bond is deemed present when the O– distance is ±2.5 Å and if the O–H---O angle is $\geq 120^0$. The C–H---O contacts have been found in a number of CD inclusion complexes. These C–H---O contacts show the stereochemical features and directionality of conventional hydrogen bonds (H bonds), called C–H---O bonds. These can be assigned by $–CH_2$ and –CH groups of the HS-β-CD inner cavity to guest molecules, or from a guest molecule to O atom of a glucopyranose unit. These bonds are relatively weak but considered as a third, cohesive host–guest interaction adding on O–H---O hydrogen bonds and van der Waals contacts. They contribute to the overall stability and structure of the inclusion complexes. On comparing the total energy of (+)-catechin and (-)-catechin in Table 6.1, it is clear that there are no meaningful differences outside the cavity. Between the different contributions, the H-bonding term practically does not influence the enantiodiscrimination due to its small contribution to the total potential. Inside the cavity the main differences between (+)-catechin and (-)-catechin are observed from the narrow rim to the broader cavity and are due to the Lennard-Jones potential.

6.3 Pharmacokinetic Evaluation by Capillary Electrophoresis-Frontal Analysis (CE-FA)

The interaction studies of the HSA protein with CE-FA included (+)-catechin and (±)-catechin. External calibrations were initially performed with concentrations ranging from 20 to 200 µM for both compounds. It is worth mentioning that the peaks for the Frontal Analysis (FA) are flat at the apex, unlike EKC or CZE. This is mainly due to the larger volume of sample injected in the capillary. All standards for (+)-catechin showed good reproducibility of the migration times resulting in well superimposed peaks observed in Figure 6.11.

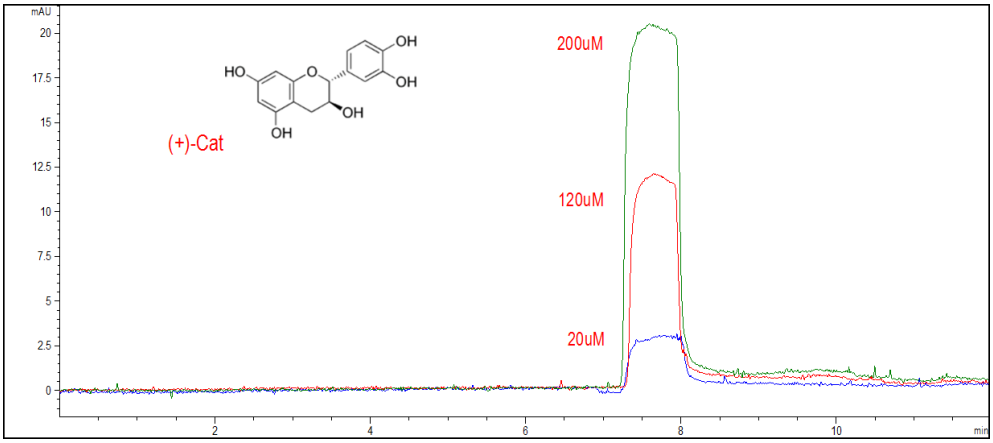


Figure 6.11: CE-FA electropherogram showing calibration standards of 20, 120, and 200uM of (+)-catechin injected by applying 1000 mbar for 60 s. Capillary thermostated at 37 ºC. BGE 30 mM Sodium dihydrogen phosphate buffer at pH 7.4. Normal polarity of 15 kV. The UV-detection wavelength was set at 220 nm

In Figure 6.12, it is observed that (±)-catechin eluted as a single compound at 220 nm but the migration time was not consistent, with an increase concentration. This change in migration time was taken into account during the preparation of the calibration curve by performing peak normalization. The correlation coefficients for compounds (+)-catechin and (±)-catechin were 0.9963 and 0.9875 respectively. The external calibration curves for both compounds are shown in Appendix 3A.

In general, CE-FA binding studies involve either measuring the change in the solute electrophoretic mobility caused by the addition of another entity to the system,

or by estimating the unbound concentration of the analyte in a mixed solution of associated and non-associated materials [22].

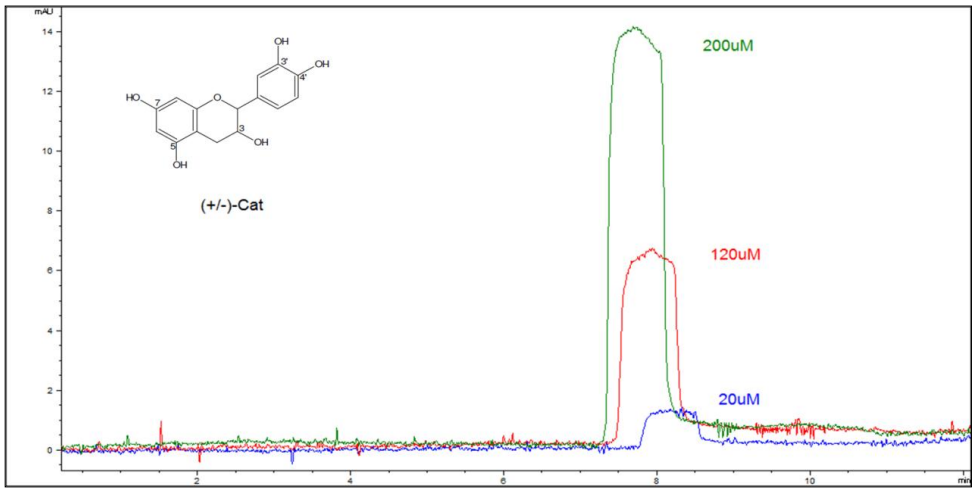


Figure 6.12: CE-FA electropherogram showing calibration standards of 20, 120, and 200 µM of (±)-catechin injected by applying 1000 mbar for 60 s. Capillary thermostated at 37 ºC. BGE 30 mM Sodium dihydrogen phosphate buffer at pH 7.4. Normal polarity of 15 kV. The UV-detection wavelength was set at 220 nm

In fact, based on the law of conservation of mass, unless the mobility of the protein and the drug–protein complex are equal, the free drug/ligand peaks shown in Figures 6.13 and 6.14 does not reflect the equilibrium concentrations of the free drug/ligand, but reflects a steady state of binding affinity [100].

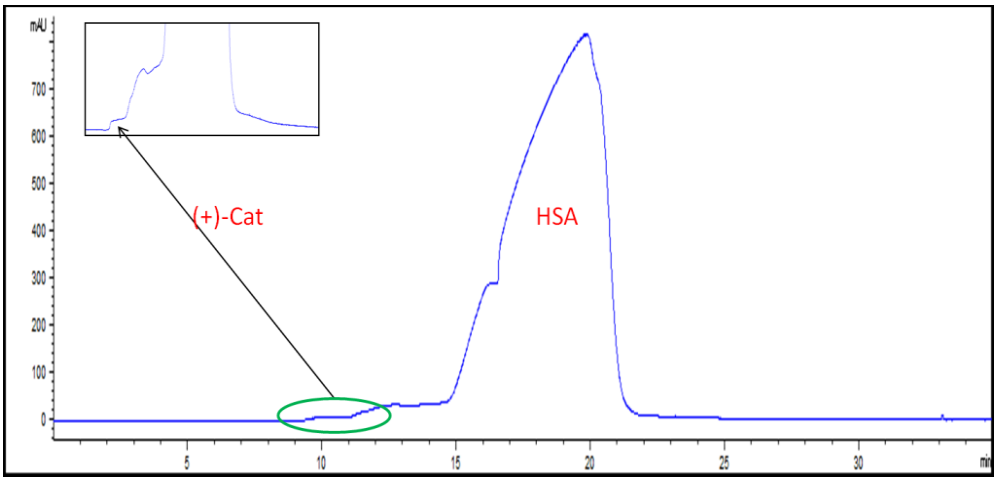


Figure 6.13.: CE-FA electropherogram of 80uM(+)-catechin and 525uM HSA mixture obtained using 50 mM Sodium dihydrogen phosphate buffer, 60s injection time, and 15 kV.

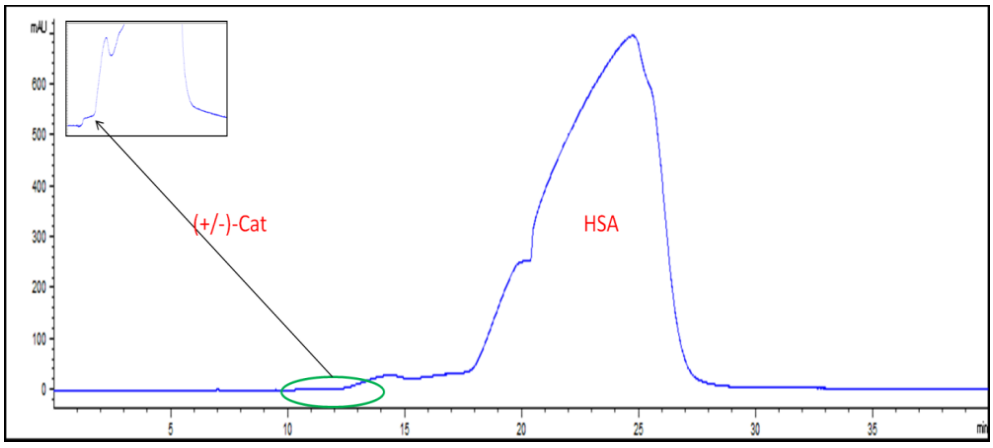


Figure 6.14: CE-FA electropherogram of 80 µM (±)-catechin and 525 µM HSA mixture obtained using 50 mM Sodium dihydrogen phosphate buffer, 60 s injection time, and 15 kV.

One of the most important factors affecting CE-FA (i.e., a mixture of drug and protein samples injected into a capillary containing only neat buffer) is in the binding equilibrium during electrophoresis. During the analysis, bound drugs remained within the protein zone. Injection of the larger sample plugs gave rise to the appearance of plateau peaks. The height of the free drug plateau peak is proportional to the free drug concentration in the original sample and the degree of binding can be determined with the aid of an external calibration curve of the absorbance signals as shown in Figures 6.11 and 6.12. It is also interesting to note that the absolute migration time of compounds progressively decreased resulting in the application of relative migration times throughout the analysis results being more accurate measure of the binding properties. As HSA binding mostly affects the activity and distribution of the compound above 95% binding, it is therefore important that strongly bound compounds elute within a reasonable time from the separation column in a reproducible order [143]. This can be observed in Figures 6.13 and 6.14, ligands eluted before the protein and thus were easily identified. The apex of the two peaks started around 9.52 and 10.65 min for (+)-catechin and (±)-catechin respectively while HSA was observed around 19.21and 24.61 min. In addition, peak widths of ± 6 and ± 9 are observed. This indicates a greater interaction occurring between HSA and

the (±)-catechin, as compared to (+)-catechin. Another important parameter that requires careful attention for CE-FA analysis is the pH of both the preparation and separation medium. Experimental binding affinities for the ligands measured at pH ≤ 7.4 could produce a tendency to over predict pharmacokinetic parameters relative to those reported in the ligand–protein database. Although this is an attractive explanation, it is to be expected that the pH dependency of ligands and receptors will be highly dependent on the nature of the ligand and the receptor. It is also important to consider the behaviour of each chiral form of a compound since their binding to HSA can be stereoselective in nature. When any immobilization methodology like High Performance Affinity Chromatography (HPAC) is used, a relatively low activity is observed which probably is due to denaturing, steric hindrance or improper orientation of the HSA during the immobilization process. Moreover the immobilized protein has non-binding regions and thus does not mimic the physiological conditions. Fortunately, the CE-FA approach used in this study was not affected by these non-binding regions, because both the ligand and protein were in solution.

6.3.1 D and P constant experiments

For interaction studies, the use of the exact concentrations is preferred because the mass used is crucial for the understanding the nature of the interaction. Therefore the exact concentrations calculated and represented as D_{exp} in Tables 6.2 and 6.3. These data suggests that the experimental design structured is such that either the D or P parameter was held constant while the other was varied. On comparing the two designs for (+)-catechin and (±)-catechin it was observed that P constant design was not consistent as can be seen in Figures 6.15 and 6.16 for (+)-catechin and (±)-catechin respectively. Detailed results are shown in Appendix 3A (for (+)-catechin-HSA) and 2B (for (±)-catechin-HSA) respectively. On comparing the Klotz and varifaction plots of D and P constants (see Figures 6.15 and 6.16), it is observed that the experiments for D constant design have determination points that are more scattered.

Table 6.2: Experimental design (ID 1 to 6 P constant and 7 to 12 D constant design) for (+)-catechin and pharmacokinetic parameters obtained using MATLAB calculations

ID	D, µM	D_{exp},µM	P, µM	P_{exp},µM	D, µM	logK_1-PB	logK_1	log(K_1n_1)
1	20.00	18.26			11.10		3.052	3.057
2	**40.00**	**36.52**			**25.00**		**2.908**	**2.917**
3	80.00	73.04	525.0	513.2	45.50	3.2027	3.041	3.063
4	120.00	109.55			58.50		3.211	3.255
5	160.00	146.07			71.60		3.286	3.353
6	200.00	182.59			99.00		3.199	3.275
7			300.0	293.2	26.10	3.446	3.086	3.100
8			375.0	366.5	26.40	3.348	2.969	2.980
9	40.00	36.52	450.0	439.9	15.40	3.269	3.464	3.485
10			**525.0**	**513.2**	**19.00**	**3.203**	**3.226**	**3.240**
11			600.0	586.5	21.70	3.269	3.464	3.485
12			675.0	659.8	30.50	NaN	NaN	NaN

Table 6.3.: The experimental design (ID 1 to 6 P constant and 7 to 12 D constant design) for (±)-catechin and pharmacokinetic parameters obtained using MATLAB calculations

ID	D, µM	D_{exp},µM	P, µM	P_{exp},µM	*d,µM*	logK_1-PB	logK_1	log(K_1n_1)
1	20.00	18.95			18.10		NaN	NaN
2	**40.00**	**37.90**			**30.90**		**2.637**	**2.643**
3	80.00	75.79	525.00	522.95	49.60	3.194	3.004	3.027
4	120.00	113.69			58.50		3.259	3.307
5	160.00	151.58			76.30		3.278	3.346
6	200.00	189.48			87.70		3.344	3.438
7			300.00	298.83	36.90	3.437	1.957	1.959
8			375.00	373.53	35.00	3.340	2.346	2.349
9	40.00	37.90	450.00	448.24	28.90	3.262	2.842	2.851
10			**525.00**	**522.95**	**30.50**	**3.194**	**2.666**	**2.673**
11			600.00	597.65	35.80	3.136	1.992	1.993
12			675.00	672.36	36.70	3.086	1.687	1.688

D: Total drug/ligand concentration, D_{exp}: Total drug/ligand concentration used during the experiment.

P: Total concentration of protein, P_{exp}: Total concentration of protein used during the experiment.

d: free drug/ligandconcentration

Both Figures 6.15D2 and 6.16D2 have negative gradiants (line in green), while P constant design in Figures 6.15D1 and 6.16D1 shows promising plots with positive gradiants. Therefore only experiments of the P constant parameter were considered for the varification of binding sites. On the other hand, experiments of D constant parameter were omitted for the discussion since they provided negative Klotz plots, negative varification plot, and scattered points for SIMPLEX model (see Figures 6.15 and 6.16).

On comparing the interaction experiments (for P constant design) of (+)-catechin-HSA to (±)-catechin-HSA some similarities and discrepancies are observed in terms of the binding constants and the number of binding sites estimated. In Table 6.3, it observed that the log*K* values in both experiments are lower in the first three (ID 1 to 3) experiments than the last three (ID 4 to 6) experiments. These observations clearly suggest the crucial nature for the selection of the total concentration of the ligand/drug to be used. However in (±)-catechin-HSA experiment, the log*K* value was reported as NaN (not available number) for ID 1, because a negative integer (negative binding) is practically impossible. This is probably due to the presence of a small quantity of each enantiomer available in the (±)-catechin standard as the mass weighed is equivalent to the sum of the enantiomers. If the total mass of the total drug is small, it is therefore difficult to quantify the free drug (*d*) using CE-FA approaches, because it is determined in the presence of the protein (no ultracentrifugation to remove the protein complex). Thus such points are considered as outliers and are therefore eliminated from the design prior to the implementation of the SIMPLEX model calculations. All points left were subjected to a Z-score and Grubbs test for the determination of outliers. Consequently there were five remaining points in the design (P constant of (±)-catechin-HSA experiment). Results are shown in Appendix 3B.

Plots from the simplex model relating d to P parameters are shown in Figures 6.15E2 and 6.16E2. The reliability of the simplex model arises from the simultaneous consideration of variables, whereas the linear plots (Klotz, Scatter and Y-regressions) forces the system to be linear based on the assumption that the number of binding site is unity (n = 1).

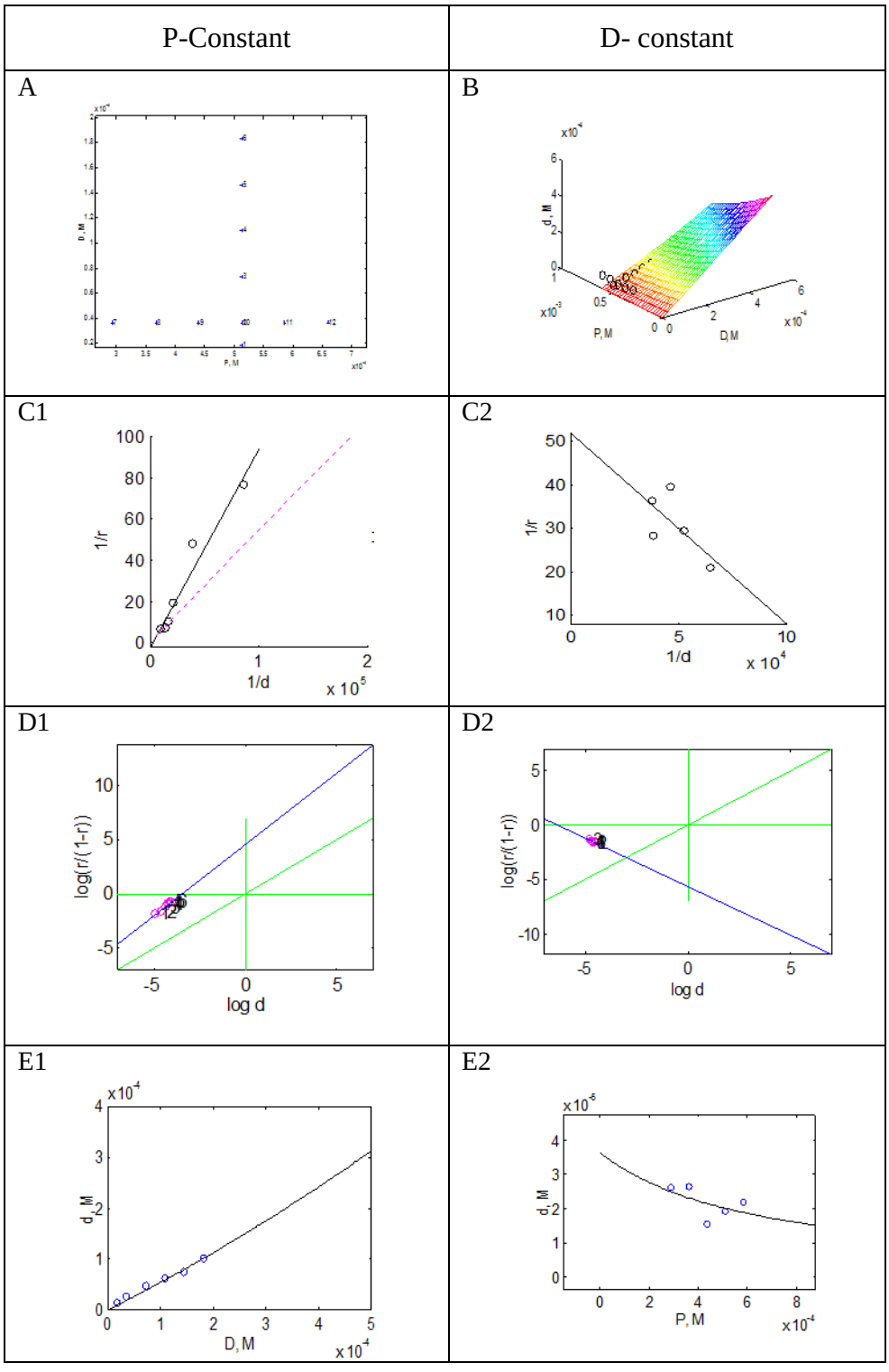


Figure 6.15: Frontal Analysis results for (+)-catechin, A and B; structure of the experimental design showing D and P constant designs simultaneously, C1 and C2 Klotz plot for estimation of log*K*:, D; plot showing the varification of n=1 where n is the slope of the blue line. If the two lines are parrallel it means the value of n is 1. E; shows the nonliner plot obtained using the SIMPLEX algorithm (Script in Appendix 5).

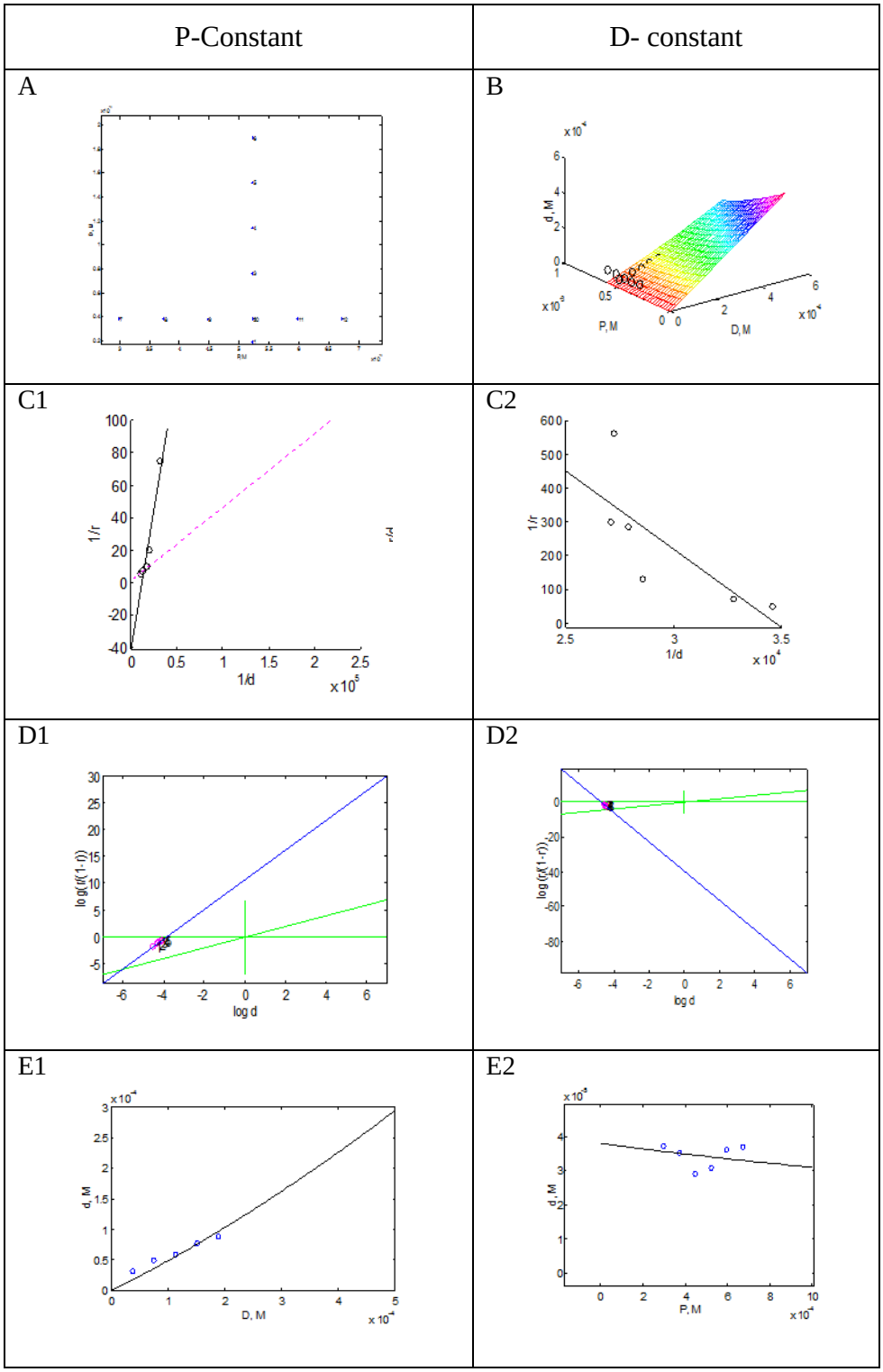


Figure 6.16: Frontal Analysis results for (±)-catechin, A and B; structure of the experimental design showing D and P constant designs simultaneously, C1 and C2 Klotz plots for estimation of log*K*:, D; plot showing the varification of n=1 where n is the slope of the blue line. If the two lines are parrallel it means the value of n is 1. E; shows the nonliner plot obtained using the SIMPLEX algorithm (Script in Appendix 5).

6.3.2 Number of Binding Sites (n)

The number of binding sites obtained from the two D and P experiments, is another interesting parameter to be considered due to the presence of racemic and non-racemic compounds. In (+)-catechin-HSA experiment, the value of n is relatively constant and closer to 1 as can be seen for the product of log*Kn* values shown in Table 6.2. This suggests that there is a binding site responsible for the interaction. However in (±)-catechin-HSA experiment, there is a variation in the product of log*Kn* values as shown in Table 6.3, due to the fact that only one enantiomer can be accommodated in the binding site. Thus forcing other compound to go into the secondary binding site as the value of n increases with an increase in concentration. Interestingly, this interaction or binding mechanism requires the verification of the n stoichiometry, as discussed in Chapter 3. Overall the CE-FA results are not conclusive for the estimation of n where a racemic compound is used, therefore an enantioselective binding of (±)-catechin to HSA was evaluated with a slightly different approach which included three major stages namely; incubation, ultracentrifigation and EKC.

6.4 Enantioselective Binding of (±)-Catechin to HSA

6.4.1 Mathematical Deficiencies and a Potential Alternative

According to the law of mass action, in the case of the present study, the mathematical model in Eqn.3.4.1 (chapter 3) is applicable for each enantiomer. These assumptions should be better satisfied for low *D*/*P*-values which is the normal situation in physiological conditions (*in vivo*), but not always followed in the *in vitro* experiments [4]. Eqn.3.4.1 is too complex to provide reliable estimations for the parameters n_i and K_i. However, HSA normally exhibits a single-high affinity class site when binding to smaller molecules, thus $m = 1$ is often adopted (without further attention), although other secondary lower-*K* class sites ($m > 1$; multiple binding) could be responsible for bad estimations [18]. In the case of $m = 1$, Eqn.3.4.1 can be simplified providing a two-parameter (n_1 and K_1) model, which has extensively been approximated in the past, transformed into linear regression equations which are mathematically inconsistent and lack of robustness [4, 18]. In order to avoid the simultaneous estimation, most authors assume a 1:1 stoichiometry ($n_1 = 1$) without further verification. However, as an alternative to these deficient approaches, the following two unexplored equations based on assuming $m = 1$ have been postulated in this work:

$$K_1 = \frac{1}{d}\frac{r}{n_1 - r} \tag{6.1a}$$

and

$$n_1 = r\frac{(1 + K_1 d)}{K_1 d} \tag{6.1b}$$

Both equations reflect the fact that n_1 and K_1 are dependent variables. The combined use of Eqn. 6.1a and 6.2b, proposed in this work as an alternative to the approaches used in the past, requires a particular algorithm based on the following criteria: (i) previous assumption of n_1 ($n_1 = 1$ is the default value); (ii) estimation of K_1 by means of Eqn. 6.1a (note: each *D*-*P*-*d* experimental data provides an individual K_1-estimate); (iii) calculate an average K_1 from the individual K_1estimates; (iv)

estimate n_1 values by means of Eqn. 6.1b using the average K_1 value; (v) verify the assumption of n_1. If incorrect, go to step (i). Accordingly, the validity of the proposed strategy was investigated by performing different simulations in this study.

6.4.1.1 Simulation-1: Validity of *m* = 1 Assumption Connected to Experimental Design.

Since the proposed approach assumes $m = 1$, it is convenient to explore the impact of multiple binding ($m > 1$) on logK_1 estimations. This study has been performed for the first time in connection with the experimental design (i.e. the ratio D/P). The protocol for the simulation study was: (i) $P = 530$ μM and D from 10 to 1060 μM were fixed; (ii) For each P-D pair, d was obtained using Eqn. 6.2 for the $m = 2$-model where $n_1 = 1$, log$K_1 = 4$, $n_2 = 1$ and log$K_2 = 2$ (i.e. $K_1/K_2 = 100$, which could be considered a normal between-sites affinity ratio); (iii) d-values calculated were used in Eqn. 6.1a to estimate K_1 values. Figure 6.17 shows the logK_1 values estimated as a function of D/P. As can be seen, for $D/P < 0.5$ logK_1-estimates are close to 4 (the fixed value), but large deviations were obtained using D/P values larger than 1. For the eventual case of log$K_2 = 3$ (i.e. $K_1/K_2 = 10$), the estimated logK_1 varied from 4.1 ($D/P = 0.5$) to 5.1 ($D/P = 1.5$). In other words, estimates of logK_1 based on an $m = 1$-model (as Eqn. 6.1a) are still valid, in the case of multiple binding, adjusting the experimental design to low D/P values. These results highlight the essential role of D/P in protein-binding studies.

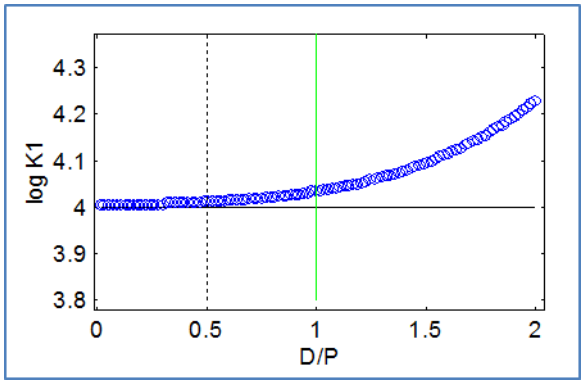


Figure 6.17: Impact of assuming m = 1 on logK_1 estimates (from Eqn. 6.1a) as a function of D/P, when an m = 2-model exists. Simulated results for the m = 2-model, using Eqn. 3.4.1, corresponds to P = 530 µM and D from 10 to 1060 µM, n_1 = 1 and logK_1 = 4, for the main site, and n_2 = 1 and logK_2 = 2, for the secondary site. The fixed value logK_1 = 4 (expected in the absence of multiple binding; m = 1) is included as a horizontal line.

6.4.1.2 Simulation-2: Intrinsic model sensitivity connected with experimental design.

In order to study the effect of D- and P-ranges on d estimations, a simulation study was performed according to the following equation.

$$d = \frac{-(1 - K_1 D + n_1 K_1 P) + \sqrt{(1 - K_1 D + n_1 K_1 P)^2 + 4K_1 D}}{2K_1} \qquad (6.2)$$

Equation 6.2 is a d-isolated expression obtained from Eqn. 6.1 for the m = 1 case. It provides error-free d values (whenever m = 1 is true or the impact of m > 1 could be considered negligible; e.g. using D/P< 0.5 according to simulation-1) associated to a given experimental design (D and P) and parameters (n_1 and K_1).

Figure 6.18 shows the simulated d vs. D and P surface for n_1 = 1 and logK_1 = 3.5 using this equation. The shape of the surface changes with n_1 and K_1 values. A new concept is introduced here; the intrinsic d-sensitivity of the model, $\Delta d/\Delta D$ or $\Delta d/\Delta P$ (irrespective of the analytical technique used). Figure 6.18b shows the d-sensitivity (to D), $\Delta d/\Delta D$ vs. D relationship, obtained using a P-constant design (P =

530 µM, D = 0-390 µM) for different logK_1 values. As can be seen the d-sensitivity decreases with logK_1 (also with n_1, not shown) but increases with D (for logK_1> 2).

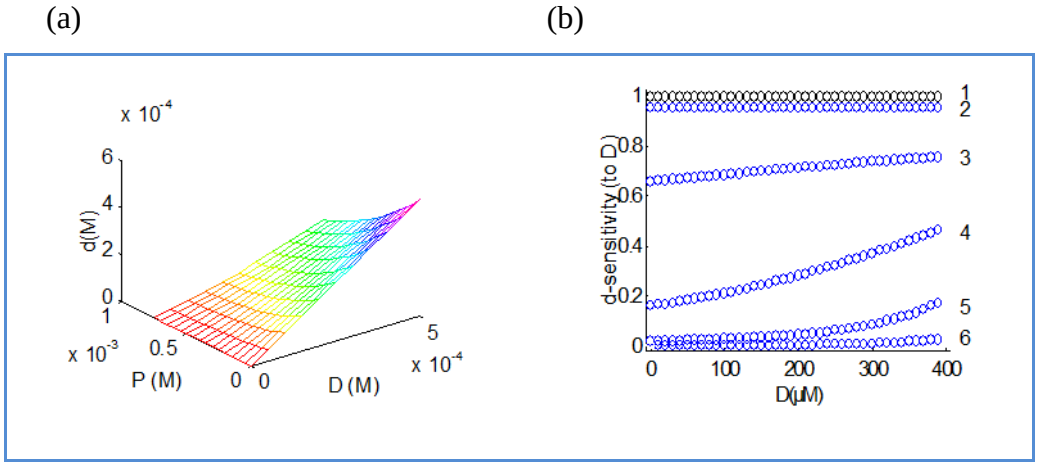


Figure 6.18: simulated data corresponding to Eqn. 6.2 (for an m = 1-model). (a) d vs. D and P surface for n_1 = 1 and logK_1 = 3.5. (b) d-sensitivity (to D), $\Delta d/\Delta D$ vs. D, in the D-range 0 - 390 µM, using a P-constant design (P = 530 µM; D/P up to 0.74), for n_1 = 1 and different logK_1 values (indicated in the right part of the plot).

In contrast, the d-sensitivity (to P), $\Delta d/\Delta P$ (using a D-constant design; not shown), is comparatively much lower (e.g. the highest $\Delta d/\Delta P$, which was found for logK_1 = 4 at P = 200 µM, is approximately 0.06; that is 4.7 times lower than $\Delta d/\Delta$D at the same conditions and much lower for other conditions as shown Figure 6.18b). This means that for binding parameter estimations based on regression (up to now the most used approach), a P-constant design varying D in ranges of relatively high D-values, would be preferable in view of the intrinsic d-sensitivity. Detailed results are shown in Appendix 4B.

6.4.1.3 Simulation-3: Verification of the assumed stoichiometry connected to experimental design.

The simulation study combined a P-constant design (P = 530 µM and five spaced D-levels from 50 to 250 µM), accounting for higher intrinsic d-sensitivity values, with a reduced D-constant design (D = 100 µM and three P values, 300, 530 and 700 µM), covering the common HSA concentration range in the body [4, 23].

Through the design *D*/*P* values under 0.5 were used, thus minimizing the eventual impact of secondary class sites ($m > 1$) participating in the binding process. The simulation was conducted in four stages: (i) a reference situation with parameters $n_1 = 1$ and $\log K_1 = 3.5$ ($m = 1$), was fixed to calculate reference *d*-values (using Eqn. 6.2) for comparison; (ii) three different n_1, $\log K_1$ combinations (see Figure 6.18a caption) providing similar *d*-values (i.e. absolute mean relative error < 0.35 % respect to the reference *d*-values) were generated, simulating different apparent enantiomer: HSA stoichiometries in the 0.5–3 range; (iii) Assuming *a priori* $n_1 = 1$ (stoichiometry 1:1 as hypothesis) in all cases, $\log K_1$ were obtained using Eqn. 6.1a. The estimated median value ($\log K_1 \sim 3.50$ in all cases) were then used in Eqn. 6.1b to estimate n_1 values, in order to verify the apparent stoichiometry in each case. Except in the reference situation, n_1 estimates were different, thus revealing a failure in the stoichiometry assumption (the purpose of the simulation study). Accordingly, alternate approaches making use of such differences to alert us on the failure of the $n_1 = 1$ assumption, when in fact $n_1 \neq 1$ were investigated.For instance, Figure 6.19a shows then$_1$ vs. *d* relationships. The almost linear trends with slopes different from zero reveal the inadequacy hypotheses, therefore serving as a 'stoichiometry verification plot'.

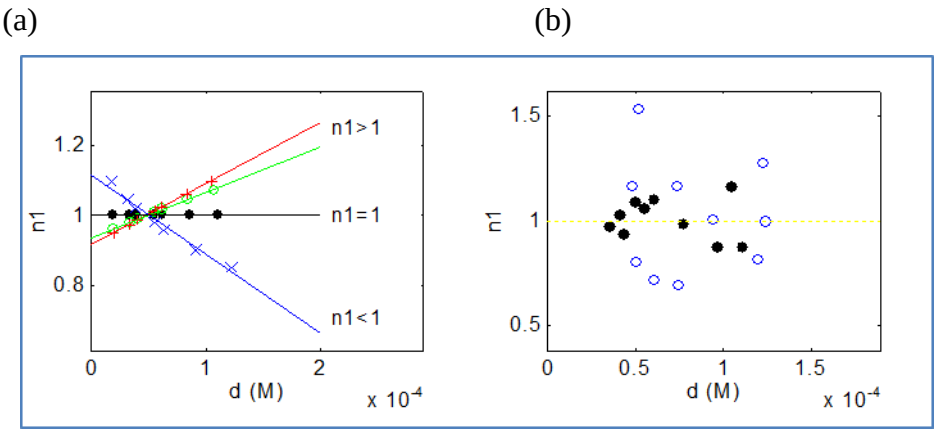


Figure 6.19: 'Stoichiometry verification plots' (n_1 vs. *d* relationships). In all cases $n_1 = 1$ was prefixed (to be verified) for Eqn. 6.1a and the $\log K_1$*median* was selected for Eqn. 6.1b (a) Simulated data generated from (●) $n_1 = 1$ and $\log K_1 = 3.500$ (reference), (×) $n_1 = 0.5$ and $\log K_1 = 3.882$, (o) $n_1 = 2$ and $\log K_1 = 3.164$, (+) $n_1 = 3$ and $\log K_1 = 2.976$. (b) Experimental data for (+)-catechin (open circles), from a *median* $\log K_1 = 3.281$, and (-)-catechin (filled circles), from a *median* $\log K_1 = 3.471$.

This protocol was verified to be valid to test any prefixed n_1 value (Appendix 4A and 4B). Unfortunately, the experimental protein binding processes is complex, with clearly identifiable sources of uncertainty; for instance in the case of enantiomers (analytes) from a racemate-HSA incubation process it could be mentioned the preparation of standards and mixtures, incubation, ultrafiltration, analytical separation and measurement, etc (where presumably, the binding degree during the incubation and the separation of the unbound fraction during ultrafiltration become the most critical factors).Consequently, negative impact on the precision of *d*-data is expected, also relating to the proposed 'stoichiometry verification plot'. Therefore, decisions based on qualitative observations, i.e. an 'increasing trend' or a 'decreasing trend' for the $n_1 > 1$ or $n_1 < 1$ cases (as Figure 6.19a indicates), respectively. It should be noticed that the 'stoichiometry verification plot', which uses *d* as x-axis, suggests experimental designs allowing *d*-variation as large as possible, so, *P*-constant designs becomes again preferable (see simulation-2).

Simulation-3 plays a key role in providing reliability to the proposed algorithm centred on the combined use of Eqns. 6.1a and 6.1b. Anticipated advantages of its use over the traditional strategies are: (i) each *D-P-d* individual data provides an individual K_1 estimate (Eqn. 6.1a) for both enantiomers, and then, an individual enantioselectivity (*ES*) estimation (as the ratio between K_1 data of enantiomers); (ii) each set of K_1 and *ES* estimates (from several experiments) represent independent (like replicate) data vectors, which can be submitted to univariate statistical analysis; e.g. outlier evaluation; (iii) the imprecision corresponding to this data vectors can be used as a reliable source of uncertainty, since it reflects the variability of the entire process and (iv) the K_1 estimates of the enantiomers constitute paired data vectors (each data pair comes from the same mixture and procedure applied), so they could be compared statistically.

6.4.2 (±)-Catechin-HSA Enantioselective Binding.

Table 6.4 illustrates the experimental design, i.e. *D* and *P* in the mixtures (the experimental values after preparation are shown); similar to that used in the simulation-3 study. Concentrations lower than 200 μM of the racemate did not

provide proper quantifiable analytical signals, so it was used as the lower concentration limit (D = 100 µM for enantiomers).The P values cover the physiological range. D/P < 0.47, favours the m = 1 assumption (see simulation-1) allowing the application of Eqns. 6.1a and 6.1b. The use of up to seven d-estimations varying D (in the 100 – 250 µM range; P-constant) exploits the intrinsic d-sensitivity (see simulation-2), also favouring the 'stoichiometry verification plot' use (simulation-3).Calibration standard solutions containing racemic (±)-catechin in the range 25 - 250 µM were prepared. The linear regression equations were *Corrected peak area* = - 0.0164 + 0.0087 (+)-catechin (µM), R^2 = 0.996, and *Corrected peak area* = - 0.1164 + 0.0138 (-)-catechin (µM), R^2 = 0.991. These calibration models were used to estimate d, by interpolating the analytical signals (corrected peak areas) corresponding to ultrafiltrates from the (±)-catechin-HSA mixtures, providing the required D-P-d values. In agreement with the simulation studies, Eqn. 6.1a (with n_1 = 1 as initial trial, to be verified) was selected for K_1direct estimations, while Eqn. 6.1b (with K_1 = *median* of previous estimates) was selected to perform the 'stoichiometry verification plot' for each enantiomer.

Figure 6.19b shows the 'stoichiometry verification plot', which reflects the uncertainty of the estimations; however, no ascendant/descendent trends were observed, suggesting that the n_1 = 1 assumption (1:1 enantiomers-HSA stoichiometry) is acceptable. After verifying and accepting n_1 = 1, we also accepted the logK_1 estimations, as well as the ES values. Table 6.4 shows the individual estimates for each experiment. As mentioned, some degree of dispersion and the possibility of outliers must be *a priori* expected. The estimates were submitted to the Grubbs test (α = 2.5%, 2 tails [144] and none of the individual values (K_1 or logK_1 and ES) were found to be outliers according to this criterion. On the other hand, robust statistics are more convenient when data are expected to follow a near-normal distribution, but are suspected to contain a small proportion of unrepresentative errors like the present data [145]. Thus, the *median* and *MADe* (a robust estimation of the standard deviation) were used to estimate a robust z-score, z_i for each independent estimate (z_i= (estimate$_i$– *median*) / *MADe*) [4], $|z|$>3 data are considered unsatisfactory.

Table 6.4: Experimental design and individual estimates for each experiment.

IDe[a]	*D*, μM	*P*, μM	logK_1 (+)Cat[b]	LogK_1(-)Cat[b]	*ES* [c]
1	100.60		3.353	3.481	1.344
2	125.74		3.489	3.512	1.054
3	150.89		3.356	3.519	1.455
4	176.04	530.64	3.284	3.460	1.500
5	201.19		3.172	3.395	1.671
6	226.34		3.277	3.391	1.298
7	251.49		3.415	3.555	1.382
8		300.36	3.102	3.497	- [d]
9	100.60	530.64	3.123	3.434	2.047
10		700.84	3.178	3.453	1.881

[a] The identification number of the experiment (*IDe*) does not correspond to the order of processed sample, since the entire experiment, including calibration standards, was randomized. *IDe* 1 to 7 correspond to a *P*-constant design (with high intrinsic *d*-sensitivity) while *IDe* 8 to 10 to a *D*-constant one. The concentrations were transformed into M units for calculations.
[b] Estimated from Eqn. 6.1a, assuming $n_1 = 1$, after verifying this assumption via the 'stoichiometry verification plot' (Figure 6.17b), based on n_1-estimates from Eqn. 6.1b.
[c] Estimated as K_1(-)-Catechin /K_1(+)-Catechin; K_1 in M^{-1} units.
[d] Eliminated after applying the robust $|z| > 3$ criterion

We have pre-fixed this criterion to directly eliminate a value in a round. None of the K_1 or logK_1 values were found outliers, but one *ES* estimate showed a z-score above 3 and was eliminated (see Table 6.4).

Statistical comparison of enantiomers logK_1 paired data-vectors was performed. Both the parametric t-test for the means or the non-parametric ones for the medians, the sign test and the signed rank test, indicated significant differences ($p < 0.05$) between the averaged affinity estimates, so enantioselective binding is demonstrated. Alternatively, the 95.0% confidence interval for *ES* mean (1.28, 1.75) also confirms the significant enantioselectivity (*ES* > 1) of (±)-catechin to HSA.

6.4.3 Comparison with Literature Results

Table 6.5 shows a comparison of the available information (experiments *IDE1* to *E5*) [18-21] in chronologic order, including this work. As can be seen, the present work provides more complete results than the previous ones, in part because here we have incubated the racemate and analyzed the separated enantiomers. The results in Table 6.5 are from different incubation processes and designs, analytical techniques and mathematical approaches for estimations. Some drawbacks (footnote a in Table 6.5) detected in previous experiments (*IDE1* to *E5*) are outlined (and avoided here). According to them, the results provided here should be *a priori* considered more robust, more consistent and closer to those corresponding to physiological conditions than those published up to now. For instance, all the previous results are from the incubation process using phosphate buffer (with non-adjusted ionic strength; far from physiological conditions), so the estimated affinity constants (which depend on the ionic strength) should be considered partially biased

In spite of this, a 1:1 stoichiometry to HSA has been reported for individual (±)-catechin enantiomers (*ID E2* and *ID E5*;without verification), coinciding with that assumed and further verified here. Regarding logK_1, the robust estimates obtained here, are close to that of *ID E2* for (+)-catechin and lower (but not too far) than that of *ID E5* for (-)-catechin. The estimations in ID E2 were based on Eqn. 6.2 using the simplex algorithm enforced for searching only integer n_1 values [19] (a constraint, combined with n_1 = 1 and high K_1 starting values, to avoid strange n_1 and K_1 estimates; a limitation of this approach as consequence of the multiple 'good' solutions). So, they should be considered as forced estimates. In the case of *ID E5*, the estimations are from a linear equation based on the fluorescence intensities (after logarithmic calculations). In accordance with the Scatchard approach, the experimental variables appears simultaneously as dependent and independent variable (mathematically inconsistent), so extremely accurate data are needed to obtain reliable estimates. The case *ID E1* deserves more attention; we were unable to reproduce the reported calculations and the 'stoichiometry verification plot' (not shown) suggests problems with this estimate. In regards to the enantioselectivity to HSA, a qualitative estimation can be derived from the *ID E3* information, based on

the ratio between the reported HSA partition coefficients for the enantiomers, calculated from retention time data corresponding to a mobile phase consisting in a modified phosphate buffer (pH 5.0; 20% acetonitrile), thus affecting the affinity between the enantiomers and the immobilized HSA and consequently the enantioselectivity (so it must be considered as a pseudo-*ES* estimation).Thus, although the result cannot be compared to the *ES* obtained here, it indicates a pseudo-*ES* in favour of the (-)-catechin, as here. With regard to protein-binding (*PB*) to HSA, a surprising 100 % value for (+)-catechin is reported in *ID E4*.This estimation is based on a previous relationship between the retention factor data and bibliographic protein binding (not error free) values for some 'reference' compounds, regarded as an indirect estimation. In addition, the mobile phase consisted of a modified phosphate buffer (pH 7.0; 5% isopropanol), should affect the equilibrium between (+)-catechin and the immobilized HSA onto the stationary phase. This value is too high, considering the *PB* = 95 % (plasma) reported by the authors (using ultrafiltration) or other literature values (e.g. 45% for HSA and 90% for plasma [19].The results obtained in the present work for (+)-catechin (53%) are slightly higher than those of *ID E2*, however been calculated at the physiological *D* (see footnote c in Table 6.5) and *P* values, taking into account the *PB* dependence on *D* and *P* (not at the experimental *D* and *P* values; usual practice in most published work), justifying the higher result.

Table 6.5: Comparison between (±)-catechin-HSA binding studies. Data between square brackets (not reported) are estimated here form the available data. The symbol "?" indicates that estimations cannot be reproduced, while the symbol "!" reflects difficult-to-compare results.

Experiment	Incubation process (and experimental design)				Analytical process		Protein binding estimations				
ID (ref.) (Drawback) [a]	Mixture	*D*, µM	*P*, µM	Max. *D*/*P*	Method (Signal)	Analyte	Approach	n_1	logK_1	*ES*	*PB* (%)
E1[18] (A, B, C)	Racemate (*D*) and HSA (*P*) in PB	5 – 32.5	40	[0.8]	HPFA (Peak area)	Racemate	Scatchard (linear regression)	[0.36] **?**	4.12 **?**	-	-
E2[19] (A, B, D, E)	(+)C (*D*) and HSA (*P*) in PB	200	[75] - 475	[2.7]	CEFA (Peak area)	(+)C	SIMPLEX (non-linear fitting)	1	3.34	-	45
E3[20] (A, F)	(-)C (*D*) and HSA (*P*) in mPB	100	Immobilized	-	HPAC (Retention time)	(-)C	-	-	-	[1.07] **!**	-
	(+)C (*D*) and HSA (*P*) in mPB	100				(+)C	-	-	-		-
E4[21] (A, F)	(+)C (*D*) and HSA (*P*) in mPB	103	Immobilized	-	HPAC (Retention time)	(+)C	Indirect linear relationship	-	-	-	100**!**
E5[23] (A, B, C)	(-)C (*D*) and HSA (*P*) in PB	0 – 20	4	[5]	Fluorimetric titration (Intensity)	(-)C	Linear regression	1	4.04	-	
This work	Racemate (with (-)C and (+)C (*D*)) and HSA (*P*) in PBS	100 – 250	300 – 700	0.47	UF, CD-EKC (Peak area, corrected)	(-)C	Direct (Eqn. 6.1a)	1	3.47 ± 0.06 [b]	1.5 ± 0.2 [b]	64 [c]
						(+)C		1	3.28 ± 0.16 [b]		53 [c]

Acronyms: PB: Phosphate buffer (pH = 7.4); mPB: Modified phosphate buffer including organic modifiers and pH different from 7.4 (influencing the affinity); PBS: Phosphate buffer saline (pH = 7.4; ionic strength adjusted to physiological conditions); HPFA: High Performance liquid chromatography in the frontal analysis modality; CEFA: Capillary Electrophoresis in the frontal analysis modality; HPAC: High Performance Affinity Chromatography; UF: Ultrafiltration; CD: Cyclodextrin; EKC: Electrokinetic chromatography.

[a] Drawbacks: (A) Ionic strength far from physiological conditions should provide biased K_1 estimates (respect to *in vivo* ones). (B) *D*/*P*> 0.5 increases the risk of invalid model assumption (e.g. impact of *m*> 1). (C) Mathematically inconsistent and risky regression estimates [4, 17]. (D) Insecure/forced estimates (multiple apparent valid solutions) (E) *D*-constant design involves low intrinsic *d*-sensitivity to *P*. (F) Immobilized HSA could alter the interaction capability (i.e. HSA flexibility). Undefined *D*/*P* (and model validity). Estimates based on the retention time (not comparable with those based on the concentration); they should be considered as qualitative estimates respect to those in this work.

[b] Robust estimates, *median* ± *MADe*.

[c] *PB* (= 100 (*D*-*d*)/*D*), where *d* was calculated from Eqn. 6.2, estimated at *D* = 5 µM and *P* = 600 µM (physiological conditions)

6.4.4 Molecular docking on (±)-catechin enantioselectivity to HSA.

Preliminary tests were performed in order to explore the potential of molecular docking methodologies in connection with the experimental estimation of enantioselectivity (chiral recognition). For this purpose the commercially available software GLIDE was used [84, 146, 147]. It has been claimed that GLIDE scoring functions are not perfect for the estimation of parameters such as binding affinities [148]. However, a semi-quantitative study centered on *ES* (a relative magnitude involving affinity data of two enantiomers; with the same molecular structure) could be a valuable guide (e.g. GLIDE protocols, other docking programs). A recent GLIDE-strategy, Prime MM-GB/SA (see Table 6.6 for details), was used, which provides relative binding free energies (ΔG [149]), parameter related to the affinity constant K_1. For comparison, an alternative protocol, Induced Fit Docking (IFD) [150], providing docking Scores, was used in this study.

In this work it was assumed as an hypothesis, that the ratio between ΔG for enantiomers should represent a relative enantioselective binding quantity, implying that ΔG (-)-catechin.ΔG (+)-catechin represent a pseudo-*ES* measure, which can be compared with the experimental *ES* estimation (1.5 ± 0.2). Several sub-domains in the HSA were investigated with the GLIDE-Prime MM-GB/SA protocol through the complexes available in the Protein Data Bank (PDB; IIA from 2BXD, IIIA from 2BXF, IB from 2VUF, IIIB from 1HK1, IIA-IIB from 2BXG and cleft from 2BXN). Only in the first two cases, a pseudo-*ES*> 1, consistent with the experimental *ES* value, was encountered (in the remaining cases, values close or below the unity were observed). Table 6.6 shows some detailed GLIDE-Prime MM-GB/SA results for the subdomains IIA and IIIA (sites I and II, respectively). Another hypothesis, related to the (±)-catechin features was assumed. Since the molecule showed low hydrophobicity (e.g. calculated logP = 0.4 [151]) and at pH = 7.4 is almost neutral, [152] we therefore assumed that the observed *ES* to HSA (i.e. the slight affinity differences between enantiomers, in favour of (-)-catechin) could be mainly related to the hydrogen bonding interaction (rather that to other effects such as hydrophobic contacts). The H-bond interactions predicted by Glide-Prime MM-GBSA are

included in Table 6.6 (see Figure 6.20a for structure of catechin with atom numbering).

Table 6.6: Molecular docking results for the best pose predicted by Glide-Prime MM-GB/SA[a] consistent with the experimental *ES* value (1.5 ± 0.2). The hydrogen bonding (H-bond) interaction (ordered by distance) is indicated.

HSA site [b] (subdomain; PDB ID)	Enantiomer	Group (Ring, atom; Fig. 17a)	HSA residues (contact)	H-bond distance(Å)	ΔG (kcal/mol)	Pseudo-*ES* [c]
I (IIA; 2BXD)	(-)C	OH (B3') OH (B4') O (Glycoside)	Gln196 (NH) Ser192 (CO) Tyr150 (OH)	1.738 1.971 2.430	-27.25	1.60
	(+)C	O (C3) O (A5) OH (B3') O (C3) O (A5)	Lys199 (NH) Tyr150 (OH) Glu292 (O^-) Hie242 (NH) Arg257 (NH)	1.905 2.003 2.091 2.134 2.153	-17.01	
II (IIIA; 2BXF)	(-)C	OH (B3') OH (A3) O (Glycoside)	Ser489 (CO) Ans391 (CO) Ser489 (OH)	1.762 1.808 2.153	-25.47	1.25
	(+)C	OH (A3) O (B3')	Ser489 (O) Lys414 (NH)	1.815 2.054	-20.41	

[a] Schrödinger's Maestro 9.1 software. (±)-catechin enantiomers and HSA were prepared at pH 7.4 to mimic physiological conditions. Docking calculations were undertaken by using Glide 5.6. [146]. The docked possess (GlideSP [84]) were re-docked (GlideXP [147]) and post-processed, involving molecular mechanics (MM) based scoring functions with the Generalized Born (GB) model as the implicit solvation model (MM-GB/SA), to calculate relative binding free energies, ΔG (Prime MM-GB/SA [149]) The HSA flexible region was chosen as any residue within 12 Å of the ligand in each active site.
[b] PDB database (http://www.rcsb.org) was used to obtain the computational information for HSA complexes. Site I was set from the PDB Warfarin-HSA complex (2BXD), Site II from the PDB Diazepam-HSA complex (2BXF) [43].
[c] The ratio ΔG (-)-Catechin/ΔG (+)-Catechin was used as a pseudo-*ES* estimation (assumed to represent a relative enantioselectivity measure).

Site I appears as the subdomain with a greater degree of enantioselectivity (pseudo-*ES* = 1.6). For site I, this protocol predicts that the prepared ligand (Figure 6.20b) undergoes a conformational change during docking, mainly on the B ring, for (-)-catechin (Figure 6.20c), allowing a strong or H-bond Gln196-B4' (exhibiting the shortest distance, 1.738 Å; thus the highest degree of interaction; Table 6.6) to be formed. This could be the key in explaining the favourable pseudo-*ES* for this enantiomer, in agreement with the proposed hypothesis. The number of H-bonds seemed to be less relevant, as well as the number of hydrophobic contacts predicted

by GLIDE-Prime MM-GB/SA (159 and 149, for (-)-catechin and (+)-catechin, respectively, for site I, and 114 and 132, for (-)-catechin and (+)-catechin, , for site II respectively), although in the case of site I, it is consistent with the high pseudo-*ES* associated with this subdomain.

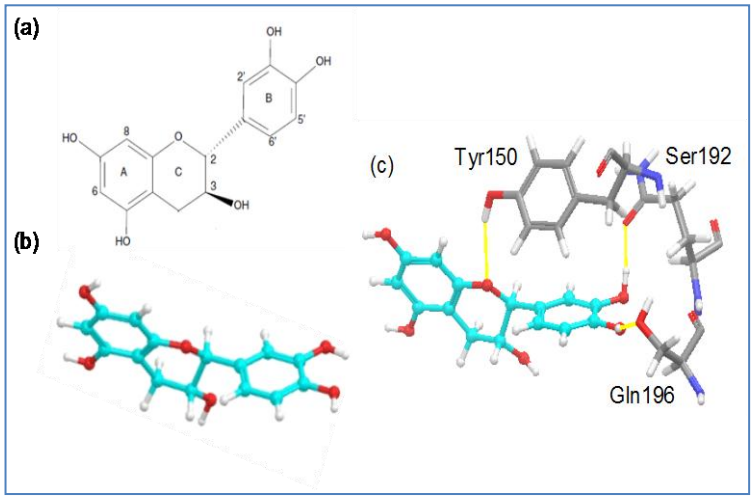


Figure 6.20: (a) 2D structure of catechin with atom numbering. (b) Prepared 3D structure of (-)-catechin (before docking). (c) 3D structure of (-)-catechin after docking (best pose predicted from Prime MM-GB/SA) to site I of HSA. Three HSA residues forming H-bonds are included.

Software-dependent docking outputs are frequently described. For this purpose, the GLIDE-IFD protocol was extended to the protocol-dependent results. The new estimation based on the ratio between the scores, indicated the same trend for site I (ratio = 1.08), but an opposite trend for site II (ratio < 1). For the rest of the subdomains, results were similar to those obtained with GLIDE-Prime-MM-GB/SA. However, in all cases, even for site I, the HSA-residues and enantiomer groups/atoms participating in H-bonds, as well as their distances, or hydrophobic contacts predicted by GLIDE-IFD are considerably distinct to those depicted in Table 6.6 (not shown). For instance, for site I, the hypothesis that *ES* can mainly be explained in terms of H-bonding could not be supported, since the strongest H-bonds predicted belong to (+)-catechin (distance < 1.6 Å). In this case, an explanation based on the number of hydrophobic contacts predicted by GLIDE-IFD (178 and 110, for (-)-catechin and (+)-catechin, for site I respectively) would be more plausible. This result suggests

that the reliability of the GLIDE-estimations (protocol-dependent) still deserves more attention. However, the preliminary study suggests (mainly from Prime MM-GB/SA, partially combined with IFD, outputs) that the interaction of enantiomers in the site I of HSA can be postulated as the most probable situation. In the case of (-)-catechin,hydroxyl group of the B3' atom in the B-ring (see Figure 1.7 and 1.8), participated in the most stable H-bonding interaction, after a conformational change due mainly to the environment of HSA residues, could be a key feature in explaining the moderate *ES* observed experimentally. Although our results have to be viewed with caution, they however provide an initial finding, to be contrasted with more computer programs/protocols. However molecular docking is would be considered an accurate or fully-validated methodology to estimate affinity constants, but the results of this study suggests that the relative magnitudes such as pseudo-*ES* values, connected to experimental *ES* values, involving estimations on two enantiomers (eliminating the molecular structure difference factor), could broaden the possibilities for synergy between-strategies in two senses: (i) docking could help to explain at the molecular level the *ES* results found experimentally and (ii) experimental *ES* values could serve to validate docking estimations.

CHAPTER 7

CONCLUSSIONS

Enantioresolution

In general, the chromatographic retention and electrophoretic migration order predicted by the lowest computed energies using rigid molecules are unable to accurately reproduce experimental results. Therefore a statistical analysis of the minimized energies or a further full geometry optimization of the complexes was necessary to obtain an appropriate elution order in agreement with the experimental observation. From the results obtained it was evident that differential interactions between each enantiomer and the chiral host resulted in different configurations for the corresponding inclusion complexes which also gives rise to enantioresolution. Our results also suggest that the van der Waals term is the main contributor to the total energy and its minimum is located at the center of the CD ring for each enantiomer. Due to limitations of the experimental methods, our findings suggest that molecular modeling can be used to rationalize experimental findings based on the molecular and chiral recognition by CDs. Molecular modeling methods have proven to be valuable tools for detailed information on the geometry and interaction energies of the inclusion compounds. Overall, molecular docking methods helped to elucidate the mechanism thus providing insight into the influences of the substrate structure to enantioresolution [141].

Pharmacokinetic Evaluation

Overall the CE-FA approaches were able to adequately provide information on the pharmacokinetic parameters of enantio-pure catechin, but unable to study the interaction of each enantiomer available in racemic forms, since they have to be separated prior to analysis. However, this has been overcome by the use of Affinity Electrokinetic Chromatography (AEKC) giving more information at nearly physiological conditions. The *in vitro* incubation of racemic (±)-catechin-HSA mixtures at resembling physiological conditions, followed by ultrafiltration and

separation/determination of enantiomers by CD-EKC makes it possible to access information unobtainable by *CE-FA* experiments. This novel strategy is based on two equations (Eqns. 6.1a and 6.1b),which allows for a simple yet direct pharmacokinetic parameter estimations (avoiding the insecure traditional regression/fitting approaches) that facilitates the stoichiometric assessment through a simple 'verification plot'. This eliminates the challenges due to the conversion of the protein-binding data analysis into a univariate problem by incorporating the robust statistics that takes into consideration: the treatment of outliers, hypothesis testing and uncertainty assessment. Apart from the scientific interest of the results obtained, the entire process is easily transferable with the null or moderate adjustment to investigations involving most enantiomeric ligand/drug interacting with HSA or any other proteins/bio-macromolecules. Furthermore, the molecular docking studies provided a deeper insight into the interaction energy of selector-enantiomer and provided useful information to support the enantiomeric separations mechanism.

Future Work

Unfortunately, the majority of the comments and drawbacks related to Table 6.5, of the literature results, can be extended to most of the protein binding results obtained in the past. The possible risk of inaccurate estimates should carefully be checked before accepting them to derive further conclusions. Data comparison is at least recommendable, and if possible (when data available in the literature) the estimations should be verified (e.g. using approaches like those proposed in this work based on the combination of Eqn. 6.1a + Eqn. 6.1b). In this sense, Eqn. 6.2 could be used to derive *d*-data when at least *D-P* data (design) and n_1 and K_1 (estimates) data are provided (e.g. as in Table 6.4). In view of the future work in this research area, the findings from the study is intended to influence the traditional schemes in order to improve the quality of the different processes involved in protein binding and related pharmacokinetic evaluations.

REFERENCES

[1] Y. Duan, B.V.B. Reddy, and Y.N. Kaznessis, Physicochemical and residue conservation calculations to improve the ranking of protein–protein docking solutions, *Protein Science,* 14 **(2005)** 316–328

[2] H.A. Gabb, R.M. Jackson, and M.J. Sternberg, Modelling protein docking using shape complementarity, electrostatics and biochemical information. *J. Mol. Biol.* 272 **(1997)** 106–120.

[3] L. Escuder-Gilabert, M.A. Martínez-Gómez, R.M. Villanueva-Camañas, S. Sagrado, and M.J. Medina-Hernández, Microseparation techniques for the study of the enantioselectivity of drug-plasma protein binding, *Biomed. Chromatogr.*, 23 (**2009)** 225-238.

[4] L. Asensi-Bernardi, Y. Martin-Biosca, R.M. Villanueva-Camañas, M.J. Medina-Hernández, and S. Sagrado Vives, Evaluation of enantioselective binding of fluoxetine to human serum albumin by ultrafiltration and CE Experimental design and quality considerations, *Electrophoresis*, 31 **(2010)** 3268–3280

[5] T.J. Ward and B.A. Baker, Chiral Separations, *Anal. Chem*., 80 (2**008)** 4363–4372.

[6] N.W. Frost, M. Jing, and M.T. Bowser, Capillary Electrophoresis, *Anal. Chem*., 82 (**2010)** 4682–4698.

[7] D. A. Smith, H. Van de Waterbeemd, and D. K. Walker, Pharmacokinetics and Metabolism in Drug Design, *Wiley–VCH, Weinheim, Germany*, **2001**

[8] H. van de Waterbeemd and E. Gifford, Admetin silico modelling: Towards prediction paradise?, *Nature Reviews Drug Dicovery*, 2 **(2003)** 192-204

[9] T. Knjazeva and M. Kaljurand, Capillary electrophoresis frontal analysis for the study of flavonoid interactions with human serum albumin, *Anal. Bioanal. Chem*., 397 **(2010)** 2211–2219

[10] V.A. Davankov, Analytical chiral separation methods, *Pure and Appl. Chem*., 69 **(1997)** 1469-1474.

[11] S. Fanali, Enantioselective determination by capillary electrophoresis with cyclodextrins as chiral selectors, *Journal of Chromatography A*, 875 **(2000)** 89–122

[12] W. Wang, J. Lu, X. Fu, and Y. Chen, Carboxymethyl-β-Cyclodextrin for chiral separation of basic drugs by capillary electrophoresis, *Analytical Letters*, 34 **(2001)** 569-578,

[13] T.J. Ward, Chiral Separations, *Anal. Chem*., 78 **(2006)** 3947

[14] K.B. Lipkowitz, Applications of Computational Chemistry to the Study of Cyclodextrins, *Chem. Rev*., 98 **(1998)** 1829-1873

[15] E. Alvira, J.I. García, and J.A. Mayoral, Molecular modelling study for chiral separation of equol enantiomers by b-cyclodextrin, *Chemical Physics*, 240 **(1999)** 101–108

[16] K. Vuignier, J. Schappler, J.L. Veuthey, P.A. Carrupt, and S. Martel, Drug-protein binding: a critical review of analytical tools, *Anal. Bioanal. Chem*., 398 (**2010)** 53–66.

[17] L. Fielding, S. Rutherford, and D. Fletcher, Determination of protein-ligand binding affinity by NMR: observations from serum albumin model systems, *Magn. Reson. Chem*., 43 (**2005)** 463-470.

[18] M. Tian, H. Yan and K.H. Row, Protein binding study of echin hydrate and genistein by high-performance frontal analysis, *Korean J. Chem. Eng*., 25 **(2008)** 1473-1476.

[19] A. Diniz, L. Escuder-Gilabert, N.P. Lopes, R.M. Villanueva-Camañas, S. Sagrado and M.J. Medina-Hernández, Characterization of interactions

between polyphenolic compounds and human serum proteins by capillary electrophoresis, *Anal. Bioanal. Chem*., 391 (**2008)** 625–632.

[20] M. Kurlbaum and P. Högger, Plasma protein binding of polyphenols from maritime pine bark extract (USP), *J. Pharm. Biomed. Anal.*, 54 (**2011)** 127–132.

21] L. Trnková, I. Boušová, V. Staňková and J. Dršata. Study on the interaction of catechins with human serum albumin using spectroscopic and electrophoretic techniques, *J. Mol. Struct*., 985 **(2011)** 243-250.

[22] Y.S. Ding, X.F. Zhu and B.C: Lin, Capillary electrophoresis study of human serum albumin binding to basic drugs, *Chromatographia*, 49 **(1999)** 343-346

[23] R. Mehvar, Role of Protein Binding in Pharmacokinetics., *Am. J. Pharm. Educ*., 69 (**2005)** 108

[24] M. Kofink, M. Papagiannopoulos and R. Galensa, Enantioseparation of catechin and epicatechin in plant food by chiral capillary electrophoresis, *Eur. Food Res. Technol.*, 225 **(2007)** 569–577

[25] http://www.beckmancoulter.com/literature/Bioresearch/360643-CE Primer1.pdf (Accessed in 29 April **2011**)

[26] E. Machtejevas, A. Maruška, A new approach to human serum albumin chiral stationary phase synthesis and its use in capillary liquid chromatography and capillary electrochromatography. *J. Sep. Sci.*, 25 **(2002)** 1–6

[27] K.T. Nguyen, D.P. Stephens, M.J. McLeish, D.P. Crankshaw, and D.J. Morgan, Pharmacokinetics of thiopental and pentobarbital enantiomers after intravenous administration of racemic thiopental., *Anesth Analg.,* 83 **(1996)** 552-558

[28] M.C. Millot., Separation of drug enantiomers by liquid chromatography and capillary electrophoresis, using immobilized proteins as chiral selectors., *J Chromatogr B.*, 797 **(200 3)** 131–159.

[29] J. Kraly, M. Fazal, A. Schoenherr, R. M. Bonn, R. Harwood, M. M. Turner E., M. Jones. and N. J. Dovichi, Bioanalytical Applications of Capillary Electrophoresis, *Anal. Chem.*, 78 **(2006)** 4097

[30] F. Tagliaro, G. Manetto, F. Crivellente and F.P. Smith, A brief introduction to capillary electrophoresis, *Forensic Science International*, 92 **(1998)** 75–88

[31] A. Benno, (**1997**), Chiral Separations Using Capillary Electrophoresis, Technische Universiteit Eindhoven, Thesis, ISBN 90-386-0958-2

[32] B. Chankvetadze, G. Endresz and G. Blaschke, Enantiomeric resolution of chiral imidazole derivatives using capillary electrophoresis with cyclodextrin-type buffer modifiers., *J. Chromatogr. A*, 700 **(1995)** 43-49

[33] M.A. Martínez-Gómez, S. Sagrado, R.M. Villanueva-Camañas and M.J. Medina-Hernández, Enantioseparation of phenotiazines by affinity electrokinetic chromatography using human serum albumin as chiral selector. Application to enantiomeric quality control in pharmaceutical formulations, *Analytica Chimica Acta*, 582 **(2007)** 223–228

[34] J.M. Berg, J.L Tymoczko, and L. Stryer. Biochemistry, W H Freeman publishers (**2002)** 5th edition

[35] A.R. Leach, 'Molecular Modelling: Principles and Applications', *Addison Wesley Longmanlimited,* **(1996)** pp 413-436.

[36] S.K. Panigrahi and G.R. Desiraju, Strong and Weak Hydrogen Bonds in the Protein–Ligand Interface, *Proteins: Structure, Function and Bioinformatics*, 67 **(2007)** 128–141

[37] Mathews and Van Holde, Biochemistry (Text book), 2nd edition, pp. 165-171

[38] P.A. Zunszain, J. Ghuman, T. Komatsu, E. Tsuchida and S. Curry, Crystal structural analysis of human serum albumin complexed with hemin and fatty acid., *BMC Structural Biology*, 3 **(2003)** doi:10.1186/1472-6807-3-6.

[39] J.R. Hoffman and M.J. Falvo., Protein–Which is Best? *Journal of Sports Science and Medicine.*, 3 **(2004)** 118-130

[40] K.L. Hein, U Kh-Hansen, J.P. Morth, M.D. Jeppesen, D. Otzen, J.V. Møller, and P. Nissen, Crystallographic analysis reveals a unique lidocaine binding site on human serum albumin, *Journal of Structural Biology*, 171 **(2010)** 353–360

[41] F. Bellon and C. Lapresle, Chemical Structure of Two Fragments of Human Serum Albumin and their Location in the Albumin Molecule, *Biochem. J*., 147 **(1975)** 585-592

[42] C.D. Kanakis , P.A. Tarantilis, M.G. Polissiou, S. Diamantoglou and H.A. Tajmir-Riahi, Antioxidant flavonoids bind human serum albumin, *Journal of Molecular Structure*, 798 **(2006)** 69–74

[43] J. Ghuman, P.A. Zunszain, I. Petitpas, A.A. Bhattacharya, M. Otagiri and S. Curry, Structural Basis of the Drug-binding Specificity of Human Serum Albumin, *J. Mol. Biol*., 353 **(2005)** 38–52

[44] F. Zsila, Z. Bikádi, and M. Simonyi, Probing the binding of the flavonoids, quercetin to human serum albumin by circular dichroism, electronic absorption spectroscopy and molecular modelling methods, *Biochemical Pharmacology*, 65 **(2003)** 447-456.

[45] V.T.G. Chuang and M. Otagiri, Stereoselective binding of human serum albumin, *Chirality*, 18 **(2006)** 159–166

[46] S. Curry, H. Mandelkow, P. Brick, and N. Franks, Crystal structure of human serum albumin complexed with fatty acid reveals an asymmetric distribution of binding sites. *Nat. Struct. Biol.*, 5 **(1998**) 827 – 835.

[47] J. Østergaard and N.H.H. Heegaard, Capillary electrophoresis frontal analysis: Principles and applications for the study of drug-plasma protein binding, *Electrophoresis*, 24 **(2003)** 2903–2913

[48] Y. Tang, H. Zhu, Y. Zhang and C. Huang, Determination of human plasma protein binding of baicalin by ultrafiltration and high-performance liquid chromatography, *Biomed. Chromatogr.*, 20 **(2006)** 1116–1119.

[49] W. Ren, Z. Qiao, H. Wang, L. Zhu and L. Zhang, Flavonoids: Promising Anticancer Agents, *Medicinal Research Reviews*, 23 (4) **(2003)** 519-534

[50] O.J. Rolinski, A. Martin, and D.J.S. Birch, Human Serum Albumin-Flavonoid Interactions monitored by means of Tryptophan kinetics, *Ann. N.Y. Acad. Sci.*, 1130 **(2008)** 314–319.

[51] N.A. Siddique, M. Mujeeb, A.K. Najmi and M. Akram, Evaluation of antioxidant activity, quantitative estimation of phenols and flavonoids in different parts of Aeglemarmelos, *African Journal of Plant Science*, 4 (1) **(2010)** 001-005

[52] S. Soares, N. Marteus and V. Freitas, Interaction of Different Polyphenols with Bovine Serum Albumin (BSA) and Human Salivary r-Amylase (HSA) by Fluorescence Quenching, *J. Agric. Food Chem.,* 55 (**2007)** 6726-6735

[53] B. Bartolomé, Alvaro Peña-Neira and C. Gómez-Cordovés, *Eur. Food Res. Technol.,* 210 **(2000)** 419-423

[54] V. Adam, R. Mikelova, J. Hubalek, P. Hanustiak, M. Beklova , P. Hodek, A. Horna, L. Trnkova, M. Stiborova, L. Zeman and R. Kizek, Utilizing of Square Wave Voltammetry to Detect Flavonoids in the Presence of Human Urine, *Sensors 7*, **(2007)** 2402-2418

[55] K.E. Heim, A.R. Tagliaferro and D.J. Bobilya, Flavonoid antioxidants: chemistry, metabolism and structure-activity relationships**,** *Journal of Nutritional Biochemistry*, 13 **(2002)** 572–584

[56] J. Koyama, I. Morita, N. Kobayashi, T. Konoshima, M. Takasaki, T. Osakai and H. Tokuda, Correlation between oxidation potentials and inhibitory effects on Epstein–Barr virus activation of flavonoids, *Cancer Letters*, 263 **(2008)** 61–66.

[57] M.J.T.J. Arts, G.R.M.M. Haenen, L.C. Wilms, S.A.J.N. Beetstra, C.G.M. Heijnen, H.P. Voss and A. Bast, Interactions between flavonoids and proteins: effect on the total antioxidant capacity., *J. Agric. Food Chem*., 50 (**2002)** 1184-1187

[58] M.J.T.J. Arts, G.R.M.M. Haenen, H.P. Voss, and A. Bast, Masking of antioxidant capacity by the interaction of flavonoids with protein, *Food and Chemical Toxicology*, 39 **(2001)** 787–791,

[59] T Ishii, K. Minoda, M. Bae, T. Mori, Y. Uekusa, T. Ichikawa, Y. Aihara, T. Furuta, T. Wakimoto, T. Kan, and T. Nakayama, Binding affinity of tea catechins for HSA: Characterization by high-performance affinity chromatography with immobilized albumin column, *Mol. Nutr. Food Res*., 54 **(2010)** 816–822

[60] C. Lin, W. Fang, C. Kuo, W. Chang, Y. Liu, W. Lin, J. Wu and C. Lin, Chiral separation of hydroxyflavanones in cyclodextrin-modified capillary zone electrophoresis using sulfated cyclodextrins as chiral selectors, *Journal of Chromatography A*, 1188 **(2008)** 301–307

[61] F. Nyfeler, U.K. Moser, and P. Walter., Stereospecific effects of (1)- and (2)-catechin on glycogen metabolism in isolated rat hepatocytes. *Biochim. Biophys. Acta*; 763 **(1983)** 50–57

[62] H.P. Bais, T.S. Walker, F.R. Stermits, R.A. Hufbauer and J.M. Vivanco. Enantiomeric dependent phytotoxic and antimicrobial activity of (6)-catechin, A rhizosecreted racemic mixture from spotted knapweed., *Plant Physiol*., 128 **(2002)** 1173–1179.

[63] M.J. Weyant, A.M. Carothers, A.J. Dannenberg, and M.M. Bertagnolli, (+)-Catechin Inhibits Intestinal Tumor Formation and Suppresses Focal Adhesion Kinase Activation in the Min/+ Mouse., *Cancer Res*., 61 **(2001)** 118-125.

[64] H. Mangiapane, J. Thomson, A. Salter, S. Brown, G.D. Bell, and D.A. White, The inhibition of the oxidation of low density lipoprotein by (+)-catechin, a naturally occurring flavonoid., *Biochem. Pharmacol*., 43 (**1992)** 445-450.

[65] H.J. Reimann, W. Lorenz, M. Fischer, R. Frölich, H.J. Meyer, and D.A. Schmal, Histamine and acute haemorrhagic lesions in rat gastric mucosa: Prevention of stress ulcer formation by (+)-catechin, an inhibitor of specific histidine decarboxylase in vitro , *Inflamm. Res*, 7 (**1977)** 69-73.

[66] W.C. Hou, R.D. Lin, C.T. Chen and M.H. Lee, Monoamine oxidase B (MAO-B) inhibition by active principles from *Uncariarhynchophylla*, *J. Ethnopharmacol.*, 100 (**2005)** 216–220.

[67] S.B. Lotito, and C.G. Fraga, (+)-Catechin Prevents Human Plasma Oxidation, *Free Radical Bio. Med.*, 24 (**1998)** 435-441.

[68] S.Y. Cho, P.J. Park, H.J. Shin, Y.K. Kim, D.W. *Shin*, E.S. *Shin* and H.H. Lee., (-)-Catechin suppresses expression of Kruppel-like factor 7 and increases expression and secretion of adiponectin protein in 3T3-L1 cells., *Am. J. Physiol. Endocrinol. Metab*., 292 (**2007)** 1166–1172.

[69] B.O. Brandsdal, F. Österberg, M. Almlöf, I. Feierberg, V.B. Luzhkov and J. Åqvist, Free Energy Calculations and Ligand Binding., *Advances in Protein Chemistry*, 66 **(2003)** 123-158

[70] T. Peters, Jr., in: All About Albumin, Genetics and Medical Applications, Academic Press, San Diego, CA (**1996)** pp. 1–432.

[71] M.A. Martinez-Gomez, R.M. Villanueva-Camañas, S. Sagrado, and M.J. Medina-Hernández, Multivariate optimization approach for chiral resolution of drugs using human serum albumin in affinity electrokinetic

chromatography–partial filling technique, *Electrophoresis*, 26 **(2005)** 4116–4126

[72] N.A. Kratochwil, W. Huber, F. Muller, M. Kansy and P.R. Gerber, Predicting plasma protein binding of drugs: a new approach. *Biochem Pharmacol.,* 64 **(2002)** 1355-74

[73] N.K. Shahper, and U.K. Asad, Computational Simulation of Mitoxantrone Binding with Human Serum Albumin., *J Proteomics Bioinform*., S1: **(2008)** S017-S020.

[74] Z. Chen and S.G. Weber, Determination of binding constants by affinity capillary electrophoresis, electrospray ionization mass spectrometry and phase distribution methods., *Trends Analyt Chem.,* 27 (9) **(2008)** 738-748.

[75] Y.C. Cheng, and W.H. Prusoff, Relationship between the inhibition constant (Ki) and the concentration of inhibitor which causes 50 percent inhibiton IC50 of an enzymatic reaction., *Biochem. Pharm*., 22 **(1973)** 3099-108.

[76] P.J. Munson, and D. Rodbard. An exact correction to the "Cheng Prusofi" correction. *J. Receptor Res.,* 8 **(1988)** 533-46.

[77] S. Jacobs, K. Chang, and P. Cuatrecasas, Estimation of hormone receptor affinity by competitive displacement of labeled ligand: Effect of concentration of receptor and of labeled ligand, *Biochem. Biophys. Res. Commun*., 66 **(1975)** 687- 692.

[78] A Papadopoulou, R.J. Green and R.A. Frazier, Interaction of flavonoids with bovine serum albumin: A fluorescene quenching study., *J. Agric. Food Chem.,* 53 **(2005)** 158-163.

[79] *De Meyts P, Bianco AR, Roth J*: Site-site interactions among insulin receptors: characterization of the negative cooperativity., *J. Biol. Chem.,* 251 **(1976)** 1877–1888,

[80] P.J. Ballester and J.B.O. Mitchell, A machine learning approach to predicting protein-ligand binding affinity with applications to molecular docking, *Bioinformatics*, 26 **(2010)** 1169-1175

[81] Ljiljana Čučković, Ismet M. Hodžić and Svetozar R. Niketić, Computational study of the chromatographic enantioseparation of tris(acetylacetonato)cobalt(III) complexes on an arginine complex of cobalt(III) acting as a chiral selector., *J. Serb. Chem. Soc.* 67 **(2002)** 735–744

[82] N. Huang, C. Kalyanaraman, K. Bernacki and M.P. Jacobson, Molecular mechanics methods for predicting protein–ligand binding, *Phys. Chem. Chem. Phys*., 8 **(2006)** 5166–5177.

[83] X. Li, Y. Li, T. Cheng, Z. Liu, and R. Wang, Evaluation of the Performance of Four Molecular Docking Programs on a Diverse Set of Protein-Ligand Complexes, *J. Comput. Chem*., 31 **(2010)** 2109–2125,

[84] T.A. Halgren, R.B. Murphy, R.A. Friesner, H.S. Beard, L.L. Frye, W.T. Pollard, and J.L. Banks, "Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening", *J. Med. Chem*., 47 **(2004)** 1750–1759

[85] C.M. Venkatachalam, X. Jiang, T. Oldfield, and M. Waldman, LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites, *Journal of Molecular Graphics and Modelling*, 21 **(2003)** 289–307

[86] I.D. Kuntz, J.M. Blaney, S.J. Oatley, R. Langidge, and T.E. Ferrin, "A geometric approach to macromolecule-ligand interaction", *J. Mol. Biol*., 161 **(1982)** 269–288,

[87] G.M. Morris, D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew, and A.J. Olson, Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comp. Chem*., 19 **(1998)** 1639–1662.

[88] D. Seeliger, and B.L. de Groot, Ligand docking and binding site analysis with PyMOL and Autodock/Vina, *J. Comput. Aided Mol. Des*., 24 **(2010)** 417–422

[89] C.R.W. Guimarães and M. Cardozo, MM-GB/SA Rescoring of Docking Poses in Structure-Based Lead Optimization, *J. Chem. Inf. Model*., 48 **(2008)** 958–970

[90] Z. Zhou, A.K. Felts, R.A. Friesner, and R.M. Levy, *J. Chem. Inf. Model*.**,** 47 (**2007)** 1599-1608

[91] N. Brooijmans and C. Humblet, Chemical space sampling by different scoring functions and crystal structures, *J. Comput. Aided Mol. Des.*, 24 **(2010)** 433–447

[92] S.Y. Huang, S.Z. Grinter and X. Zou, Scoring functions and their evaluation methods for protein–ligand docking: recent advances and future directions, *Phys. Chem. Chem. Phys.*, 12 **(2010)** 12899–12908

[93] H. Gohlke and G. Klebe, Statistical potentials and scoring functions applied to protein–ligand binding, *Current Opinion in Structural Biology*, 11 **(2001)** 231–235

[94] P. Cozzini, M. Fornabaio, A. Marabotti, D.J. Abraham, G.E. Kellogg and A. Mozzarelli, Free Energy of Ligand Binding to Protein: Evaluation of the Contribution of Water Molecules by Computational Methods, *Current Medicinal Chemistry*, 11 **(2004)** 1345-1359

[95] P.A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, and P. Cieplak, J. Srinivasan, D.A. Case and T.E. Cheatham, *Acc. Chem. Res.*, 33 **(2000)** 889-897.

[96] C.R.W. Guimarães and A. M. Mathiowetz, Addressing Limitations with the MM-GB/SA Scoring Procedure using the Watermap Method and Free-Energy Perturbation Calculations, *J. Chem. Inf. Model*., 50 **(2010)** 547–559

[97] A. Krammer, P.D. Kirchhoff, X. Jiang, C.M. Venkatachalam and M. Waldman, *LigScore*: a novel scoring function for predicting binding affinities. *J. Mol. Graph. Model*., 23 (**2005)** 395-407

[98] S.L. Mayo, B.D. Olafson and W.A. Goddard, DREIDING: A Generic Force Field for Molecular Simulations., *J. Phys. Chem.,* 94 (**1990)** 8897-8909.

[99] M. Pattanayak, P.K. Seth, S. Smita, and S.K. Gupta, Geraniol and Limonene Interaction with 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) Reductase for their Role as Cancer Chemo-preventive Agents, *J. Proteomics Bioinform*, 2 (11) **(2009)** 466-474

[100] K. Vuignier, J. Schappler, J. Veuthey, P. Carrupt, and S. Martel, Improvement of a capillary electrophoresis/frontal analysis (CE/FA) method for determining binding constants: Discussion on relevant parameters, *Journal of Pharmaceutical and Biomedical Analysis*, 53 **(2010)** 1288–1297

[101] H.S. Bloom, The Core Analytics of Randomized Experiments for Social Research, MDRC Working Papers on Research Methodology, August 2006

[102] M.J. Rodríguez Cuesta, **(2006)**, Limit of detection for second-order calibration methods,. Doctoral Thesis, Rovira I Virgili University, Tarragona

[103] IUPAC, 'Nomenclature, symbols, units and their usage in spectrochemical analysis–II,' *Spectrochim. Acta, Part B,* 33 **(1978)** 242

[104] M.M. Sanagi, S.L. Ling, Z. Nasir, W.A. Ibrahim and A.A. Naim, Comparison of Signal-to-noise, Blank Determination, and Linear Regression Methods for the Estimation of Detection and Quantification Limits for Volatile Organic Compounds by Gas Chromatography., *J AOAC Int.* 92 **(2009)** 1833-8.

[105] D.A. Skoog, D.M. West, F.J. Holler, S.R. Crouch, Fundamentals of Analytical Chemistry, 8th edition.

[106] L.A. Currie, IUPAC communication or Analytical Nomanclature, Recommendation in Evaluation of Analytical including Methods and Quantification capabilities., *Pure Appl. Chem.,* 67 **(1995)** 1699

[107] L.A. Currie, Limits for qualitative and quantitative determination — application to radiochemistry, *Anal. Chem.,* 40 **(1968)** 586–593.

[108] K. Lauk, **(2010)**, Statistical methods for evaluation of data from inter-laboratory comparisons with small number of participants, Master's Thesis, University of Tartu

[109] W. Horwitz, Protocol for the design, conduct and interpretation of method performance studies. *Pure & Appl. Chem.*, 67 **(1995)** 331-343.

[110] Official methods of analysis of AOAC international 16th ed., **1999** Gaithersburg (Md.): AOAC International,

[111] S. Burke, Missing values, outliers, robust statistics and non-parametric methods, *LC*GC Europe Online Supplement,* 22 **(1999)** 19–24.

[112] G. Scatchard., The attraction of proteins for small molecules and ions, *Ann NY Acad Sci* 51 **(1949)** 660.

[113] I.M. Klotz., Equilibrium constants and free energies in unfolding of proteins in urea solutions, *Proc Natl Acad Sci USA,* 93 **(1996)** 14411–14415

[114] W.E. Lindup and M.C. L'e Orme, Plasma protein binding of drugs., *British Medical Journal,* 282 **(1981)** 212 - 214

[115] T.W. Traut, Physiological concentrations of purines and pyrimidines, *Mol Cell Biochem.* 140 **(1994)** 1-22.

[116] Z. Jia, T. Ramstad and M. Zhong, Determination of protein–drug binding constants by pressure-assisted capillary electrophoresis (PACE)/frontal analysis (FA), *Journal of Pharmaceutical and Biomedical Analysis* 30 **(2002)** 405–413

[117] E. Di Cera, F.A. Bassi and S.J. Gill, Information theory and the analysis of ligand-binding data., *Biophys Chem*., 34 **(1989)** 19-28.

[118] H. Zhou, G. Rivas, and A. Minton., "Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences". *Annual review of biophysics*, 37 **(2008)** 375–397.

[119] M. Azad, L. Hernandez, A. Plazas, M. Rudolph, and I-. A. Gornez., Determination of Binding Constants between the Antibiotic Ristocetin A and D-Ala-D-Ala Terminus Peptides by Affinity Capillary Electrophoresis, *Chromatographia*, 57 **(2003)** 339- 305

[120] Primer., High Perfommance Capillary Electrophoresis, **1 January 2010**, 5990-3777EN

[121] H.H. Lauer and J.B. Ooms, Advances in capillary electrophoresis: the challenges to liquid chromatography and conventional electrophoresis, *Analytica Chemica Acta*, 250 **(1991)** 45-60

[122] P. Camilleri, Capillary Electrophoresis: Theory and Practice, 2nd ed., CRC Press, Boca Raton, FL, **1997**.

[123] N.H.H. Heegaard, *J. Mol. Recogn*, 11 **(1998)** 141

[124] M.H.A. Busch, H.F.M. Boelens, J.C. Kraak, and H. Poppe, Vacancy affinity capillary electrophoresis, a new method for measuring association constants, *Journal of Chromatography A*, 775 **(1997)** 313–326

[125] Y. Tanaka and S. Terabe, Estimation of binding constants by capillary electrophoresis, *Journal of Chromatography B*, 768 **(2002)** 81–92

[126] Y. Tanaka and S. Terabe, Partial separation zone technique for the separation of enantiomers by affinity electrokinetic chromatography with proteins as chiral pseudo-stationary phases, *Journal of Chromatography A*, 694 **(1995)** 277-284

[127] Maestro, version 9.1; LigPrep, version 2.4; Schrödinger, Protein Preparation Wizard; Epik version 2.1; Impact version 5.6; Prime version 2.2; Glide, version 5.6. LLC, New York, **2010**.

[128] LigPrep, version 2.4, Schrödinger, LLC, New York, NY, **2010**.

[129] Schrödinger Suite **2010,** Protein Preparation Wizard

[130] Epik version 2.1, Schrödinger, LLC, New York, NY, **2010**

[131] Impact version 5.6, Schrödinger, LLC, New York, NY, **2010**

[132] Prime version 2.2, Schrödinger, LLC, New York, NY, **2010**.

[133] Glide, version 5.6, Schrödinger, LLC, New York, NY, **2010**.

[134] R.A. Friesner, J.L. Banks,; R.B. Murphy, T.A. Halgren, J.J. Klicic,; D.T. Mainz,; M. P. Repasky, E. H. Knoll; D. E. Shaw, M. Shelley; J. K. Perry, P Francis and P. S. Shenkin, Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy, *J. Med. Chem.*, 47 (**2004)** 1739-1749

[135] T.A. Halgren, R.B. Murphy, R.A. Friesner, H.S. Beard, L.L. Frye, W.T. Pollard, and J.L. Banks, Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening., *J. Med. Chem.*, 47 **(2004)** 1750–1759.

[136] S. Kawatkar, H. Wang, R. Czerminski, and D. Joseph-McCarthy, Virtual fragment screening: an exploration of various docking and scoring protocols for fragments using Glide., *J. Comput. Aided Mol. Des,* 23 **(2009)** 527–539.

[137] M.D. Eldridge, C.W. Murray, T.R. Auton, G.V. Paolini, and R.P. Mee,. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput. Aided Mol. Des.***,** 11, (**1997)** 425–445.

[138] J. Østergaard, H. Jensen and R Holm, Use of correction factors in mobility shift affinity capillary electrophoresis for weak analyte – ligand interactions, *J. Sep. Sci*., 32 **(2009)** 1712–1721

[139] K.D. Altria, Essential Peak Area Normalisation for Quantitative Impurity Content Determination by Capillary Electrophoresis, *Chromatographia*, 35 **(1993)** 117-182.

[140] S.M. Han and N. Purdie, Cyclodextrin complexation of barbiturates in aqueous solution., *Anal. Chem*., 56 **(1984)** 2825

[141] C. Yan, Z. Xiu, X. Li, and C. Hao, Molecular modeling study of β-cyclodextrin complexes with (+)-catechin and (-)-epicatechin, *Journal of Molecular Graphics and Modelling*, 26 **(2007)** 420–428

[142] X.H. Wen, Z.Y. Liu, and T.Q. Zhu, Mass spectrometry and molecular modelling studies on the inclusion complexes between a, b-cyclodextrins and simvastatin, *Chem. Phys. Lett.*, 405 **(2005)** 114–117.

[143] K. Valko, S. Nunhuck, C. Bevan, M.H. Abraham and D.P. Reynolds, Fast Gradient HPLC Method To Determine Compounds Binding To Human Serum Albumin. Relationships with Octanol/Water and Immobilized Artificial Membrane Lipophilicity, *Journal of Pharmaceutical Sciences*, 92 **(2003)** 2236–2248

[144] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. De Jong, P.J. Lewi, and J. Smeyers-Berbeke, in Handbook of Chemometrics and Qualimetrics, Part A (Eds: B.G.M. Vandeginste, S.C. Rutan), *ELSEVIER, Amsterdam*, (**1997)** pp 112-114.

[145] RobStat.xla 1.0, Robust statistic toolkit. The Royal Society of Chemistry. AMC technical brief number 6.

[146] Maestro, version 9.1, Schrödinger, LLC, New York, NY, **2010**

[147] R.A. Friesner, R.B. Murphy, M.P. Repasky; L.L. Frye, J.R. Greenwood;, T.A. Halgren, P.C. Sanschagrin, and D.T Mainz., Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes. *J. Med. Chem.*, *49*, (**2006)** 6177–6196.

[148] H. Koldso, K. Severinsen, T.T. Tran, L. Celik, H.H. Jensen, O. Wiborg, B. Schiøtt, and S. Sinning,The two enantiomers of citalopram bind to the human serotonin transporter in reversed orientations., *J. Am. Chem. Soc.*, 132 **(2010)** 1311-1322.

[149] S. Kawatkar, H. Wang, R. Czerminski, and D. Joseph-McCarthy, Virtual fragment screening: an exploration of various docking and scoring protocols for fragments using Glide, *J. Comput. Aided Mol Des*, 23 **(2009)** 527–539.

[150] W. Sherman, T. Day, M.P. Jacobson, R.A. Friesner, and R. Farid, "Novel Procedure for Modeling Ligand/Receptor Induced Fit Effects," *J. Med. Chem.*, *49* **(2006)** 534-553.

[151] *http://pubchem.ncbi.nlm.nih.gov/summary/summary.cgi?cid=9064* (Accessed **18 May 2011**).

[152] J.M. Herrero-Martínez, M. Sanmartin, M. Rosés; E. Bosch, and C. Ràfols, Determination of dissociation constants of flavonoids by capillary electrophoresis, *Electrophoresis*, 26 **(2005)** 1886–1895.

Chiral

"Handed'', having the characteristic of "handedness", which is having the potential to exist as two non-superimposable structures that are mirror images; not synonymous with the terms enantiomerically pure or optically active.

Enantiomers

Two stereoisomers whose molecules are non-superimposable mirror images of one another.

Diastereoisomerism

Stereoisomerism other than enantiomerism and *cis-trans* isomerism. Diastereoisomers (or diastereomers) are stereoisomers not related as mirror images. Diastereoisomers are characterized by differences in physical properties, and by differences in chemical behaviour towards achiral as well as chiral reagents.

Displacer

Is a drug which displaces the other when two drugs compete for the binding site.

Docking

Is a computational scheme that attempt to find the "best" match between two molecules: a receptor and a ligand.

Enantiomeric excess (ee)

Percent by which one enantiomer, E_1, is in excess in a mixture of the two.

Enantioselectivity

Degree to which a chemical reaction produces one rather than the opposite enantiomer.

R or S:

Absolute configuration about a disymmentric center, assigned by drawing the molecule according to specific conventions and finding whether a circle passes through certain atoms surrounding the center clockwise to the right (R, rectus) or counter clockwise to the left (S, sinister).

Racemic

Existing as a racemate, or 50-50 mixture of the two enantiomers; also denoted dl (discouraged) or (+/-) (preferred). Racemates are also called racemic mixtures or racemic modifications.

Receptor

A biological macromolecule that interact with a pharmaceutical agent.

Stereoisomers

Compounds whose molecules consist of the same number and types of atoms with the same connectedness but different configurations.

Pharmacokinetics

Describes how the body affects a specific drug after administration. This includes the study of the mechanisms of absorption and distribution of an administered drug, the rate at which a drug action begins and the duration of the effect, the chemical changes of the substance in the body and the effects and routes of excretion of the metabolites of the drug.

Pharmacophore

A pharmacophore is the ensemble of steric and electronic features that are necessary to ensure the optimalsupramolecular interactions with a specific biological target structure and to trigger (or to block) its biological response.

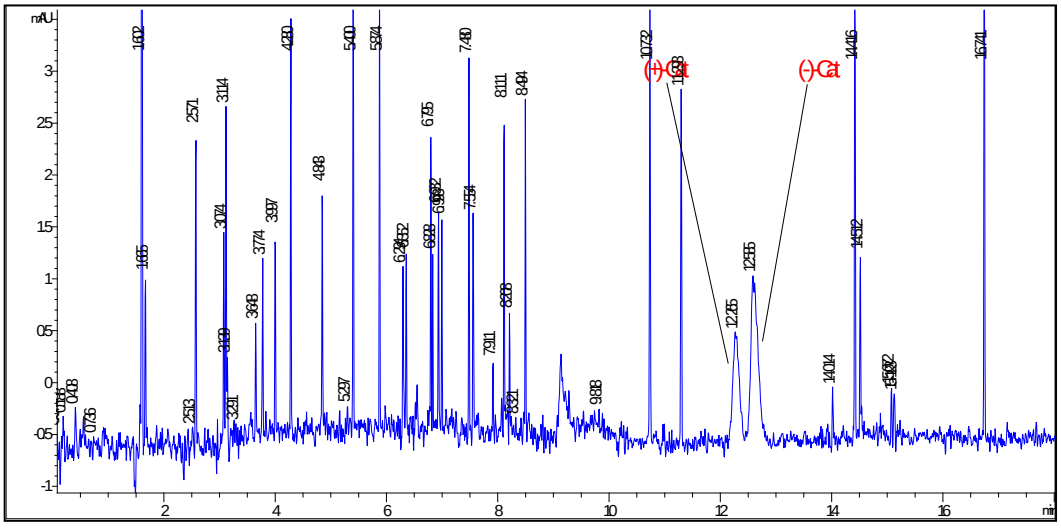
Ultrafiltration

Is a technique where two or more compounds that were mixed together to an extent that they reach equilibrium are separated by filtration at very high revolutions per minute (rpm). Filtration is based on the size different sizes of compound present in the sample, small molecules are filtered through while bigger molecules are left behind. There are different types and sizes of filters (e.g. cellulose or nylon filter 45 μm) that are available for this technique; the selection can be based on the nature of the sample and quantity.

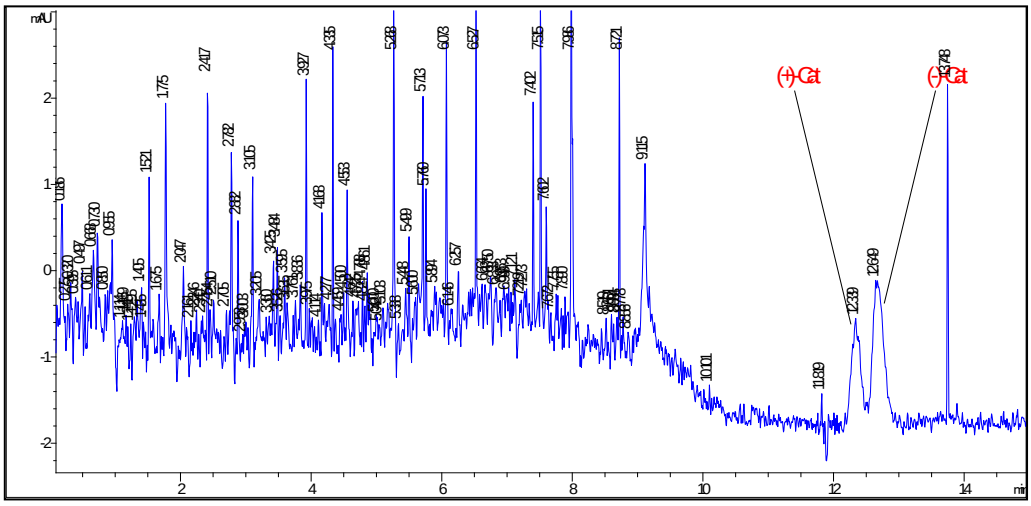
Results for the study of change in enantiomeric excess of (±)-Catechin with time using 200 µM standard

		(+)-Catechin			(-)-Catechin			
Day	Time	$t_{m,}$/min	P/Area	Corr.Area	$t_{m/}$min	P/Area	Corr. Area	Ratio
Day 1	AM	12.27	11.70	0.95	12.58	18.60	1.48	1.55
Day 2	AM	12.34	10.10	0.82	12.65	17.40	1.38	1.68
Day 3	AM	12.66	11.10	0.88	12.69	20.20	1.59	1.82
Day 4	AM	10.98	5.80	0.53	11.19	10.70	0.96	1.81

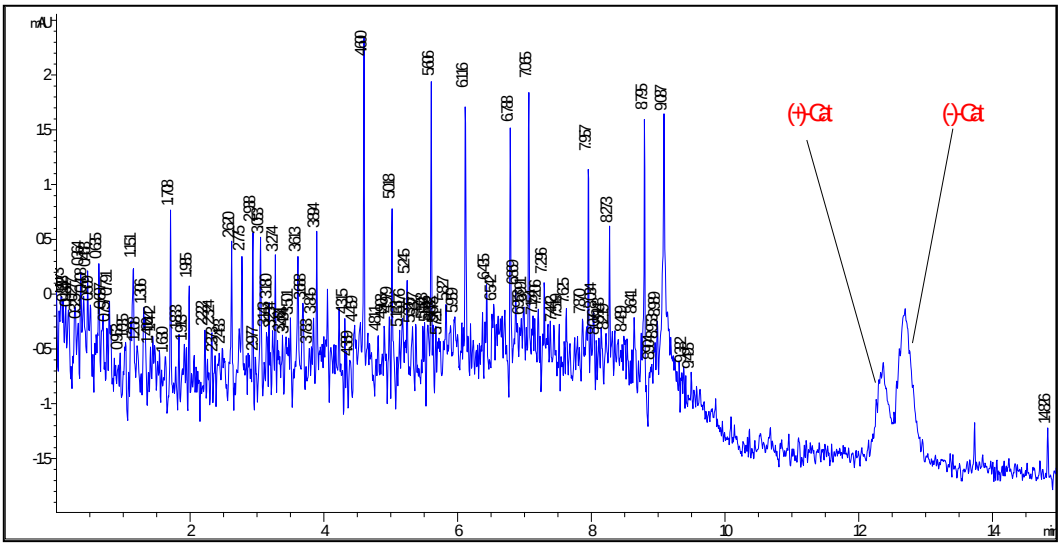
Day 1



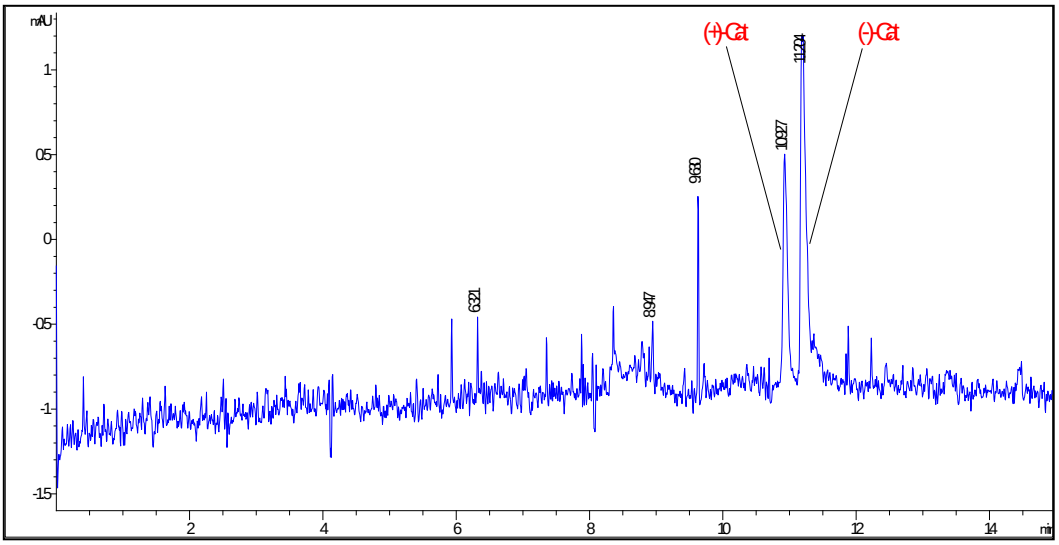
Day 2



Day 3

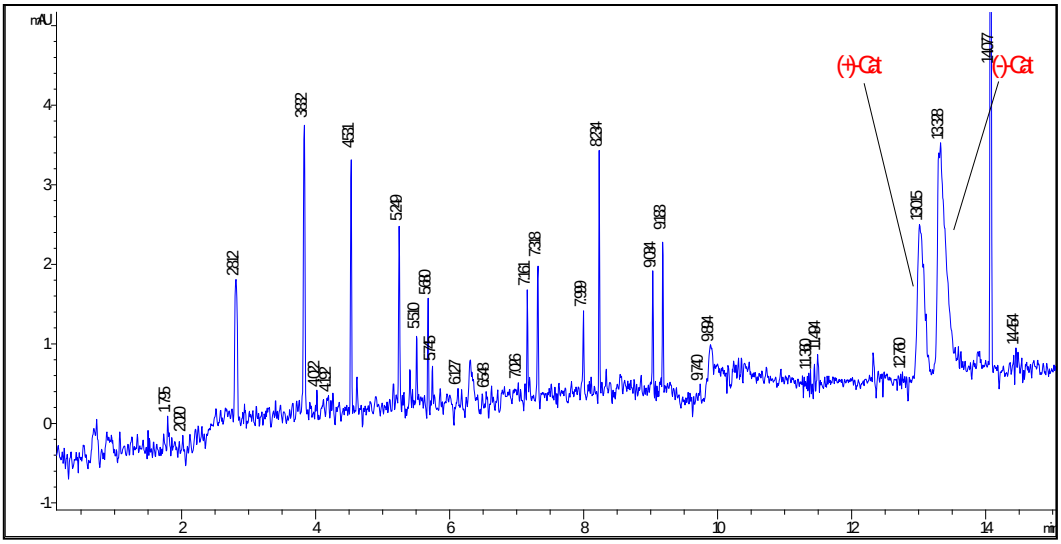


Day 4

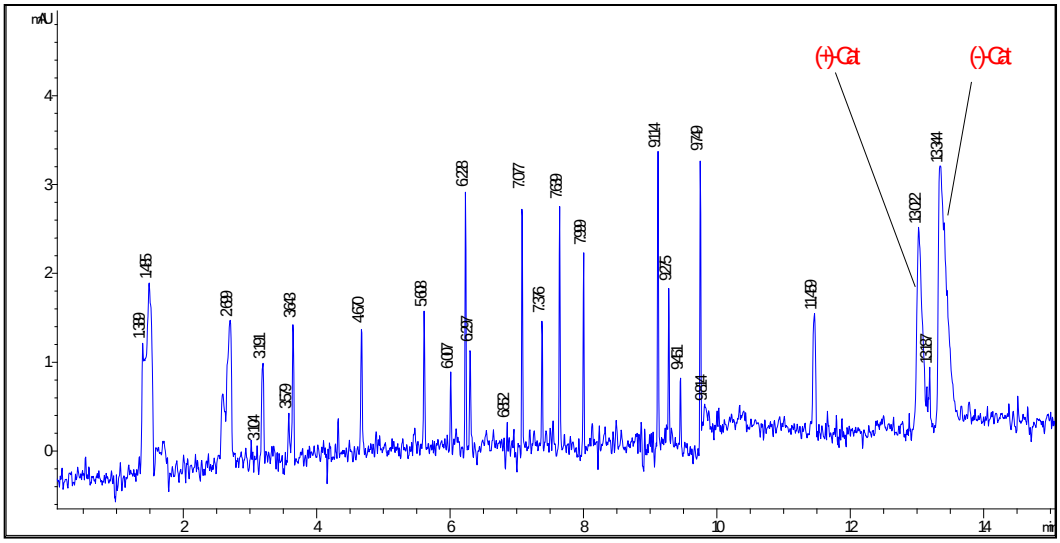


Study of the pH effect on chiral resolution of 200 µM (±)-Catechin

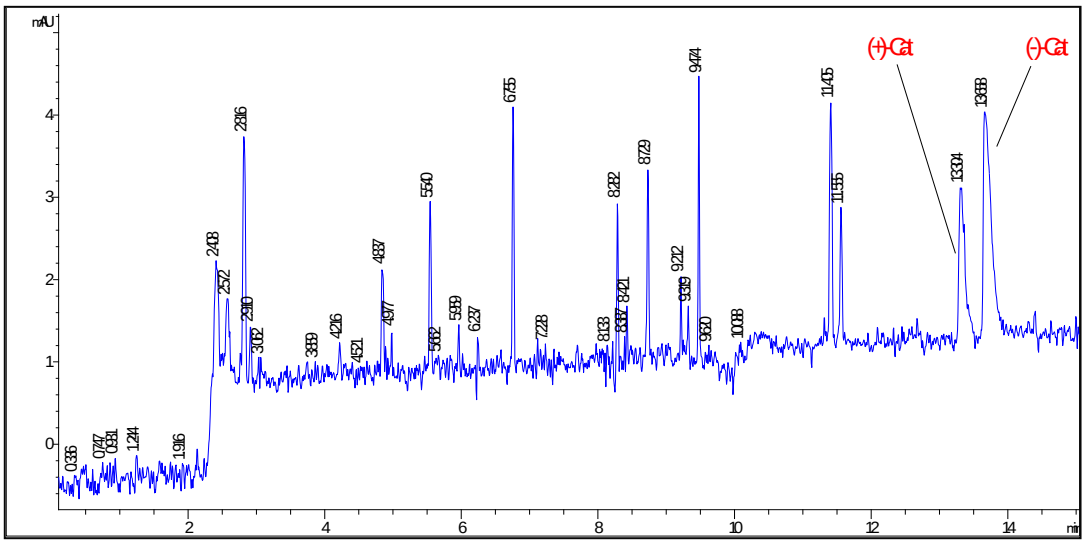
pH 6.0



pH 6.5

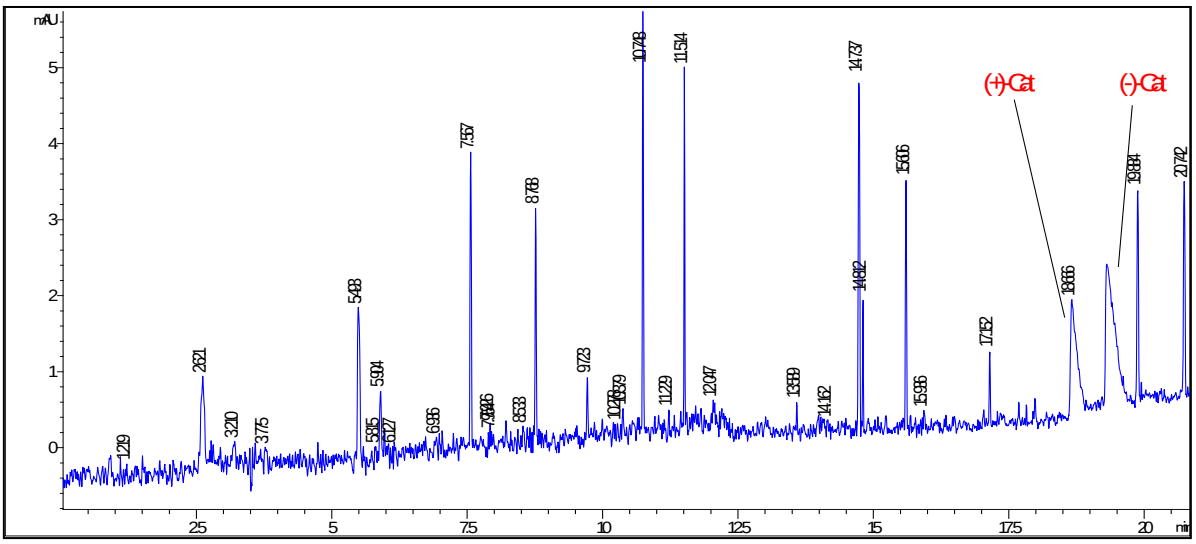


pH 7.0

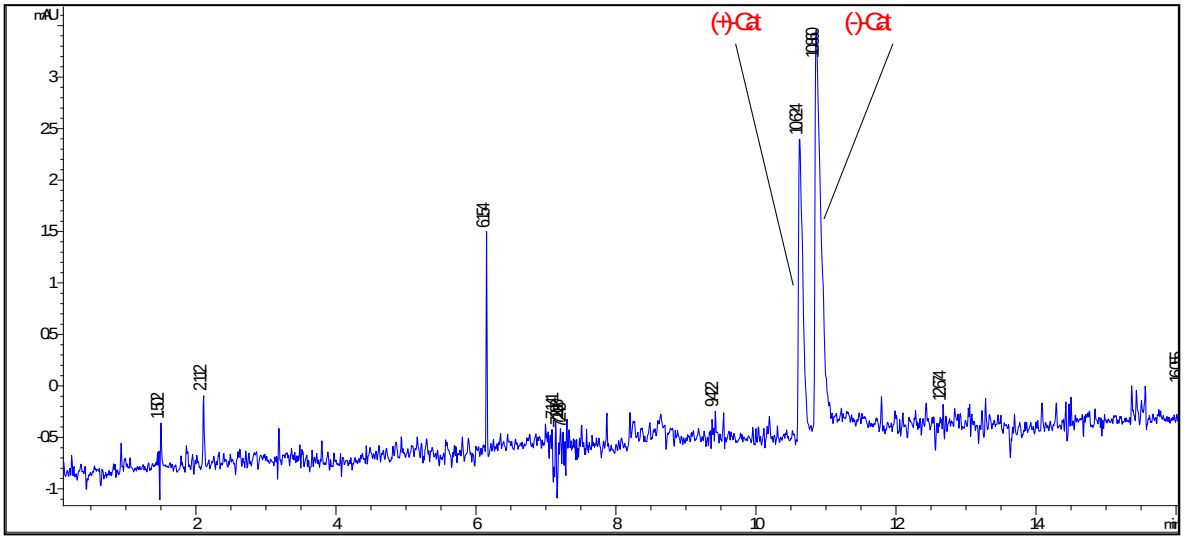


Study of the effect of temperature on chiral resolution 200 µM (±)-Catechin

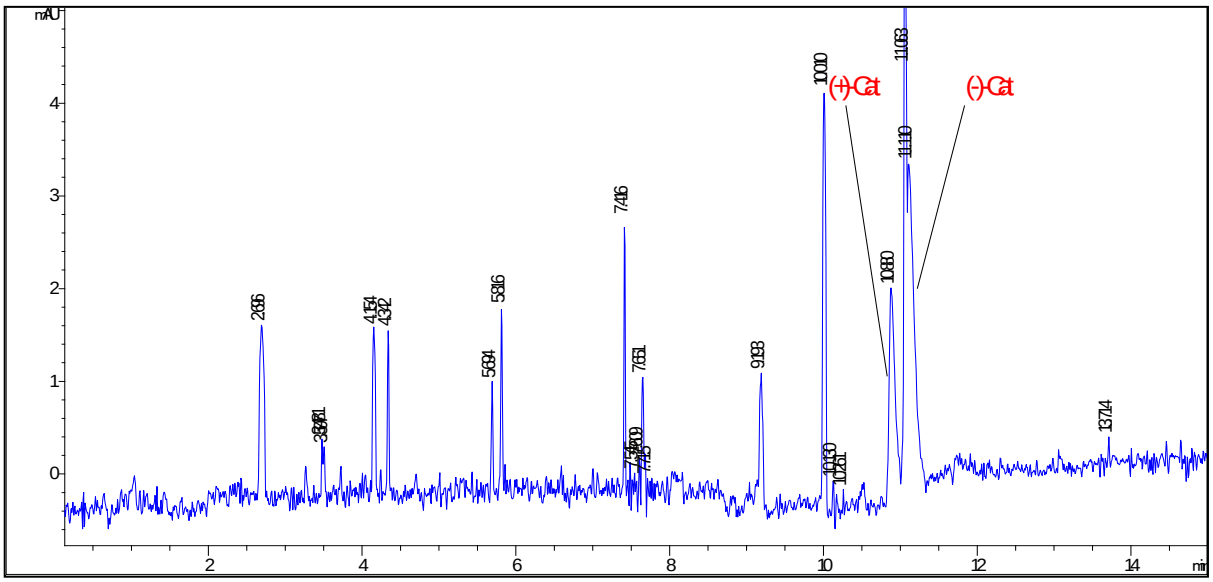
Temperature = 15 ^{o}C



Temperature = 20 ^{o}C



Temperature = 25 ^{o}C

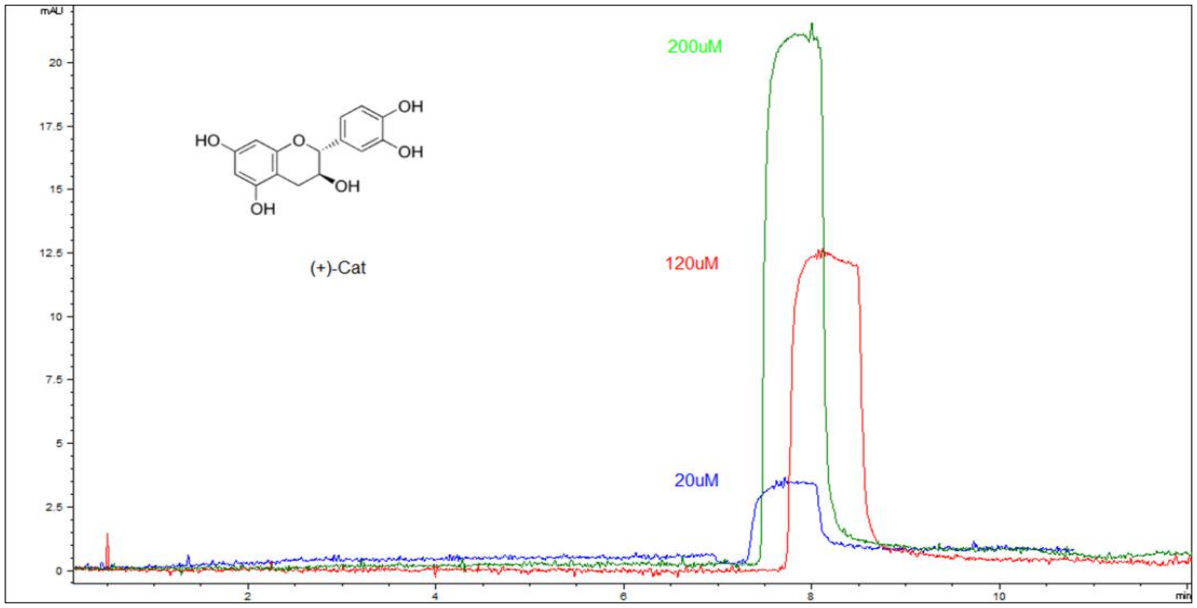


Study of Ligand-Protein Interaction with FA (Discussed in Chapter 6)

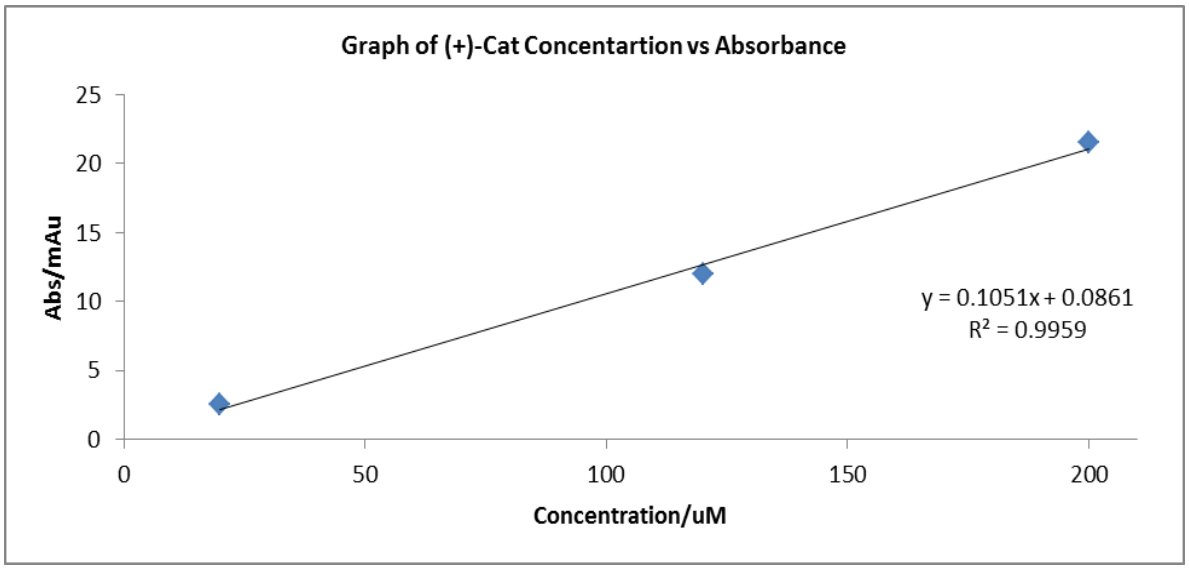
P and D-constant Experimental Design for (+)-Catechin: The experiment was performed using a two buffers (Incubation with 67 mM Phosphate buffer and FA with 50 mM phosphate buffer). Responses measured, 60 s injection time, and 15 kV at 220 nm

ID	Type	HSA/µM		Drug/µM											
						Baseline				P/height					
		P	P_{Exp}	D	$[D]_{exp}$	1	2	3	Mean	1	2	3	Mean	Abs	d
1		525.0		20.00	18.26	0.00	0.00	0.00	0.00	1.60	1.82	2.22	1.88	1.88	11.10
2		**525.0**		**40.00**	**36.52**	**0.23**	**0.30**	**0.12**	**0.22**	**3.03**	**3.53**	**4.17**	**3.58**	**3.36**	**25.00**
3		525.0	**513.2**	80.00	73.04	0.20	0.12	0.12	0.15	4.79	5.79	6.48	5.69	5.54	45.50
4		525.0		120.0	109.6	0.41	0.38	0.35	0.38	6.40	7.41	8.09	7.30	6.92	58.50
5		525.0		160.0	146.1	0.36	0.30	0.24	0.30	7.60	8.78	9.45	8.61	8.31	71.60
6	Mixture	525.0		200.0	182.6	0.40	0.31	0.28	0.33	10.50	11.54	12.63	11.56	11.23	99.00
7		300.0	293.2			0.22	0.14	0.08	0.15	3.20	3.69	3.98	3.62	3.48	26.10
8		375.0	366.5			0.01	0.04	0.07	0.04	3.08	3.48	4.09	3.55	3.51	26.40
9		450.0	439.9			0.26	0.15	0.13	0.18	2.03	2.55	2.99	2.52	2.34	15.40
10		**525.0**	**513.2**	**40.00**	**36.52**	**0.14**	**0.11**	**0.18**	**0.14**	**2.30**	**2.80**	**3.50**	**2.87**	**2.72**	**19.00**
11		600.0	586.5			0.18	0.15	0.10	0.14	2.43	3.20	3.82	3.15	3.01	21.70
12		675.0	659.8			0.18	0.11	0.20	0.16	3.25	4.20	4.87	4.11	3.94	30.50
16	Cal 1			20.00	18.26	0.17	0.10	0.02	0.0959	2.77	3.070	2.94	2.93	2.831	2.831
17	Cal 2			120.0	109.6	0.20	0.22	0.16	0.1933	11.44	12.07	11.63	11.71	11.52	11.52
18	Cal 3			200.0	182.6	0.17	0.16	0.08	0.1367	20.07	20.47	19.88	20.14	20.00	20.00

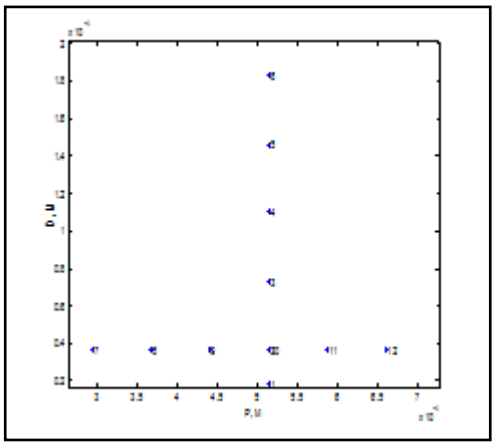
Overlaid signals of (+)-Catechin 20 µM, 120 µM and 200 µM



Calibration curve of 20 µM, 120 µM and 200 µM (+)-Catechin



Structure of the experimental design (D and P constant)



Outliers on K1

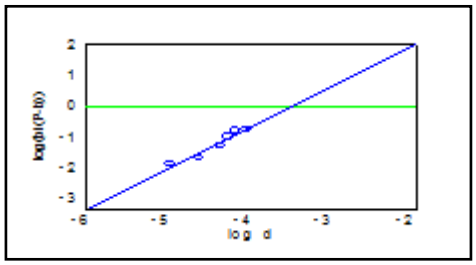
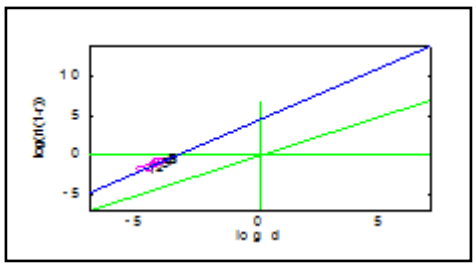
Outliers on K1	**Grubbs' Test Results**
z-critic : 2.5 : z-score based on median±MADe (Eq. 7.2a and 7.2b) 1345.4 ± 697.9 index of eliminable-points: 9 **Eliminating 9** z-critic : 2.5 : z-score based on median±MADe 1281.9 ± 541.3	Value number 9 IS NOT AN OUTLIER (alpha= 2.5%, 2 tails) Value number 9 and 12 ARE NOT OUTLIERS (alpha= 2.5%, 2 tails)

Outliers on log K1

z-critic : 2.5 : z-score based on median±MADe 3.128 ± 0.214 index of eliminable-points: 12 **Eliminating 12** z-critic : 2.5 : z-score based on median±MADe 3.148 ± 0.178	**Grubbs' Test Results** Value number 12 IS AN OUTLIER (alpha= 2.5%, 2 tails).

<u>P-Constant</u> (Points 1 to 6)

Verification of n

n1 ± IC/2 : 1.31595 ± 0.440 (verification) logK1 ± IC/2 : 4.569 ± 1.932 F = 62.66 ; p = 0.001378 se = 0.1374	 

Summary of Linear Plots

Plot	n_1	Log K_1	Log (n_1K_1)
Klotz	-0.9505	3.094	3.072
Scatter	-0.1783	3.746	2.997
y-rec	. -0.2021	3.696	3.002

Estimates from Eqn 7.2a and 7.2b

ID	logK1-PB	log(K1a3)	log(K1n1
1	3.203	3.102	3.108
2	3.203	2.952	2.962
3	3.203	3.071	3.095
4	3.203	3.231	3.276
5	3.203	3.307	3.375
6	3.203	3.216	3.294
Mean	3.203	3.159	3.192

Non-Linear SIMPLEX

Using Eqn.7.3

P-Constant

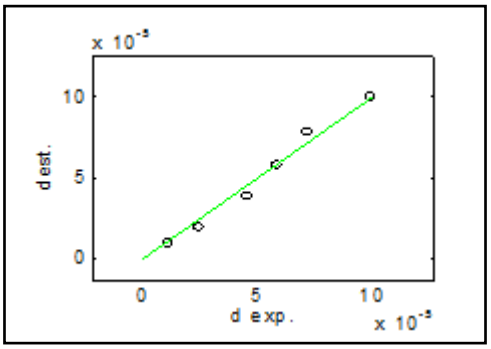
n1-integer restirction for Non-Linear SIMPLEX

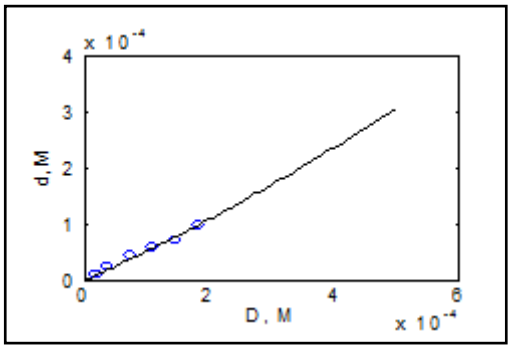
n1 , K1 , logK1 and n1K1 :

R2 = 0.971

log (n1*K1) = 3.291

PB% = 39.34





D-Constant (Points 9 to 12)

Verification of n

n1 ± IC/2 :

-1.77218 ± 2.130

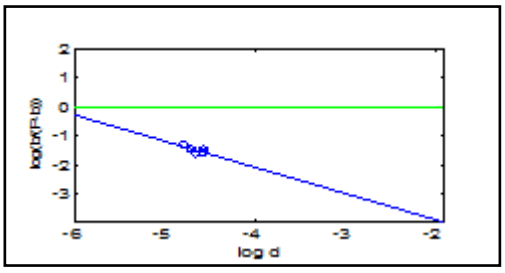
logK1 ± IC/2 :

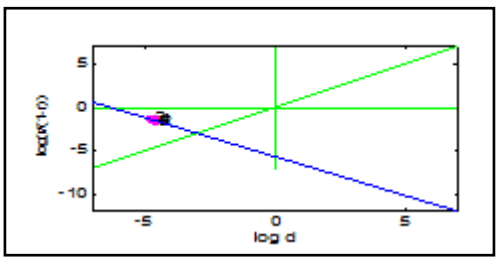
-9.79251 ± 9.896

R2 = 0.5716

F = 5.340

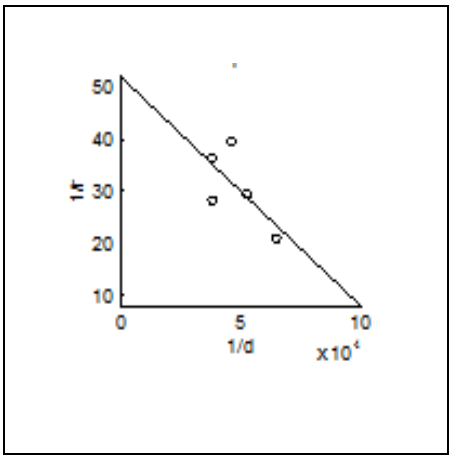
p = 0.0819





Summary of Linear Plots

Plot	n_1	Log K_1	Log (n_1K_1)
Klotz	0.008099	4.849	2.757
Scatter	0.008575	4.836	2.769
y-rec	. 0.006449	4.799	2.608



Estimates from Eqn . 7.2a, 7.2b and 7.3

ID	logK1-PB	log(K1a3)	log(K1n1)
7	3.446	3.134	3.149
8	3.348	3.018	3.030
9	3.269	3.493	3.515
10	3.203	3.254	3.269
11	3.145	3.066	3.077
12	NaN	NaN	NaN
Mean	3.269	3.134	3.149

Based on the findings of z and Grabbs' test point 12 was eliminated completely from the rest of data analysis. Therefore 5 point were left for D-constant part

Non-Linear SIMPLEX

Using Eqn.7. 3

P-Constant

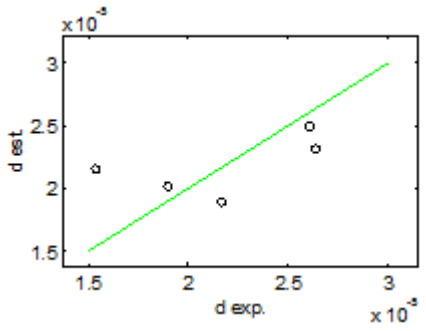
n1-integer restirction for Non-Linear SIMPLEX

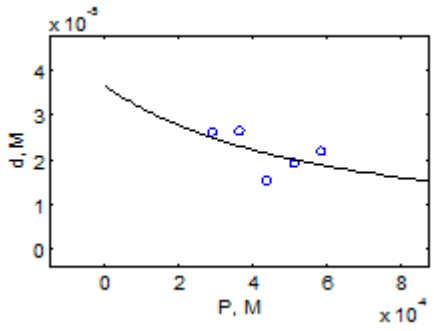
n1 , K1 , logK1 and n1K1 :

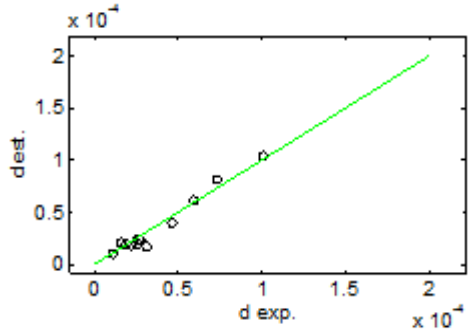
R2 = 0.3375

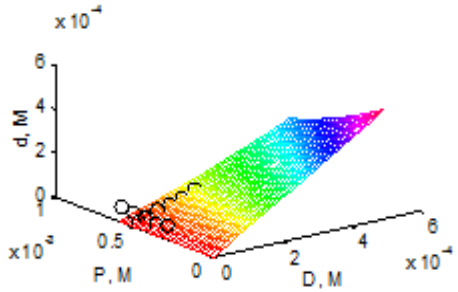
log (n1*K1) = 3.217

PB% = 54.08

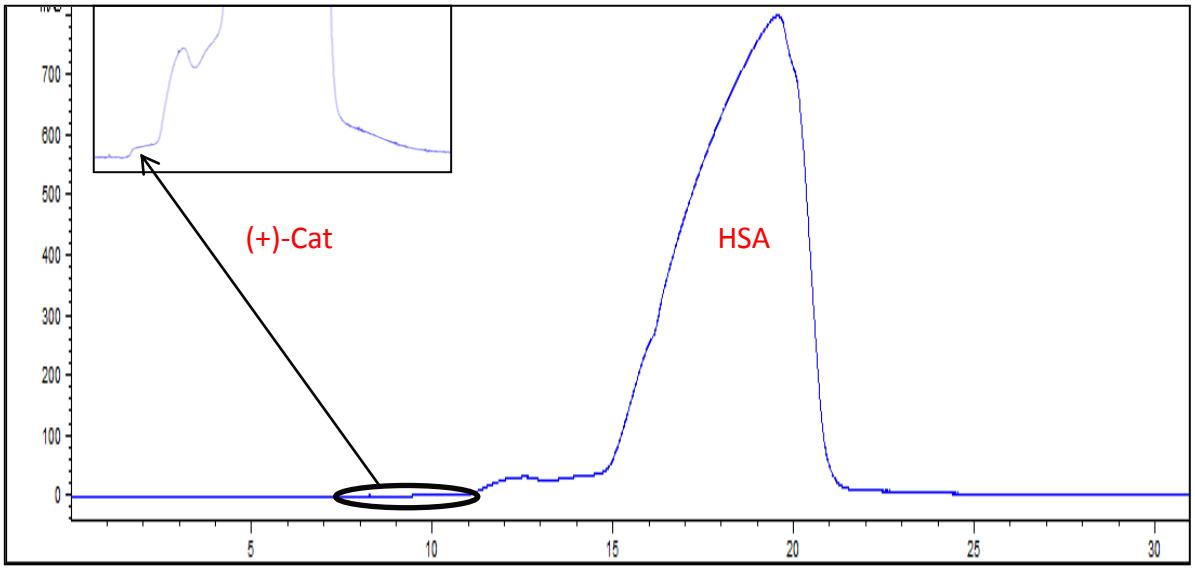




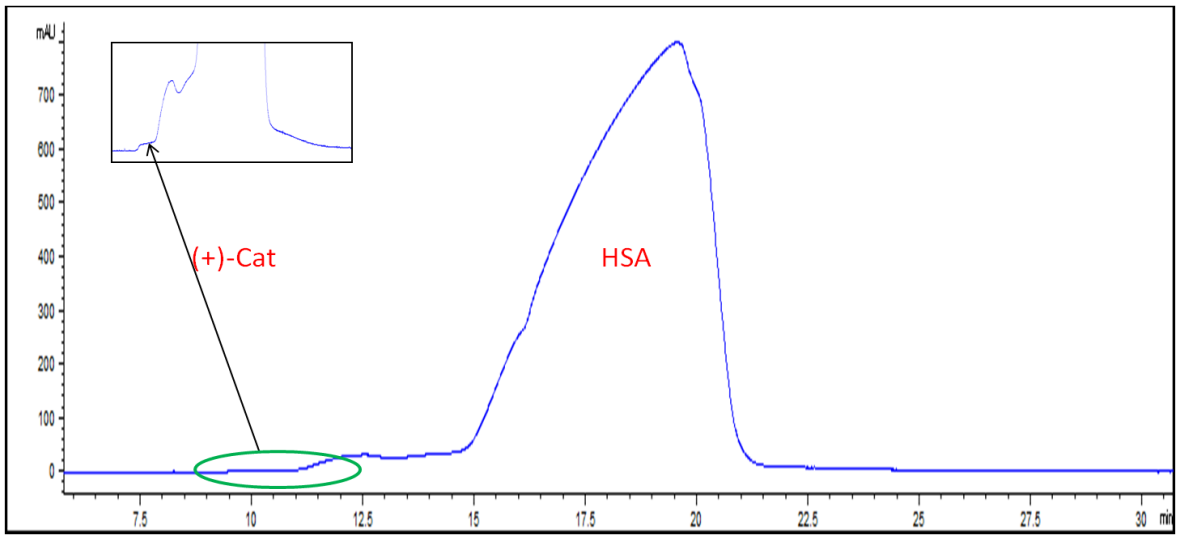




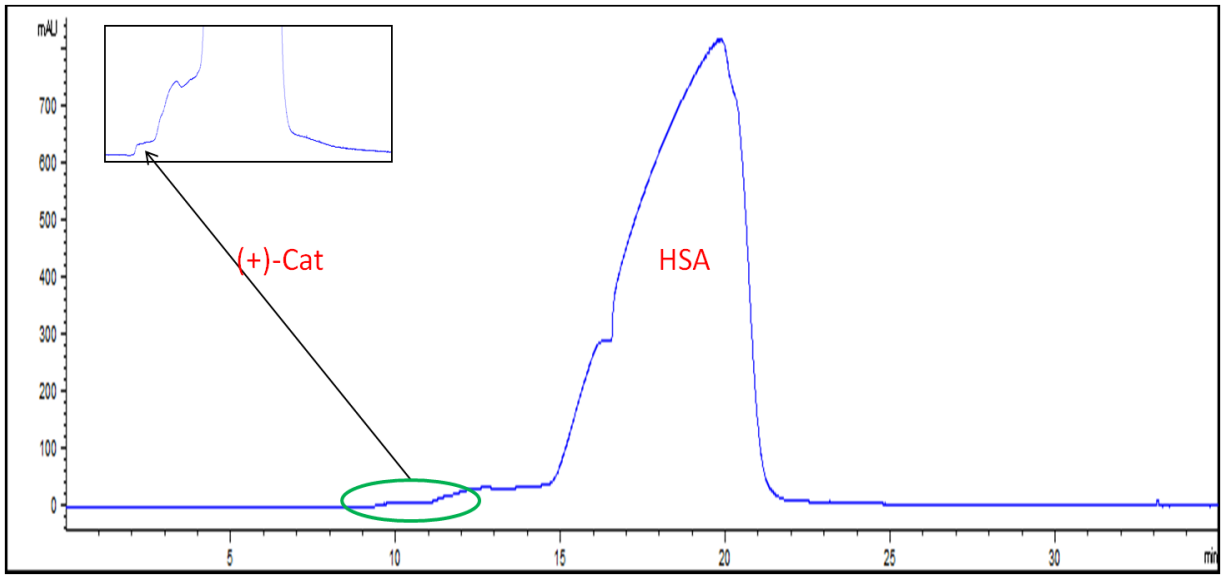
(+)-Catechin 20 µM_HSA 525 µM



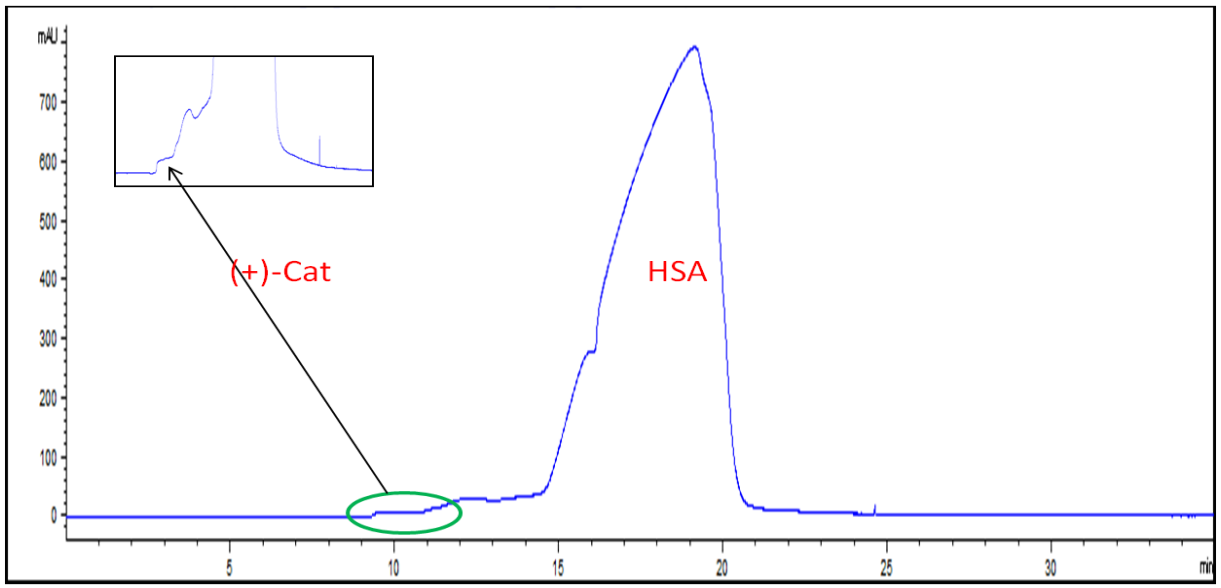
(+)-Catechin 40 µM_HSA 525 µM



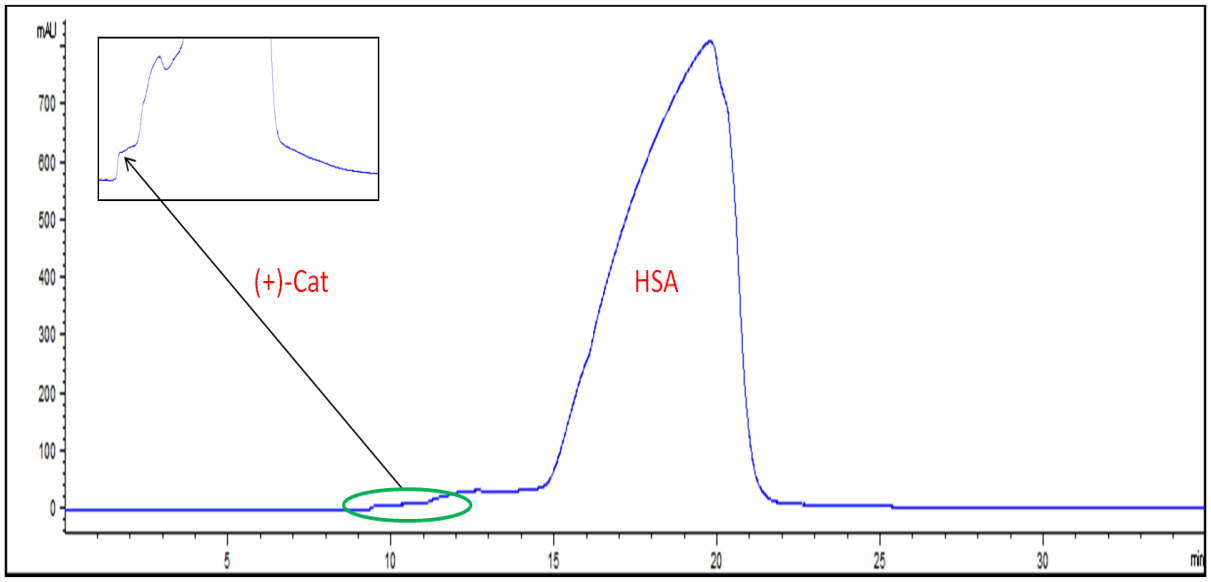
(+)-Catechin 80 µM_HSA 525 µM



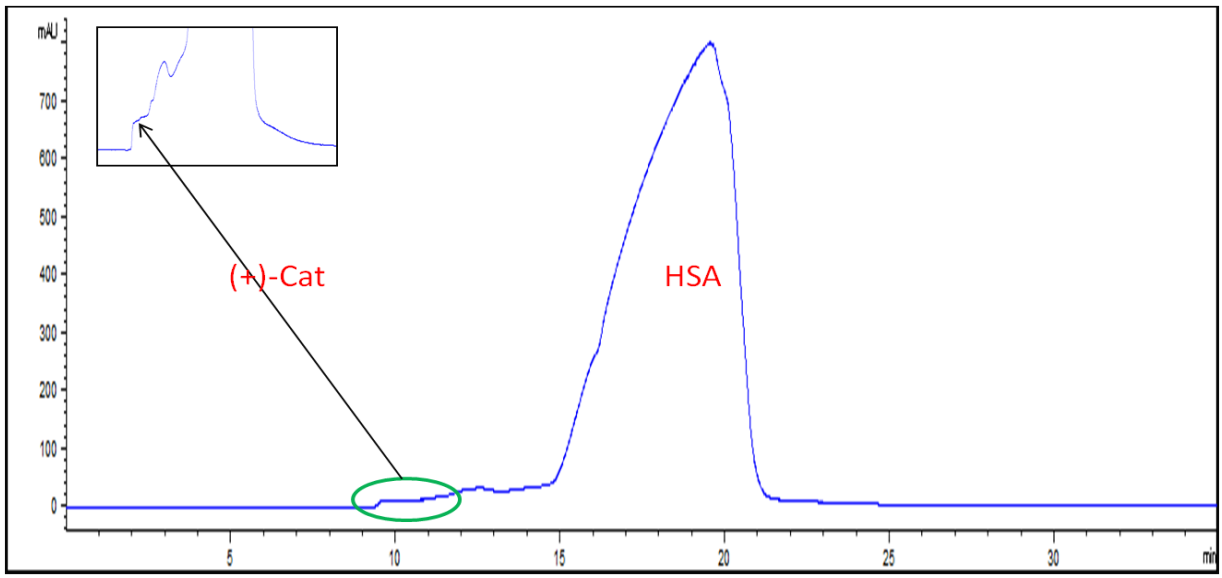
(+)-Catechin 120 µM_HSA 525 µM



(+)-Catechin 160 µM_HSA 525 µM

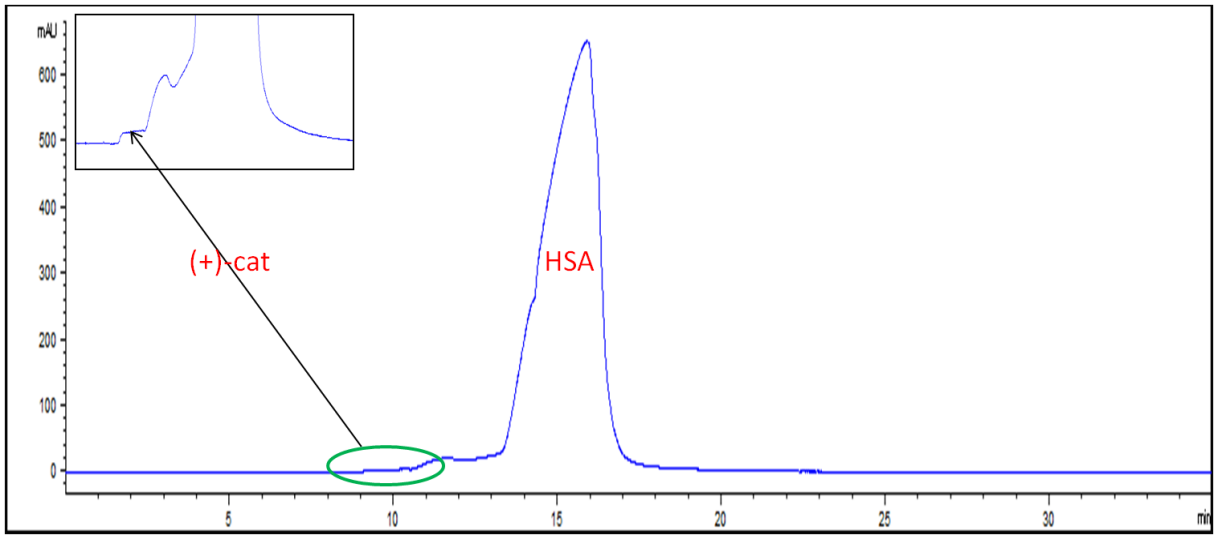


(+)-Catechin 200 µM_HSA 525µM

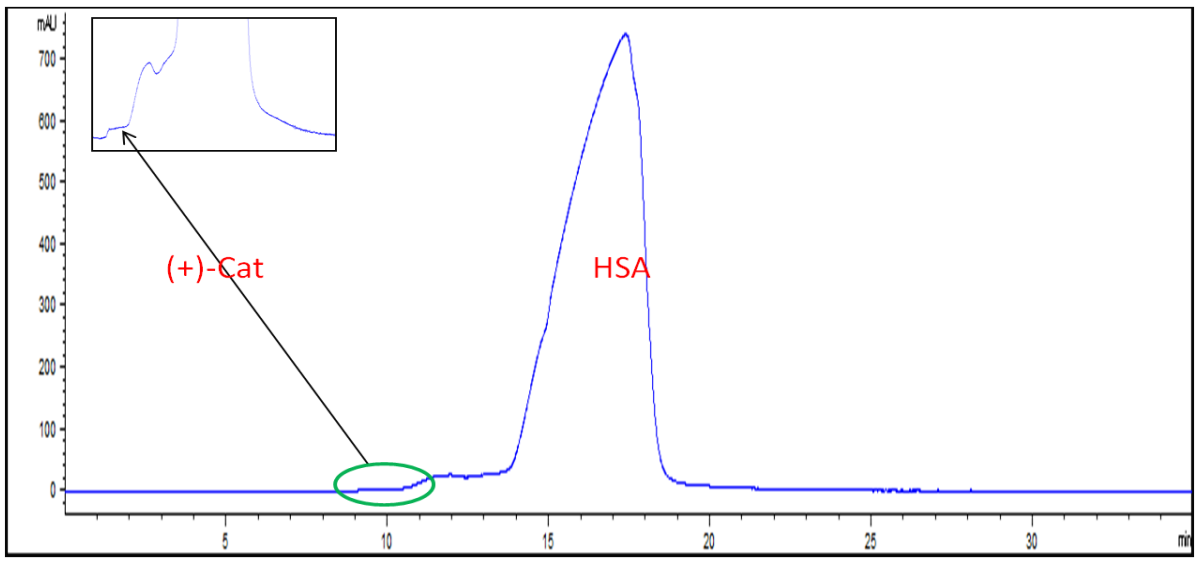


D-constant design (7 – 11)

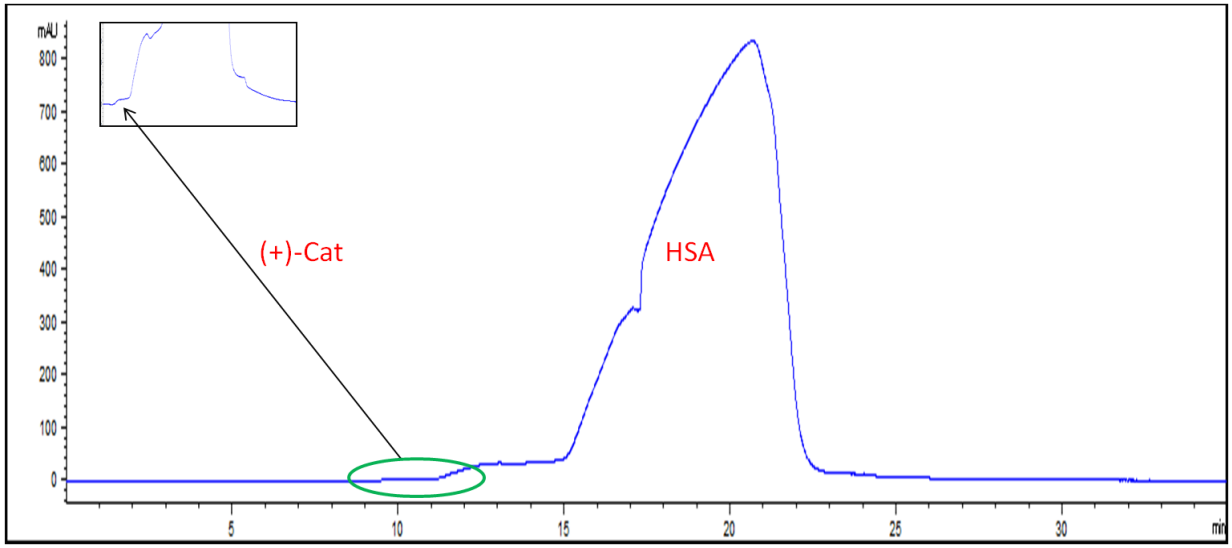
(+)-Catechin 40 µM_HSA 300 µM



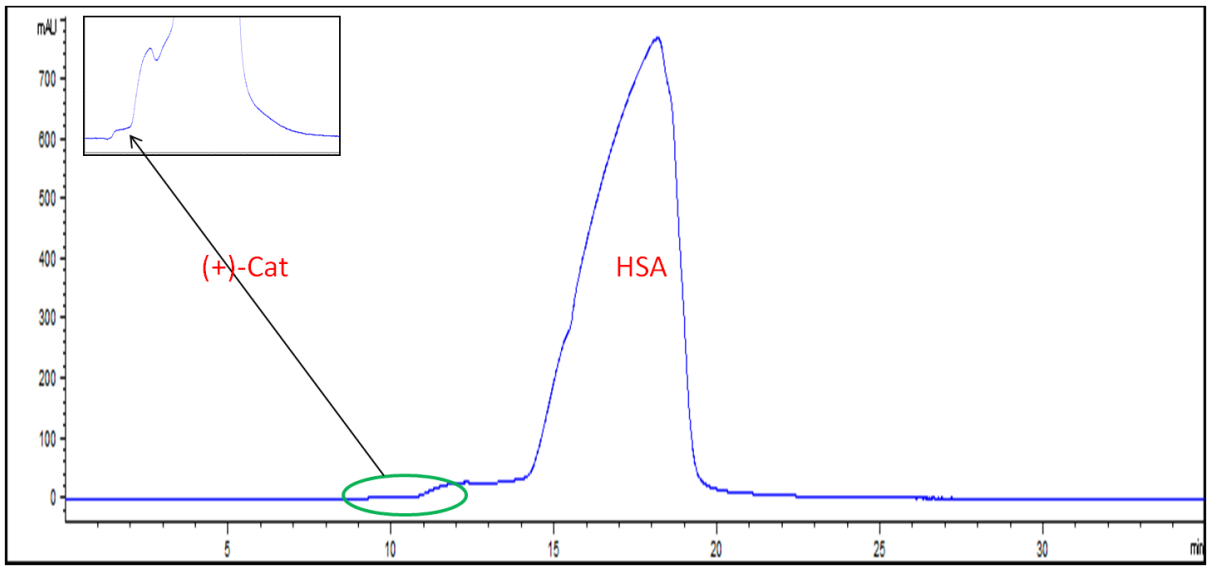
(+)-Catechin 40 µM_HSA 375 µM



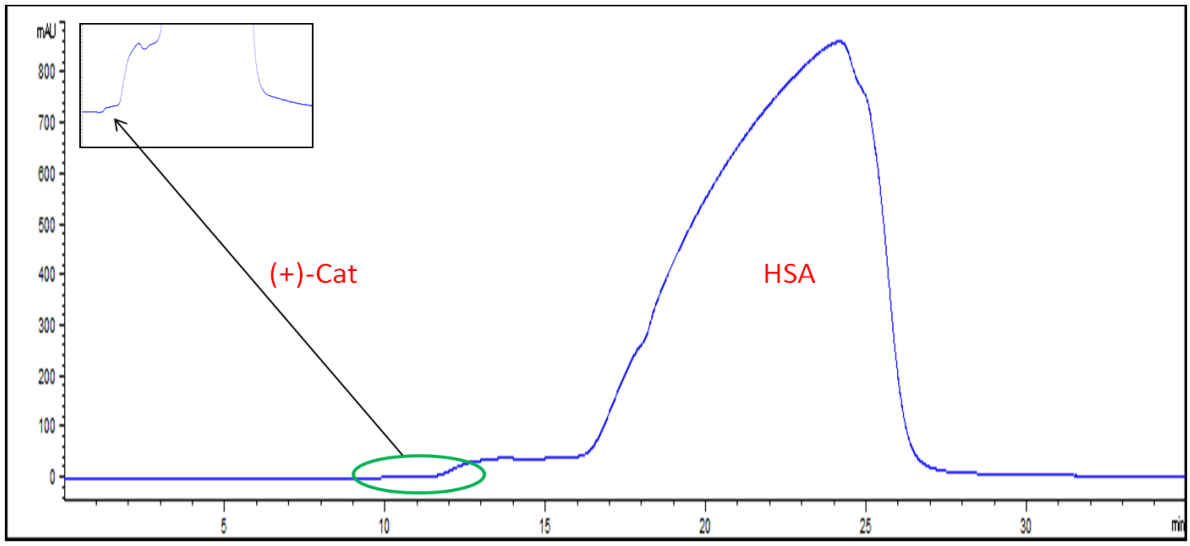
(+)-Catechin 40 µM_HSA 450 µM



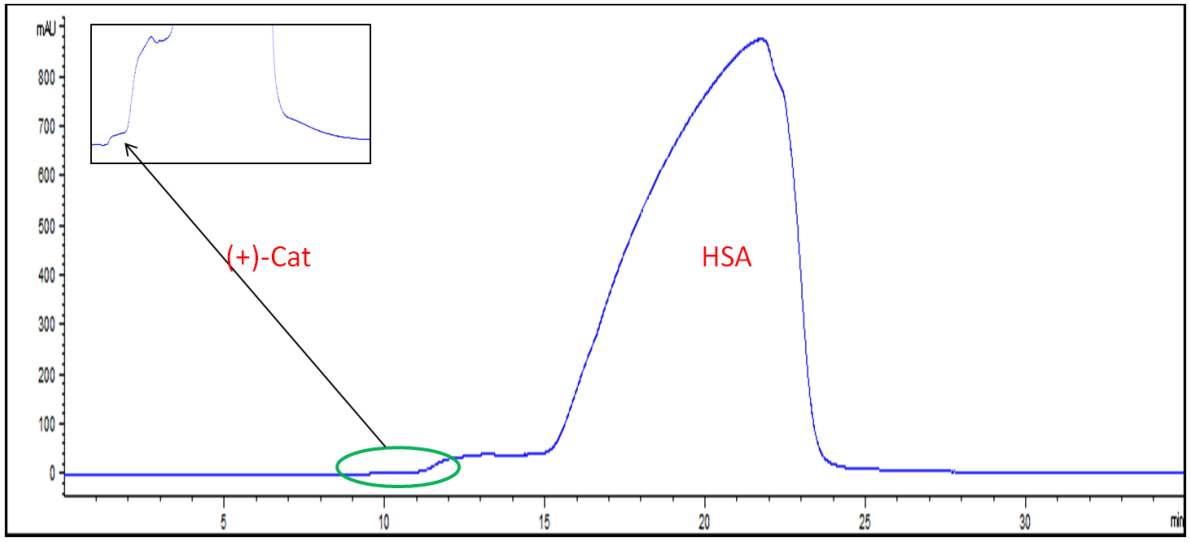
(+)-Catechin 40 µM_HSA 525µM



(+)-Catechin 40 µM_HSA 600 µM



(+)-Catechin 40 µM_HSA 675 µM

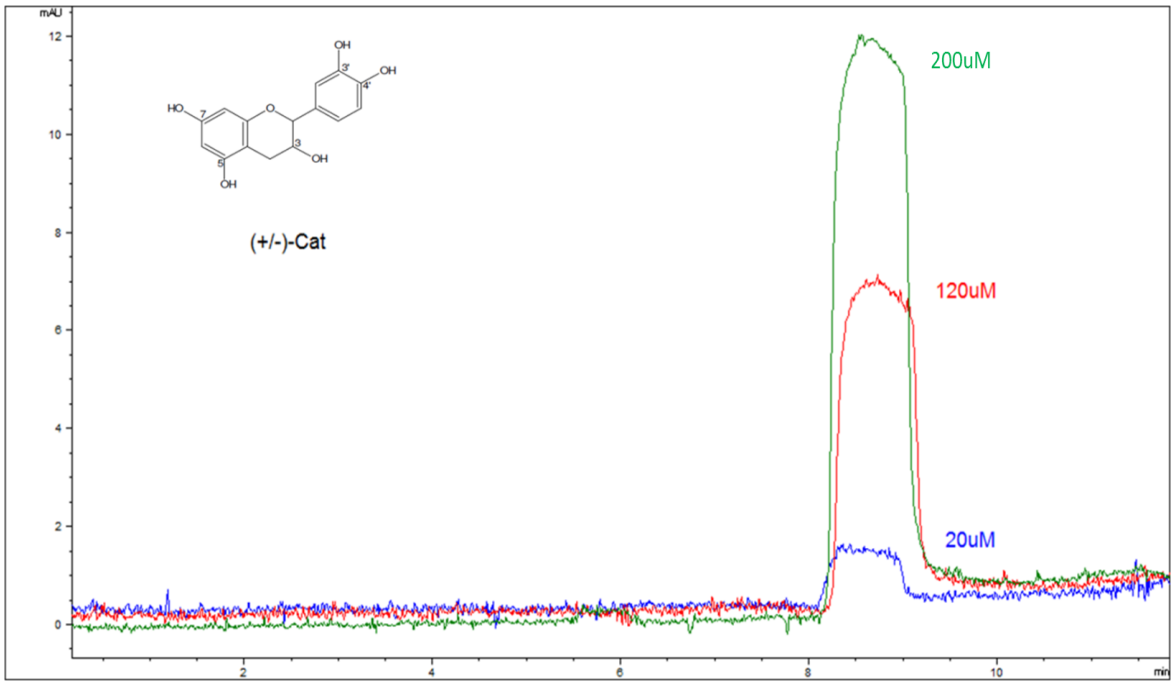


P and D-constant Experimental Design for (±)-Catechin: The experiment was performed using a two buffers (Incubation with 67 mM Phosphate buffer and FA with 50 mM phosphate buffer). Responses measured, 60 s injection time, and 15 kV at 220 nm

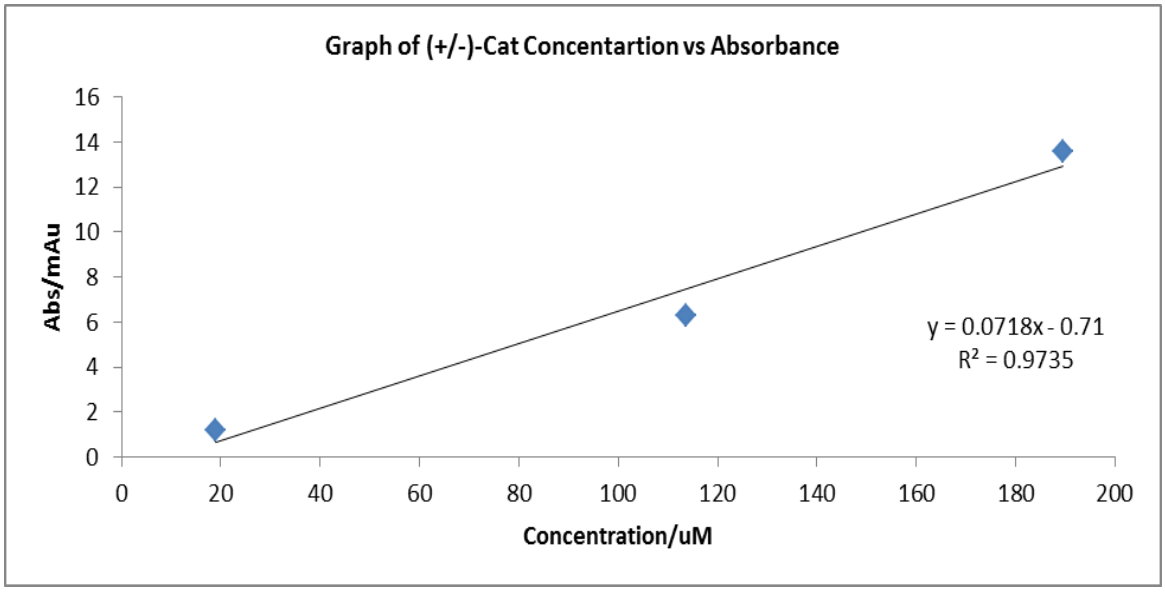
ID	Type	HSA/uM		Drug/uM											
		P	$P_{, Exp}$	D	D_{Exp}	Baseline				Peak height					d
						1	2	3	Mean	1	2	3	Mean	Abs	
1				20.00	18.95	0.08	0.03	0.01	0.04	1.21	1.58	1.90	1.56	1.52	31.00
2				**40.00**	**37.90**	**0.07**	**0.06**	**0.12**	**0.08**	**1.42**	**1.76**	**2.14**	**1.77**	**1.69**	**33.40**
3		525.00	522.95	80.00	75.79	0.23	0.26	0.28	0.26	2.64	3.06	3.56	3.09	2.83	49.30
4				120.00	113.69	0.23	0.33	0.26	0.27	3.23	3.70	4.01	3.65	3.37	56.80
5				160.00	151.58	0.00	0.00	0.00	0.00	3.75	4.45	5.19	4.46	4.46	72.00
6	Mixture			200.00	189.48	0.47	0.36	0.41	0.41	4.53	5.33	6.86	5.57	5.16	81.70
7		300.00	298.83			0.11	0.18	0.23	0.18	1.96	2.18	2.56	2.23	2.06	38.50
8		375.00	373.53			0.11	0.15	0.26	0.17	1.87	2.13	2.34	2.11	1.94	36.90
9		450.00	448.24			0.04	0.06	0.02	0.04	1.32	1.70	1.79	1.60	1.57	31.70
10		**525.00**	**522.95**	**40.00**	**37.90**	**0.31**	**0.16**	**0.23**	**0.23**	**1.57**	**1.93**	**2.19**	**1.90**	**1.67**	**33.00**
11		600.00	597.65			0.13	0.01	0.02	0.05	1.60	2.08	2.44	2.04	1.98	37.50
12		675.00	672.36			0.04	0.06	0.07	0.06	1.63	2.00	2.67	2.10	2.05	38.40
13	Cal 1			20.00	18.95	0.13	0.04	0.13	0.10	1.27	1.30	1.24	1.27	1.17	
14	Cal 2			120.00	113.69	0.25	0.17	0.13	0.18	6.37	6.76	6.28	6.47	6.29	
15	Cal 3			200.00	189.48	0.21	0.13	0.35	0.23	13.88	14.10	13.36	13.78	13.55	

*ID 2 and ID 10 were used as control samples/points

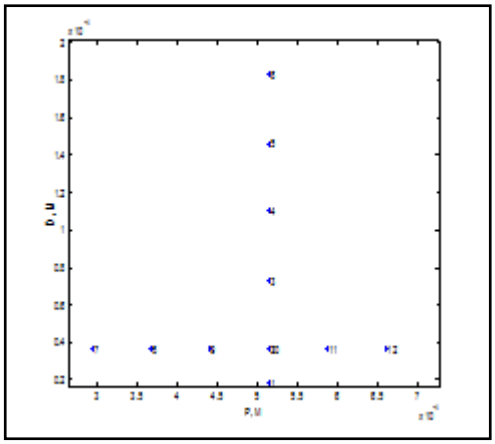
Overlaid signals of 20 µM, 120 µM and 200 µM (±)-Catechin



Calibration curve of 20 µM, 120 µM and 200 µM (±)-Catechin



Srtucture of the experimental design (D and P constant)



Outliers on K1

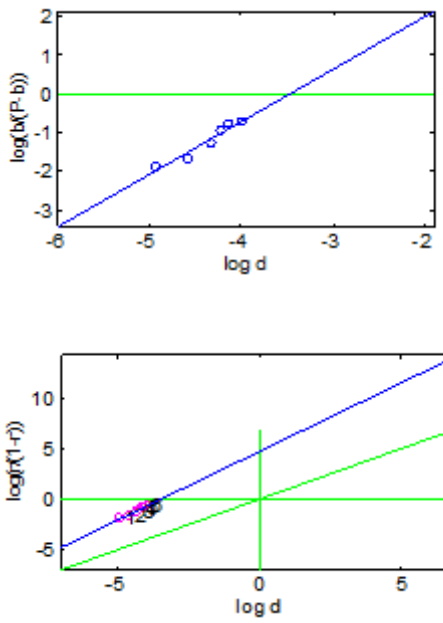
z-score analysis on x	**Grubbs' Test Results**
z-critic : 2.5 : z-score based on median±MADe (Eqn 7.2a and 7.2 b) 1208.2 ± 676.54 index of eliminable-points: 9	Value number 9 IS NOT AN OUTLIER (alpha= 2.5%, 2 tails)
Eliminating 9 **z-critic : 2.5 :** z-score based on median±MADe 1157.1 ± 491.6 No outlier	Value number 9 and 12 ARE NOT OUTLIERS (alpha= 2.5%, 2 tails)

Outliers on log K1

z-critic : 2.5 : z-score based on median±MADe 3.0818 ± 0.2397 index of eliminable-points: 12 **Eliminating 12** **z-critic : 2.5 :** z-score based on median±MADe 3.100 ± 0.2075 index of eliminable-points: No outlier	**Grubbs' Test Results** Value number 12 IS AN OUTLIER (alpha= 2.5%, 2 tails).

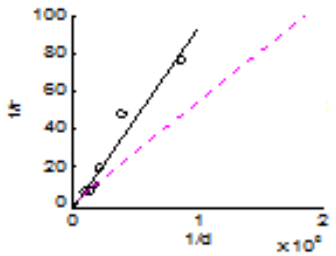
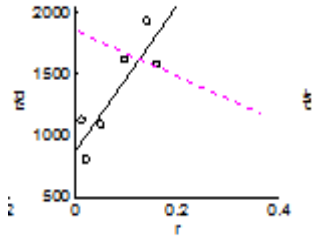
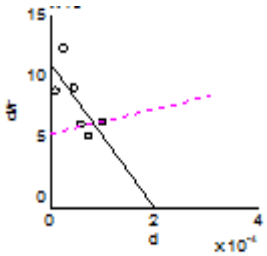
P-Constant (Points 1 to 6)

Verification of n

n1 ± IC/2 : 1.35811 ± 0.455062(verification) logK1 ± IC/2 : 4.718± 1.993 R2 = 0.94495 ; F = 68.66 ; p = 0.001158 se = 0.1255 high value of stdev	

Summary of Linear Plots

Plot	n_1	Log K_1	Log (n_1K_1)
Klotz	0.503646	3.31469	3.01682
Scatter	0.150604	3.77002	2.94786
y-rec	-0.167812	3.72738	2.9522

Klotz	Scatter	y-rec
		

Estimates from Eqn . 7.2a, 7.2b and 7.3

ID	logK1-PB	log(K1a3)	log(K1n1)
1	3.203	3.052	3.057
2	3.203	2.908	2.917
3	3.203	3.040	3.063
4	3.203	3.211	3.255
5	3.203	3.286	3.353
6	3.203	3.199	3.275
Mean	3.203	3.125	3.159

Non-Linear SIMPLEX

Using Eqn 7.3

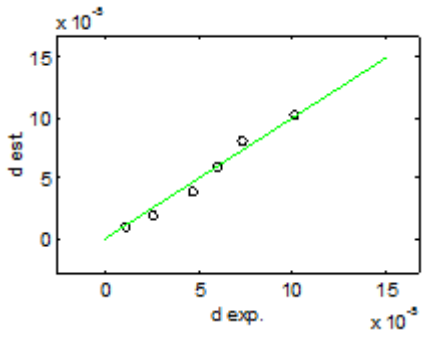
P-Constant

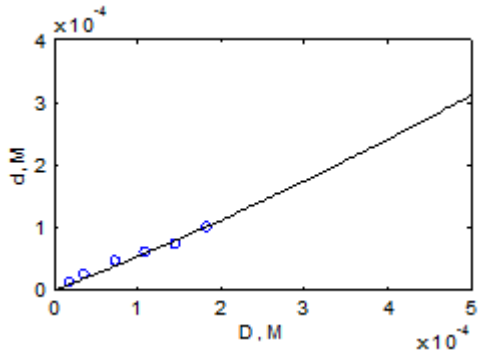
n1-integer restirction for Non-Linear SIMPLEX

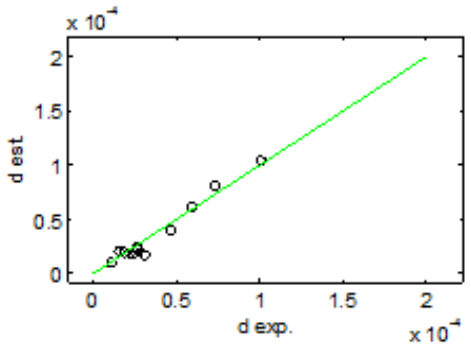
n1 , K1 , logK1 ; n1K1 :

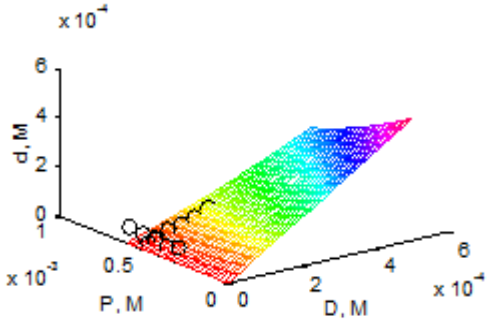
R2 = 0.969816

log (n1*K1) = 3.26975



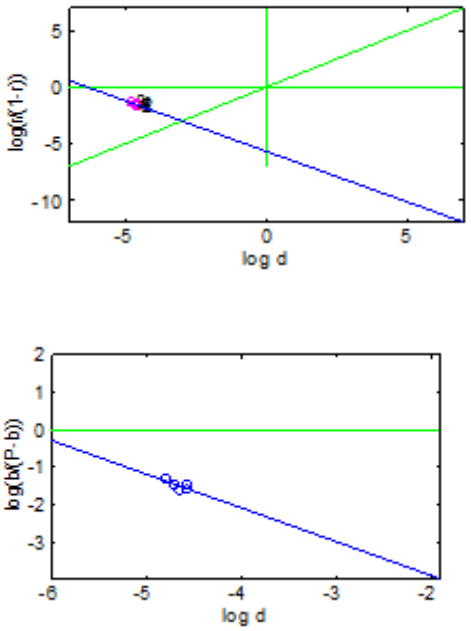






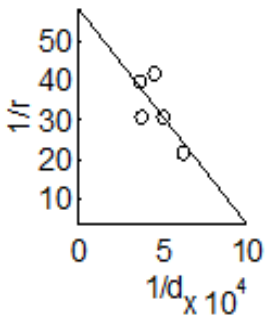
D-Constant (Points 9 to 12)

Verification of n

n1 ± IC/2 : -0.894251 ± 1.51346 logK1 ± IC/2 : -5.65364 ± 7.05065 R2 = 0.94495 ; F = 68.6608 ; p = 0.00115798 se = 0.125508 Very Poor Estimation Obtained	

Summary of Linear Plots

Plot	n_1	Log K_1	Log (n_1K_1)
Klotz	0.01723	5.026	3.262
Scatter	0.01416	4.934	3.085
y-rec	0.01877	5.081	3.355

Klotz	Scatter	y-rec
	Very poor plots obtained	Very poor plots obtained

Estimates from Eqn . 7.2a, 7.2b and 7.3

ID	logK1-PB	log(K1a3)	log(K1n1
7	3.446	3.086	3.100
8	3.348	2.969	2.980
9	3.269	3.464	3.485
10	3.203	3.226	3.240
11	3.145	3.031	3.041
12	NaN	NaN	NaN
Mean	3.269	3.086	3.100

Based on the findings of z and Grubbs test; point 12 was eliminated completely from the rest of data analysis. Therefore 5 point were left for D-constant part

Non-Linear SIMPLEX

Using Eqn 7.3

P-Constant

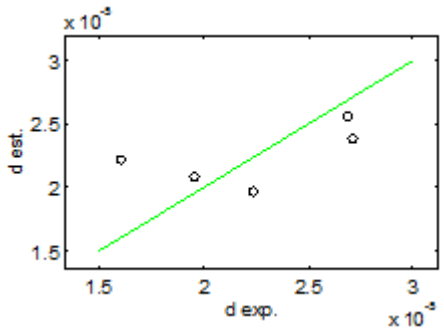
n1-integer restirction for Non-Linear SIMPLEX

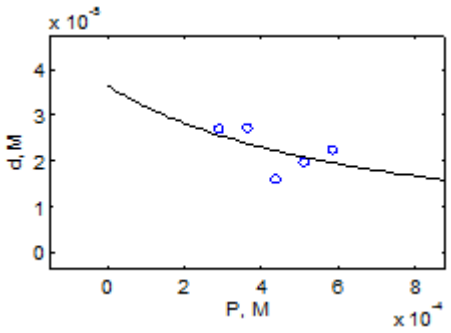
n1 , K1 , logK1 ; n1K1 :

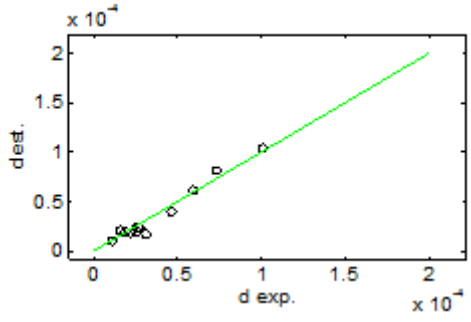
R2 = 0.334601

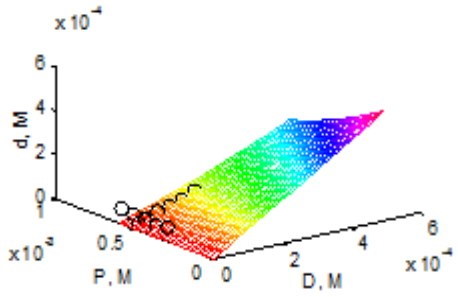
log (n1*K1) = 3.18127

PB% = 56.1644

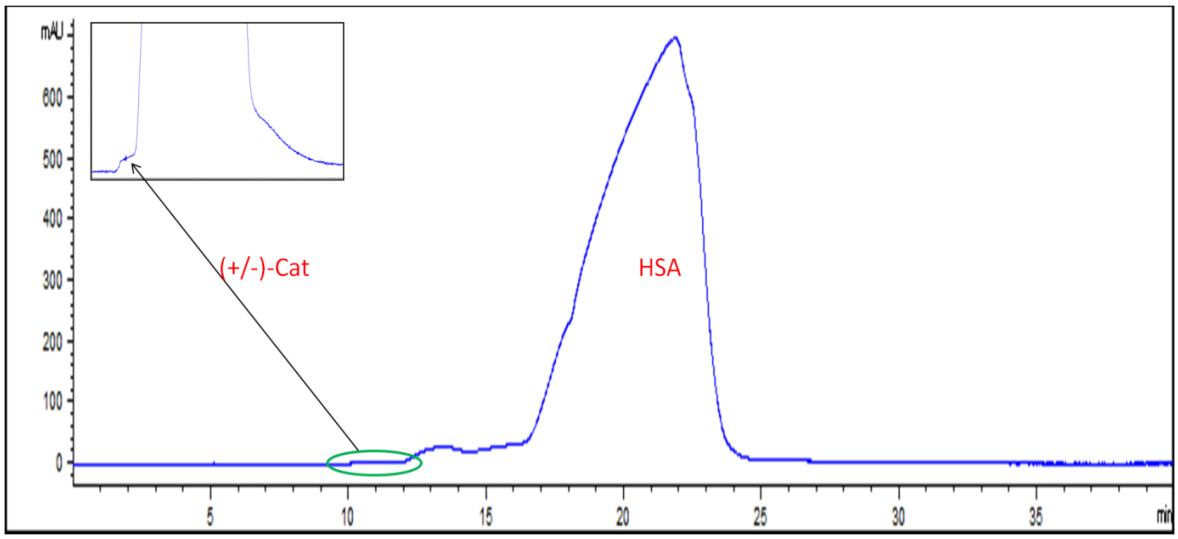




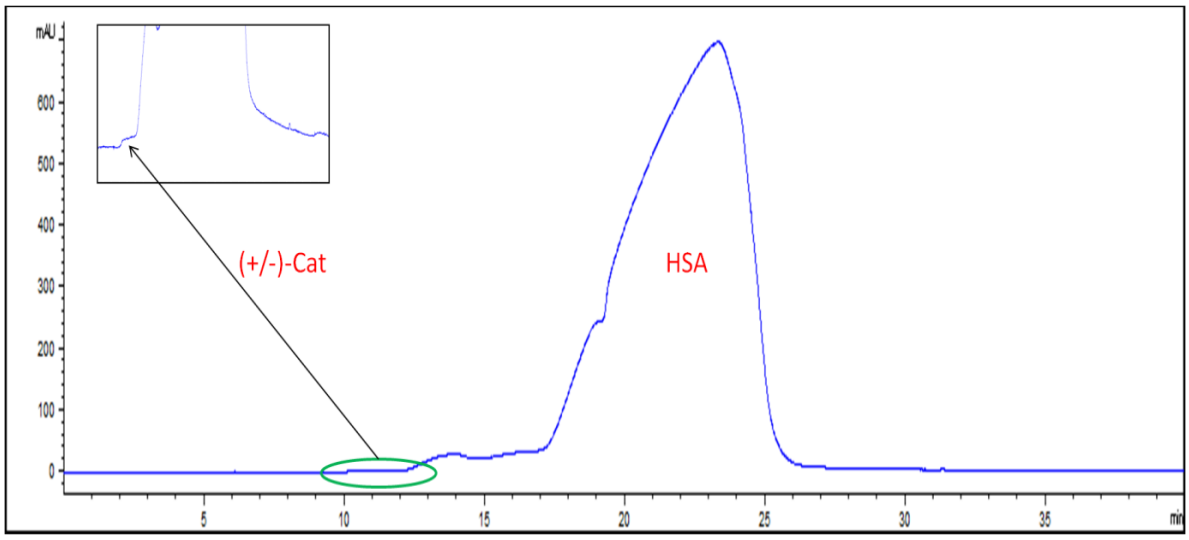




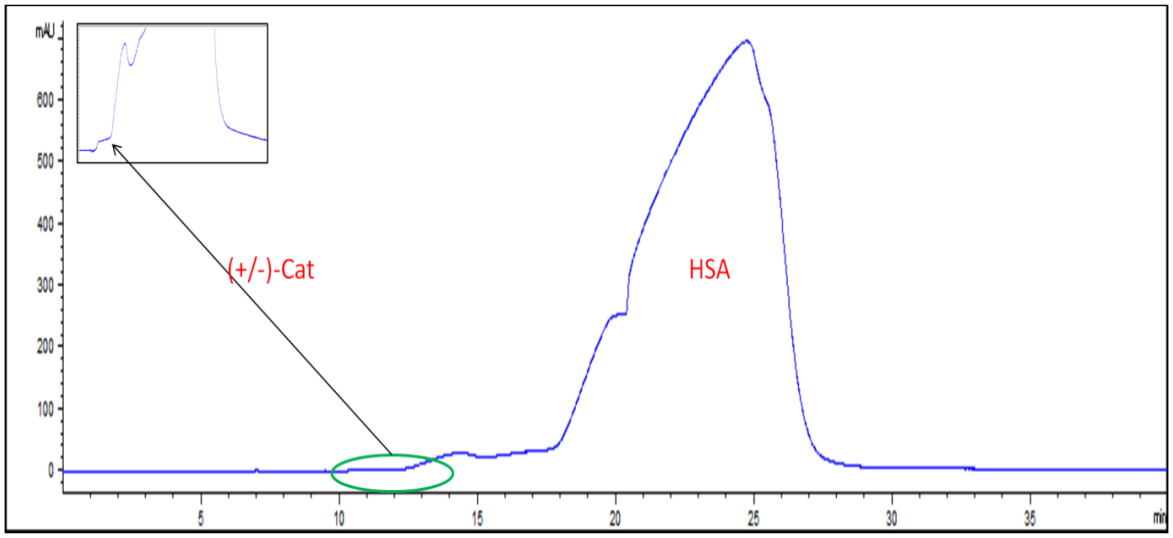
(±)-Catechin 20 µM_HSA 525 µM



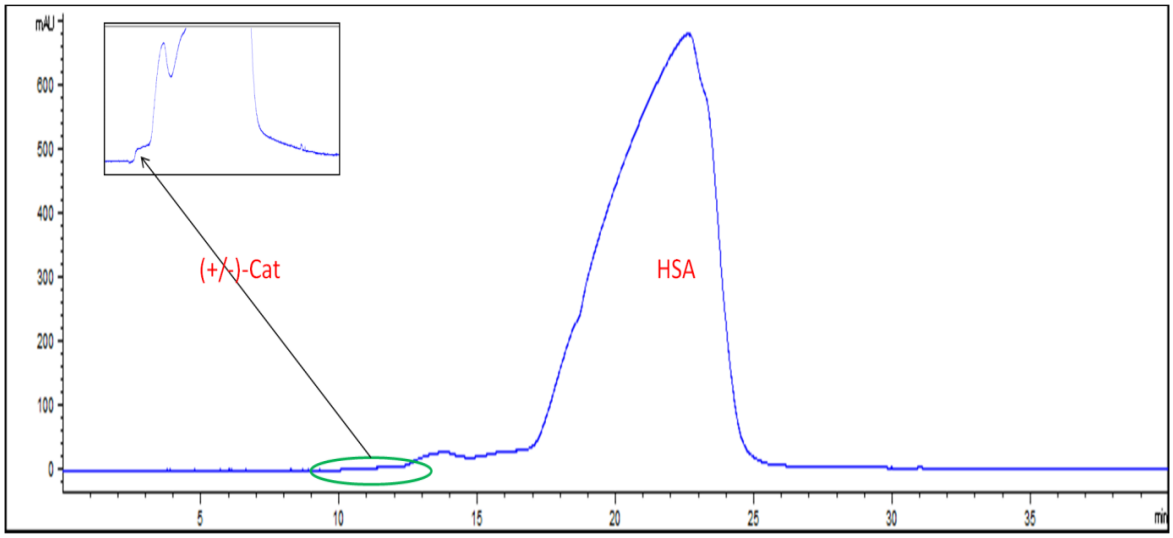
(±)-Catechin 40 µM_HSA 525 µM



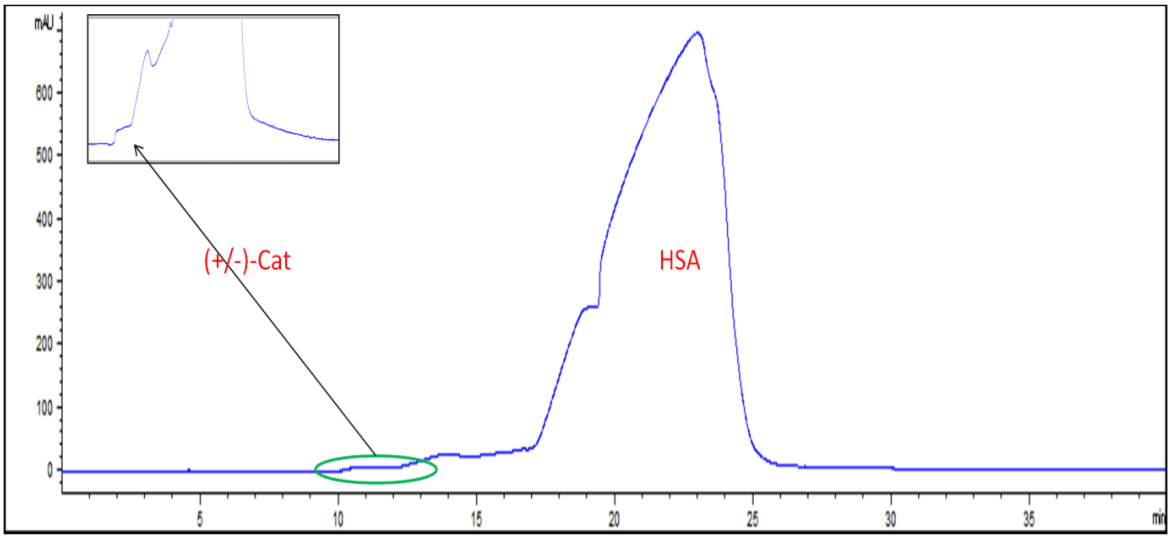
(±)-Catechin 80 µM_HSA 525 µM



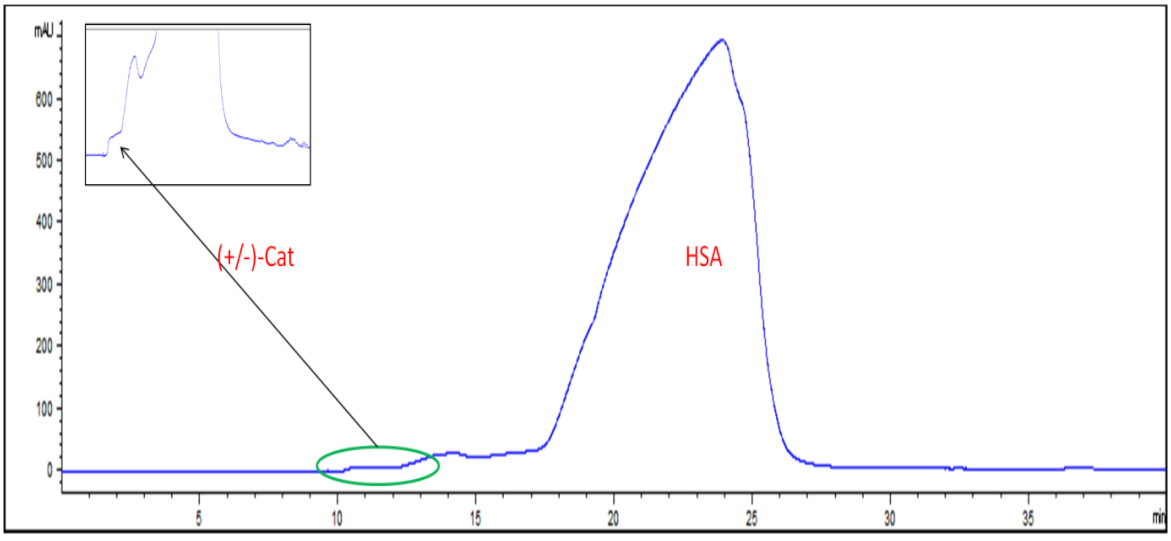
(±)-Catechin 120 µM_HSA 525 µM



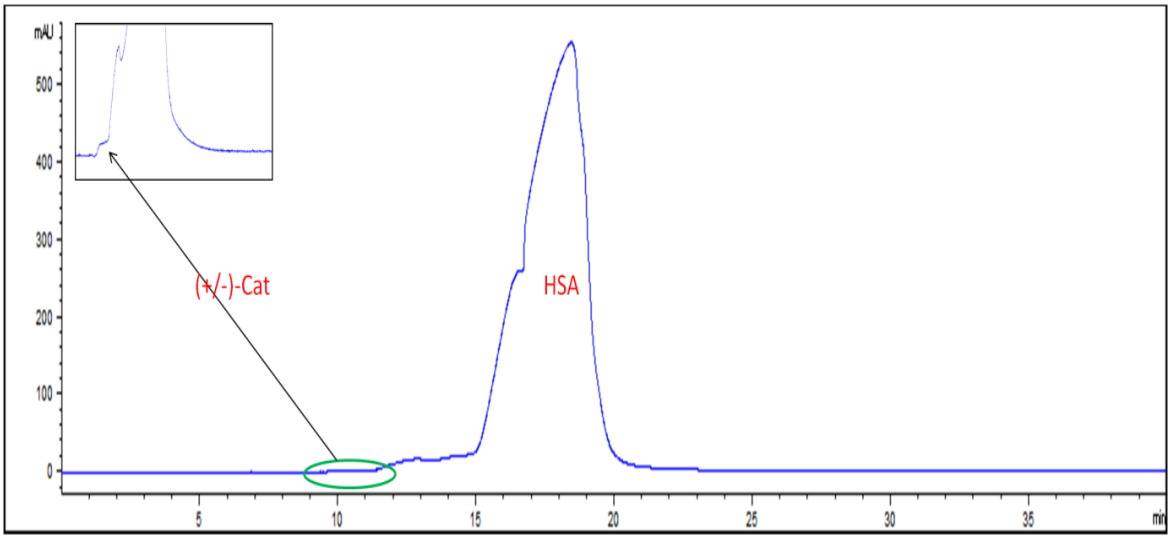
(±)-Catechin 160 µM_HSA 525 µM



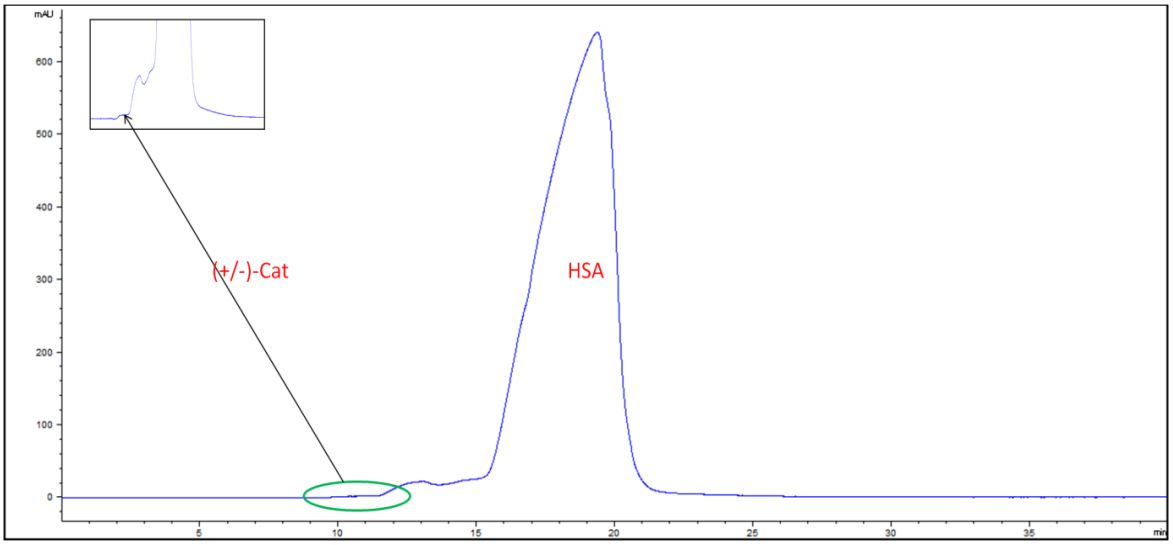
(±)-Catechin 200 µM_HSA 525 µM



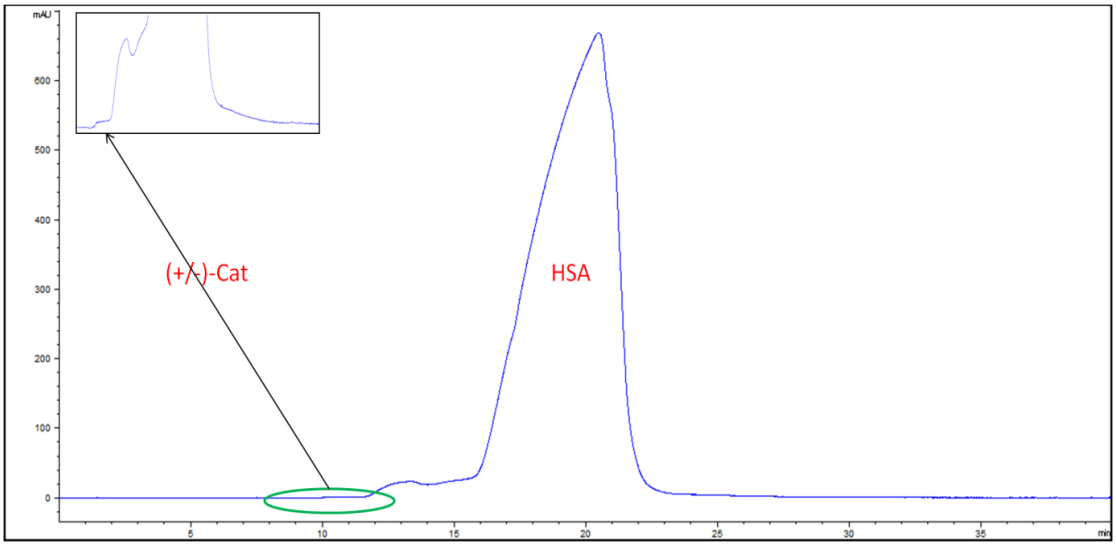
(±)-Catechin 40 µM_HSA 300 µM



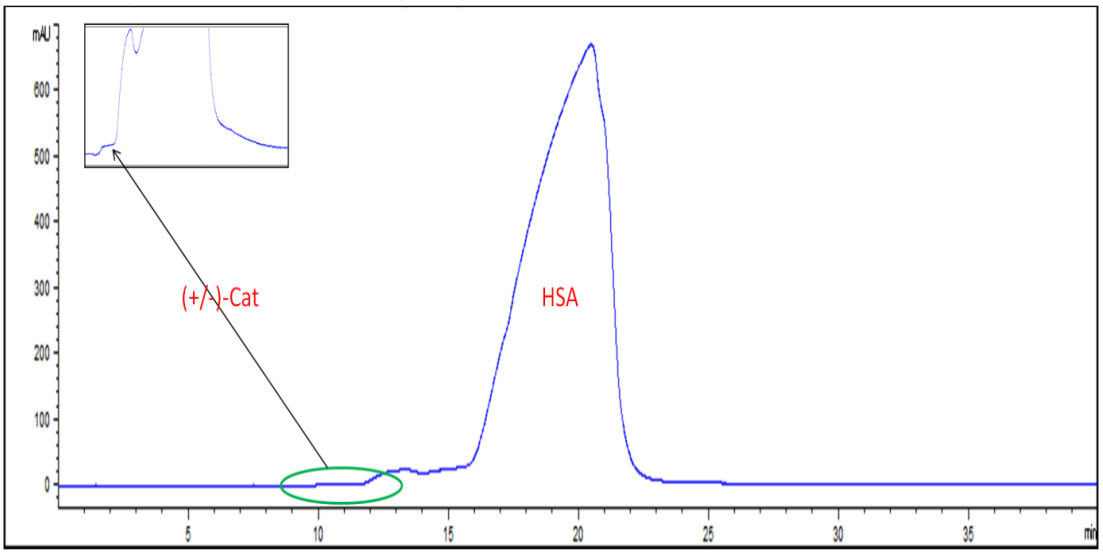
(±)-Catechin 40 µM_HSA 375 µM



(±)-Catechin 40 µM_HSA 450 µM

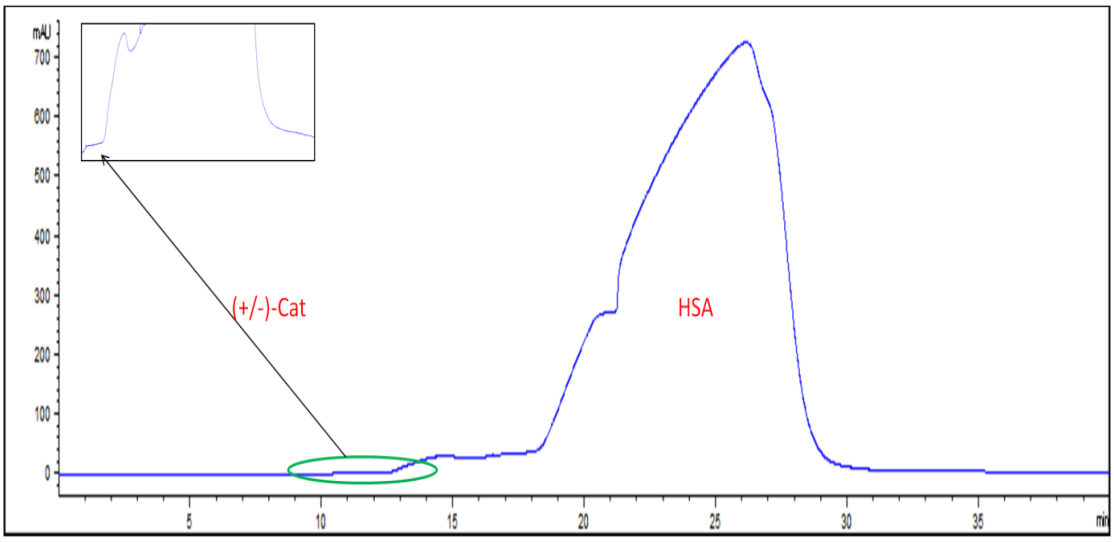


(±)-Catechin 40 µM_HSA 525 µM



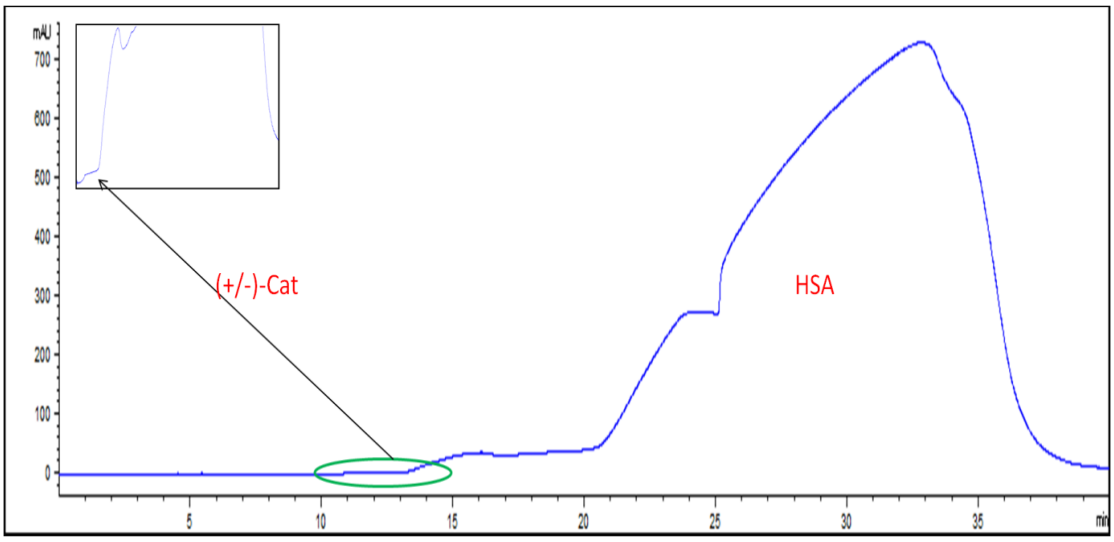
(±)-

Catechin 40 µM_HSA 600 µM



(±)-

Catechin 40 µM_HSA 675 µM



Study of ligand Protein Interaction (Discussed in Chapter 7)

P-constant Experimental Design: The experiment was performed using a two buffers 67 mM Phosphate buffer + 9.0 g/L NaCl for incubation and 30mM phosphate buffer for EKC separation

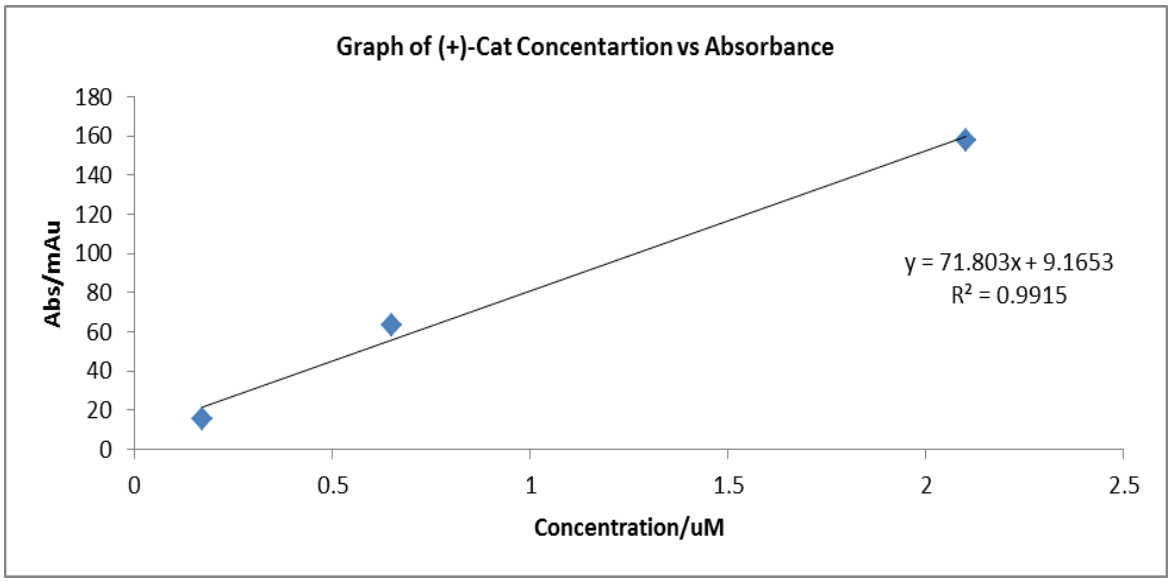
Data for (+)-Catechin

	(+)-Catechin, Design (µM)					Experimental (µM)						Esimated Control (<1)
Injection Order	***ID***	***type***	***D***	**Dexp**	***P***	***D+***	***Pexp***	**Time/min**	**Area**	**Corr Area**	***d+***	***D / P***
9	1		100	100.6		37.35		ND	ND	ND	ND	0.19
2	2		200	201.2		74.70		11.11	4.500	0.405	36.0	0.38
10	3		250	251.5		93.40		11.06	4.840	0.438	38.7	0.47
1	4		300	301.8		112.1		11.55	7.280	0.630	55.2	0.57
6	5	mixture	350	352.1	530	130.7	530.6	11.05	8.900	0.805	70.1	0.66
7	6		400	402.4		149.4		10.63	10.98	1.033	89.6	0.76
11	7		450	452.7		168.1		10.56	11.30	1.070	92.7	0.85
3	8		500	503.0		186.8		12.72	13.40	1.054	91.3	0.95
4	9	cal-1	25	25.15		9.340		11.03	1.300	0.118	9.34	
8	10	cal-2	100	100.6	0.0	37.35	0.0	10.73	4.120	0.384	37.3	
5	11	cal-3	250	251.5		93.38		10.64	11.60	1.090	93.4	

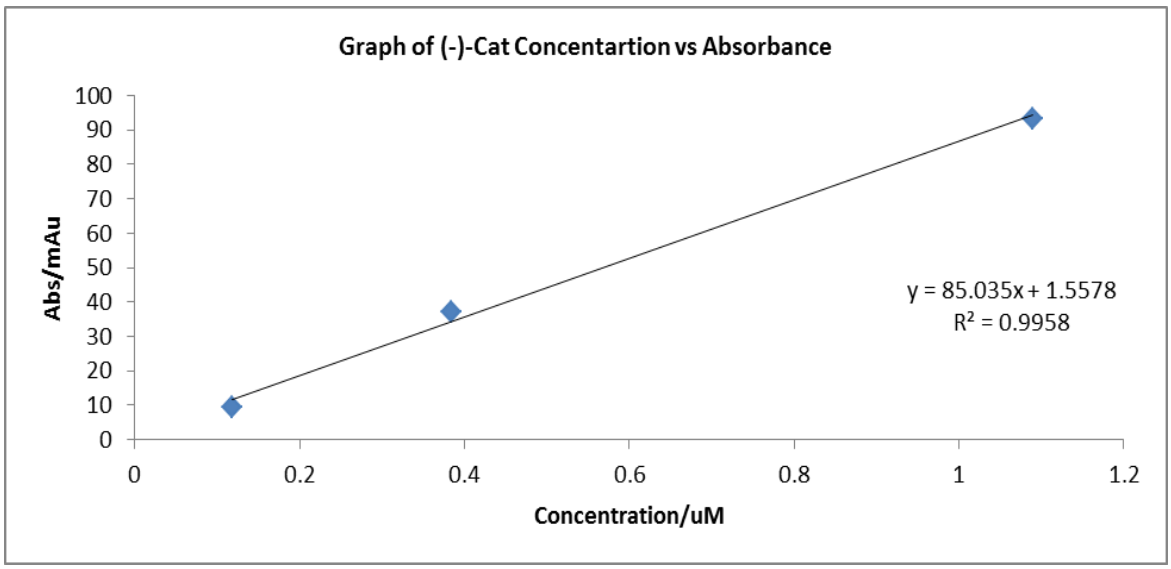
Data for (-)-Catechin

	(-)-Catechin Design (µM)					Experim. (µM)						Estimate Control (<1)
Injection Order	ID	type	D	Dexp	P	D(-)	Pexp	tm,Min	Area	Corr_Area	d-	D / P
9	1		100	100.6		63.24		ND	ND	ND	ND	0.130
2	2		**200**	201.2		126.49		11.40	6.870	0.60	52.10	0.238
10	3		250	251.5		158.11		11.31	8.630	0.76	63.70	0.298
1	4		300	301.8		189.73		11.87	11.27	0.95	77.20	0.358
6	5	mixture	350	352.1	530.0	221.36	530.60	11.31	14.04	1.24	98.40	0.417
7	6		400	402.4		252.98		10.86	17.14	1.58	123.0	0.477
11	7		450	452.7		284.60		10.80	19.70	1.82	141.0	0.536
3	8		500	503.0		316.22		13.06	22.40	1.72	133.0	0.596
4	9	cal-1	25	25.15		15.81		11.26	1.940	0.17	15.81	
8	10	cal-2	100	100.6	0.00	63.24	0.00	10.98	7.140	0.65	63.24	
5	11	cal-3	250	251.5		158.11		10.87	22.83	2.10	158.11	

Calibration curve for the (+)-catechin experiment



Calibration curve for the (-)-catechin experiment



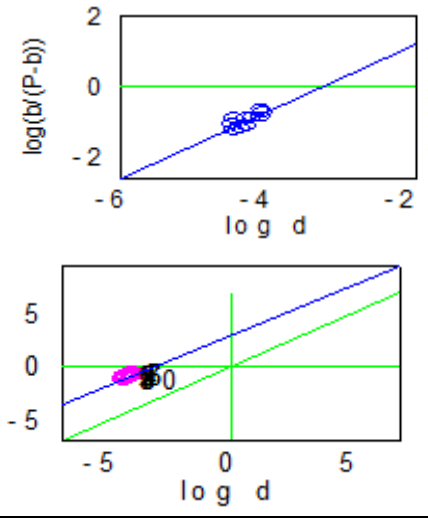
Outliers on K_1

z-score analysis on x	**Grubbs' Test Results**
z-critic : 2.5 : z-score based on median±MADe 1867.9 ± 575.5 (Eqn. 7.2a & 7.2b) index of eliminable-points: none	NO OUTLIER

Outliers on log K1

STATISTICS on Log K1 : (Eqn. 7.2a & 7.2b) z-score based on median±MADe 3.27 ± 0.146 index of eliminable-points: z-critic : 2.5 :	**Grubbs' Test Results** Value number 2 IS NOT AN OUTLIER (Grubbs test alpha= 2.5%, 2 tails) **Estimation:** STATISTICS on Log K1 : E1: m ± s (RSD) & md: 3.27 ± 0.124 (3.78176); 3.271

Verification of n

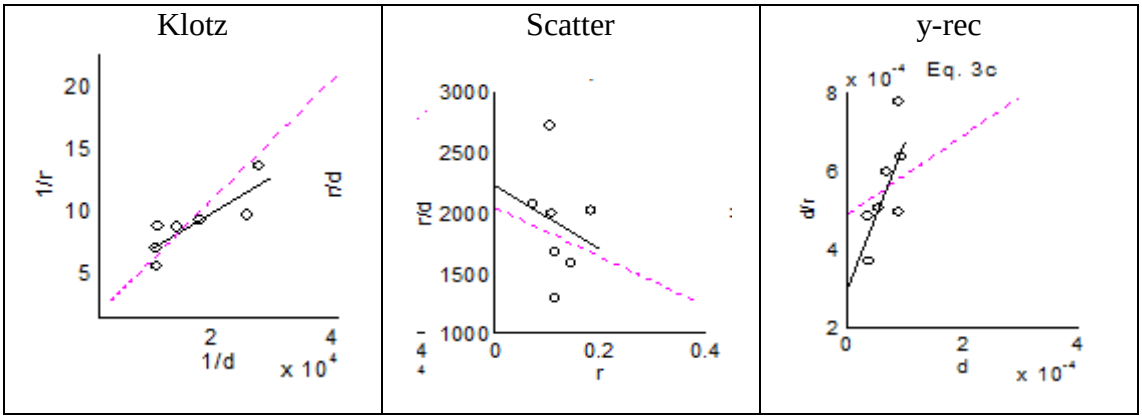
P-Constant n1 ± IC/2 : 0.59173 ± 1.795 logK1 ± IC/2 : 1.36648 ± 7.798 R^2 = 0.946074 ; F = 17.54 ; p = 0.149197 se = 0.01678	

Estimates from Eqn 7.2a and 7.2b

ID	logK1-PB	log(K1a3)	log(K1n1
1	3.188	3.317	3.350
2	3.188	3.435	3.483
3	3.188	3.299	3.349
4	3.188	3.223	3.277
5	3.188	3.112	3.166
6	3.188	3.198	3.266
7	3.188	3.305	3.393
8	3.188	3.305 -	3.393
Mean	3.188	3.299	3.349

Summary of Linear Plots

Plot	n_1	Log K_1	Log (n_1K_1)
Klotz	0.258	4.127	3.539
Scatter	0.848	3.419	3.347
y-rec	0.263	4.110	3.530



Non-Linear SIMPLEX

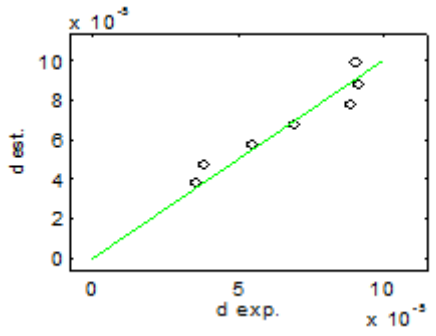
Using Eqn. 7.3

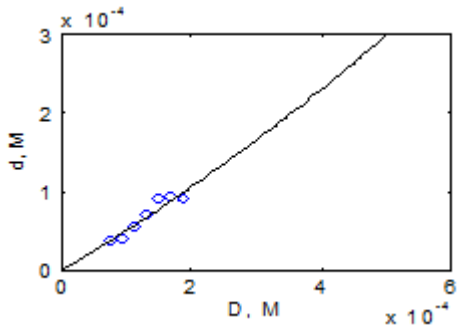
P-Constant

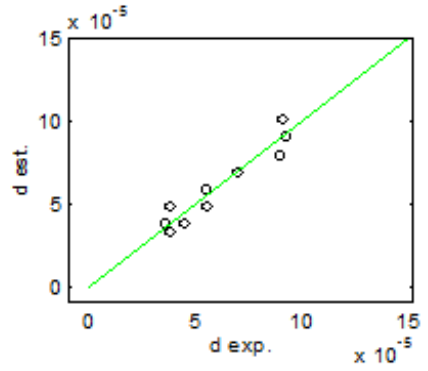
n1-integer restirction for Non-Linear SIMPLEX

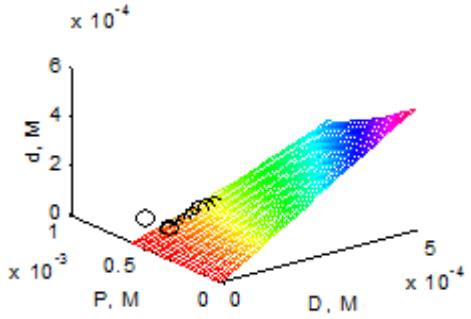
log (n1*K1) = 3.311

R^2 = 0.916









Non-Linear SIMPLEX

Using Eqn 7.3

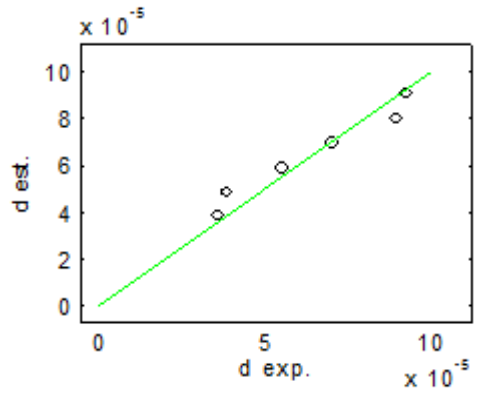
P-Constant: Eliminating 7

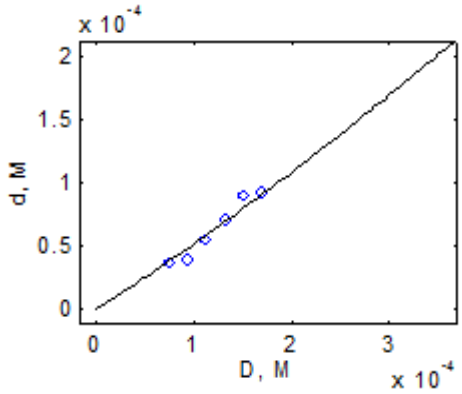
n1-integer restirction forNon-Linear SIMPLEX

log (n1*K1) = 3.283

R^2 = 0.927

PB%: 52.39





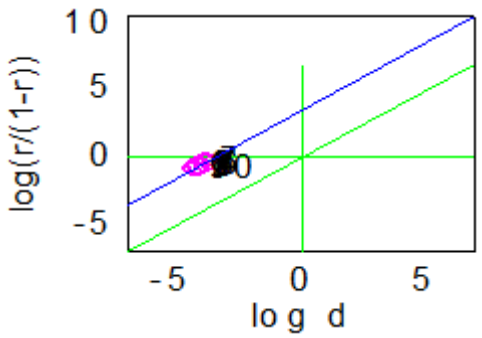
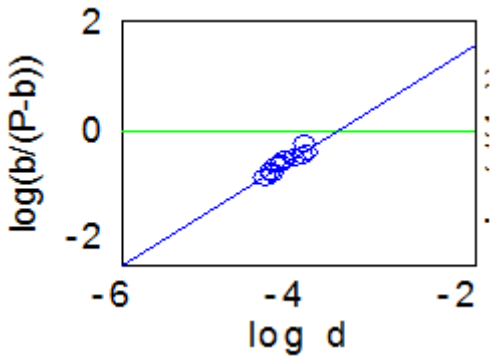
Outliers on K1

z-score analysis on x	**Grubbs' Test Results**
z-critic : 2.5 : z-score based on median±MADe 3087.9 ± 440.6 (Eqn. 7.2a & 7.2b) index of eliminable-points: No Out lier	Value number 7 IS NOT AN OUTLIER (Grubbs test alpha= 2.5%, 2 tails) **Estimation:** STATISTICS on Log K1 : E1: m ± s (RSD) & md: 3.491 ± 0.05681 (1.627); 3.490

Outliers on log K_1

STATISTICS on Log K1 : (Eqn. 7.2a & 7.2b) z-critic : 2.5 : z-score based on median±MADe 3.490 ± 0.06213	GRUBBS' TEST RESULTS Value number 7 IS NOT AN OUTLIER (Grubbs test alpha= 2.5%, 2 tails)

Verification of n

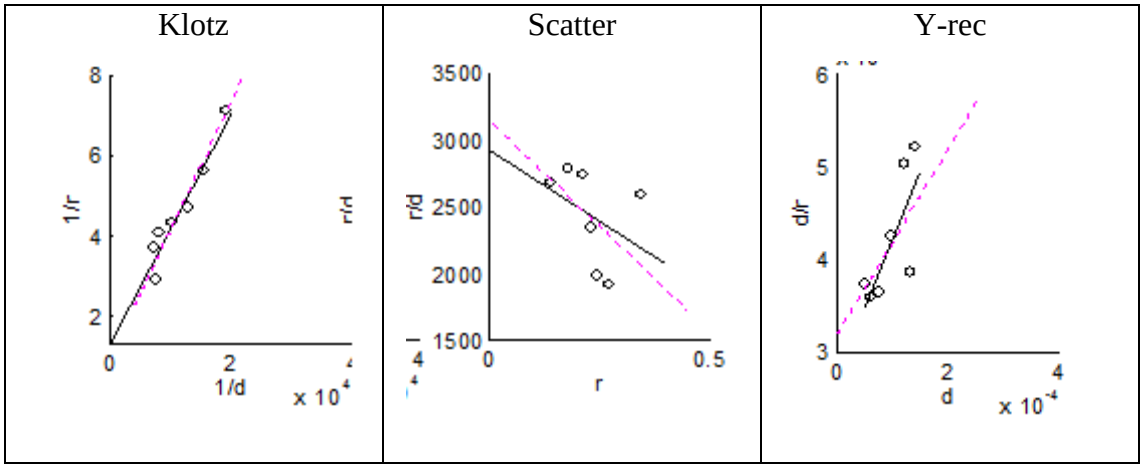
Pconstant n1 ± IC/2 : 0.746 ± 0.219 logK1 ± IC/2 : 2.454 ± 0.894	
	

Estimates from Eqn. 7.2a & 7.2b

ID	logK1-PB	log(K1a3)	log(K1n1
1	Eliminated		
2	3.188	3.428	3.494
3	3.188	3.445	3.530
4	3.188	3.438	3.542
5	3.188	3.374	3.485
6	3.188	3.298	3.420
7	3.188	3.282	3.419
8	3.188	3.413	3.597
Mean	3.188	3.413	3.494

Summary of Linear Plots

Plot	n_1	Log K_1	Log (n_1K_1)
Klotz	0.769	3.651	3.537
Scatter	1.376	3.328	3.467
y-rec	0.693	3.717	3.557



Eliminating point number 7

Estimates from Eqn. 7.2a & 7.2b

ID	logK1-PB	log(K1a3)	log(K1n1
1	Eliminated		
2	3.188	3.428	3.494
3	3.188	3.445	3.530
4	3.188	3.438	3.542
5	3.188	3.371	3.485
6	3.188	3.298	3.420
7	Eliminated		
8	3.188	3.282	3.419
Mean	3.188	3.400	3.490

Summary of Linear Plots

Plot	n_1	Log K_1	Log (n_1K_1)
Klotz	0.588	3.808	3.577
Scatter	0.567	3.833	3.586
y-rec	0.501	3.923	3.622

Non-Linear SIMPLEX

Using Eqn. 7.3

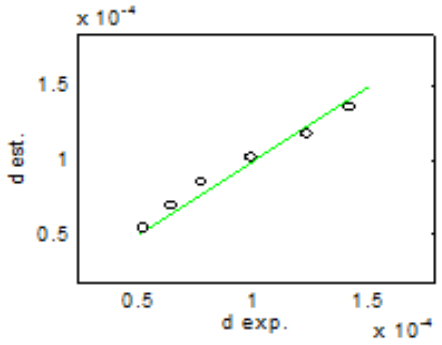
P-Constant: Eliminating 7

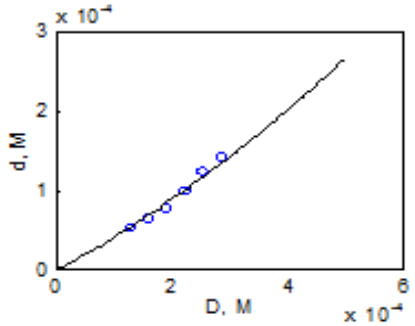
n1-integer restirction forNon-Linear SIMPLEX

log (n1*K1) = 3.497

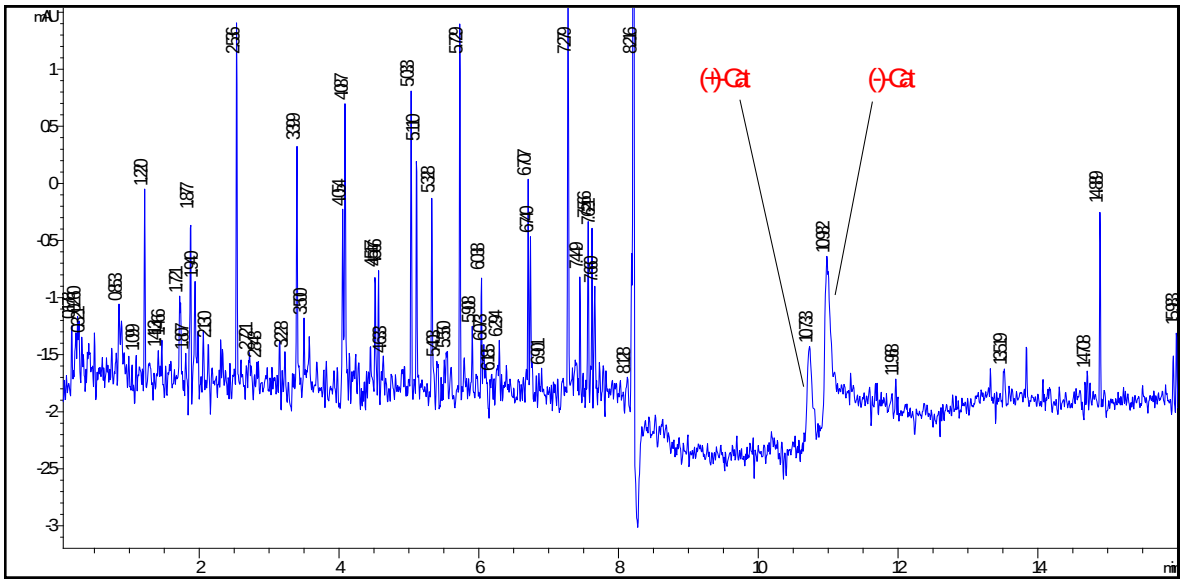
R2 = 0.939

PB%: 58.72

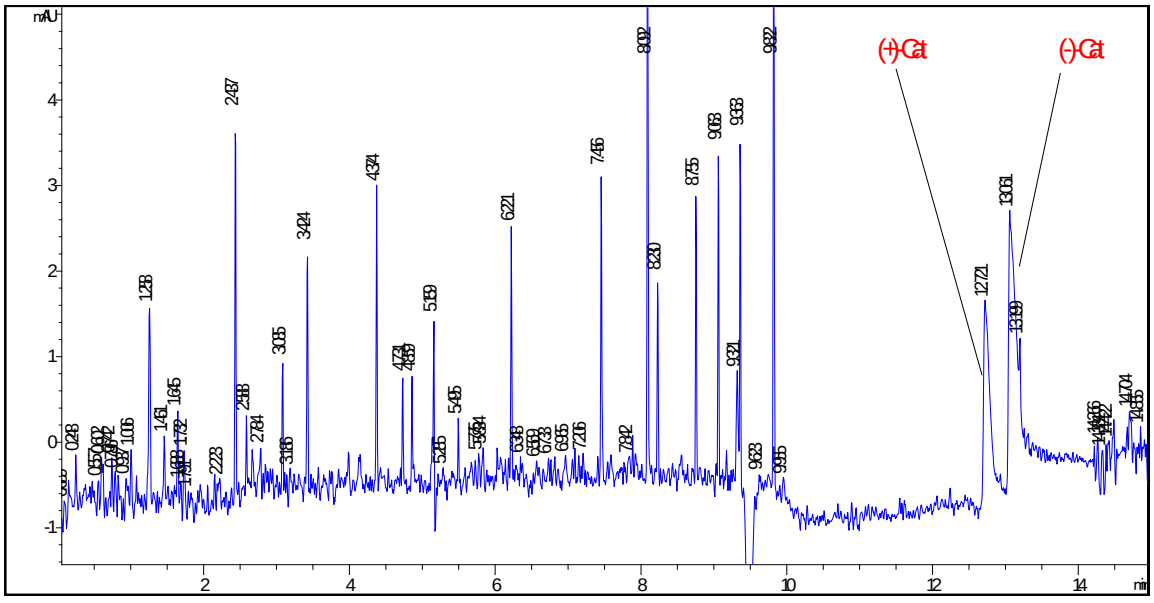




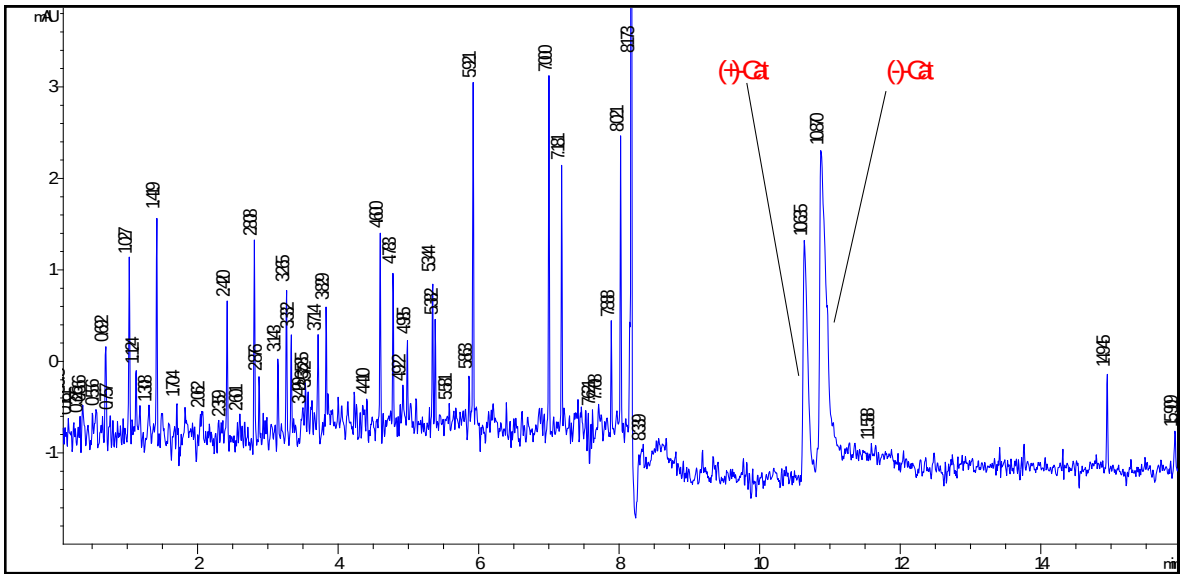
(±)-Catechin 50 µM



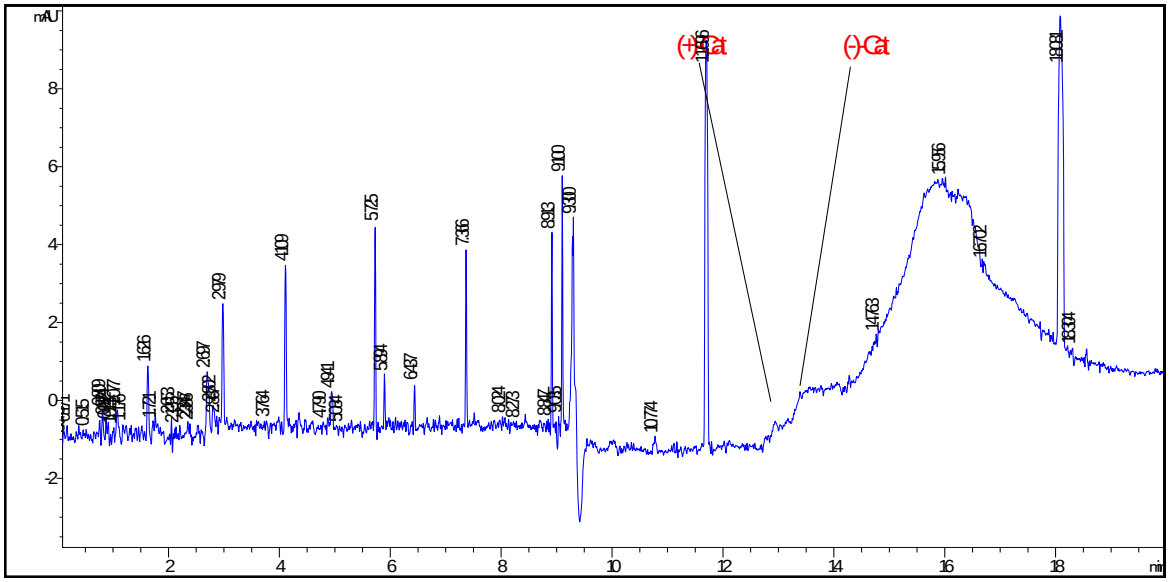
(±)-Catechin 100 µM



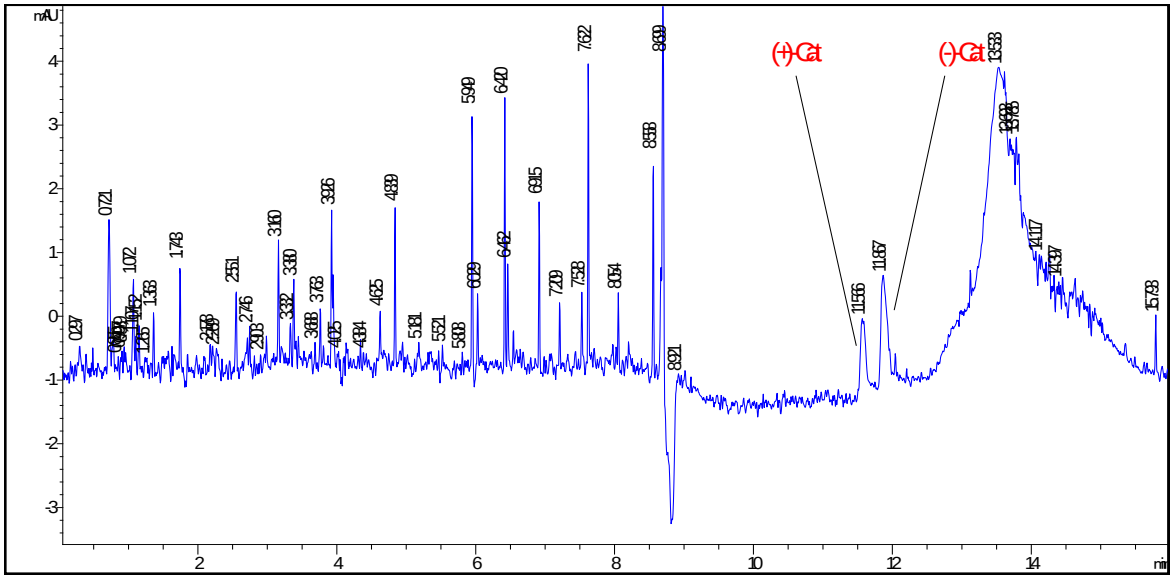
(±)-Catechin 250 µM



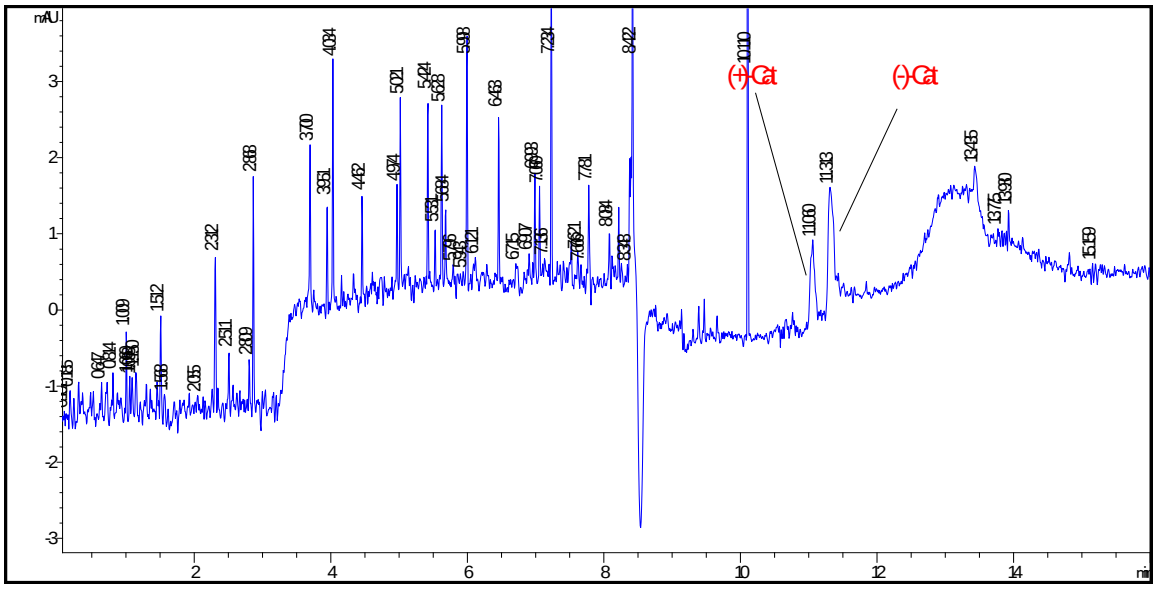
(±)-Catechin 100 µM _HSA 530 µM



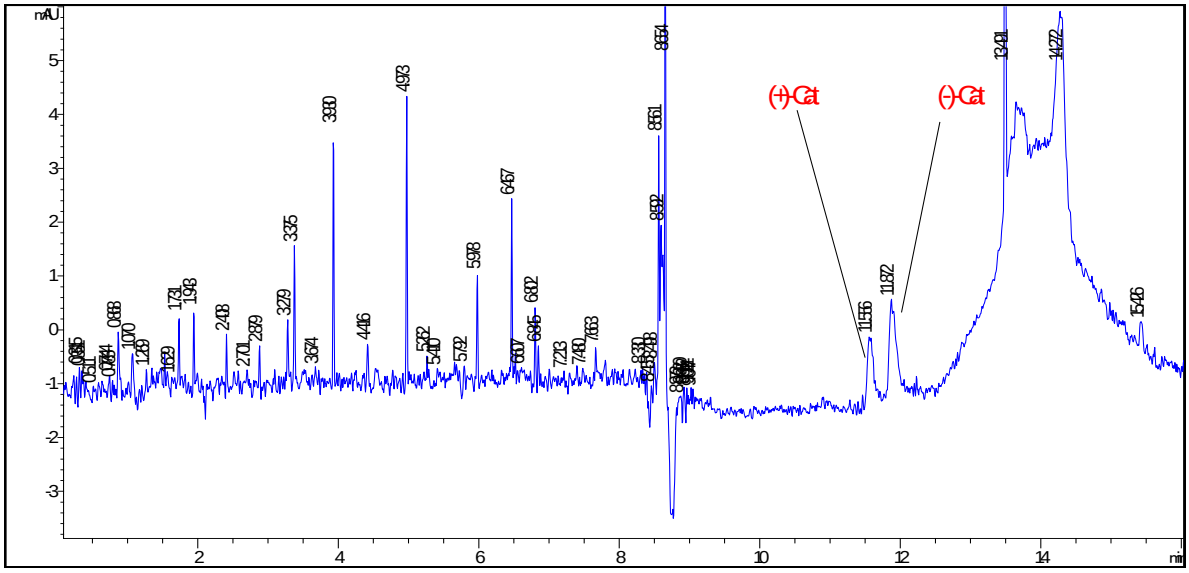
(±)-Catechin 200 µM _HSA 300 µM



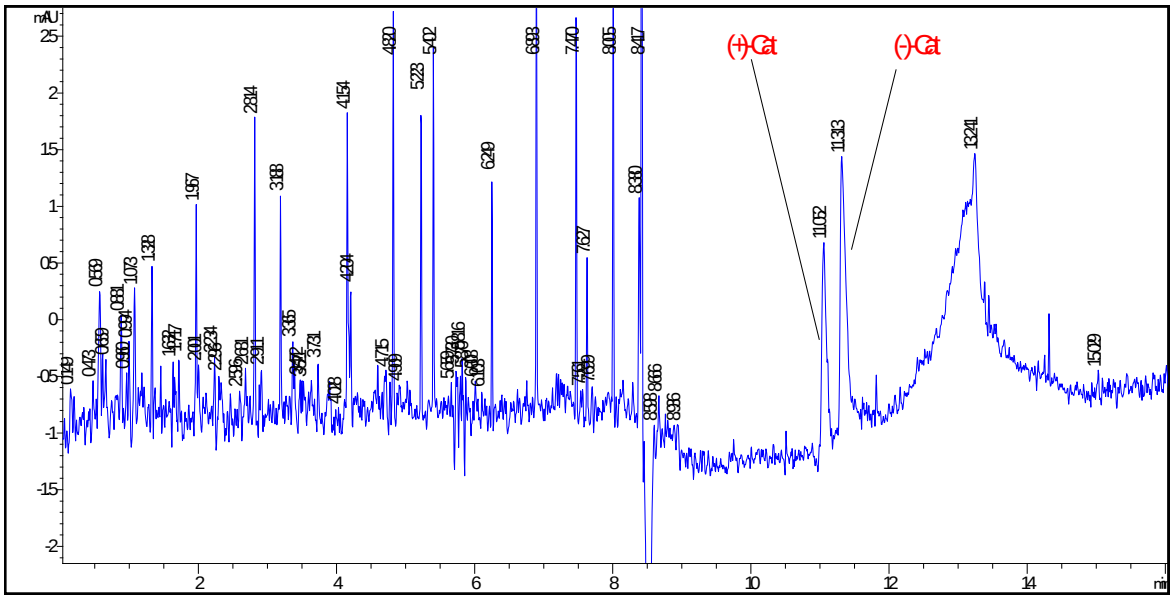
(±)-Catechin 250 µM _HSA 530 µM



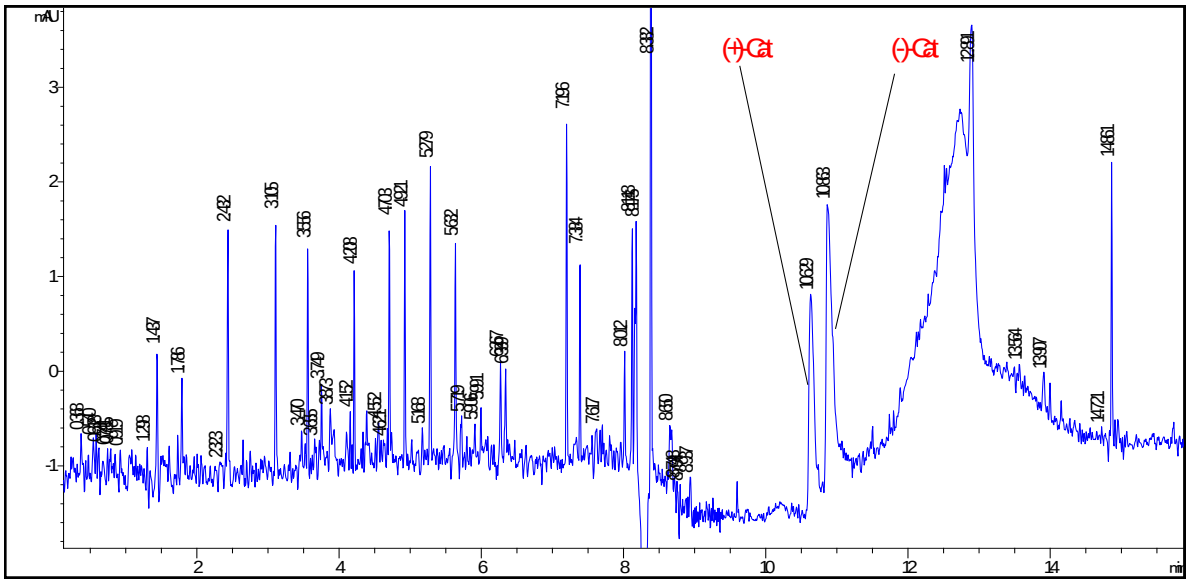
(±)-Catechin 300 µM _HSA 530 µM



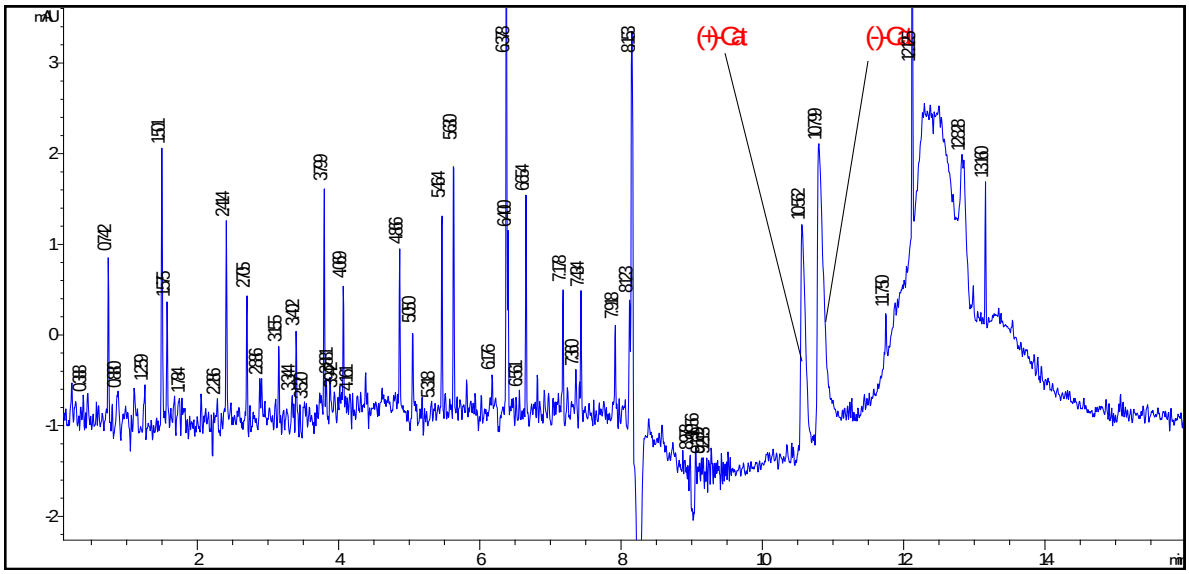
(±)-Catechin 350 µM_HSA 530 µM



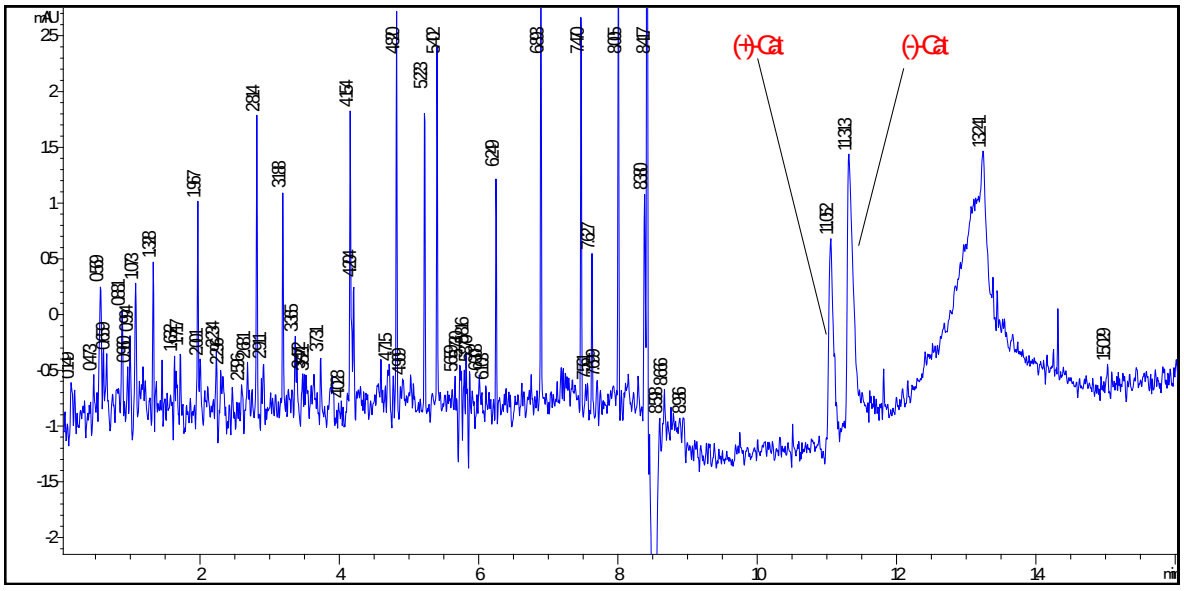
(±)-Catechin 400 µM _HSA 530 µM



(±)-Catechin 450 µM _HSA 530 µM



(±)-Catechin 500 µM_HSA 530 µM



APPENDIX 5

MATLAB SCRIP

*In this script (+)-catechin is used as an example

```
% PapFCat6.m (frontal analysis_ + catechin_calibration 2)
% ...adapted from PapMolec.m (EDITED, SAVED AS (Changing 'Molec' by ABREV.
of your molecule, e.g. Papcat.m )
% ------------------------------------------------------------------------
--
% PROTOCOL (to see it put in MATLAB help PapMolec) :
% 1.- Click on PapMolec.m and modify (Include your data after line 40 in
the same format that for the example catechin-lines 22-38)
% 2.- Save as : Pap[+ABREVIATURE-Molecule up to 5 letters].m (e.g.
PapZPN.m)
% 3.- Open MATLAB and write the new file (without .m; e.g. Papfcat). For a
normal Study select Normal mode
% 4.- You can EXIT if you do not want to run an option
% 5.- In the ENANTIOMER: Model... menu be sure you select the right option
according you data
% 6.- In the ESTIMATIONS ... menu be sure you select the right option
according step %5
% 7.- You can copy and paste Figures and Numeric outputs into word
% Enjoy this & good papers!... (Salva'10)
% ------------------------------------------------------------------------
--

arranca
format compact

mm=menu('Enantioselective PB Analysis:','[Normal mode]','Advanced... ask
Salva');

% Include here NEW data sets for a new small molecule with the same format
than in example data (before)(in M not uM) !!!

   %(+)CATECHIN-FRONTAL ANALYSIS : [Code: E1]   no existe E2=> add NaN
column
    % 1-ID   2-D(Etot)       3-P         4-dE1       5dE2
   X=[
         1   1.83E-05    5.13E-04    1.16E-05
         2   3.65E-05    5.13E-04    2.58E-05
         3   7.30E-05    5.13E-04    4.67E-05
         4   1.10E-04    5.13E-04    6.00E-05
         5   1.46E-04    5.13E-04    7.33E-05
         6   1.83E-04    5.13E-04    1.01E-04
         7   3.65E-05    2.93E-04    2.69E-05
         8   3.65E-05    3.67E-04    2.72E-05
         9   3.65E-05    4.40E-04    1.60E-05
         10  3.65E-05    5.13E-04    1.96E-05
         11  3.65E-05    5.86E-04    2.24E-05

      ];
   %    12  3.65E-05    6.60E-04    3.14E-05
      %Adding extra-column:
      X=[X NaN*ones(length(X),1)];

     %Optional previous selection in this case:
```

```
      menu(' Anulate (=convert to NaN) Data? : ',' ID= no ; K1-outliers in
the low sensible D-cte part ', ' ID= no; log K1-outliers (Grubbs, z-score)
in the low sensible D-cte part ', ' [No] ');
       % if ans==1%Anulating 12, 9
        %    X(9,[4])=[NaN]
         %   X(12,[4])=[NaN]
        if ans==2 %Anulating 12
           X(12,[4])=[NaN]
        else
        end

 % End of NEW data ----------------------------------------------------
--
    Xor=X; %original data
      %Data-vectors: -----------------------------------------------------
--
            ID=X(:,1);
            D=X(:,2);
            P=X(:,3);
            d1=X(:,4);
            d2=X(:,5);
            d=d1+d2; %sum = racemate
            %--------------------
            b1=D-d1;
            b2=D-d2;
            r1=b1./P;
            r2=b2./P;
       %-------------------------------------------------------------------
--
%2-Exp. design:
++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
    plotnum1(P,D)
    axis([.9*min(P) 1.1*max(P) .9*min(D) 1.1*max(D)])
    xlabel('P, M')
    ylabel('D, M')

   if mm==2 %advanced...

     figure(1),clf,op
      subplot(121)
        plotnum2(D,d1)
        hold on
        plotnum1(D,d2)
        %plotnum3(IDdis1(:,2),IDdis(:,6))
        ylabel('d, M')
        xlabel('D, M')
        %axis([-.1e-4 2.6e-4 -.1e-5 6.1e-5])
        %text(1e-4,5.5e-5,'(a)')

      subplot(122)
        plotnum2(P,d1)
        hold on
        plotnum2(P,d2)
        %plotnum3(IDdis2(:,3),IDdis(:,6))
        ylabel('d, M')
        xlabel('P, M')
        %axis([-.1e-4 6.1e-4 -.1e-5 6.1e-5])
        %text(5e-4,5.5e-5,'(b)')
   end
```

```
%3-log K1 - Outliers? :
++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
        %Eq.4 %K1(if n1=1) to Check for OUTLIERS
            K1n1E1=1./d1.*(r1./(1-r1));
            K1n1E2=1./d2.*(r2./(1-r2));

     %-------------------------------------------------------------new
Sept10

     menu('Eq. 4 - K-INCONSISTENCES / BIG OUTLIERS ?:',' YES',' NO');
     if ans==1

       %LOOKING FOR K-INCONSISTENCES (or BIG OUTLIERS)
            for i=1:length(K1n1E1)
                if K1n1E1(i)<0
                    disp('      ID       K1E1***   K1E2')
                    disp('  ---------  --------  ---------')
                    disp([ID K1n1E1 K1n1E2])
                    menu('WARNING: KE1<0 detected','OK','Acabar?');
                    if ans==2
                        acabar
                    end
                end
            end

            for i=1:length(K1n1E2)
                if K1n1E2(i)<0
                    disp('      ID       K1E1      K1E2***')
                    disp('  ---------  --------  ---------')
                    disp([ID K1n1E1 K1n1E2])
                    menu('WARNING: KE2<0 detected','OK','Acabar?');
                    if ans==2
                        acabar
                    end
                end
            end
          %z-score test over  K1 estimations
               ww4=1;
                 while ww4==1
                                        menu('z-score test:','On K1E1','On
K1E2','EXIT');
                                        if ans==1
                                        figure(1)
                                            [xz,ind,em,es] =
fz1(felinan2(K1n1E1),1,1); %orig. sin NaN  %median y MADE experimental
                                        elseif ans==2
                                        figure(2)
                                            [xz,ind,em,es] =
fz1(felinan2(K1n1E2),1,1); %orig. sin NaN  %median y MADE experimental
                                        else
                                            ww4=0;
                                        end
                 end
                         %Grubbs test over log K1 estimations
                                   ww3=1;
                                   while ww3==1
                                        menu('Grubbs test:','On K1E1','On
K1E2','EXIT');
                                        if ans==1
```

```
                                        grubbs2(felinan2(K1n1E1));
%orig. sin NaN
                                    elseif ans==2
                                        grubbs2(felinan2(K1n1E2));
%orig sin NaN
                                    else
                                        ww3=0;
                                    end
                                end

          %Manual observation
            ww=1;
                   %Provisional:
                   pID=ID;
                   pd1=d1;
                   pd2=d2;
                   pK1n1E1=K1n1E1;
                   pK1n1E2=K1n1E2;

            while ww==1
             menu('Show extra between-E1-E2-INCONSISTENCES?...
','YES','No');
                if ans==1

                   figure, op
                       subplot(232)
                                            plotnum3(pd1,pd2)
                                            title('d2 vs. d1')
                       subplot(233)
                                            plotnum3(pID,pd1./pd2)
                                            title('d1/d2 vs. ID')
                                            hline(1)
                       subplot(234)
                                            plotnum1(pID,pK1n1E1)
                                            hold on
                                            plotnum2(pID,pK1n1E2)
                                            hold off
                                            hline(0)
                                            title('KE1 & KE2 vs. ID')
                       subplot(235)
                                            plotnum3(pK1n1E1,pK1n1E2)
                                            title('KE2 vs. KE1 ')
                       subplot(236)
                                            plotnum3(pID,pK1n1E1./pK1n1E2)
                                            title('KE1/KE2 vs. ID')
                                            hline(1)
              % end

                   %del=input('Delete pID (one or various)= ...or 0 to
exit... :  ');
                   del=input('Nan-pID (one by one)= ...or 0 to exit... :
');
                   if del==0
                       ww=0;
                       acabar
                   else
                        %pID=delsamps(pID,del);
                        %pd1=delsamps(pd1,del);
                        %pd2=delsamps(pd2,del);
                        %pK1n1E1=delsamps(pK1n1E1,del);
```

```
                        %pK1n1E2=delsamps(pK1n1E2,del);

                        pID(del)=NaN;
                        pd1(del)=NaN;
                        pd2(del)=NaN;
                        pK1n1E1(del)=NaN;
                        pK1n1E2(del)=NaN;
                    end
                else
            ww=0;
                    end

                end
    end
    %-------------------------------------------------------------------new
             % Verify n1=1
            ww=1;
            while ww==1
               menu('Eq. 4-log K :',' Log K & Outliers','Verify n1=1 for E1
(*Advanced!)','Verify n1=1 for E2 (*Advanced!)','EXIT');
               if ans==1
                                    disp('-------------------- Eq. 4-log K :
to look for OUTLIERS')
                                    % logK:
                                       LoK1n1E1=log10(K1n1E1);
                                       LoK1n1E2=log10(K1n1E2);
                                       IDs=ID;

                                       %copy to work on it:
                                       LoK1n1E1s=LoK1n1E1;
                                       LoK1n1E2s=LoK1n1E2;
                                       md1=nanmdn(LoK1n1E1s);
                                       md2=nanmdn(LoK1n1E2s);

                                     figure(3),clf,op
                                      subplot(222)
                                       plotnum1(ID, LoK1n1E1s)
                                       hold on
                                       plotnum2(ID, LoK1n1E2s)
                                       hline(md1,'b:')
                                       hline(md2,'m:')
                                       title('E1(+), E2(o)')

                                 if mm==2 %advanced...
                                       ww1=1;
                                       while ww1==1

                                           menu('Eliminate data-ID for : ','
E1 (+) ','  E2 (o) ','Regenerate data','EXIT');
                                           if ans==1
                                               clc
                                               vectel1=input('ID to delete:
e.g. [1 6] :');

LoK1n1E1s(vectel1,1)=NaN*ones(length(vectel1),1);
                                                  figure(3),clf,op
                                                  subplot(222)
                                                   plotnum1(ID, LoK1n1E1s)
                                                   hold on
                                                   plotnum2(ID, LoK1n1E2s)
```

```
                                                hold off
                                                hline(md1,'b:')
                                                hline(md2,'m:')

                                            elseif ans==2
                                                clc
                                                vectel2=input('ID to delete:
e.g. [14 16] :');

LoK1n1E2s(vectel2,1)=NaN*ones(length(vectel2),1);
                                                 figure(3),clf,op
                                                subplot(222)
                                                 plotnum1(ID, LoK1n1E1s)
                                                 hold on
                                                 plotnum2(ID, LoK1n1E2s)
                                                 hold off
                                                 hline(md1,'b:')
                                                 hline(md2,'m:')

                                            elseif ans==3
                                                LoK1n1E1s=LoK1n1E1;
                                                LoK1n1E2s=LoK1n1E2;
                                                 figure(3),clf,op
                                                   subplot(222)
                                                    plotnum1(ID, LoK1n1E1s)
                                                    hold on
                                                    plotnum2(ID, LoK1n1E2s)
                                                    hold off
                                                    hline(md1,'b:')
                                                    hline(md2,'m:')

                                            else
                                                    ww1=0;
                                            end
                                        end %ww1
                                    end %of mm.........

                                %Statistics for logK:

                                        m1=nanmean(LoK1n1E1s);
                                        s1=nanstd(LoK1n1E1s);
                                        RSD1=100*s1/m1;
                                        md1=nanmdn(LoK1n1E1s);

                                        m2=nanmean(LoK1n1E2s);
                                        s2=nanstd(LoK1n1E2s);
                                        RSD2=100*s2/m2;
                                        md2=nanmdn(LoK1n1E2s);

                                        disp('STATISTICS on Log K1 :')
                                        disp(sprintf('E1: m ± s (RSD) & md: %g
± %g (%g); %g', m1, s1, RSD1, md1))
                                        disp(sprintf('E2: m ± s (RSD) & md: %g
± %g (%g); %g', m2, s2, RSD2, md2))
                                        %PB estimation
                                        PBest

                                        subplot(221)
                                        plot(ID, LoK1n1E1s,'+b')
                                        hold on
```

```
                                    plot(ID, LoK1n1E2s,'om')

                                    menu('Central value for log
K','Mean','Median');
                                    if ans==1
                                        hline(m1,'b:')
                                        hline(m2,'m:')
                                    else
                                        hline(md1,'b:')
                                        hline(md2,'m:')
                                    end
                                    xlabel('ID')
                                    ylabel('log K')
                                    %axis([-1 19 3.3 5.1])
                        ww2=1;
                         while ww2==1
                            ot=menu('Outlier test:','Grubbs','z-
score','EXIT');
                              if ot==1
                                  %Grubbs test over log K1 estimations
                                    ww3=1;
                                    while ww3==1
                                        menu('Grubbs test:','On log
K1E1','On log K1E2','EXIT');
                                        if ans==1
                                            grubbs2(felinan2(LoK1n1E1));
%orig. sin NaN
                                        elseif ans==2
                                            grubbs2(felinan2(LoK1n1E2));
%orig sin NaN
                                        else
                                            ww3=0;
                                        end
                                    end

                              elseif ot==2
                                   %z-score test over log K1 estimations
                                    ww4=1;
                                    while ww4==1
                                        menu('z-score test:','On log
K1E1','On log K1E2','EXIT');
                                        if ans==1
                                        figure(1)
                                            [xz,ind,em,es] =
fz1(felinan2(LoK1n1E1),1,1); %orig. sin NaN  %median y MADE experimental
                                        elseif ans==2
                                        figure(2)
                                            [xz,ind,em,es] =
fz1(felinan2(LoK1n1E2),1,1); %orig. sin NaN  %median y MADE experimental
                                        else
                                            ww4=0;
                                        end
                                    end
                              else
                                   ww2=0;
                              end
                         end %ww2 ........
                    elseif ans==2 %Verify n1=1 on E1...
                        if mm==2
                        %All data E1
                                    figure(2)
```

```
                                    %Xall1=[D P d1];
                                    Xall1=felinan2([D P d1]);  %sin NaN
                                    [n1, loK1]=fDPHill(Xall1,1); %A slope=1
means n1=1 is OK
                                   %menu('Eliminating ID=outliers 14:16
','YES','NO');
                                    vecton=input('Vector of ID to maintain:
e.g. [1:12 16:17] :');
                                    [n1, loK1]=fDPHill(Xall1(vecton,:),1);
%A slope=1 means n1=1 is OK
                    end
                elseif ans==3 %Verify n1=1 on E2...
                    if mm==2
                    %All data E2
                                     %Xall2=[D P d2];
                                    figure(2)
                                     Xall2=felinan2([D P d2]);  %sin NaN
                                     [n1, loK1]=fDPHill(Xall2,1); %A slope=1
means n1=1 is OK
                                     %menu('Eliminating ID=15 ','YES','NO');
                                     vecton=input('Vector of ID to maintain:
e.g. [1:12 16:17] :');
                                     [n1, loK1]=fDPHill(Xall2(vecton,:),1);
%A slope=1 means n1=1 is OK
                    end
                else
                                ww=0;
                end
            end
%4- Variables, Models & Desings:
++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
 %NOTE initial data X selection to provide XX (in the DPd format)
 %Loop
 w4=1;
 while w4==1
    X=Xor; %Regenerate original data
    %default limits for sesigned data:
        lPctelim=1;
        hPctelim=6;
        lDctelim=7;
        hDctelim=11;

    menu('WARNING: The following default limits for designed data must be
verified/changed: P-cte ID from 1 to 6 & D-cte ID from 7 to
12','Accept','Modify limits');

    if ans==2
        lPctelim=menu('P-cte ID FROM:
','1','2','3','4','5','6','7','8','9','10','11','12','13','14');
        hPctelim=menu('P-cte ID TO:
','1','2','3','4','5','6','7','8','9','10','11','12','13','14');
        lDctelim=menu('D-cte ID FROM:
','1','2','3','4','5','6','7','8','9','10','11','12','13','14');
        hDctelim=menu('D-cte ID TO:
','1','2','3','4','5','6','7','8','9','10','11','12','13','14');
    end

    menu('WARNING: Only some options has sense: In this case...: "NO E2-
options" ','Get it!...');
    md=menu('ENANTIOMER: Model VARIABLE-SET... Eq./Algorithm:','1> E1:
Indep. [D Pcte d1] ...Eq.3d/n1k1De ','2> E2: Indep. [D Pcte
```

```
d2]...Eq.3d/n1k1De ','3> E1: Indep.  [Dcte P d1]...Eq.3d/n1K1Pe ','4> E2:
Indep.  [Dcte P d2]...Eq.3d/n1K1Pe ','5> E1: Indep.  [D P d1]...Eq.3/n1K1e
(Surface)','6> E2: Indep. [D P d2]...Eq.3/n1K1e (Surface) ','7> E1 & E2:
Compet. [D Pcte d1 d2]...Eq.3d1&3d2/CBED ','8> E1 & E2:  Compet.  [D Pcte
d1 d2]...Eq.3d1&3d2/CBEP ','9> E1 & E2: Compet. [D P d1 d2]...Eq
3d(Eq.3d1&3d2)/CBE (Surface) ','10> RACEMATE:   [D P d(sum)]......',' Exit
');

    if md==1
        %D  P=cte   d1; Independent model
    disp('1> E1: Indep. [D Pcte d1] ...Eq.3d/n1k1De*****************')
        XX=X([lPctelim:hPctelim],[2 3 4]);
   elseif md==2
        %D  P=cte   d2; Independent model
    disp('2> E2: Indep. [D Pcte d2]...Eq.3d/n1k1De*****************')
        XX=X([lPctelim:hPctelim],[2 3 5]);

    elseif md==3
        %D=cte  P   d1; Independent model
    disp('3> E1: Indep.  [Dcte P d1]...Eq.3d/n1K1Pe*****************')
            XX=X([lDctelim:hDctelim],[2 3 4]);

    elseif md==4
    %D=cte  P   d2; Independent model
    disp('4> E2: Indep.  [Dcte P d2]...Eq.3d/n1K1PeE2
results*****************')
            XX=X([lDctelim:hDctelim],[2 3 5]);

    elseif md==5
    % All-d1; Independent model
    disp('5> E1: Indep.  [D P d1]...Eq.3/n1K1e
(Surface)*****************')
            XX=X(:,[2 3 4]);

    elseif md==6
    %All-d2; Independent model
    disp('6> E2: Indep. [D P d2]...Eq.3/n1K1e (Surface)*****************')
            XX=X(:,[2 3 5]);

    elseif md==7 %'D Pcte d1 d2'; Competitive model
    disp('7> E1 & E2: Compet. [D Pcte d1
d2]...Eq.3d1&3d2/CBED*****************')
            XX=X([lPctelim:hPctelim],[2 3 4 5]);

    elseif md==8 %'Dcte P d1 d2'; Competitive model
    disp('8> E1 & E2:  Compet.  [Dcte P d1
d2]...Eq.3d1&3d2/CBEP*****************')
            XX=X([lDctelim:hDctelim],[2 3 4 5]);

    elseif md==9 %'D P di '; Competitive model
    disp('9> E1 & E2: Compet. [D P d1 d2]...Eq 3d(Eq.3d1&3d2)/CBE
(Surface)*****************')
            XX=X(:,[2 3 4 5]);

    elseif md==10 %'D-drug P d-(sum) '
    disp('10> RACEMATE:   [D P d(sum)]....*****************')
            XX=[X(:,[2 3]) d];
            XX=[XX(:,1)*2 XX(:,2) XX(:,3)]; %[D]=2*[E]
```

```
    else
        w4=0;
        acabar

    end

 %5- AnnexII-Eq. based estimates:
++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

    % Desinged/Experimental/Calculated data:
        D=XX(:,1);
        P=XX(:,2);
        d=XX(:,3);
        b=D-d;
        r=b./P;
        PB=100*(b./D);
        D2P=D./P; %D/P-ratio

    %Fixing X for data Analysis---------------------------------------
        X=XX;

    w5=1;
    while w5==1

        %m=1 assumed, Parameters: n1 & K1

        est=menu('ESTIMATIONS','1> Initial Estimates (*Advanced!) ','2> h-
Plot to check if [n1=1] (*Advanced!) ','3> Indep.-(P-cte), n1K1De','4>
Indep.-(D-cte), n1K1Pe','5>  Indep.-Surface, n1K1e','6> Compet.- CBED','7>
Compet.- CBEP','8> Competitive CBE','9> EXIT');

        if est==1
          %if mm==2 %advan
            disp(' ---------------------- Initial estimates : ')

            %Annex II. Eq. 6a:
                menu('PB available in bibliography?:','YES','[NO]');
                if ans==1
                        bPB=input('Bibliographic PB (%) = ');
                else
                        bPB=NaN;
                end
                    K1PB=bPB/(100-bPB).*(1./P);

            %Anexo II. Eq. 6:
                K1a3=r./d;
                K1a3xd=K1a3.*d;
                Ea3=100*(K1a3xd-1); %if Ea3>10% use K1a3 as n1K1 (eq. 5 in
Annex II)
                %Also check h-plot...

            %Annex II. Eq. 4:
                K1n1=(1./d).*(r./(1-r));
                %Also check h-plot...

                    RES2=[K1PB, K1a3, Ea3, K1n1];
                    %        D, P, d, b, r, PB%, log(K1PB), log(K1a3),
Ea3%, log(K1n1)
                    clc
```

```
                    disp('-----------------------------------------')
                    disp('      ID      D(M)       P(M)      d(M)      b(M)
r      PB%-at-D   D/P-ratio')
                     RES1=[[1:length(D)]' D, P, d, b, r, PB, D2P];
                    disp(RES1);

                    disp('-----------------------------------------')
                    disp('Eq. 6a 6 & 4-estimates:')
                    disp('-----------------------------------------')

                    disp(' ----------  [Eq. 6a]  [Eq.6   &  (E%)]
[Eq.4]')
                    disp('     ID      logK1-PB  log(K1a3), Ea3%
log(K1n1)... & their Median')
                    loRES2=[[1:length(D)]' log10(K1PB), log10(K1a3), Ea3,
log10(K1n1)];
                    disp(loRES2);
                    disp('-----------------------------------------')
                    mdloRES2=median(loRES2);
                    disp(mdloRES2);
          %end
        elseif est==2 %h-plot AnnexII. footnote h
          %if mm==2 %advan
          disp(' ---------------------- %h-plot AnnexII. footnote h : ')
                [n1, loK1]=fDPHill(felinan2(X),1); %A slope=1 means n1=1 is
OK
                menu('Anulate data-ID... to NaN : ',' YES ',' NO ');
                  if ans==1
                      clc
                      vectel=input('Vector of ID to NaN: e.g. [14 16] :');
                      Xd=X;
                      Xd(vectel,1)=NaN*ones(length(vectel),1);
                      [n1, loK1]=fDPHill(felinan2(Xd),1); %A slope=1 means
n1=1 is OK
                  end
          %end
        elseif est==3 %Indep. n1K1De, dat=[D Pcte d]
            if md>2
                menu('Probable ERROR in Algorithm selection
','OK','Acabar');
                disp('ERROR in Algorithm selection ?')
                if ans==2,acabar,end
            end
                n1K1De
        elseif est==4 %Indep. n1K1Pe, dat=[Dcte P d]
            if md<3 | md>4
                menu('Probable ERROR in Algorithm selection
','OK','Acabar');
                disp('ERROR in Algorithm selection ?')
                if ans==2,acabar,end
            end
                n1K1Pe
        elseif est==5 %Indep. n1K1Pe, dat=[D P d]
            if md<5 | md>6
                menu('Probable ERROR in Algorithm selection
','OK','Acabar');
                disp('ERROR in Algorithm selection ?')
                if ans==2,acabar,end
            end
                n1K1e
        elseif est==6 %Compet. CBED, dat=[D Pcte d]
```

```
            if md<7 | md>7
                menu('Probable ERROR in Algorithm selection
','OK','Acabar');
                disp('ERROR in Algorithm selection ?')
                if ans==2,acabar,end
            end
                CBED
        elseif est==7 %Compet. CBEP, dat=[Dcte P d]
            if md<8 | md>8
                menu('Probable ERROR in Algorithm selection
','OK','Acabar');
                disp('ERROR in Algorithm selection ?')
                if ans==2,acabar,end
            end
                CBEP
        elseif est==8 %Compet. CBE, dat=[D P di]
            if md<9 | md>9
                menu('Probable ERROR in Algorithm selection
','OK','Acabar');
                disp('ERROR in Algorithm selection ?')
                if ans==2,acabar,end
            end
                CBE
        elseif est==9
            w5=0;
        end
    end % Loop w5 ------------------------------------------------
 %else
      %w4=0;
 %end
end % Loop-w4
```

PUBLICATIONS