UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
DEPARTAMENTO DE QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA

“Applications of NMR techniques: Hyphenations (LC-SPE-NMR), affinity (DOSY) and NOE based (STD and Tr-NOESY) to probe the binding interactions of ligands (synthetic and natural) towards protein”

Sheraz Ahmad*

Tese apresentada como parte dos requisitos para obtenção do título de DOUTOR EM CIÊNCIAS, área de concentração: QUÍMICA ORGÂNICA...

Orientador: Prof. Dr. Antonio Gilberto Ferreira

*** bolsista (TWAS/CNPq)**

São Carlos - SP
2014

Ficha catalográfica elaborada pelo DePT da
Biblioteca Comunitária/UFSCar

A286an

Ahmad, Sheraz.
Applications of NMR techniques : Hyphenations (LC-SPE-NMR), affinity (DOSY) and NOE based (STD and Tr-OESY) to probe the binding interactions of ligands (synthetic and natural) towards protein / Sheraz Ahmad. -- São Carlos : UFSCar, 2014.
219 f.

Tese (Doutorado) -- Universidade Federal de São Carlos, 2014.

1. Química orgânica. 2. Ressonância magnética nuclear. 3. Interação ligante-proteína. 4. STD NMR. 5. Tr-NOESY. 6. STD-TOCSY. I. Título.

CDD: 547 (20ª)

UNIVERSIDADE FEDERAL DE SÃO CARLOS
Centro de Ciências Exatas e de Tecnologia
Departamento de Química
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA
Curso de Doutorado

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a defesa de tese de Doutorado do candidato **Sheraz Ahmad,** *realizada em 17 de janeiro de 2014:*

Prof. Dr. Antonio Gilberto Ferreira

Cardoso.
Prof. Dr. Daniel Rodrigues Cardoso

Prof. Dr. Claudio Francisco Tormena

Profa. Dra. Maria Fátima das Graças Fernandes da Silva

Kleber Thiago de Oliveira
Prof. Dr. Kleber Thiago de Oliveira

In the name of Allah, the Compassionate, and the most Merciful.

Recite! (or read!) In the name of your Lord who created (1) Created man from
clots of blood (2) Recite!, your Lord is the most Gracious (3) Who taught by the
pen (4) Taught the man what he knew not (5).

(The first verses of the holy Quran that revealed to Prophet Muhammad (P. B. U. H) from the sky)

This work is dedicated to my (Late) Father: Ghazi Khan, Brother: Mushtaq Ahmad Khan and to my Family

Memorandum

The research presented in this thesis has been carried out at the Department of Chemistry Federal University of São Carlos, São Carlos SP Brazil, between November 2010 and December 2013. This thesis presents the original work of the author unless stated elsewhere and otherwise. Additionally, none of this work has been submitted for any other degree prior to this.

The copyright of this thesis rests with the author and the university. No single quotation from this thesis can be published or presented without the prior written consent to the author, and more importantly, information derived from it should be acknowledged properly.

SUMMARY

Table of Contents

List of abbreviations used in this thesis

Ala	Alanine
Arg	Arginine
Asn	Asparagine
Asp	Aspartate
BSA	Bovine Serum Albumin
CD	Circular Dichroism
COSY	COrrelation SpectroscopY
CP	Cross Polarization
Cys	Cysteine
1D	One-dimensional
2D	Two-dimensional
d1	Relaxation delay
DOSY	Diffusion Ordered SpectroscopY
D_2O	Deuterium oxide
DQ	Double Quantum
DQF	Double Quantum Filtered
DMSO	Dimethylsulfoxide
FID	Free Induction Decay
Glu	Glutamate
Gly	Glycine
His	Histidine
HPLC	High Performance Liquid Chromatography
HSQC	Heteronuclear Single Quantum Coherence
HMBC	Heteronuclear Multiple Bond Correlation
IR	Infrared

lb	Line broadening
k_{on}	Association rate
k_{off}	Dissociation rate
K_D	Dissociation constant
LC	Liquid Chromatography
LC-MS	Liquid Chromatography - Mass Spectrometry
Lys	Lysine
MeOH	Methanol
MHz	Megahertz
min	minute
MW	Molecular Weight
NOE	Nuclear Overhauser Effect
NOESY	Nuclear Overhauser Effect and Exchange SpectroscopY
NMR	Nuclear Magnetic Resonance
NS	Number of scans
PBS	Phosphate Buffer Saline
Phe	Phenylalanine
Pro	Proline
SAR	Structure Activity Relationship
Ser	Serine
STD	Saturation Transfer Difference
S/N	Signal to noise ratio
SPE	Solid Phase Extraction
SW	Spectral Width
T	Tesla
Thr	Threonine
TOCSY	Total Correlation SpectroscopY

TMS	Trimethylsilane
TPPI	Time Proportional Phase Increment
Trp	Tryptophan
Tyr	Tyrosine
T_1	Longitudinal relaxation time
T_2	Transverse relaxation time
TD (F1)	Size of FID in F1 (Increments)
TD (F2)	Size of FID in F2 (Increments)
$T_{1\rho}$	Relaxation time in rotating coordinates
Tr-NOE	Transferred Nuclear Overhauser Effect
τm	Mixing time
UV	Ultra-Violet
v/v	volume by volume
WATERGATE	Water suppression by Gradient Tailored Excitation
WaterLOGSY	Water Ligand Observation with Gradient SpectroscopY

List of tables

List of figures

.

Acknowledgement

All praises for Almighty ***ALLAH,*** *Who is the most Omnificent, Omniscient, Omnipotent, and Omnipresent, heartily thanks for giving me the will and strength to accomplish this job.*

I am showing my humble submission to the heart and soul of the world and the world hereafter; The ***Holy Prophet Hazrat Mohammad*** *(Peace be upon him), whose life is an ideal pattern for us.*

The work presented in this thesis was impossible without the invaluable guidance, keen interest and valuable suggestions of my ingenious and determined research supervisor ***Prof. Dr Antonio Gilberto Ferreira*** *to whom, wish to acknowledge my indebtedness and sincere thanks. I also pay my thanks to* ***Prof. Dr. Tiago Venancio****, faculty here in the laboratory of NMR UFSCar; the work embodied here could never be accomplished without the encouragement and precious attention.*

I wish to express my thanks to who have extended their co-operation in providence of research assistance to me during the entire course of this work, prominent among them are ***Ms. Luciana Vizotto*** *our lab technician,* ***Mr. Paulo R. Lambertuci*** *the electrical technician and my lab members for their cooperation and nice hospitality during the stay of Sao Carlos. Here I would like to specially thank to the Ms. Tatiani M. Bondancia and* ***Prof. Dr. João Batista Fernandes,*** *Ms. Barbara S. Bollete and Her Supervisor* ***Prof. Dr. Maria Fátima das Graças Fernandes da Silva*** *from natural product research group for providing of natural product Samples to me. Moreover, I owe my deepest gratitude to Mr. Zia Uddin and his supervisor* ***Prof. Dr. Edson Rodrigues Filho*** *for providing me their new synthetic compounds.*

I owe my deepest gratitude, to referees; ***Prof. Dr. Cláudio Francisco Tormena, Prof. Dr. Daniel Rodrigues Cardoso*** *for their generosity to accept my request and sparing time for review my thesis. In addition, especially, thanks to* ***Prof. Dr. Maria Fátima das Graças Fernandes da Silva*** *and* ***Prof. Dr. Kleber Thiago de Oliveira*** *for accepting my request as referee* *during my qualification and general seminar also for providence of useful suggestion during this course of my PhD.*

*I offer my utmost gratitude to my funding Agencies (**TWAS/CNPq**) who made me able to come here and supported me from beginning of the research to the end of this thesis.*

The whole work was carried out, completed, and made accessible with the excellent facilities at Laboratory of NMR, department of chemistry Federal University of São Carlos SP Brazil. I owe my deepest gratitude to all faculty members for their cooperation, valuable suggestions, and the deepest thanks for the technical and nontechnical staff of Federal university of São Carlos.

I offer my sincere thanks for sincere company of my Pakistani friends working in USP São Carlos, UFSCar, and Embarapa; who made some memorable trips and parties during vacations to have fun and joy during this stay.

*How can I forget and definitely it is my duty to mention the person who been involved in this research during the whole course and who deserve for much attention and time in my newly marriage life, is my sweet wife **Ms. Nazish Urooj Tanoli.** I offer deepest gratitude for her patience because I have no as such wording of thanks for her sufferings and sacrifices for her making my life happy.*

At last but not least, my acknowledgment will never be completed without thanking my parents especially my mother, brothers and sisters for their prays, moral support and encouragement.

SHERAZ AHMAD KHAN TANOLI

List of Publications

1. **Sheraz Ahmad Khan Tanoli[†*],** Nazish Urooj Tanoli[†], Tatiani Mesquita Bondancia[‡], Saman Usmani[ǂ], Zaheer-Ul-Haq[ǂ], João Batista Fernandes[‡], Sérgio Scherrer Thomasi[†] and Antonio Gilberto Ferreira[†]. Direct NMR mediated ligands screening for a specific target (Human Serum Albumin): Exemplified by using a highly bioactive *Strypnodendron polyphyllum* fully bloomed flower extract. *Analytical chemistry*—submitted
2. **Sheraz A. K. Tanoli[†*],** Nazish U. Tanoli[†], Zia-ud-din[‡], Saman Usmani[ǂ], Zaheer Ul-Haq[ǂ], Edson Rodrigues Filho[‡], Antonio G. Ferreira[†]. Mapping binding sites of anti-leishmanial and anti-trypanosomiasis synthetic library towards the Human Serum Albumin complex as probed by STD NMR, Tr-NOESY Experiments. (In Preparation)
3. **Sheraz Ahmad Khan Tanoli[†*],** Nazish Urooj Tanoli[†], Sérgio Scherrer Thomasi[†], Barbara sayuri bellete[‡], Maria Fatima das Gracas Fernandes de silva[‡] and Antonio Gilberto Ferreira. Structure assignments of Gibaline A and B two acridone alkaloids from *Conchocarpus marginatus* (Rutaceae) by 1D and 2D NMR spectroscopy. (prepared for *Phytochemistry Letters*)

4. **Sheraz Ahmad Khan Tanoli[†*],** Nazish Urooj Tanoli[†], Sérgio Scherrer Thomasi[†], Barbara sayuri bellete[‡], Maria Fatima das Gracas Fernandes de silva[‡] and Antonio Gilberto Ferreira[†]. STD and Tr-NOESY binding study of *Conchocarpus marginatus* (Rutaceae): An alkaloid rich specie. (In Progress)

5. **Tanoli, S. A. K.**; Tanoli, N. U.; Usmani, S.; Ul-Haq, Z.; Ferreira, A. G. The exploration of interaction studies of smaller size, mostly ignored yet

intrinsically inestimable molecules towards BSA; An example of STD and DOSY NMR. *Central European Journal of Chemistry* **2014,** *12* (3), 332-340.

6. **Tanoli, S. A. K.**; Tanoli, N. U.; Bondancia, T. M.; Usmani, S.; Kerssebaum, R.; Ferreira, A. G.; Fernandes, J. B.; Ul-Haq, Z. Crude to leads: a triple-pronged direct NMR approach in coordination with docking simulation. *Analyst* **2013,** *138* (17), 5137-5145.

7. Zia Ud Din, **Sheraz Ahmad Khan Tanoli**, Nazish Urooj Tanoli, Sajjad Hussain, Saima Gul, Sabir Khan, Pir Muhammad khan, Shabnum. Nutritional potential and antioxidant activity of solanum nigram and oenothera speciosa from northern area of Pakistan, *International Journal of Biological & Pharmaceutical Research.* **2012**, 3(8):974-979.

8. Fouzia Sultana Rehmani, Abdulahad Shafique, Asadullah Shafique, **Sheraz Ahmad Khan Tanoli**, Sabahat Ambreen, Muhammad Inam-Ul-Haq, Muhammad Ibrahim and Syed Sammar Abbas.Compettition of Aluminium on Iron Binding Site in the Biological System. *Journal of Pakistan Chemical Society*, **2010**, 32(4):467-470.

9. Nazia Bibi, **Sheraz Ahmad. K. Tanoli** , Sadia Farheen , Nighat Afza , Salman Siddiqi, Ying Zhang , Shahana U. Kazmi , Abdul Malik.In vitro antituberculosis activities of the constituents isolated from Haloxylon Salicornicum, *Bioorganic & Medicinal Chemistry Letters* 20, **2010,** 4173–4176.

10. Abiodun Falodun, Rizwana Siraj, Muhammad Irfan Qadir, **Sheraz Ahmad Khan Tanoli**, Muhammad Iqbal Choudhary. Chemical Composition and Insecticidal Activity of Volatile oil of Khaya grandifoliola, *Medicinal and Aromatic Plant Science and Biotechnology*: 3, (1) **2009**, 61-63.

11. Tabopda Turibio Kuiate; Ngoupayo Joseph; **Khan Tanoli Sheraz A**; Mitaine-Offer Anne-Claire; Ngadjui Bonaventure Tchaleu; Ali Muhammad Shaiq; Luu Bang; Lacaille-Dubois Marie-Aleth. Antimicrobial pentacyclic triterpenoids from Terminalia superba. *Planta medica*: **2009**, 75(5), 522-7.

12. T. Kuiate, J. Ngoupayo, J. Liu, **Sheraz A. Khan Tanoli**, Khan S. N., Ali, M. S., Ndadjui, B. T., Tasamo, E., Luu, B; Bioactive aristolactams from *piper umbellactum*, *phytochemistry*, 69, **2008**, 1726-1731.

RESUMO

O foco principal desse trabalho foi a implementação, otimização e aplicações práticas de métodos de ressonância magnética nuclear (RMN) com o propósito de avaliar as interações entre moléculas de diferentes massas molares, sendo que essas técnicas foram implementadas pela primeira vez no laboratório de RMN do DQ-UFSCar. Existem várias abordagens que podem ser utilizadas com esse propósito, e dentre elas, destacamos: STD NMR, Tr-NOESY, WaterLOGSY, SALMON, INPHARMA, DOSY, SAR. Esses métodos são muito úteis para detectar mudanças de comportamento, a nível molecular, quando adicionamos macromoléculas em um meio contendo somente micromoléculas. O entendimento desse comportamento molecular ajuda a desvendar sistemas complexos de interações moleculares existentes no corpo humano, e que, são muito importantes para o descobrimento de novos medicamentos. O primeiro passo para a implementação das técnicas foi a utilização da proteína de soro bovino (do inglês BSA) e proteína de soro humano (do inglês HSA) como fonte de macromoléculas e micromoléculas orgânicas isoladas da fração etanólica do extrato bruto (1 mg) de *Rauia resinous* e da fração acetato de etila de *Strypnodendron polyphyllum*, utilizando cromatografia líquida de alta eficiência, extração por fase sólida e a ressonância magnética nuclear (CLAE-EFS/RMN) para a completa elucidação estrutural quando necessário. As técnicas utilizadas foram: *saturation transfer difference* (STD), *transfer nuclear Overhauser spectroscopy* (Tr-NOESY) e STD-TOCSY (*total correlation spectroscopy*). Essa mesma metodologia além de representar um importante mecanismo para avaliar as interações entre moléculas, também pode ser utilizada para outras matrizes variando tanto as macro quanto as micromoléculas.

Abstract

The main focus of this work was the implementation, optimization and practical applications of nuclear magnetic resonance (NMR) methods for the purpose of evaluating the interactions between molecules of different masses, and these techniques were implemented for the first time in the laboratory NMR DQ - UFSCAR. There are a number of ligand-based screening approaches available, such that, STD NMR, Tr-NOESY, WaterLOGSY, SALMON, INPHARMA, DOSY, SAR by NMR etc. These methods are sensitive towards the perturbations as results of the macromolecular addition in a medium containing the small molecules. The molecular understanding of this behavior helps to uncover the complex systems of molecular interactions existing in the human body, which are very important for the discovery of new medicines. In the first step while implementing these techniques, Bovine Serum Albumin (BSA) and Human Serum Albumin (HSA) as a source of organic macromolecules were used. While, for the small organic molecules, a 1 mg crude extract from the hydroethanolic fraction of the *Rauia resinous* and ethyl acetate fraction of *Strypnodendron polyphyllum* was utilized, however, for the complete structural characterization, solid phase extraction following high-pressure liquid chromatography in an integrated fashion and then nuclear magnetic resonance (LC-SPE/NMR) was employed when necessary. On the other hand, the ligand screening techniques used were the saturation transfer difference (STD), Nuclear Overhauser Transfer Effect SpectroscopY (Tr- NOESY), Diffusion-Order Spectroscopy (DOSY) and STD-TOCSY (TOtal Correlation SpectroscopY). More importantly, this methodology also represents an important mechanism to evaluate the interactions between molecules or the first hand detection of the active constituents/inhibitors, which can also be used for other matrices varying both the macro and the small molecules.

➊

Introduction

1. Introduction

1.1. Drug discovery: A historical perspective

As far as life is concerned, health care is the most important aspect of life. Therefore, the necessary prevention or cure develops on the apprehension of medicinal requirements after the breakthrough of some disease. The branch of science responsible for the discovery, development, and production of drugs is known as pharmaceutics. Since the human birth, human kind has been curious to study and analyze the surrounding natural products to get require remedies to treat sufferings. These sufferings may be due to some infections, inflammations, and some traumas or because of ageing etcetera. There is no documented report that could tell exactly, how and when use of drugs came into being. However, the earliest papyrus report that describes the use of herbals and other natural remedies in compensation to diseases dates to back fifteenth BC (RANG, H. P, 2006). Moreover, the findings of those remedies were not scientific at all until the mid-nineteenth century however, was more based on some judgements and accidental experimentations. Nevertheless, the scientific interest in utilizing the herbals as a source of medicines began in the eighteen century. Still, it is just a century back; it became more believable to develop the drugs in a rational way, and improvements particularly in biomedicine and chemistry lead the real foundation of the pharmaceutical industry.

Initially, drug research got its roots from chemistry after the maturity of chemistry principles and their application outside of chemistry sphere. In addition, pharmacology science also added to it after being accepted as a scientific discipline in itself (JÜRGEN DREWS, *et al.*, 2000). However, the first rational target directed drug came out at the end of the eighteenth century as a result of the postulate that put forward by Paul Ehrlich (Figure 1) about the existence of therapeutic effect (EHRLICH P., 1910).

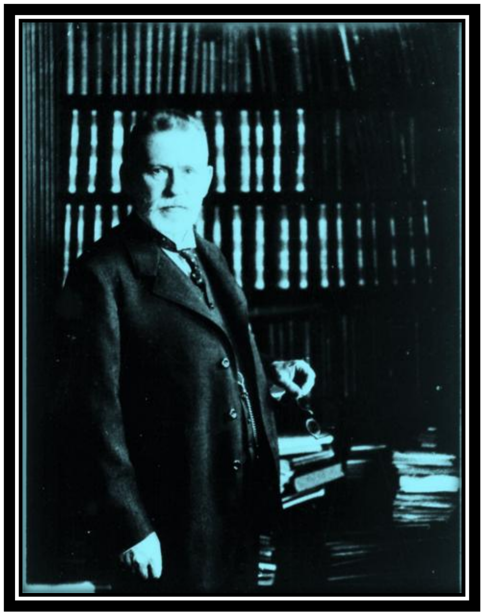


Figure 1. Paul Ehrlich, a medical student university of Strasburg, who postulate about the existence of chemoreceptor. [Image reported by JÜRGEN DREWS, *et al.,* 2000 as collected by the National Library of Medicine]

On the other hand, analytical chemistry, particularly the isolation and purification of active components from the medicinal plants played a vital role in promoting this most promising newly born and valuable branch of science. Ever since antiquity, plants have been the wellspring of the majority of therapeutics, demanded an increase in qualitative and quantitative approach that laid the basis for isolations and separation of active organic content of natural mixtures. The important discoveries of late eighteenth century like, isolation of morphine from opium extract (LUCHA L., 2009), Papaverin (MERCK, G., 1884) and ephedrine (PALAMAR, J., 2011), etc., which are still in use. These noteworthy reports affected the accompanying eras and expedited the disclosure of numerous drugs.

Early twentieth century dedicated to flourish the pharmaceutics and provided a tremendous boost up to this field by adding new technologies that revealed novel drugs in multiple areas, which are using nowadays like, sedative, antibiotics, and antiepileptic etcetera. Similarly, the discovery of the penicillin in 1929 by

Alexander Fleming was the blockbuster of this era (FLEMING A., 1929) that opens the new think of bacterial infection treatments. Conversely, biochemistry influenced drug research in numerous ways but the most overwhelming ideas offered is based on the study enzymes and receptor, which provided an excellent drug targets. The accidental discovery of the sulphonamide during the characterization of carbohydrates clued that this compound not only inhibit but also impact on water discharge (SCHWARTZ, W. B., 1949). As, Paul Ehrlich was the first who proposed the idea about the chemotherapeutic agents, however, J. N. Langley brought a new concept in pharmacology that, the receptor serves as a "switch" that not only accepts but also produces particular signals that could be either hindered by antagonist or activated by agonists (LANGLEY J. N., 1905). These findings led the further fruitful outcomes in the form of adrenergic receptors (AHLQUIST, R. P., 1948), beta-blockers (WOLSTENHOLME *G* and PORTER R., 1967) and other benzodiazepine receptors (HUNKELER, W. *et al.*, 1981) by the late fifties. This very short and inclusive historical background shows how and in which era the development process took place in the pharmaceutical industry from the blind folk medication to modern drugs.

1.2. Stages involve in drug development

Drug discovery is extremely long and exceptionally cost intensive process (LOMBARDINO, J. G., and LOWE, J. A., 2004) and in a normal run of 10 to 12 years of time with an average of 1 billion of dollar money (ADAMS, C. P and BRANTNER, V. V., 2006). However, three primary stages might be characterized that are involved in this process are, drug discovery, development and then marketing. Therefore, each primary phase relay on interlinked sub-stages. The schematic representation of various important steps involved in drug discovery, are represented in Figure 2. Nevertheless, the steps selected here for understanding this process but, not necessary to follow the sequence shown below, therefore, remains side by side.

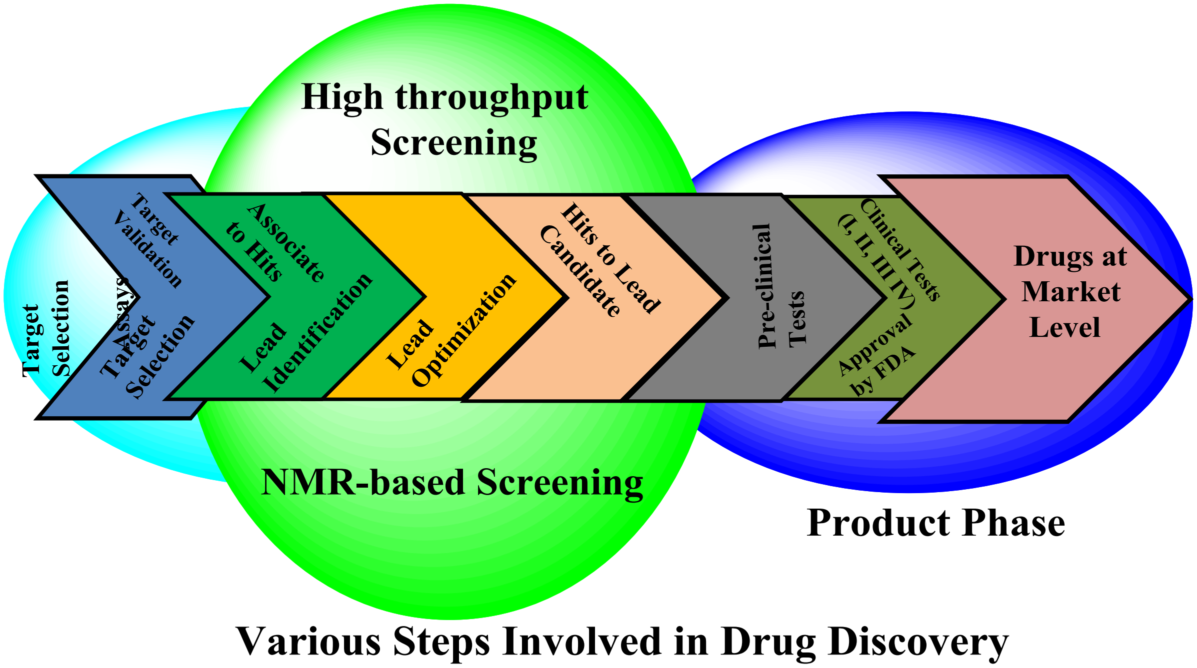


Figure 2. Overview of the various steps (target selection to market) involve in the drug discovery process. The NMR based screening process (inclosed in the green balloon) will remain the focus area of this thesis.

The need of the drug discovery starts after the distinguishing proof of a therapeutic condition, which may also include the counting and comparison of the existing treatments for sufficiency and certainty. From these earlier assessments, together with surrounding information and some basic knowledge about the target disease, will lay the foundation of some speculation, which later on can reshape into hypothesis, perhaps on how to enhance treatment. On the groundwork of some reasonable hypothesis, lead particular targets for the venture. Afterwards, the key process of finding some biological active hits, and thereby, the optimization of that begins. The optimization process may include the testing of some analogues to get better activity results—the best will try for some *in vivo* test and then toxicology screening. All data will be submitted to the Food and Drug Administration (FDA)

USA, in case if the given compound passes all critical stages. In the clinical process, drug tolerance, efficacy, safety and dose range are set thereof. In clinical steps, almost 90% of the drug candidates fail to make the future drug, (LOMBARDINO, J. G., and LOWE, J. A., 2004) conversely, the candidate making all steps clear, FDA approves as a new drug. However, drug discovery process, from target selection to market level, took almost an average of 12 years (ADAMS, C. P and BRANTNER, V. V., 2006).

The main focus of this thesis is on NMR screening methods involved drug discovery process, therefore, a little introduction will be provided for aforementioned steps in the later sections for better understanding of this process.

1.2.1 Target selection

In the drug discovery process, the term ‘target’ conveys several connotations. However, in general medicinal language target name referred to a macromolecule that is linked to some specific disease, the suppression or expression of it could eliminate or activate the particular disease or function. The process of affirming this whether the selected target has or not the particular effects is known as target validation. The target selection strongly influenced by the financial and some industrial setup (KNOWLES, J.; GROMO, G., 2003), consequently, lack of any of them make this impractical process. Nowadays, more and advanced targets become accessible due to the easy availability of 3D structures of protein—result of development in structural biology. Therefore, an even less expensive method of screening can be offered by this 3D structural information of proteins (KITCHEN, D. B. *et al.*, 2004) that could be a better choice (CLARK, D. E., 2006). The toxicity, inability, specificity measures the potency of a compound against the available macromolecules (lipids, polysaccharide, proteins, and nucleic acids) from the human body that can be interfaced by therapeutic agents, however, for protein, selection could be helpful through binding or modify the activity thereof (HOPKINS, A. L.; GROOM, C. R., 2002; IMMING, P.; SINNING, C.; MEYER, A., 2006). Moreover, they pointed out 120 drug targets in the same reports, can be used for therapeutic purpose. According to this report, enzymes are the most widely used drug targets following GPCR receptors, conversely, the DNA and Integrins are the least found targets as shown in Figure 3.

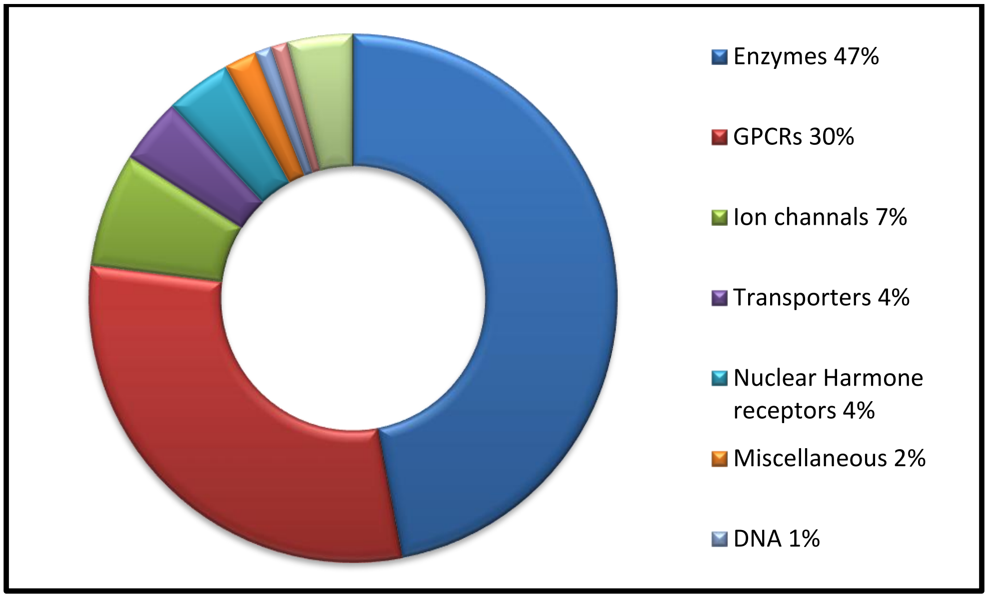


Figure 3. Division of commercially available small molecule drug target according to their biochemical classes, as defined by HOPKINS, A. L. and GROOM, C. R., in 2002 and also explained by IMMING, P.; SINNING, C.; MEYER, A., 2006)

A three year later, Lapel S., and co-workers estimated drug targets from the human genome to ~200-3000 (RUSS, A. P.; LAMPEL, S., 2005) , an even larger than earlier report. (HOPKINS, A. L.; GROOM, C. R., 2002) This number of register drug targets seem too big, however, up to four targets per year is an average increments in the existing numbers (GRAUL, A. I. *et al.*, 2010). Anyhow, a good target should posses the following characteristics like, efficacy, safety, fulfill the clinical and commercial requirements and, above all, be druggable in its behavior (HUGHES, J. P. *et al.*, 2011). Nowadays, the most important approach that uses not only for identification but also for the sequential priority for drug targets is data mining—obtained from various sources (YANG, Y., *et al.*, 2009).

1.2.2 High Throughput Screening (HTS)

The drug discovery industry is facing an unusual problem of productivity, mainly associated with two factors: a) some financial support and b) to find a modulator for the large number of drug targets, came as a result of genome studies. Thus, following target selection and validation process, high throughput screening is the obvious choice for the screening large number of compounds together (MACARRON, R., 2011). However, there are numerous screening paradigm (Table 1) exist that help in recognition of hit compound.

Screen	Description	Comments
High throughput	Large library of compounds can be analyzed through 384 or above number of plates well.	Huge library screening can be performed by using the computer programing and this setup can be applied either by some pharma and/or by academia in order to reduce the cost.
Focused screen	Already identified compounds or hitting some specific class of targets having similar structure.	It provides a cheaper means to find a hit molecule, however, suffers in identifying the completely new and novel structure.
Fragment screen	On the base of soaking compounds into crystals, a molecule with low mM activity, which can be used as building block for larger structure, can be obtained.	It requires a crystal structure, but by adding the selected fragments together, the potency of selected molecule can be increased.
Structural aided drug design	To design molecules, crystal structure is necessary	It can be used to enhance the potency on the basis of modification of different fragments and then docking. Usually large pharma follow this procedure.
Virtual screen	Crystal structure of protein or ligand can be sued to enhance the potency or activity that also depends on docking.	It is a cheaper way to use docking and other theoratical methods to find novel or potent compound from existing molecules.
NMR screening	Screening of large library is possible, soaking of targets or small molecules crystals in order to find low mM activity compound that provide a way to generate larger structure.	NMR as a structure determining tool

Table 1. Screening strategy involved in the drug discovery process, reproduced from (HUGHES, J. P. *et al.* 2011).

In HTS, large library of compound is screened either directly or indirectly against a specific target, however, in the latter case the target in turn depends on a secondary assay to observe the site of interaction, such as cell-based assay (FOX S.

et al., 2006). Moreover, the monitoring of the interaction between the receptor and compound is not necessary in the cell-based assays, but the other parameters can also be observed like, changes in cell metabolism, growth of the cells and antibacterial action etc. Another approach, which relay prior knowledge of targets that might be in the form of an article or some proceedings, but, the chemical compounds are likely to have some activity against the targets are called knowledge-based screening (BOPPANA, K. *et al.*, 2009). The knowledge-based approach becomes the basis for virtual screening (MCINNES, C., 2007) though the use of pharmacophore and molecular modeling.

Other techniques, like computational methods, Structure Activity Relationship (SAR) by NMR and more importantly the structure-based drug design are also being in screening studies. Computational methods are gaining importance because of screening hundreds of compounds rapidly, however, suffers more from the false negative results (CHEN, A., AND SHAPIRO, M. J., 1999). Crystallography is the base for the structure-based drug design, where 3D structure of the target is obtained before and then, co-crystallized with small molecule in order to get the interaction information. By taking help of this structure, new ligands are designed and re-crystallized in a similar way to get better-hit compound. The 3D crystal structure obtained as a result of this study further help the researcher in the field of molecular docking and virtual screening methods (BOHACEK, R. S., *et al.* 1996).

On the other hand, ligand screening is greatly influenced by NMR spectroscopy. NMR provides a plethora of a very simple and robust method formats for ligand screening methods against the selected target. Notable amongst are, WaterLOGSY (Water-Ligand Observed via Gradient Spectroscopy), STD (Saturation Transfer Difference, NMR), transfer-NOESY, SAR (Structure Activity Relationship) by NMR etc. These experiments provide the liberty to observe binding activity of small molecules without any prior knowledge of structural information of protein, just the observation fallow the ligands behavior. Similarly, the relative diffusion of the small molecule with and without the addition of target molecule can also be counted to see the existence of the interaction between them through DOSY (Diffusion Ordered Spectroscopy) NMR. Besides these experiments, NMR also provides information regarding changes occur on binding within the target molecule as well, like in ^{1}H-^{15}N experiment, which help in chemical shift mapping. The main focus of this thesis is the utilization of the techniques, which help to observe changes in ligands structure upon binding thus, will be discussed in details in later sections.

The molecule, which comes as a result of HTS, certainly, has some activity towards the selected target and is now called as 'hit' molecule. The step is the development of assays, where the majority of assays are performed in industries rely on cell lines expression or protein expression studies that are sometimes called bio-chemical assays. However, recent studies have shown that the involvement of the cell-based is much reliable in assay development (DUNNE, A. *et al.*, 2009). Therefore, the cell-based assays usually focused the target and provide the result as a function of compound activity. Generally, nuclear receptors, ion channels and membrane receptors are the particular class for the application of cell-based assay (MICHELINI, E. *et al.,* 2010). On the other hand, the affinity of the selected ligand for particular receptor and enzymes target can also be determined by using the bio-chemical assay (MOORE, K AND REES, S. 2001).

The drug like properties of small molecules that fail during the initial assessment stages mainly because of low bioavailability and toxicity can be dealt with the Lipinski rule of five (LIPINSKI, C. A. *et al.*, 2001)−based on following statement. 1) The lipophilicity LogP of drug should be below 5, 2) weight less than 400 DA, 3) less than 10 hydrogen donors and 5 accepters bond. After having the hit compound in hand, as a result of HTS screening, the next step is a competition with analogues. After this stage, the compound, which stands tall in all SAR tests, is known a ‘Lead’ compound.

1.2.3 Lead optimization

Lead optimization is the last and multifaceted process before the kick off of clinical trails. To assess the drug metabolism and pharmacokinetics (DMPK), typically, lead compounds have to surpass various *in vivo* and *in vitro* screening, in order to check safety measures (KORFMACHER, W. A., 2003, BAKHTIAR, R., 2006). However, dose proportionality, efficacy and selectivity for a lead compound are important parameters besides, others essential ones (Table 2). It is necessary for a lead compound that it should meet the five essential requirements simultaneously, mentioned in Table 2 (CHENG, K. C *et al.*, 2007), in contrast, lacking in any one of them, may become a cause of rejection. Thus, it is the toughest part of the drug discovery after clinical trails.

Three (safety issues, duration oral and bioavailability) of the five essential parameters, necessary to enter into the clinical stage can be placed on DMPK process and can be dealt separately. Therefore, DMPK process can further be said-divided accordingly. a) Drug, in most of cases, would be orally administered, b) duration of drug in body be increased to reduce the dose quantity, C) reduce in toxicity, if exist any. Nowadays, DMPK environment associated with the lead optimization process, but, in the past these studies was performed soon after the lead recognition (HSIEH, Y *et al.*, 2009). A number of well-established assays, including *in vivo* and *in vitro*, are present that are well-known for lead optimization in the process of DMPK studies like Caco-2, plasma protein binding, CYP P450 inhibition, metabolite profiling and transporter profiling etc. (VAN DE

WATERBEEMD, H., *et al.*, 2001; JENKINS, K. M. *et al.*, 2004; HARTMANN, T. *et al.*, 2006).

Property	Definition/Requirement
Potency	HTS measured by invitro methods, by using the inherent ability of a compound in order to generate good response.
Safety	The existance of the targeted response versus non-targeted response of the compound is measured.
Pharmaceutical Acceptability	Different pharmaceutical properties of targeted compound has been checked e.g., stability, synthetic rutes, solubility and dissolution etc.
Oral Bioavailability	In this step, abaility to reach the target and rute to tract pathways like, GI or liver pathways is varified.
Duration (Half-life)	The ability of compound to stay in circulation system available to provide a meaningful response is verified.

Table 2. Typical characteristics of drug-like lead compounds, regenerated from (CHENG K. C *et al.*, 2007).

1.2.4 Drug development and clinical trials

From the last two decades, biomedical studies have advanced the drug development pathway by adding more and advanced new molecular entities (NMEs), in contrast, therefore, the rejection by the clinical trials and FDA make this too hard to generate a single drug up to market (TWOMBLY, R., 2006). Consequently, a single drug takes 10-15 years and extraordinary cost to reach in the market, before passing these difficult but necessary stages (DIMASI, J. A. *et al.*, 2003). According to estimate about, 75 % of the new drug, costs are directly linked to the initial development process and clinical trials (KOLA, I. LANDIS, J., 2004; lorusso, P. M., 2009).

The drug development process comprises all events take place from lead candidate to showcasing a drug into the market, including all transformation necessary to enhance the activity, durability, half life, efficacy and of course, the safety concerns. Therefore, a huge criticism exists about the creation of an absolute model, which hinders the researchers during this developmental process. The most pointed out factor are; a) the wrong selection of animal model to be used in trials (JOHNSON, J. I. *et al*., 2011), b) the use of traditional drug development models without any new innovation (POCARD, M., 2010), c) lake of assimilation between the early activity tests and pharmacodynamic endpoints (CALVERT, A. H. PLUMMER, R., 2008) and improper validation of biomarker necessary for drug approval.

The pre-clinical tests are long enough which may take up to two years where the main concerns is to observe the reproducibility of the test under the good laboratory practice. During this process, most of the time, in order to enhance the drug permeability, stability, access to the target and solubility etc., some additive contents are also used like, the most widely used components is cyclodextrins, and other dextran units (ALLEN, T. M. CULLIS, P. R., 2004). The clinical stage is comprised a few sub-phases, although some contradiction exists about the absolute number of sub-phases, the roughly speaking therefore, major four phases are present (KOLKMAN, A. J. 2013) as given below.

- Phase I Clinical Trials
- Phase II Clinical Trials
- Phase III Clinical Trials
- Phase IV and Beyond

Each phase has its specificity to deal the drug-like lead candidate, and also different number of sampling from other counterpart. In all phases of clinical trials practical applicability of the drug-like compound is tested against patients, in contrast to the phase I, where, it is not necessary to be a patient, a healthy common person can be a part of these tests. Usually, a small sampling is used to deal with phase I may have up to 100 peoples. The main purpose of the phase I, is to check the drug tolerance, safety and some pharmacokinetics of this drug-like lead candidate, that may happen as results of low dose injection. The gradual success of the drug-like candidate, therefore, become more risky, during the clinical stages because, the rejection rate by failure at any stage, is likely—that's why the drug discovery becomes too expensive, time inefficient and full of risk (ADAMS, C. P & BRANTNER, V. V., 2010; DIMASI, J. A. *et al.*, 2003). However, in the phase II, the test limit may extend to 300 patients, where the further acute attrition took place, and further safety and dose range are the major concerns (ORLOFF, J. *et al.*, 2009). Similarly, in the phase III and IV sampling further extend to thousands of patients, and the major concerns are side-effects and mode of action related issues. The phase IV is also associated with other events such as marketing, licensing, and patenting etc. A lead candidate after passing all the clinical trials consider as a new drug and well documented report, as an outcome of clinical trials, is sent to the FDA for the approval and thereby, for launching into the market.

The initial screening for a target starts with a library of thousands of compounds, after overcoming all obstacles of this long drug discovery process, one or two drugs become available to the market. If we consider this drug discovery process as funnel having different sieve-size within it, at intervals then the process can be easily understood (Figure 4).

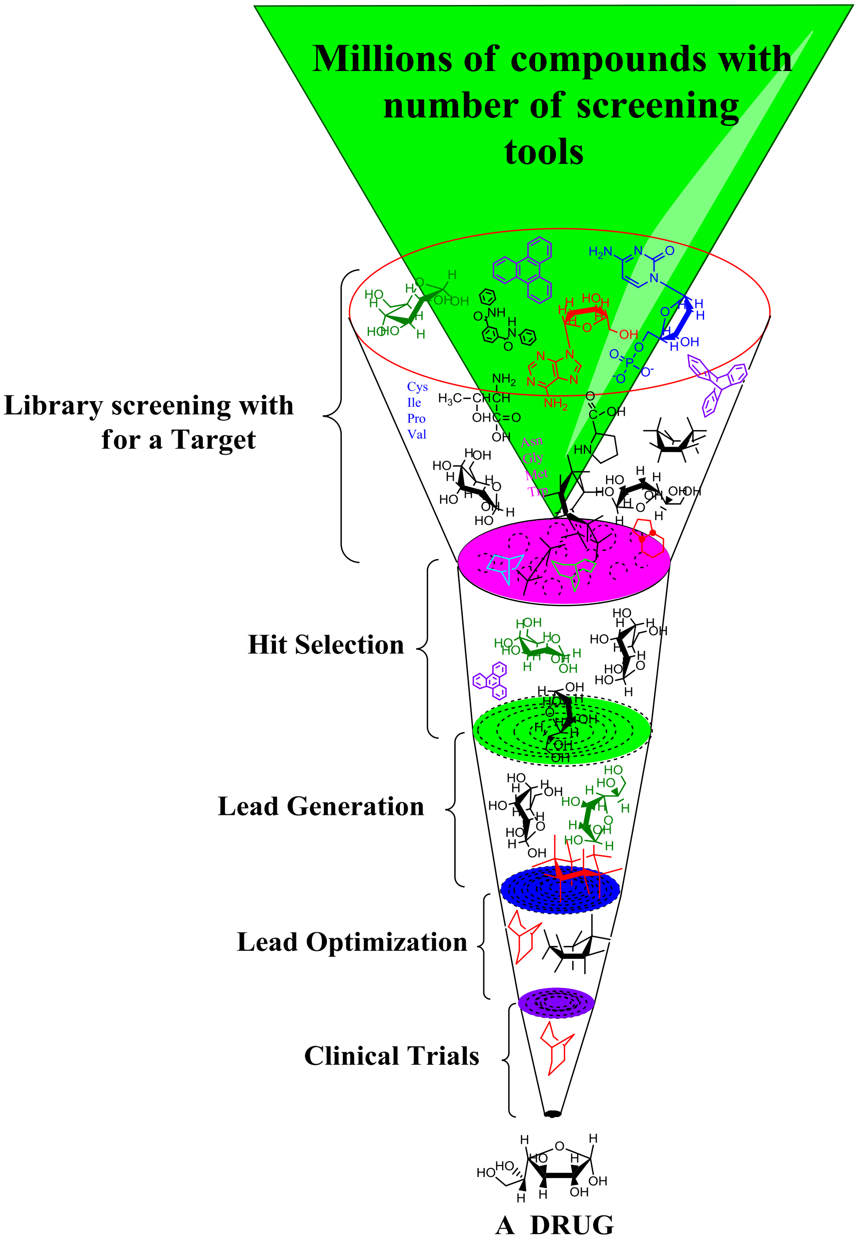


Figure 4. The pictorial presentation of drug discovery process along with some major stages.

1.3 Role of theoretical approaches in drug discovery

The high failure rate of the lead compounds in clinical testing, consequently, makes the drug discovery process enormously costly, and time inefficient. However, a little compensation to this process was provided by high throughput screening (HTS) in identifying the initial lead compound more rapidly. Therefore, in some cases, HTS screening shows a huge failure rate in the identification of leads, provides the false positive result (RISHTON, G. M., 1997), and obviously opens a way to other techniques to fill this space. However, a process that reduces the cost during initial assessments can lead the drug discovery process ahead and can benefit the process. The developments bioanalysis provided a huge support in the initial screening; unfortunately, bioanalysis alone cannot provide all information, necessary for this process work efficiently (WHITE, R. H., 2006). However, still the need of technique that can readily respond, and can screen an array of compounds rapidly, was present. Thus, the computational methods such as molecular modelling and docking came into being to driven the slow drug discovery process towards fast track.

1.3.1 Molecular docking

Molecular Docking (MD) is a computational tool that predicts the favourable binding sites and binding conformations of small molecule within the molecule-receptor complex (HALPERIN, I. *et al*., 2002; WARREN, G. L. *et al*., 2005). This is the best tool based on energetic calculations that predict the interaction conformations within the complex. In addition, re-docking suggest the accurate results after refining a molecular pose with a number of times (HARTSHORN, M. J. *et al*., 2007). Since proteins are macromolecules—possessing thousands of atoms with numbers of interactions types which require a good time and excellent processing power, that’s why, most of the time docking analysis targets only a single system (TAUFER, M. *et al*., 2005). The most common forms of interactions that make the ligand-complex stable are hydrogen bonding, Van der Waals interactions, electrostatic interactions, dipole interactions, hydrophobic and hydrophilic interactions, but, on the other hand, solvent effect and water effects are also important. To start this MD process, the 3D protein (apo and holo form) structure is necessary—demanding preparation like, protonation, structure repair, etc., before undergoing analysis. Thus, before posing and scoring, small molecules are added into the protein structure and hence, posing (systematic or random) correctly can lead the excellent dock (KITCHEN, D. B. *et al.*, 2004; HALPERIN, I. *et al.,* 2002). Therefore, the posing can be implemented differently on different systems, for example., Glide, FlexX (RAREY, M. et al., 1996), TrixX (SCHLOSSER, J.; RAREY, M., 2009), Enovo (PEARCE, B. C. *et al*., 2009) and Dock (EWING, T. A. *et al*., 2001) program can approach the systematic posing while, Gold (JONES, G. 1995) or Autodock (STARK, J.; POWERS, R., 2007) can do random posing. In addition, the rigid and fixed structure can also simplify the docking process. Usually five types of scoring functions: Knowledge-based, empirical, consensuses, shape-based, and force field-based, are used. More details of the docking is not provided, however, can be in the earlier published reviews (KUNTZ, I. D. *et al*., 1994; COLE, J. C. *et al.,* 2005). In short, docking studies help in understanding the binding sites and different conformational effects that make the complex stable, very recently an excellent example shown by Ferreira A. G., and co-workers, explaining the flavonoidal interaction within serum albumin complex, as shown in figure 5 (Tanoli S. A. K. *et al*., 2013).

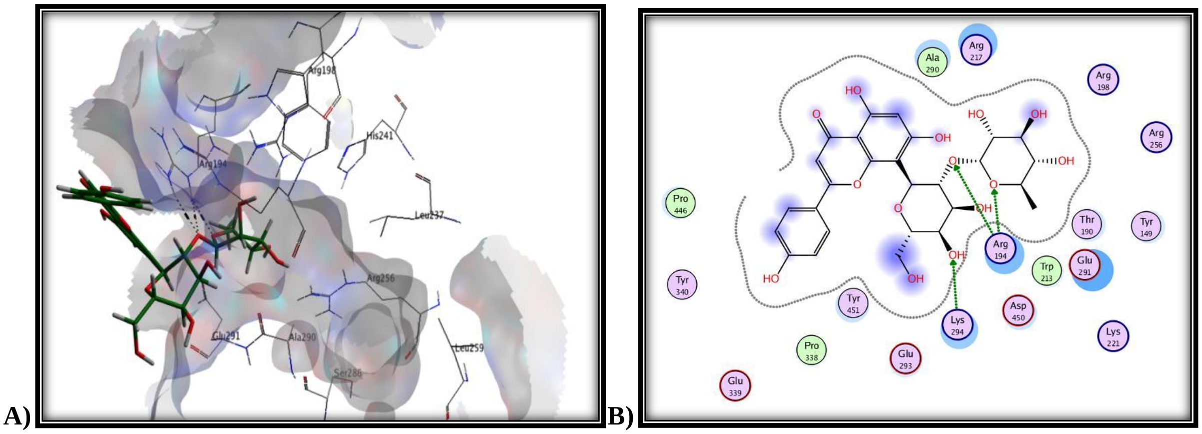


Figure 5. Molecular Docking model of interactions between the Bovine Serum Albumin (BSA) and vetexine-2-O-rhamnoside, (A) and classical interaction types (B) with different colors as generated by using Molecular Operating Environment (MOE) 2011.10. Adopted from Tanoli S. A. K. *et al*., 2013.

Molecular docking depends on the X-ray crystallography and NMR, for the availability of 3D structures. However, homology model (HILLISCH, A. *et al.*, 2004) can also be used in the absence of any 3D structure but, need extra care and high skill to process, and some case wrong selection of amino acids from protein may cause a divergence in the same results at different occasions (KUFAREVA, I. *et al.*, 2011).

1.3.2 Virtual screening

In the recent years, virtual screening becomes the most widely applicable technique in the drug discovery process (SHOICHET, B. K., 2004). In the structure-based drug design, a number of protocols available for screening large libraries against some targets and this quest still is going on. Computer aided drug design helps in utilization of minimal resources like money, time and space etc., compared to the traditional high throughput screening. Therefore, the screening a large library of compound against some disease through docking and thereby, ranking of compounds based on binding affinity is called virtual screening (KLEBE, G. *et al*., 2006; SCHNEIDER, G.; BÖHM, H. J., 2002). Some time, the virtual screening is called the neighbourhood behavior, because derives the information from closely related structures (2D or 3D) from the database, with similar activity (VIDAL, D. *et al*., 2006) thus, it provides a mean to get a potent ligands from all similar structures (LENGAUER, T., *et al*., 2004). Virtual screening involved various methods including, ligand based screening, structure based screening (KUNTZ, I. D., *et al*., 1982), machine learning techniques (JORISSEN, R. N. GILSON, M. K. 2005) and virtual combinatorial libraries.

1.4 Analytical tools used in drug discovery

A robust, precise and accurate analytical method that fit to uphold all phases of drug discovery including purity assessment, ligand recognition for specific targets, preclinical trials to drug formulation and quantification, is the basic need for the accomplishment of a drug discovery process. Therefore, the drug discovery process is a multidisciplinary analysis approach that covers cell-based assays to highly sophisticated technologies of a new era. However, the strong quality control methods that can deal bio-analysis with new and advanced instrumental setup can cope with drug discovery demands (RODA, A., 2010). From last three decades, analytical chemistry benefited drug discovery process by owing various technologies having a number of methods. In short, broad spectrum of analytical methods seems to be appealing tool with a number of applications: from finger printing metabolism studies (PRAKASH, C. 2007) to impurity finding in lead candidates (WRIGHT, P. 2001; MÜHLEBACH, A. 2011), and from pre-clinical (DEAR, G. J. 2011; MAURER, H. 2007) to clinicians as well as in the last stages of quality controlling too (YEKKALA, R *et al.*, 2008; PARK, J. H., *et al*., 2008).

The major and most widely used analytical tools in the pharmaceutical analysis are, including high-performance liquid chromatography, capillary electrophoresis, circular dichorism, fluorescence spectroscopy, mass spectrometry, vibrational spectroscopies (Ramman and infrared) UV-visible spectroscopy, X-ray diffractometry, and nuclear magnetic resonance phenomenon. Sometime, to take extra advantages of these integrated techniques, therefore, coupling of two techniques together in an integrated fashion provided even better results for this analysis. The most common hyphenation system that are being used are LC-MS, GC-MS, LC-NMR , UV-MS, UV-MS-NMR and LC-MS-NMR systems. However, the application of the aforementioned hyphenation is much benefit to natural product analysis, metabolomic profiling, and in system like, body fluid and related complex mixtures.

Besides spectroscopic methods, a number of computational approaches have also been used in the drug discovery process, as a tool to recognize the lead candidate from the large library in very short time.

The main theme of this thesis is to target only the NMR techniques that are being used in the drug discovery process, therefore—will be detailed discussed in the last section of the introduction chapter 1, conversely, an overview of other some spectroscopic and computational methods will also be presented here bellow.

1.4.1 Fluorescence spectroscopy as a tool for drug discovery

Fluorescence Spectroscopy (FS) is an extensively technique, used as a tool in clinical treatments especially for the diagnosis of the tissue malignancies (MILLER, J. N., 1984; RUBIO, S., 1986; MASILAMANI, V., 2010). Due to inherent sensitivity, FS can provide landmark of both minute qualitative and quantitative variations in tissue cells (YUANLONG, Y., 1987; MARTIN, S. F. *et al*., 2007; ANKERST, J., *et al.*, 1884). Besides sensitivity, a true equilibrium condition during the experiment is another reason of FS to be used extensively. Generally, FS can sense in a range of nanomolar to picomolar amount in any ligand-receptor system (JAMESON, D. M., *et al.,* 2003). In additin, fluorescence correlation based assay can help to observe the membrane receptor like mechanism for example, physiological regulatory mechanisms (HEGENER, O., *et al*., 2002) up to the molecular level. And, due to its operation in real time during the Fluorescence Correlation Spectroscopy (FCS) assay, it has been proven an excellent state-of-the-art tool for the natural product and especially evaluation of their interactions towards some physiological targets. Moreover, the FCS techniques can be applied to study the binding interaction where, fluorescence anisotropy changes exist—an unusual character of fluorescence. In addition, to the aforementioned advantages, FS can also help with measurements of conformational changes (BISMUTO, E., *et al.,* 2001) and protein velocity changes as well (MAGDE, D., et al., 1978). Another fluorescence application, the polarization is also being used to observe the ligand-protein interaction (JAMESON, D. M. *et al.*, 2003; JAMESON, D. M.; SEIFRIED, S. E., 1999). Similarly, the interaction between the DNA and drugs remains a matter of discussion under the fluorescence quenching (FQ) experiments—which clearly justifies the not only the interaction but also the drug localities within the binding mode too (SUH, D.; CHAIRES, J. B., 1995). More importantly, FS is well established as tool distinguish the types of binding within the DNA and small ligands, for example, groove binding, the intercalation and external bindings by showing increase or decrease in absorption maxima (LI, W.-Y., *et al.*, 1997).

1.4.2 UV-visible spectroscopy (UV)

The UV-Visible spectroscopy is the most intensively used technique in the drug discovery to analyze the absorbance of the UV light by steady state or the time resolved experiments (spectroscopy) of ligand-protein systems. The absorbance of the UV light in protein depends upon some factors like, presence of chromophores, by the process of oxidation, by ligations and different changes of conformations. By Gibbs free energy calculations steady-state spectroscopy helps in the differentiation of ligation from unligation in complex systems, by taking account of binding equilibrium (NIENHAUS, K. NIENHAUS, G. U., 2005). In contrast, time resolve spectroscopy helps to observe the short live transition state. The use of the optical spectroscopy of the pre-clinical and clinical trials is becoming more regular (BROWN, J. Q., *et al.*, 2009). A very recent article shows that UV is the simplest techniques to observe the binding affinity, evolve as a result of interaction between the DNA and drug (SIRAJUDDIN, M., *et al.*, 2013) and their stability in the complex. The ligation to the receptor can be predicted from the unligation through the observation of absorbance maxima in both cases; the shifting indicates the difference (SUN, H.; XIANG, J., *et al.*, 2011). Nowadays, the UV technique however is extensively used in combination of some other spectroscopic methods like, circular dichroism (CD), Fluorescence spectrometry (FS), etc., as a tool to observe the interaction between the ligands and protein molecules (YU, X., *et al.*, 2013; VIGNESH, G., *et al.*, 2014; ZOLGHADRI, S., *et al.*, 2010).

1.4.3 Circular Dichroism (CD) as a tool for drug discovery

The polarized light induced by the chromophores as a result of differential absorbance of light laid the base of circular dichroism. Protein is a macromolecules and posses some chromophores, which on the absorbance of light can give raise the distinct signals—driven the conclusion, which provide useful information in the drug discovery process. So, when a small molecule ligand/drug binds to the large receptor molecule, due to the electronic rearrangements as a result of perturbations of the structure, provides an induced spectrum. However, the wavelength and intensity of interaction depend upon the ligands absorption and geometry of bindings, respectively. In contrast, if the ligands do not have any chirality, then, after addition of the protein, they acquire an induced chirality that provides in turn an induced spectrum (ICD). In most of the times, CD concerned with the structural characterization especially, as an advance tool to uncover the chirality problems (Gottarelli, G., *et al*., 2008). However, alone CD is not much helpful in the interaction studies but, somehow it helps to understand the mechanism of binding. On contrary, in coordination with other spectroscopic tool (YU, X., *et al*., 2013; VIGNESH, G., *et al*., 2014; ZOLGHADRI, S., *et al*., 2010) this widely used in interaction studies.

S. Arunachalam and coworkers, in a very recent article illustrated the use of CD to uncover the conformational changes as well as the environment changes, which occur during interaction between the BSA/HSA protein and inorganic surfactant complexes (VIGNESH, G., *et al*., 2014). Similarly, the use of CD in the DNA-drug interactions has also been reported (HEGDE, A.; SEETHARAMAPPA, J., 2013) —where, binding of an anti-leukemic drug at various concentrations, and in addition, conformational aspects were also studied towards calf thymus DNA.

1.4.4 Mass Spectrometry (MS) —and related hyphenation systems in drug discovery

Mass spectrometry (MS) is a very famous technique, especially in terms of strong detection power and high sensitivity. Mass spectrometry has too long history in the structural characterization of small natural, synthetic as well as the macromolecular proteins (GOODWIN, C. R.. *et al*., 2012; HOWE, I.; JARMAN, M., 1985; SCHOONJANS, V., *et al*., 2000). Therefore, developing interest in the structural characterization, binding modes, and binding selectivities have triggered the idea of combining some instrument together (ZHENG, J., *et al*., 2011) that can help MS in this realm. In this viewpoint, several hyphenation methods have developed that provided excellent results, for example, LC-MS—one separation and other characterization technique; LC-ESI system, ESI-MS system, MALDI-TOF-TOF system, LC-SPE-MS etc.

Hyphenation systems got special hype in the field of virology and related interaction studies (ZHENG, J., *et al*., 2011) especially, Tandem Affinity Purification Mass Spectrometry (TAP-MS) and Electrospray Ionization Mass Spectrometry (ESI-MS). TAP-MS has two important advantages, a) can be helpful to eliminate contamination and 2) can be applied to any protein there is no constraint, and having high reproducibility (PUIG, O., *et al*., 2001). The ESI-MS has a great impact in studying the binding events of the Duplex DNA/Drug Complexes (GALE, D. C., *et al*., 1994) and DNA-ligand complexes (WAN, K. X., *et al.*, 2000). In ESI-MS analyts are sprayed, nebulised and then converted into the gas form, which later on detected by the MS detectors according to mass to charge ratio, however, the conversion into ion is not too effective as compared to the Atmospheric Pressure Chemical Ionization technique (APCI).

Another common form of MS analysis is quadruple-MS, which frequently used in pharmaceutical industries (LEE, H., 2012). This quadrupole setup consists of a series of four electrodes, however, in a triple quadrupole-MS there are three pair of four sets of electrode worked jointly. The series of quadrupole electrodes helps satisfactorily in the metabolite/fluid analysis finger printings where a simple MS system some time does not provide fruitful information in term of ionization. Another hyphenation of MS with immense effect in pharmaceutical analysis is the Matrix-Assisted Laser Desorption Ionization coupled with Time-Of-Flight (MALDI-TOF) mass spectrometry (WOODS, A., *et al.*, 2002).

A long historical background and a comprehensive application of MS are avoided because the main focus of the thesis is NMR. Thus, a short overview of the

MS techniques involved in the drug discovery has been discussed here, although, there are plenty of applications of MS exist—and it's difficult to write in a single paragraph.

1.4.5 X-ray diffractometry

It would not be wrong to say that, 3D structure of the target is the basic requirement for the structure-based drug discovery. However, in terms of theoretical approach, it is the first step towards the drug discovery. Thanks to the RCSB Protein Data Bank web site: (http://www.rcsb.org) where, thousands of proteins (apo and holo form) publicly available to everyone, making a huge contribution in studying protein-protein and ligand-protein interaction phenomenon. X-ray crystallography is the major source; around 90% of available structures of the RCSB web site emerged from this marvelous technique. For the 3D structure X-ray is very important however, in the ligand-protein interaction studies usually, done by co-crystallization or by soaking the fully-grown crystals, which can provide information about the interactions and most importantly, about the bound conformations. NMR and X-ray are the only techniques those can provide full information about the ligands conformations inside the complex. For many years, X-rays has been applied only for the structure elucidation, but now therefore is established for interaction studies (DAVIS, A. M., *et al*., 2003). A few years before, Sevrioukova and coworkers, studied the crystal structure of bacterial cytochrome P450BM-3, within the reductase complex (SEVRIOUKOVA, I. F., *et al*., 1999) to observe the interactions and bound conformations of heme and FMN peptide of protein.

1.4.6 HPLC-NMR/MS and HPLC-SPE-NMR/MS

After the analytical methods and LC-MS, NMR is the second large method that has been exploited by chemists and biologists to study the structural behavior and quantitative measures of the compound. Both techniques have strong background of structural characterization but the combination of both into an integrated fashion has made the structural characterization too easy, fast for the mass limited samples. In addition, provides an easy way to overcome the problems associated with dereplications (SCHLOTTERBECK, G., *et al*., 2009). On the basis of the status of the sample during the experiments within an LC system, can be divided into three working modes. a) Continuous-flow NMR/MS measurements; if the spectra are being taken from the NMR/MS during the chromatographic run without any interruption this called continuous-flow mode, a possible schematic representation of this mode of working is shown in figure 6. It provides the short residence time for the nuclei to stay in NMR probe, hence, suffers from line broadening and mostly depend upon the flow rate and cell volume (GRIFFITHS, L., 1995; GRIFFITHS, L., 1997), on the other hand, allows normally to get 1D spectra (WARIDEL, P., *et al*., 2004).

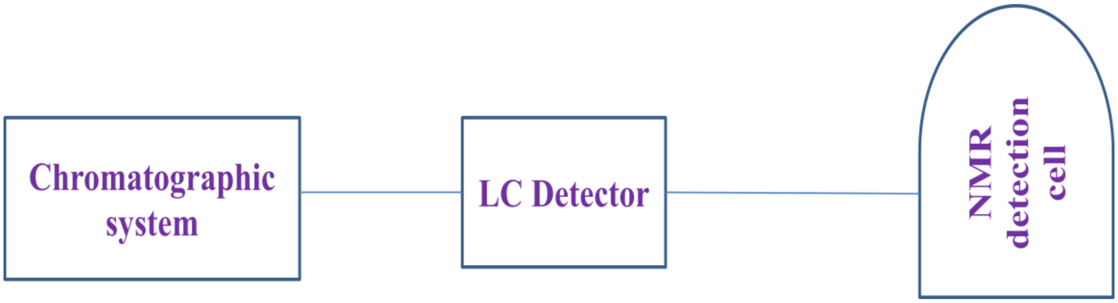


Figure 6. The On-flow/continuous-flow working mode of the LC-NMR modified from (Albert K., 2002).

b) Stopped-flow mode; in this method, the sample is directly transferred to the NMR cell but, LC conditions are interrupted in order to create static conditions during the run. Thus, to enhance the sensitivity problems associated with on-flow method, and to get 2D spectra stopped-flow method was used extensively (BRINGMANN, G., *et al*., 1999; EDRADA, R. A., *et al*., 2002). The schematic representation is given in Figure 7.

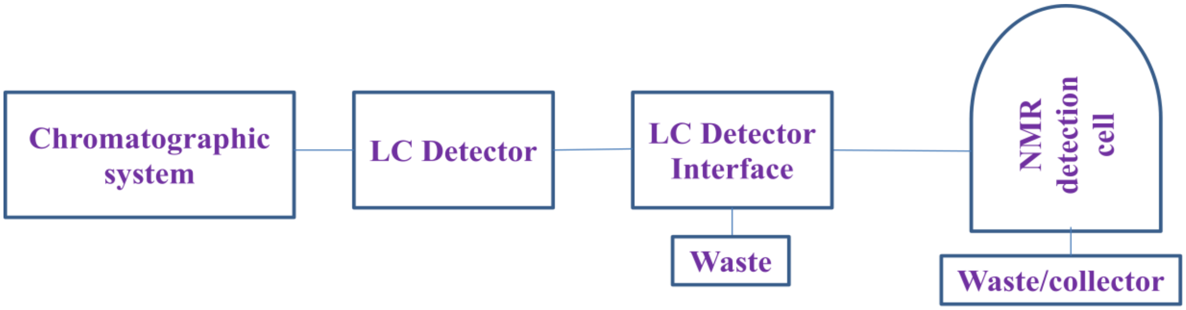


Figure 7. The stopped-flow working mode in NMR reproduced from (Albert K., 2002).

c) Loop-storage/Loop-transfer NMR/MS; sometimes the eluent from the LC system is collected in loops or other storage devices and send to NMR one by one while the chromatographic runs keeps working without any interruption. This is the most famous and widely used method to for structural characterization in the drug discovery process, where the peaks are sent to loop storage from LC and further storage can be done subsequently (TSENG, L.-H., *et al*., 2000). The schematic representation is made in Figure 8.

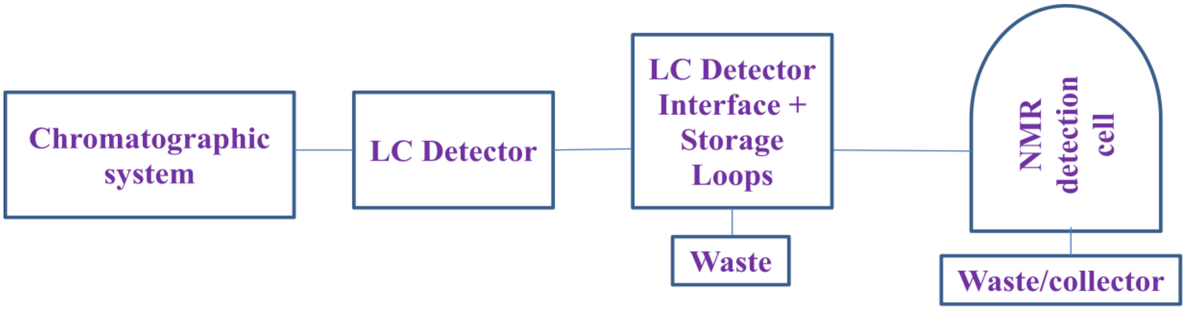


Figure 8. The Loop-storage/Loop-transfer NMR system is shown; the figure was reproduced from (ALBERT K., 2002).

Later on, the application of solid phase extraction (SPE) method as an interface between the NMR and LC system have removed all decomposition problems that were associated with the loop - storage system, and emerge as new and fast method for the characterization of mass limited samples (JAROSZEWSKI, J. W., 2005; TANOLI S. A. K. *et al*., 2013).

1.4.7 Nuclear magnetic resonance as a tool for drug discovery

NMR is frequently esteemed for its capacity to shed light on molecular structure; however, its most terrific potential in drug discovery likely lies in the data that it can uncover about atomic interaction. Nuclear magnetic resonance has a long history of its structure resolving power for small molecules, macromolecular as well as molecular dynamics. Conversely, it also helps in observation of both transient and static characteristics of the proteomics in any given organism. Another simple and straightforward parameter of NMR that is too touchy to the defining environment of the atom and subsequently concludes information whether or not a small molecule is bound the chemical shift change. In addition, group epitope mapping reveals information relating to bound conformations; what parts of the small molecule are interfaced and to which part of the macromolecular focus on the small molecule is bound. NMR has additionally turned into an important screening apparatus for the coupling of ligands to protein (targets), and has the key preferences about having the capacity to locate and quantify interaction with high affectability without demanding any earlier knowledge of protein function. Among other experiments, relaxation methods and saturation transfer difference are sensitive towards the molecular motion of the given compounds, which quite different for bound versus free conformations within the complex. Furthermore, nuclear magnetic resonance also adds in the optimization of week binding hits within the strong binding leads by providing the information on both structures (ligands and target). Thus, in this thesis, by taking help of these straightforward NMR method lies that might utilized to accept ligand tying or to distinguish potential ligands within in mixtures will be discussed in detail. A number of NMR methods named already mentioned above in the high throughput screening (section 1.2.2) therefore, will be discussed in the later sections, which are being used to help the drug discovery.

To understand a function of the complex human body, it is necessary to observe the binding events involving ligands-protein or protein-protein interaction. Moreover, with the availability of high resolution NMR, it is possible for the interaction even at the atomic level. When a sample (ligand and protein) is placed within the magnetic and RF fields, perturbations occur within both structures therefore, it is possible to observe these changes on either structure. However, both (ligand and protein) are equally important to be tracked. Therefore, on the basis of tracking resonance approach, NMR screening can be divided into categories a) receptor based screening and b) ligands based screening (WILLIAMSON, M. P. 2013; MEYER, B.; PETERS, T., 2003).

1.4.7.1 Receptor based screening

1.4.7.2 Chemical shifts mapping

Receptor or target based NMR screening include the chemical shift mapping/chemical shift perturbation method: can be applicable to any type of compound, and is a very efficient study without any protein size constraint. The chemical shift changes of the isotopically labelled target are followed by adding the ligands or other macromolecule consequently, binding locus of ligand within the protein can be observed. Usually these studies are monitored by ^{15}N-HSQC experiments (WILLIAMSON, M. P. 2013). Therefore, if the enough concentration of the protein is used then this type of titration is measurable by a day, however, the protein should be well-behaved one (WILLIAMSON, M. P. 2013). The dissociation constant (K_D) calculation is also possible up to the upper limit of ligands in milimoler to micromolar amount (FERNÁNDEZ, C.; JAHNKE, W., 2004). In contrast to advantages as mentioned before the only major disadvantage associated with chemical shift mapping is demanding isotopic labelled protein.

1.4.7.3 Structure Activity Relationship (SAR) by NMR

The process of identification, optimization of small ligand that binds to the protein, and linked together to produce high affinity ligands, through the NMR is called SAR by NMR (SHUKER, S. B., *et al*., 1996). In this approach, discovery of a second molecule that strongly binds to the targeted compound (receptor) is performed through the help of chemical shift mapping of an already weak binding ligand (PELLECCHIA, M., *et al*., 2008). Therefore, the site of attachment is marked from the already available outline of weak binder. This NMR study provides a way to produce a high rank lead as results of combination from all chemical linkers (PELLECCHIA, M., *et al*., 2008). SAR by NMR functional approach is schematically represented in figure 9. However, this approach suffers from protein size limit (30-50 KD) a major draw back after labelling. The added further advantages in drug based design approach, by these SAR by NMR methods, therefore, are augmentation in affinity of the ligands and the generation of bi-dentate ligands (BECATTINI, B.; PELLECCHIA, M., 2006). There are some other methods present that help in detection of low affinity ligands in contrast to the HTS approaches where, only strong affinity ligands are detected. There are some other methods present that help in the detection of low affinity ligands in contrast to the HTS approaches where, only strong affinity ligands are detected. In addition, based on some paramagnetically labelled ligands, these approaches also aid in detection of ligand—ligand magnetization transfer through protein (BECATTINI, B., *et al*., 2006).

All these approaches are out of the scope of this thesis, therefore, are mentioned in the table 3 below, to just take an overlook.

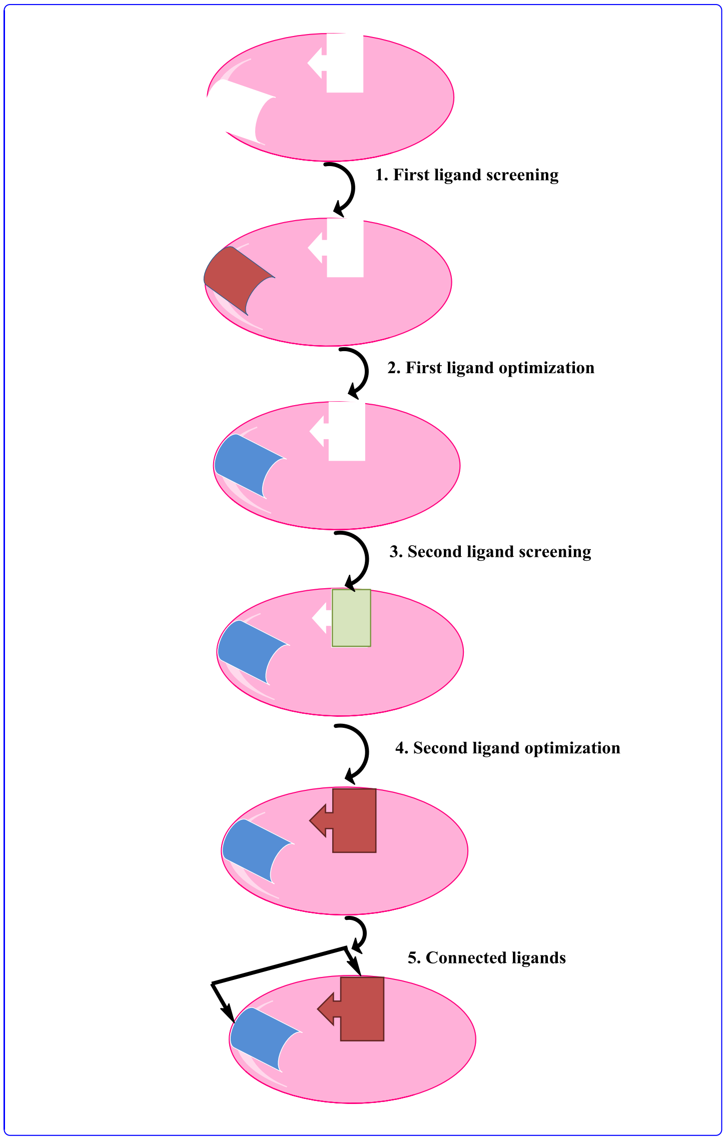


Figure 9. The schematic presentation of the working principle of a SAR by NMR reproduced from (SHUKER, S. B., *et al.*, 1996).

Approach	**Observation**	**Use**	**Description**
SAR by NMR	Ligand	Optimization of compounds on FBDD screening	Designing bi-dentate ligand
SLAPSTIC	Ligand	Optimization of compounds on FBDD screening	Second-site interacting compounds with weakly binding though fragments based detection.
SAR by ILOEs	Ligand-to-ligand	Optimization of compounds on FBDD screening	Detects protein arbitrated ligand–ligand interactions for compounds occupying adjacent sites only
Pharmacophore by ILOEs	Ligand-to-ligand	Optimization of compounds on FBDD screening	For the search of bi-dentate ligand, protein mediated ligand-ligand interaction
H_2O/D_2O exchange-rate measurements	Target	Characterization	Characterize epitope results
INPHARMA	Ligand-to-ligand	Characterization	Compition for same site for protein mediated ligand-ligand interaction

Table 3. Target/ Receptor-based NMR methods for used for lead optimization, reproduced from PELLECCHIA, M., *et al.*, 2008.

1.4.8. Ligand- based screening

Nowadays, there has been a strong emphasized on the ligand based screening methods, because they are neither demanding any isotopic labelling nor some structural knowledge of protein. Therefore, the relaxation properties and hence, associated correlation time are the key factors to understanding the ligand-based interaction that persists because of the large size of protein usually in (10 x KDa) and small ligands up to 1 KDa. Smaller size molecule posses fast tumbling rates and slow relaxation and in contrast, macromolecule posses fast relaxation rate and slow tumbling rate. Moreover, capabilities of selecting low affinity ligands from a huge library make them a method of choice for most routinely interaction experiments. Conversely, high affinity ligands are therefore considered as non - binders because of low off-rate. The well-known ligand-based methods include, T_1 and T_2 relaxation measurements, waterLOGSY, diffusion experiments, saturation transfer difference (STD) NMR, Nuclear Overhauser Effect (NOE) pumping, SALMON (Solvent Accessibility, Ligand binding, and Mapping of ligand Orientation by NMR spectroscopy), transferred-NOE and INPHARMA (Interligand NOE for PHARmacophore MApping). The complete list along description of each method is provided in Table 4 below.

NMR Method	Competition experiments possible?	Epitope mapping	Description
STD NMR	Yes	Yes	Saturation transfer protein→weakly binding ligands GEP and K_D calculation is possible.
WaterLOGSY	Yes	No	Binding of ligands identified through negative NOEs resulting from the water resonance.
SALMON	Yes	Yes	Binding of ligands identified, GEP also possibl through negative NOEs resulting from the water resonance.
Transferred NOEs	No	No	Interaction and binding conformation is possible through cross peaks.
INPHARMA	Yes	Yes	Inter-ligand NOEs arising from protein mediated complex, where the competitive binding ligands can be searched.
SAR by ILOE	No	Yes	Direct inter-ligand NOEs of simultaneous binding ligands and binding constants can be enhanced by chemically linkage of the two ligands.
Pharmacophore by ILOE	No	Yes	Inter-ligand NOEs of simultaneous binding ligands using pharmacophore-based bi-dentate molecules.
FABS	Yes	No	19F chemical shift perturbation of a CF3-group, where the introduction of CF3 through enzyme.
FAXS	No	No	19F chemical shift perturbation of a fluorinated media, where weakly binding reporter molecule upon binding of a high affinity ligand.
T_2 and T_{1rho}relaxation	Yes	Yes	Resonance broadening upon ligandbinding, binding epitope also possible.
Selective T_1relaxation	Yes	Yes	Change longitudinal relaxation rate upon ligand binding can be determind.

Table 4. Summary of Ligand-based methods with complete assay description, reproduced from LUDWIG C.; GUENTHER U. L., 2009.

1.4.8.1 Relaxation experiments (T_1 and T_2) and related methods

In NMR parlance, the relaxation refers to the phenomenon of restoring the equilibrium magnetization and arbitrary phase of the nucleus. Where, the sign T_1 usually mean the longitudinal while the T_2 describe the transverse relaxation. The relaxation experiments based on the tumbling correlation time, internal motion of the molecules under observation—thereby, electric field fluctuations. The transverse relaxation T_2 of the macromolecule is large, because of slow tumbling and long correlation time (τ_c) in solution and hence, large line width in NMR. Therefore, in interaction studies small molecule shows increase in T_2 when in contact with macromolecule (LUDWIG C.; GUENTHER U. L., 2009). Similarly, with or without the addition of protein a comparison is also helpful with the T_2 filter experiments (HAJDUK, P. J., *et al*., 1997).

Another very sensitive relaxation based approach, which recently employed by Segal and co-workers, by using the solid support (sepharose beads or glass) to make protein immobilized in a solution, and difference of spectra (beads with and without protein) were made (SIEGAL, G., *et al*., 2007). The resultant spectrum with line broadening having the signals of ligands only those showed interaction towards protein. They named this method as “Target Immobilized NMR Screening” (TINS). Similarly, SLAPSTIC (Spin Labels Attached to Protein Side chains as a Tool to identify Interacting Compounds) is also based on relaxation method, which determines line broadening of that ligand which present close to the protein (JAHNKE, W., *et al.*, 2001).

1.4.8.2 Diffusion based experiments

Over the past two decades or so, NMR has been using as a tool for mixture analysis by adding up a supplementary dimensions—accounts for the diffusion behavior beside the conventional chemical shift (MORRIS, K. F.; JOHNSON, C. S., 1992; MORRIS, K. F.; JOHNSON, C. S., 1993). As a result of this additional diffusion component, now each molecule depends upon own his diffusion properties that depend upon the some of the physical characteristics of molecules, such as the size, mass, shape, and charge as well as its surrounding environment like temperature, aggregation state, and more importantly the solution where these molecules stay (HUO, R., et al., 2003). Therefore, ones can separate the mixture of different components on the basis of diffusion changes. Thus, this principle can be done by using the pulsed field gradient spin echo (PFG-SE) (STEJSKAL, E. O.; Tanner, J. E., 1965) and pulsed field gradient stimulated echo (PFG-STE) (Gibbs, S. J.; JOHNSON JR, C. S., 1991). The DOSY pulse sequence (ledbpgp2s) most frequently used, within this thesis is shown below in figure 10, in the interaction studies.

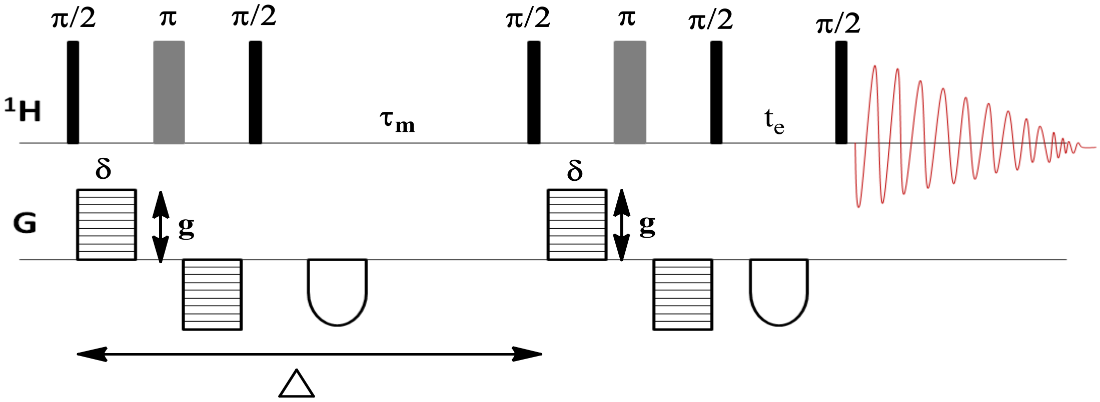


Figure 10. A schematic diagram of the ledbpgp2s (2D-DOSY) pulse sequences, where, π/2, π, and τ_m represent the 90°, 180° and delay time respectively, while, δ and g indicate the length and strength of the gradient used. The shaded rectangular and unfilled crowns are the representation of gradients and spoiler gradient, respectively. The big delta is total diffusion time.

This aforementioned sequence offers advantages over simple LED sequence by using the double gradient strength. In contrast to STE sequence, also helps in reduction of eddy-current, and an extra delay time adds in current settling and restoration of magnetization at the end.

DOSY NMR is non-invasive technique, provides the physical separation that is why few authors called this NMR chromatography (JAYAWICKRAMA, D. A., *et al*., 1998; COLBOURNE, A. A., *et al.*, 2011), and is viable in recognition of weakly bound ligands in a mixture (LIN, M., *et al*., 1997; DERRICK, T. S., *et al*., 2002). As, the diffusion-ordered spectroscopy provides two information at a time; a) chemical shift and b) diffusion coefficient, because of this it is now method of choice with a number of applications in different areas. The most practical examples are; diffusion measurements of large protein complexes (DIDENKO, T., *et al*., 2012), an analysis of herbal mixtures (BALAYSSAC, S., *et al*., 2009), in generic or counterfeit drug analysis (TREFI, S., *et al*., 2008; TREFI, S., *et al*., 2008) analysis of interaction between the ligand-protein (DERRICK, T. S., *et al*., 2002; DINGLEY, A., *et al*., 1995), analysis of polymers and small molecular complex (HODGE, P., *et al*., 2001) etc. Another important expect of DOSY NMR is, tracking the reaction intermediates (LI, D., *et al*., 2009).

1.4.9 Methods involving magnetization transfer

In the past two decades, methods involving the magnetization transfer lie in the heart of the drug discovery, because of the small amount of target use as well as larger applicability to vast complex systems. The binding events, which involve in the ligand-protein complex exploits a number of properties associated with the changes happening through the fast and slow tumbling rates. On the other hand, lot of other assays are being used to determine the K_D values, chemical shift mapping, and molecular weights etc. Therefore, the changes, which affect the physical appearance of the spectra mainly, devote to the exchange system that exists between the free and bound state of the ligands. In contrast, macromolecule, due to the fast relaxation and short tumbling rate show comparatively broad signals and sometimes not any detectable signal in NMR spectra. The possible changes that occur during this complexation can be summarized as shown in scheme (Figure 11).

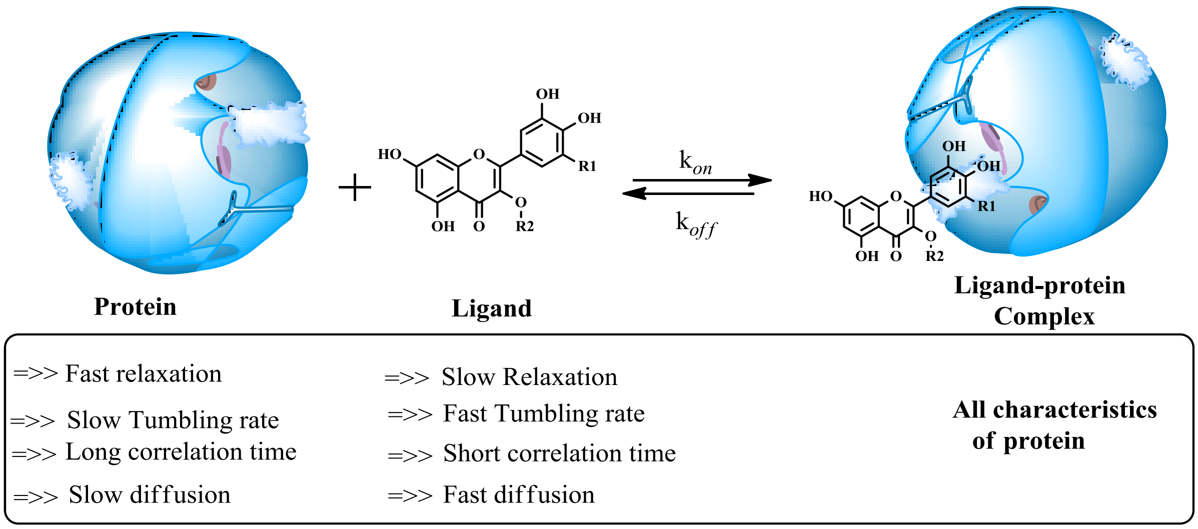


Figure 11. The Schematic representation of the molecular characteristics, which are responsible for recognition process, of small and macromolecule, when they are free and bound state.

The information provided in the scheme (Figure 11), can be applied in both screening mean (ligand or receptor based), therefore, this sequence is a general way of observing ligand-receptor interactions. The relaxation studies and line width are devoted to target-based screening while, magnetization transfer methods rely on ligand-based screening. Here, in the equilibrium state, the k_{off} rate plays an important role in the distinction of the free and bound state of the ligand. Therefore, within the NMR time slice the fast k_{off} rate is the more favorable rate for signal distinction. Usually, two experiments, with the addition and without the addition of protein, are performed to see the changes shown in figure 11. The experiments, which depends on the magnetization transfer include; Saturation Transfer Difference (STD), Tr-NOESY, WaterLOGSY etc., demanding the fast exchange rates within a limited range (millimolar to 1 nanomolar) of dissociation constant (MEYER, B.; PETERS, T., 2003).

1.4.9.1 Transfer Nuclear Overhauser Effect Spectroscopy (Tr-NOESY)

The transfer NOE phenomenon was first established by Botherner-By about four decades ago (BALARAM, P., *et al.*, 1972) but, now it is well-established for protein structural analysis (BRAUN, W., *et al*., 1981; MEYER, B.; PETERS, T., 2003). Upon binding a ligand to protein cavity several changes undergo, that can be determined by NMR. Two types of couplings (scalar or dipolar) arise upon binding to a receptor which can clue about the interaction, in contrast to the relaxation changes. The scalar coupling arises through the bonding electrons, on the other hand, dipolar coupling are as a result of nuclear spins of two groups, adjacent or nearby in space (CAVANAGH, J., 2007). Therefore, the former is responsible for the nuclear assignments and the later one for the magnetization exchange (Ludwig C.; GUENTHER U. L., 2009). The perturbations as the results of this exchange between the different spins system is known as NOE. Directly, this NOE is dependent on the tumbling rates and hence correlation time τ_c of the free and bound small nucleus (MEYER, B.; PETERS, T., 2003). The small molecules (1-2KDa) posses the fast tumbling rates and short correlation time τ_c in contrast to the macromolecules (10 x KDa), which posses long τ_c and slow tumbling rate. Thus, small molecule has small NOE or positive NOE while, macromolecule posses a strong negative (MEYER, B.; PETERS, T., 2003). When a small molecule binds within the binding cavity of large macromolecule then, it adopts the NOE effect of macromolecules and turns to negative NOE signals, this is called Tr-NOE (MEYER, B.; PETERS, T., 2003; (LUDWIG C.; GUENTHER U. L., 2009), this mechanism is quite clear within figure 12. Tr-NOESY and INPHARMA (Interligand NOE for PHARmacophore MApping) method has the same purpose, however, the principle difference therefore, exist in magnetization transfer; in Tr-NOESY it remains on the same molecule, conversely, the magnetization is transferred from one molecule to another through protein, which provides additional cross peaks(SÁNCHEZ-PEDREGAL, *et al.*, 2005).

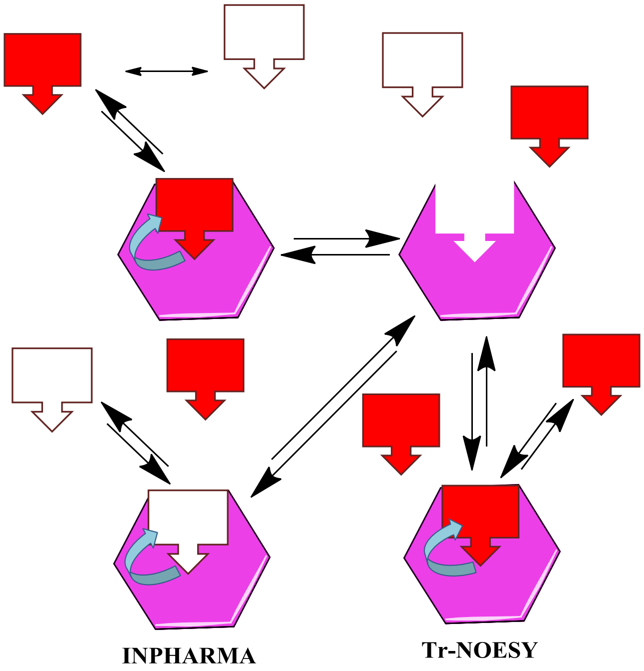


Figure 12. The schematic presentation mechanism of Tr-NOESY, and INPHARMA. The only difference is in Tr-NOESY magnetization is transferred via protein then back to same molecule, while in INPHARMA this magnetization is transferred to another molecule with same binding cavity. Adopted from (LUDWIG C.; GUENTHER U. L., 2009).

Thus, to observe the binding events with respect to ligands structure, it is necessary to perform experiments with and without the addition of protein, that not only helps in binding epitope mapping but also in recognition of bound conformations (NI, F., 1994; POST, C. B., 2003).

NOESY experiment is a very sample pulse sequence (Figure 13) almost similar to DFQ-COSY sequence, however, the difference exists the during the second 90°, in NOESY magnetization is restore on –Z axis while in COSY it shifts to X-axis. Tr-NOESY is best for the larger protein where, sometimes to reduce the background signals T_2 or T_{1rho} filters are used (HENRICHSEN, D., *et al.*, 1999).

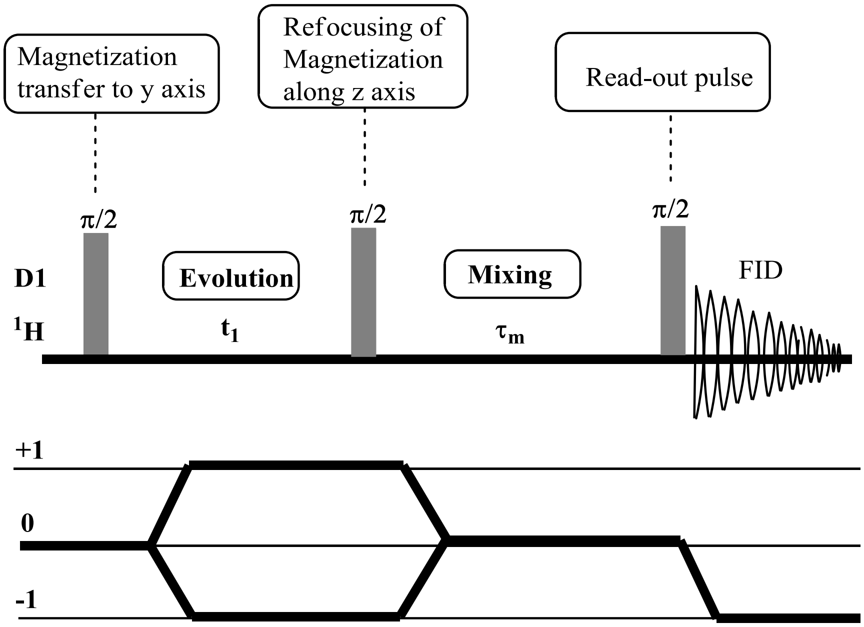


Figure 13. Standard NOESY pulse sequence with associated phase cycling, the sequence is similar to COSY sequence.

Apart from the interaction studies, a number of Tr-NOESY applications with different schemes have been coming forward illustrating the use of this technique quantitative purpose (LONDON, R. E., 1999; NI, F., 1994) which provided even better conformational outputs. Another variant of the Tr-NOESY, is called ILOEs (interligand-NOE), this case the original transfer of magnetization between the two ligands take place especially, where the binding cavity of the protein is too large (LI, D., *et al.*, 1999). This effect is particularly common where the spin diffusion is low as results of short mixing time (BECATTINI, B., *et al.*, 2006).

1.4.9.2 Saturation Transfer Difference (STD) NMR

Bernd Meyer has introduced the STD NMR in 1999, however, in general principle, it is the most robust, and more routinely used techniques for the ligand-based screening method (MAYER, M.; MEYER, B., 1999). The almost sample preparation is required, as that for NOE, however, in use of target amount and time required for completion of the experiment, it is even smarter than NOE approach (LUDWIG C.; GUENTHER U. L., 2009).

The general principle of saturation transfer difference is summarized in figure 14 below. For this purpose, the selective irradiation of protein signal, which normally not overlapping with any ligand signals, is done by selecting the train of soft Gaussian shaped pulses through RF, to saturate the protein signals. This is called on-resonance. This saturation transfer from one proton to others through spin diffusion and consequently, spreads on whole protein, however, depends upon the saturation time (TANOLI, S. A. K., *et al*., 2013; TANOLI, S. A. K., *et al*., 2014; MEYER, B.; PETERS, T., 2003). Then, through cross relaxation this saturation transfers to the ligands in contact with protein (as shown in figure 14). Similarly, the for the off-resonance, a region far from any protein or ligand signal is selected for irradiation to saturate the only protons from the ligands not in contact or freely moving without binding in the solution. The STD NMR is a pseudo 2D sequence so, therefore, two kinds of spectra are the outcome of signal experiments but, can extend to even higher number of spectra (XIA, Y., *et al*., 2010). At the last step during the processing, the subtraction of on-resonance from the off-resonance through phase cycling results the STD difference spectrum—the only signals which involved in binding interactions (Figure 14).

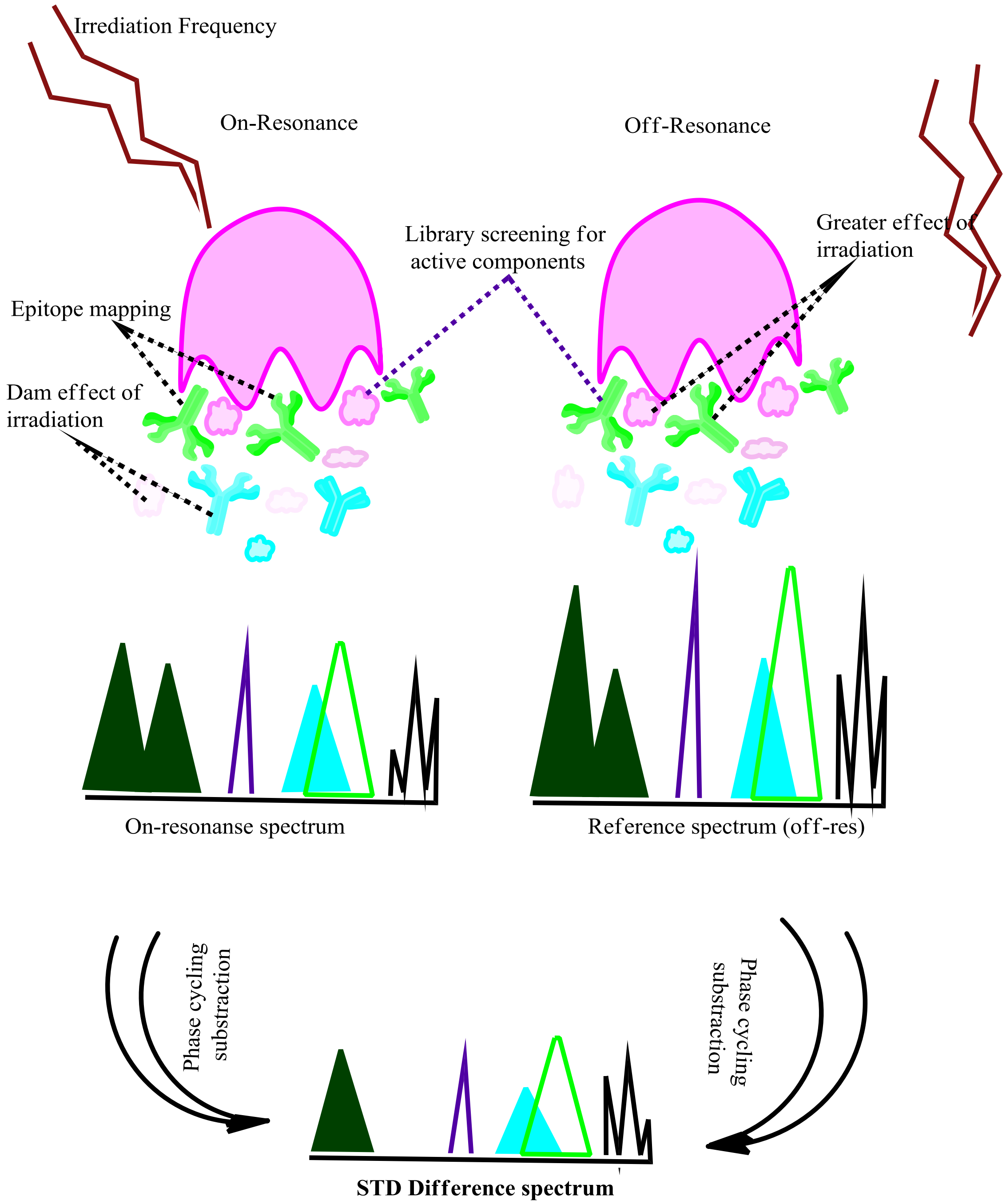


Figure 14. The schematic representation of STD mechanism, and methodology, which can be expresed on molecular basis. The area of contact provides information of epitope mapping, the larger the integration greater will be the contact with protein and vice versa. As, STD is a pseudo 2 D sequence provides, two spectra, then the process phase cycling is necessary to get the STD spectrum

with the signals from interacted molecule. During the phase, cycling on-resonance spectrum is subtracted from off-resonance.

A number of pulse sequences are produced by the spectrometer making companies to work on STD NMR, however, the basic principle is same for all (Figure 15), here two examples are shown A) a general method and B) mostly used to get spectra for the sake of this thesis. In principle, every pulse has the same initial components like, relaxation delay followed by the selective trains (20-100) of soft pulses (Gaussian shape) for the irradiation of protein and the region far from any signal. The next is 90° readout pulse, however, there can be some purging gradients before this readout pulse and the last part of the sequence contains the T_2 filter for the removal of background protein signals (Figure 15 A) and then acquisition. Moreover, for some special cases, this sequence was linked with other solvent suppression sequences to remove the solvent signals e.g., STD with WATERGATE sequence (Figure 15 B), similarly, excitation sculpting is also merging with STD sequence and these provided good results (TANOLI, S. A. K., *et al.*, 2013).

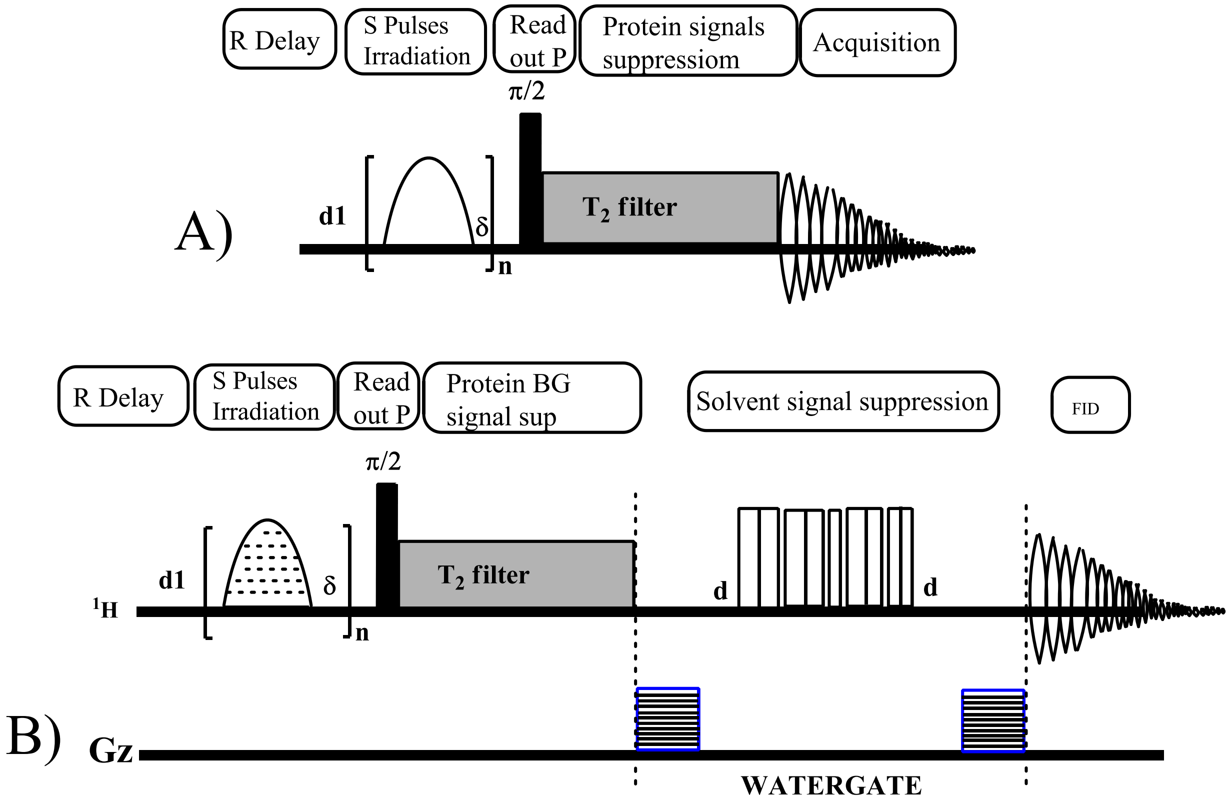


Figure 15. Conventional STD pulse sequence (A) and WATERGATE coupled STD sequence (B) for solvent suppression, (MAYER, M.; MEYER, B., 2001).

The wide applications of STD NMR are the consequences of the simple procedure, no necessary instrumentations, small protein use, no size limits on protein, can also work with the system with chemical exchange etc., (MEYER, B.; PETERS, T., 2003; WAGSTAFF, J. L., *et al.*, 2013). Therefore, a variety of reviews available illustrating the utility of STD in different areas likes, applications fragment-based drug design (BHUNIA, A., *et al.*, 2012), comparison with other methods (LUDWIG C.; GUENTHER U. L., 2009; PENG, J. W., *et al.*, 2004) and alone as robust method (MEYER, B.; PETERS, T., 2003). Apart from this, STD NMR is particularly being utilized for the screening of carbohydrates as illustrated in the review (POVEDA, A.; JIMENEZ-BARBERO, J., 2012).

STD NMR provides a mean for the calculation of dissociation constant of small molecule those bind loosely with protein (MEYER, B.; PETERS, T., 2003). As mentioned earlier in the 1.4.9 section, that is interaction depends on k_{off} rate of the molecule within the complex ligand equilibrium. The ligands those binds too tightly they have a low k_{off} rate hence, NMR can´t detect them, and similarly, the very week binder where the transfer of magnetization is difficult, NMR is also not helpful to screen them. Thus, the dissociation constant limit for STD as illustrated by the Meyer and Peter, is between 1mM to 100 nM (MEYER, B.; PETERS, T., 2003; MAYER, M.; MEYER, B., 2001). Recently, Angulo *et al*. have provided an absolute method for of binding constant calculation, by taking an example of WGA as receptor and saccharides as a ligand (ANGULO, J., *et al.*, 2010).

To further enhance the applications of STD NMR in screening studies, Meyer and co-workers in 2004, another variant know as STDD (saturation transfer double difference) was introduced by making an example of cyclic-peptide as ligand and integrin α_{IIb}ß3 as a receptor directly on living cell (CLAASEN, B., *et al.*, 2004). In this new variant, they used the cells alone STD spectra and cells plus ligands (cyclic peptides) STD spectra then, a double difference were calculated containing only the bound ligands information. Following this method, another article published on marine natural product screening towards G protein coupled receptors CB1 and CB2 (PEREIRA, A., *et al.,* 2009).

STD NMR technique became one of the most utilizing tool for the ligand-based screening methods, for that reason, to discuss a complete application is not possible, therefore, with few example this topic is closed. In a number of articles STD NMR is used as a quantitative tool (JI, Z.; YAO, Z.; LIU, M., 2009; (ANGULO, J., *et al.*, 2010). Moreover, the usefulness of STD has also been applied in two-dimensional NMR, for example, ^{13}C STD-HSQC and in ^{15}N STD-HSQC (WAGSTAFF, J. L., *et al.*, 2010), in DOSY NMR (KRAMER, M.; KLEINPETER, E., 2010), in TOCSY (TANOLI, S. A. K. *et al.,* 2013) etc.

1.4.9.3 WaterLOGSY & SALMON

Water-Ligand Observed via Gradient SpectroscopY (WaterLOGSY) is the variant of STD NMR where, the information about the bound ligands can be obtained through bulk water on the surface of ligand-protein complex. Instead of protein irradiation, the selective pulses for the irradiation have been selecting to water, in contrast to STD NMR. The magnetization transfer can take place through the water in between the ligand and protein or water surrounding the ligand alone (DALVIT, C., *et al*., 2000; DALVIT, C., *et al*., 2001). In order to enhance sensitivity and to reduce the radiation damping effect and solvent suppression, bulk water and application of gradients were employed respectively. The signals of non-binders are turn out to be inverted and binding ligands are remain normal. The magnetization is the transfer of water to protein and then to ligands or directly from water to ligand. The sign of signals can observe the distinction between the interactions in the final spectrum. The direct interaction between ligand and water generate small NOE due to the smaller correlation time of both, on the other hand, the interaction between the ligand and protein provide negative NOE with larger correlation time lead the normal upward signals in the final spectrum (LUDWIG C.; GUENTHER U. L., 2009). For the primary screening for the ligands it is the method of choice, conversely, when the epitope is required STD can help better. Therefore, WaterLOGSY can provide only information whether the ligands are showing binding or not.

Another variant of waterLOGSY which can provide binding as well as the epitope mapping is called SALMON (Solvent Accessibility and protein Ligand binding studied by NMR Spectroscopy). In SALMON, the gradients were applied for the transfer pathway, in contrast to waterLOGSY where, gradient used was for phase cycling. Moreover, in order to control the blurring as a result of massive spin diffusion from solvent, shorter mixing time was applied (LUDWIG, C., *et al.*, 2007; LUDWIG, C., *et al.*, 2008).

1.4.9.4 Experiments containing fluorinated reporter molecule

1.4.9.4.1 FABS & FAXS

Both experiments are demanding the isotopic labelling substrate molecules. FABS (3 Fluorine Atoms for Biochemical Screening) or 3-FABS is helping in controlling the inhibitions and kinetics with the help substrate with a CF_3 group by tracking the chemical shift change with and without reporter molecule (DALVIT, C., *et al.*, 2003). On the other hand, Fluorine chemical shift Anisotropy and eXchange for Screening (FAXS) use the ^{19}flurorine labelled spy molecule (weak binder) and a control (non-binder). The working principle is very simple, when the reported molecule remains in contact with protein then its signal disappear due to larger tumbling from protein and chemical shift anisotropy from ^{19}F. On the other hand, if the new ligand is stronger binder then the spy, it replaces the spy and consequently, spy chemical shift restores.

❷

Aims of the thesis

2. Aims of the thesis

Over the past three decades, there has been expanded accentuation set on drug discovery process to find the new methods that can help in reducing the cost and to cut back of time in the generation of new drugs. Currently, the importance and relevance of protein dynamics for cooperative binding and molecular recognition is a matter of interest for pharmacologist, biologist and chemist. Therefore, there are two major sources to generate a drug a) synthesis and b) natural products. On the other hand, pharmaceutics and chemists are looking for new methods that could help in time cutback used to find the leads and a method that should be environmental friendly and above all should be cost effective. The drug discovery process contains a number of phases, where the applicability of NMR lies in the lead recognition and optimization phase. In this viewpoint, the aim of the experimental work presented in this thesis is the investigation of the molecular recognition between protein and ligands using a combined separational, spectroscopic, and docking approach can be divided into three parts;

- ➢ The development and optimization of ligand-based NMR screening techniques, like the Saturation Transfer Difference (STD) NMR, Transfer Nuclear Overhauser Spectroscopy (tr-NOESY), Diffusion-ordered Spectroscopy, 2D STD-TOCSY etc., in the NMR Laboratory at the Federal University of São Carlos, SP Brazil.

- ➢ To generate a new method that can help in rapid characterization, recognition, and optimization of lead compounds within the natural products or the folk medicines e.g., Barbatimão extract used in this thesis.

- For the screening of large library optimization of these ligand-based NMR methods, which are most robust in their applications and widely applicable to every system, in order to find the binding characterization at atomic level. That in future may serve as a lead compound.

Thus, we have devised most routinely used NMR techniques like, STD, Tr-NOESY and STD-TOCSY in coordination with a hyphenation system that could not only recognized the ligands for natural extract but also their unequivocal structure characterization. This recognition process followed by docking simulation that afforded leads directly from crude. This could be a more economically favorable application for further investigation of natural products.

❸

Materials and methods

3. Materials and methods

3.1 Equipments used

3.1.1 NMR spectrometers

All the practical experiments, except the theoretical calculations, were performed in the laboratory of Nuclear Magnetic Resonance, Department of Chemistry, Federal University of São Carlos, SP Brazil whereas, the theoretical part (Docking simulation) was performed by the Dr. Zaheer Ul-Haq Lab´s and co-workers (Dr. Panjwani Center for Molecular Medicine and Drug Research, International Center for Chemicals and Biological Sciences, University of Karachi, Karachi Pakistan).

Most of the NMR spectra were recorded on a Bruker AVANCE III 600.23 MHz—14.1 T, (hydrogen nucleus) spectrometer at 296-302K temperature, equipped with a 5mm cryo-probe TCI (triple resonance $^{13}C/^{15}N/^{1}H$) having pulse field gradients along z direction, with the gradient strength of 53.5 G/cm and with automatic tuning and matching unit (ATMA). Data acquisition and processing was performed with the Bruker software Topspin 3.0 and 3.2 version installed on the spectrometer.

However, some times Bruker AVANCE III-9.14 T with 400.13 and 100.62 MHz for hydrogen and carbon nucleus respectively, have also been used for recording two-dimensional NOESY spectra. The spectrometer is equipped with multinuclear broadband inverse detection probe (BBI) with 5mm of diameter, and gradients are along the z direction having the maximum strength of 50 G/cm and ATMA unit as well.

3.1.2 LC-SPE-NMR systems

An Agilent 1200 series of High performance liquid chromatography, containing a quaternary pump (model: G1311A) with degasser (model: G1322A), an UV-DAD ultraviolet diode array detector (G1315C SL) and G1315D G1329A auto injector. During the course of this research, columns of different types have been used in quest of required resolution. Therefore, every time different column was tried and the column with best resolution was selected. For the separation of crude extract from the *Rauia resinous* and *Strypnodendron polyphyllum,* Prontosil C18 (250 mm x 2.0 mm i.d, particle size 4 μ m) column whereas, Kromasil C18 (25 cm x 4.6 mm with particle size of 5 µm) for the *Conchocarpus marginatus* was used. Moreover, Hystar 3.2 software, installed as a default on the computer system, was used for controlling all these separation method. The peak collection and then transfer of the separated pure compounds were performed in a similarly fashion to loop storage by using SPE system.

The Prospect2 solid phase extraction system (SPE) (Netherland) contains 192 HySphere resin cartridges (Spark Holland, Emmen, The Netherlands) in two trays, each having the size: 10 x 2 mm. The left hand tray comprises 96 cartridges of general phase resin while the left tray has the same number therefore; with 6 different types of resins in six rows, each row of 12 cartridges. The selection of type of resin according to the nature of sample helps better in trapping and hence in elution. Trapping of peaks on to the cartridges made possible by the use of makeup pump k-120 (Knauer, Berlin, Germany), which used the sample water with three times more flow rates to the HPLC system—reduce the organic phase and more helps in loading more sample. The commercial Agilent 1200 series, which was employed through all these separations is schematically shown (Figure 16) below, to get closer insights on the Agilent working principle.

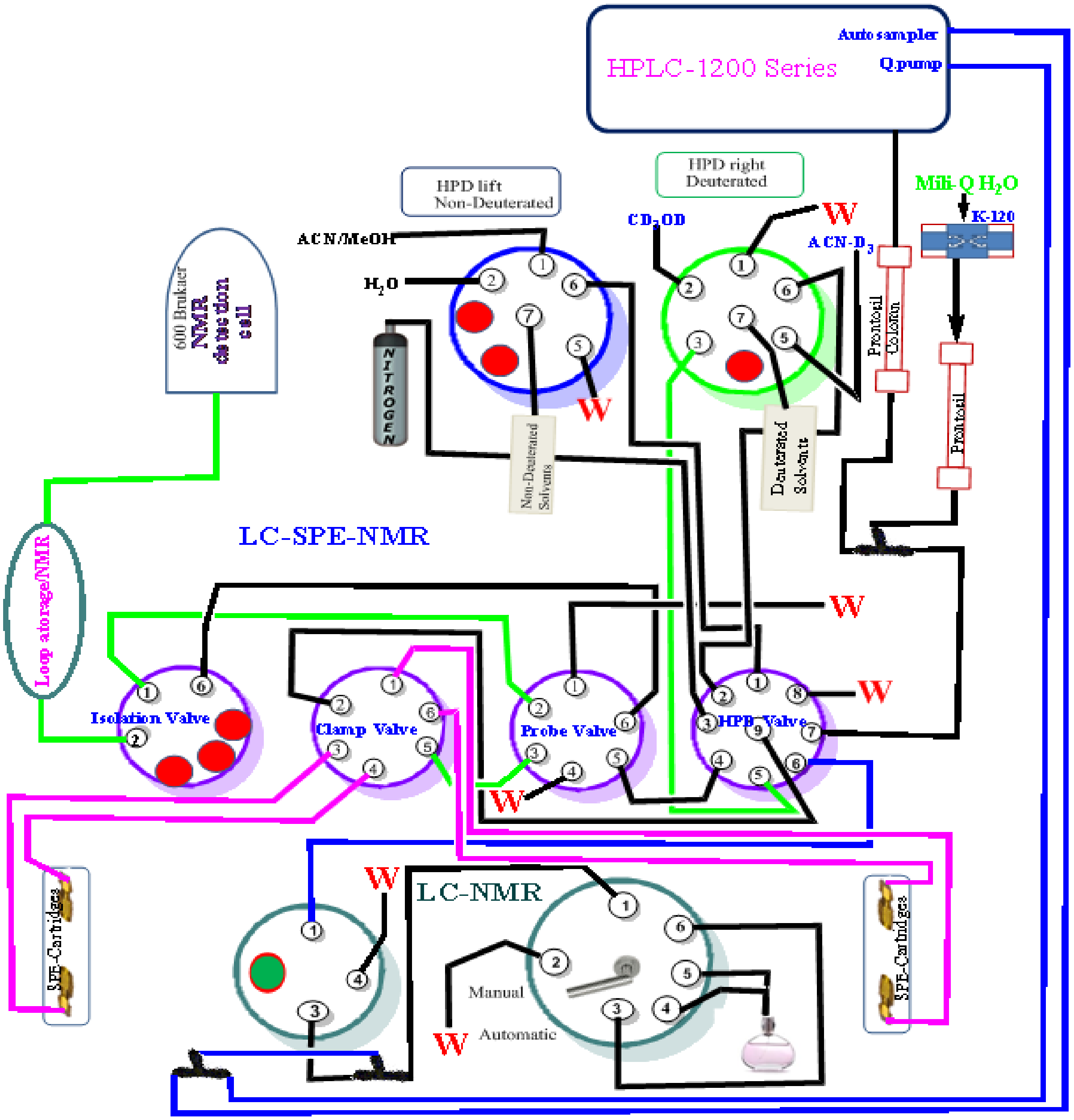


Figure 16. The commercial LC-SPE-NMR system placed in laboratory of NMR, Federal University of São Carlos, SP Brazil. The idea of valve connection was reproduced from (Bruker carporation Natherland)

3.1.3 Analyticel balance

The electronic balance Biopercisa Model (FA 2104N) with weighing limit to 0.01 mg, was used to measure the chemical compounds for the NMR experiments.

3.1.4 Biomixer

Slightly soluble ligands and sparingly soluble extract were shacked through QL-901 Biomixer (Brazil), for complete solubility and to get homogeneity.

3.1.5 Sonicator

Samples and especially HPLC solvents were degassed before use through the Symphony Model sonicator with 35 KHz operating frequency, 1-30 minutes heating time and up to 60 °C heating temperature.

3.2 Chemicals and solvents used

The solvents and commercially avilable chemicals are given in the table 5, whereas, the natural extracts and synthetic compounds were obtained from the natural product and biotransformation groups of UFSCar.

Chemical or solvent	Descriiption	Suplyier
HSA protein	Human Serum Albumin	Sigma (A9511)
BSA Protein	Bovine Serum Albumin	Sigma (A7030)
2-aminopyridine	Alphapyridine	Sigma (504-29-0)
Isovanillin	3-hydroxy-methoxybenzaldehyde	Sigma (121-33-5)
CD_3OD	Deuterated Methanol, 99%	CIL
D_2O	Deuterated water or deuterium oxide, 99%	CIL
DMSO-d_6	Deuterated dimethy sulphoxide, 99%	CIL
Water	HPLC grade	TEDIA (USA)
Methanol	HPLC grade	TEDIA (USA)
NaCl	Sodium salt	Baker
TFA	Trifluoroacetic acid	Fluka
Na_2HPO_4	Sodium monophosphate	Fluka
NaH_2PO_4	Sodium diphosphate	Fluke

Table 5. Chemicals, solvents used in this thesis with complete description of their sources.CIL= Cambridge isotope laboratories.

3.3 Sample preparation

3.3.1 Buffer preparation (PBS)

5 ml of phosphate buffer was prepare using Na_2HPO_4=146.25 mg NaH_2PO_4 = 26.25 mg and NaCl =2.63 mg in D_2O with the final pH = 7.4 (not corrected for D_2O) at room temperature, however, pD of the final buffer can be calculated with the help of the following equation (TENNANT, J. M. *et al.*, 2006).

$$\boldsymbol{pD = pH\ read + 0.40}$$

Where, pD is the power of deuterium ion. Therefore, pD of the final PBS solution is equal to 7.8, afterwards, degassed through sonication up to 2-3 mints.

3.3.2 BSA stock solution

A 50 µM BSA stock solution was prepared by dissolving 10 mg of protein in 3 ml of the D_2O phosphate buffer and solution was stored in fridge, and after the addition of protein the pH of mixture was maitained up to 7.4 (not corrected for D_2O).

3.3.3 HSA stock solution

A 25 µM HSA stock solution was prepared by dissolving 5 mg of protein in 3 ml of the D_2O phosphate buffer and solution was stored in fridge.

3.3.4 Ligand stock solution for worked illustrated in chapter 4

Isovanillin 3.8 mg was dissolved in 5ml (5: 95 v/v DMSO-d_6 and D_2O respectively) of solution (final concentration = 5 mM). Sonication was applied for dissolution.

2-aminopyridine 2.3 mg was dissolved 5 ml (5: 95 v/v DMSO-d_6 and D_2O respectively) of solution (final concentration = 5 mM). Sonication was applied for dissolution.

3.3.4.1 STD-NMR titration for K_D calculation (chapter: 4)

For the calculation of dissociation constant, amount of ligand used for different STD experiments by using 2.5 mm Norel NMR tube by making the 80 µl final volum, can be seen in Table 6.

Ligand Conc mM	**Ligand Excess**	**Volume of Ligand**		**Volume of Protein**	**Volume of Buffer**
		2aminopyridine µL	**Isovanillin µL**	**µL**	**µL**
0.310	10	5	5	50	20
0.500	20	8	8	40	24
0.620	40	10	10	25	35
0.750	60	12	12	20	36
0.875	80	14	14	17.5	34.5
1.000	100	16	16	16	32
1.560	125	25	25	20	10
2.000	200	32	32	16	--

Table 6. The graphical expression for the concentrations of the isovanillin, alphaaminopyridine and BSA protein, used for K_D calculation.

3.3.4.2 Sample for the STD build-up experiment (chapter: 4)

B – 15 µL of 2-aminopyridine and 15 µL of vanillin were added to 15 µL of BSA stock solution and 40 µL of buffer solution so the protein to ligand ratio was1:100.

A similar method of sample preparation was also adopted for the synthetic compounds studied in chapeter 7, where also the conentrations of ligands were maitained between the 5-10 mM, however, D_2O and CD_3OD was utilized for complete dissolution of samples.

3.3.4.3 Sample preparation for plant extract

STD NMR studies were performed with the 50 µM solution of BSA protein in the phosphate buffer pH 7.4 (not corrected for D_2O) in D_2O and 1mg of crude extract (*Rauia resinous*) in D_2O and CD_3OD (95:5% v/v) respectively for the chapter 5.

For the NMR experiments, the 1 mg fully bloomed flower extract of *Strypnodendron polyphyllum* was dissolved in D_2O and CD_3OD, (85: 15 % v/v ratio) while the HSA protein solution was prepared in sodium phosphate buffer (pH 7.2-4, not corrected for D_2O) in D_2O, chapter 6.

3.3.4.4 Sample preparation for LC system

The crude extract (1 mg) was dissolved in HPLC grade methanol (TEDIA USA) brand and water (ultrapure water; MilliQ) in the ratio 7:3 i.e., 700 and 200 µl respectively, then the solution was filtered through 0.45 µm Millipore filter (Tedia PVDF) to ensure sample being injected is clean from any kind of dust (chapter 5).

A 3 mg of the fully bloomed flower extract (*Strypnodendron polyphyllum*) was dissolved in methanol (HPLC grade, TEDIA USA) and water (ultrapure water; MilliQ) with 85:15 v/v %. Afterwards, the corresponding solution was filtered to promise the dust free sample through a 0.045 µm Millipore filter (Tedia PVDF) (Chapter 6).

3.3 NMR acquisition and processing

3.4.1 ^{1}H-NMR spectrum

Locked the spectrometer on respective solvent after inserting the sample inside the magnet and then shimming to get the homogeneity of the magnetic field around the sample. Adjusted the pulse program to zg and then first tried to get solvent-probe parameters by typing (getprosol). Then, after adjusting the important parameter such that spectral width, data points, relaxation delay, number of scans etc, and so on—obtained the receiver gain and then gave the first acquisition to see the spectrum.

3.4.2 90° Pulse calibration

Started from the small duration of the hard pulse, performed a series of experiments to get the null point means no signal in the spectrum—towards the Pulse calibration. The null point represents the 180°, divided this duration by 2 to get 90° pulse, which gave the signal with good intensity. To verify, repeated this procedure two or more times with multiplying the null point pulse duration by 2 and then dividing by four and so on. The acquisition after the 90° pulse calibration provided the ^{1}H-NMR with good intensity. However, in some cases, to save the time just used the automatic calibration command 'paropt' or 'pulsecal sn'.

3.4.3 Solvent suppression

New file for the solvent suppression was generated from the previously pulse calibrated and properly shimmed ^{1}H-NMR—to get rid of adjusting all parameters and especially offset. A numbers of pulse sequences have been applied and optimized for this thesis, therefore, the most important ones are; zgcppr (Figure 17) lc1pnf2 (Figure 18), p3919gp (Figure 19), excitation sculpting using a pulse train 'zggpw5' (Figure 20) etc.

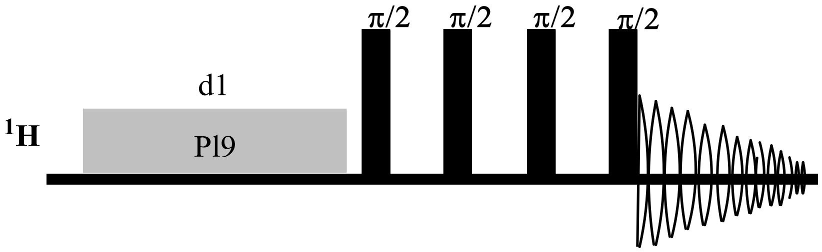


Figure 17. Schematic representations of 'zgcppr' pulse sequence, where solvent suprression was performed by using composite pulses (BAX, A., 1985).

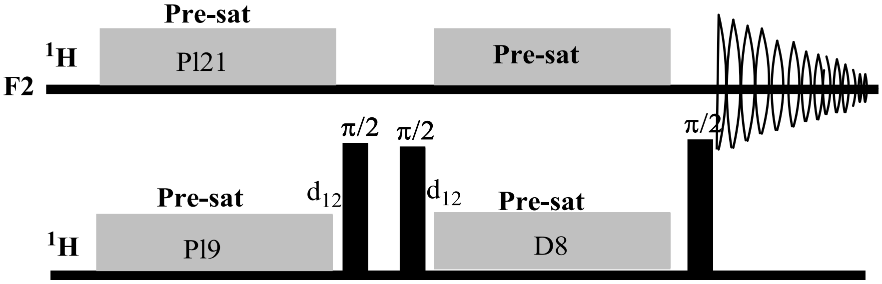


Figure 18. Schematic representation of 1D-NOESY based double pre-suppression by using two channels and two pulse power for presaturation (PARELLA, T., 2004).

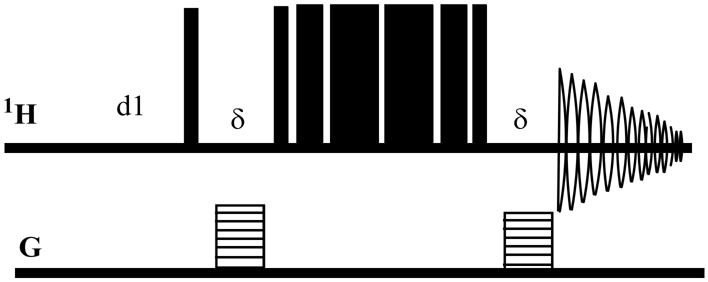


Figure 19. Schematic representation of 1D gradient based-Watergate using 3-9-19 scheme, (PIOTTO, M.; SAUDEK, V.; SKLENÁŘ, V., 1982) is shown.

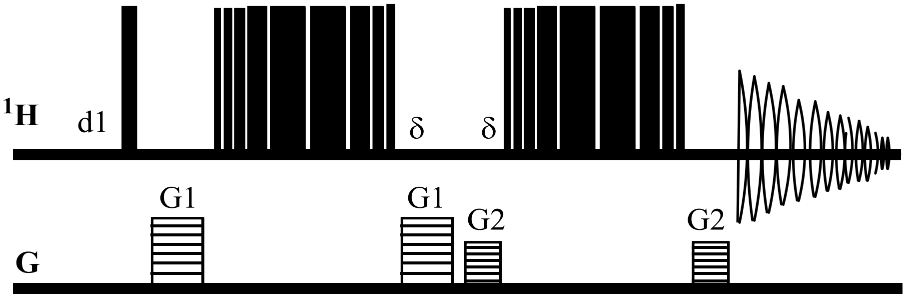


Figure 20. Schematic representation of 1D excitation sculpting using W5 pulse train, (LIU, M. *et al.*, 1998) is shown.

The best suppression of the sample can be obtained by increasing the power of the composite pulses, or mixing time in case of 1D NOESY sequence. Generally, the mixing time is pre-defined (d8=80 ms) however, the power for 'pl9' and 'pl21' was kept between the 43-50dB (31.62-100 W), both sequences were extensively used for suppression of HOD signals of the compounds obtained, after separation from LC methods. On the other hand, while using the sequence with gradients the define shape and power was adjusted as given elsewhere (BAUER, C., *et al*., 1984), and performs several experiments in measuring the delay for binomial water suppression 'd19' in case of Watergate sequence, where, 240 µs delay was found perfect that value used afterwards to obtain clear spectra.

3.4.4 Saturation Transfer Difference (STD) NMR

Saturation transfer difference NMR studies for performed by using a numbers of pulse sequence and with little modification of pulses according to the requirements of the experiments. Moreover, in all experiments there were at least two solvents involved therefore, in order to remove the solvents peaks Bruker STD pulse sequence with water suppressions are prefered all time. However, two sequences with water suppression were emplyoid extensively; STD coupled with WATERGATE sequence (Figure 15 B) and STD coupled with excitation sculpting (Figure 21) below.

In all STD experiments around 50 µM solution of BSA or HSA protein in PBS buffer pH 7.4 (not corrected for D_2O) has been used with varying concentration of ligands (chapter 4& 7) and different extracts (chapter 5 and 6). In order to get STD build-up rate, a series of STD spectra were acquired at different pre-saturation times (0.5-5 s) and relaxation delay of 1 s, with a selective train of Gaussian shaped soft pulses having a truncation of 1%, 50 ms of length and separated with 2 ms. The saturation time of ca. 2.0, 2.5 and 3.0 s (for chapter 4, 5 and 6 respectively) found to be more efficient for the STD studies in this case after calculating the STD amplification factors at different saturation times (0.5-5 s). In all spectra, selective irradiation was done at -0.5 ppm for the on-resonance and 30 ppm for the *off-resonance* (reference or STD control) with 20-40 ms of spin-lock filter for the protein signal suppression. Subtraction of spectra (*on-resonance* from the *off-resonance*) was done with phase cycling.

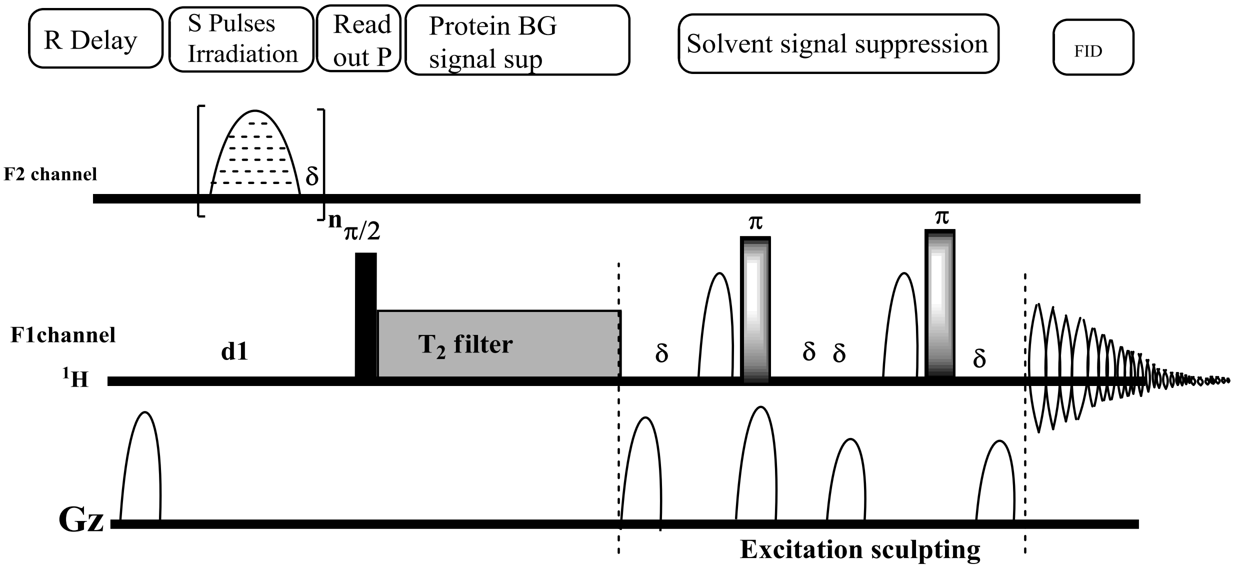


Figure 21. Conventional STD pulse sequence coupled with excitation sculpting for solvent suppression (MAYER, M.; MEYER, B., 2001).

The number of scans for each experiments are different from one another, therefore, for complete experimental details see the reference (TANOLI, S. A. K., *et al.*, 2013 & TANOLI, S. A. K., *et al.*, 2014) and in table 7 below.

Experimental conditions	
Parameter for STD with water suppression	**Values**
Instrument	600.23 MHz (^{1}H Nucleus)
Solvent	CD_3OD/D_2OD; DMSO-d_6/D_2O
Spectral width	Variable (10-60)ppm
On- resonance frequency	-0.5ppm
Off- resonance frequency	30ppm
saturation time	Variable (0.5-5) s
shaped pulse for saturation	50ms
spinlock time	20-40ms
Gradient pulse	3ms
relexatiion delay	1-3sec
delay for homospoil/gradient recovery	200µsec
power of shaped pulse for saturation	45-55 dB (31.22-316.22 W)
td2	32-64K
LB	0-1Hz
NS	16-256
Temperature	296-302K
td1	2
O1p(ppm)	At solvent freequency

Table 7. Acquisition parameters for the STD coupled with solvent supression techniques.

3.4.5 NOESY and Tr-NOESY

Gradient Phase-sensitive NOESY (nuclear Overhauser effect spectroscopy) spectra were acquired in aqueous solution (pH 7.4 in the presence of protein) D_2O and CD_3OD (95:5% v/v) for chapter 5, (85: 15%v/v) for chapter 6 for extract and sodium phosphate buffer in D_2O, pH 7.4 (not corrected for D_2O)for protein with a standard Bruker pulse sequence. Transfer NOESY with a spin lock filter of approximately 3.5 kHz after the first 90° pulse to remove the background protein signals. A relaxation delay of 2 s (chapter 5) and 3 s (chapter 6) allowed full relaxation between scans and the HDO signal was suppressed using low-power pre-irradiation. For the NOESY spectra of the hydroethanolic extract of rauia specie and *Strypnodendron polyphyllum* in the presence of protein, 16 transients were recorded in t_2 with 4K data points and 256 increments in t_1. For processing prior to Furrier transform, a line broadening function of 0.30 Hz was added in both F1 and F2 axis respectively. Processing consisted of apodization with a squared cosine window function, followed by zero-filling and Fourier transformation 4K (t_2) x 1K (t_1) data matrices were obtained. NOESY buildup experiments were performed with two different temperatures (304 and 298K), with different mixing times (100, 200, 300,400, and 500ms), and with two different magnets (400 and 600 MHz). But the measurements with 304 K temperature had problems of signal overlapping, similarly mixing times (100, 200, 300 and 500) and with 400MHz magnet resolutions and intensities were not good enough (data not shown here) (chapter 5). So, in both cases (with or without protein) mixing times were set for 400 ms at 600 MHz with temperature 298 K. Both spectra were acquired with the same spectral conditions, 7211.5 Hz spectral width in both dimensions (F_1 and F_2) with 16 transients in t_2 and with 256 increments in t_1 with a total of 2 hours and 57 min experimental time for the completion of each spectrum.

3.4.6 2D TOCSY & STD-TOCSY

Two-dimensional STD total correlation spectroscopy (TOCSY) spectrum was acquired with the same aqueous solution with protein and phosphate buffer, pH 7.4 (not corrected for D_2O) that used for in the 2D-NOESY experiments (for both chapter 5, 6 and 7). Spectrum was recorded with 7211.5 Hz spectral width in both dimensions (F_1 and F_2) with 64 transients in t_2 and with 384 interlaced mode increments for the *on* and *off* resonance in t_1 with the total experimental time of more than 16 hrs. The on-resonance saturation frequency was set to -0.5 ppm and off-resonance to 30 ppm. Saturation of the protein was achieved with the cascade of 50 Gaussian pulses having 50 ms of duration each with a delay of 200 μs in between; with a total saturation time of 2.5 s. Phase sensitive standard pulse sequence (Bruker) with Malcolm Levitt's composite-pulse decoupling sequence (MLEV-17), mixing time of 60 ms was applied to get the STD-TOCSY spectrum. For standard TOCSY MLEV-17 sequence, the spectral width was kept same as that for STD-TOCSY with a total number of 56 scans per increment with 256 increments in t_1 were collected with 2 s relaxation delay for each scan with the experimental time less than 10 hrs. Subtraction of both spectra was done after the identically processing and phasing of both. Prior to the Fourier-transform 2D spectrum was multiplied with 90°-shifted squared sine bells in both (F_1 and F_2) dimensions and zero-filled with the data point of 4K x 1K in t_2 and t_1 respectively.

3.4.7 Diffusion Ordered Spectroscopy (DOSY) studies

Bruker standard LED pulse sequence was utilized for these diffusion edited studies. The gradient length used was of 1 ms with the recovery times of 200 μs for these gradients, and longitudinal eddy-current delay of 5 ms. A total of sixty-four different gradient amplitudes were used for each (with and without protein addition) experiment with the optimized diffusion time was 0.08 s. In both cases (with and without receptor), the acquisition data were collected with gradient strengths starting from 1.07 G/cm to 53.5 G/cm with 256 numbers of scans to observe this interactions. In every case, the relaxation delay kept to 3 s with the pre-scan delay of 10 μs, and automatically baseline correction.

Processing of 2D DOSY spectra with and/or without the addition of protein were performed with the help of Topspin 3.0, by the line broadening function 1Hz and 0.03 Hz along the F_2 and F_1 dimensions respectively. The DOSY macros were run by using 64K data points and with noise sensitivity factor of 20. The Gaussian apodization function of 0.01 Hz was put only in the F_1 dimension. Zero filling up to 4K points along the F_2 and 1K points along the F_1 were employed in both cases for the 2D spectra before performing Fourier transformation. The resultants DOSY were pseudo two-dimensional spectra with NMR chemical shifts (in ppm) along one axis and calculated diffusion coefficients (in log m^2/s) along the other. Comparison of diffusion coefficients in the presence or absence of protein was measured by taking the water signal as standard, all acquisition parameters are given in the table 8.

Acquisition Parameters	Values
Instrument	600.23MHz for ^{1}H
Solvent	DMSO-d_6-D_2O
Pulse sequence	ledbpgp2s (chapter 4)
Gradient pulse (little DELTA)	1000-1400µs
Diffusion time (big DELTA) D20	80msec
gpz6 variable in 1D	100%
gpz7	-17.13
gpz8	-13.17
gpnam6	SMSQ10.100
gpnam7	SMSQ10.100
gpnam8	SMSQ10.100
Relaxation delay	3sec
Noise sensitivity factor	20
Delay for gradient recovery (d16)	200µsec
DS	04
td2	64k
LB	0.30
NS	256
Temperature	298K
td1	8

Table 8. General acquisition parameters for 2D-DOSY experiments used to obtained spectra shown in chapter 4.

3.5 Execution of separation methods

3.5.1 HPLC conditions used in LC-SPE-NMR separation

The crude extract (1 mg) was dissolved in HPLC grade methanol (TEDIA USA) brand and water (ultrapure water; MilliQ) in the ratio 7:3 i.e., 700 and 300 µl respectively, then the solution was filtered through 0.45 µm Millipore filter (Tedia PVDF) to ensure sample being injected is clean from any kind of dust. However, 3 mg of the extract used in chapter 6, was disolved 85:15 % v/v methanol and water respectively. The 1 ml filtered sample loaded to Agilent 1200 series for the separation. The LC separation was done using a column of Prontosil C18 (250 mm x 2.0 mm i.d, particle size 4 µ m) sampling 10µl of sample i-e 0.01mg/ injection at a flow rate of 1ml/min with methanol and water as mobile phase. The method and gradient solvent system used for separation can be seen in supporting information. To attain the attenuation 0.05 % of triflouroacetic acid was used in both solvents. Detection of compounds was done by DAD starting from 210 to 360 nm with five different wavelengths (210, 254, 280, 330 and 360 nm). We could find 6 and 13 peaks from the *Rauia resinous* and *Stepnodendron polyphyllum* extract respectively. HPLC methods used for both extract are presented in Table 9 and 10 for *Rauia resinous* and *Stepnodendron polyphyllum* respectively.

Time min	**Solvents gradient**	
	Methanol	**Water**
0.00	26	74
15	26	74
43	37	63
53	50	50
55	100	-----
60	100	-----
63	26	74
70	26	74

Table 9. The separation method used for the *Rauia resinous* fraction in LC-SPE-NMR.

Time min	**Solvents gradient**	
	Methanol	**Water**
0.00	25	75
8	50	75
15	75	25
20	80	20
24	100	-----
27	100	-----
29	25	75
35	25	75

Table 10. The separation method used for the *Strypnodendron polyphyllum* fraction in LC-SPE-NMR.

The above described HPLC conditions provided 6 and 13 peaks from method shown in table 9 and 10 respectively. Thus, out of 5 compounds trapped from rauia extract 3 are fully recognized and conversely, from the second *Strypnodendron polyphyllum* only 4 the larger peaks are trapped to get NMR spectra.

3.5.2 Solid Phase Extraction (SPE) conditions

The extraction was performed on the Prospect2 SPE system (Netherland) after detection of sample with DAD. Where the dilution of the sample was done with the help of a makeup pump provided by the flow of pure water (3ml/min flow rate) added to the postcolumn. The dilute solution (sample with excess of water) was passed through the HySphere resin GP cartridge (10 x 2 mm) each having the single separated peak (single compound) from HPLC. After 15 and 30 times of trapping (rauia and polyphyllum respectively) the cartridges having the sample were dried by flushing N_2 supplied from nitrogen separator for 40 minutes to ensure the residual solvents have been completely removed. Deuterated methanol was then used for each single dry analyte to elute.

3.6 Theoratical experiments

3.6.1 Molecular docking

Docking simulation studies on a small set of compounds against bovine serum albumin (BSA) and Human Serum Albumin (HSA) were conducted by using Molecular Operating Environment (MOE) 2011.10 to validate the binding affinity and interacting mode of ligands within the target's binding cavity. All structures were constructed using ChemDraw, converted into 3D form by babel program incorporated in OpenEye. Atom type correction, charge application and minimization of ligands were done on MOE 2011.10. On the other hand, protein (BSA and HSA) was retrieved from protein data bank (PDB) with the entry code

4F5S and 4L8U. BSA protein is homodiamer, chain B was extracted and prepared for docking. Protein preparation was started with its correction using auto-correction option in MOE. Furthermore, protonated and minimized structure used as receptor input for docking. Active site information was extracted from the literature (JIN, X.-L., *et al*., 2012) that used to construct a map around active site residues to indicate the binding area for new molecules. Conversely, Human Serum Albumin protein contains two distinct binding sites with amino acids; Tyr150, Glu153, Lys195, Lys199, Arg218, Arg222, His242, Arg257, Leu284, Ser287, Ala291 and Arg114, Leu115, Val116, Arg117, Ile142, Tyr138, His146, Tyr161, Asp183, Leu185, Arg186, Lys190, Ser193 for site-1 and site-2, respectively. Default MOE docking parameters were used for docking with the Triangle Matcher algorithm. London dG was selected as rescoring function 1 and GBVI/WSA dG as rescoring function 2 with the generation of 10, 20 and 30 conformations, for compounds in chapter 4, 5 and 6 respectively, of each ligand to fit better within the binding pocket. The docking results were clustered in a mdb output file that were then analyzed for evaluation of ligand-protein interactions and their binding affinity.

4

The exploration of interaction studies of smaller size, mostly ignored yet intrinsically inestimable molecules towards BSA; an example of STD and DOSY NMR

4. The exploration of interaction studies of smaller size, mostly ignored yet intrinsically inestimable molecules towards BSA; an example of STD and DOSY NMR

4.1 Abstract:

Larger size or novel structure molecules always appreciated by all fields of experimental and computational science. Conversely, molecules with smaller size and simple structures are usually ignored with no explanation for why. However, the vast majority of more diminutive molecules behaved as a cornerstone in the synthesis of the bigger structural framework. Subsequently, we planned to uncover the interactions of small molecules towards macromolecules, and successfully presented the bindings results of 2-aminopyridine, isovanillin towards BSA through NMR techniques. STD epitope mapping and also the DOSY results provided evidence that isovanillin remained closer to the binding cavity of protein. Titration experiments afforded 584 µM (0.584 mM) and 487 µM (0.487 mM) dissociation constant for isovanillin and 2-aminopyridine respectively. Furthermore, changes of diffusion coefficient (with and without protein addition in DOSY spectra) were found 0.081 log (m^2/s) and 0.096 log (m^2/s) points for isovanillin and 2-aminopyridine respectively. Docking study exhibits that these molecules can tie to site 1 (sub-area IIA) through the pi-pi interaction and hydrogen bonding with Trp213. Effective results demonstrated here that both compounds could be utilized as a part of as a transporter in the circulatory system and their extension-inspired compounds may be utilized to make the new drug design.

4.2 Introduction

In drug discovery, long history of NMR guarantees its use in a variety of experiments for portraying and distinguishing ligand-based communication. Previously, the harnessing of NMR's capability was just for the molecular structure mapping, however, with the disclosure of new high-throughput screening strategy, it comes to be more delicate and functional device for hit-to-lead drug finding far superior, then, X-ray crystallography in the utilization of small to modest amount of sample for library screening (PELLECCHIA, M., *et al*., 2008). By consolidating the structural and additionally functional information of ligand inhibitions towards protein, NMR has an edge to the next spectroscopic and non-spectroscopic strategies. The protein activity hinges on upon ligands interaction with it that form the cascade of complicated functions within the body. Therefore, the ligands-protein interaction studies are the utmost importance for understanding the functional complexities of the body at atomic level (MAYER, M.; MEYER, B., 2003). The potential capability of NMR to uncover the interactions between the macromolecules and small organic molecules in drug discovery has proceeded in nineties century, however, is thrived in the most recent decade.

In NMR, ligands and in addition protein, both structures experience perturbations inside the strong magnetic field. Consequently, absolutely NMR strategies, utilized for interaction studies divided into two sub-groups based on perturbation observed, a) receptor-based screening and b) ligands-based screening. Normally, the receptor-based screening strategies demand the isotopically label proteins, then again, longer experimental time. Nevertheless, even more generally applicable, the most overlooked and an extremely emphasized method for the screening purpose is a so-called ligand-based technique. Utilization of minor to moderate amount of protein in ligand-based screening makes it even more routinely utilized method as a part of hit-to-lead optimizations (MAYER, M.; MEYER, B., 1999); (MAYER, M.; MEYER, B., 2001). In the course of the last numerous years, a variety of ligand-based NMR strategies have been established to depict the binding affinity of many biological active substances towards protein. Such techniques incorporate STD NMR (MAYER, M.; MEYER, B., 1999), diffusion experiments (HAJDUK, P. J., *et al.*, 1997), NOE pumping (CHEN, A.; SHAPIRO, M., 1998), Tr-NOESY (HENRICHSEN, D., *et al*., 1999), waterlogsy (DALVIT, C., *et al*., 2000), SALMON (LUDWIG, C., *et al*., 2007) and INPHARMA (SÁNCHEZ-PEDREGAL, V. M., *et al*., 2005) and so forth.

The Saturation transfer difference NMR is sensitive to the sub-atomic motions in free and additionally the bound state that is the reason the perturbation created by the radio frequencies (RF) give the information about the free or bound state of a ligand to a protein. Nowadays, STD NMR is most well known, and cornerstone method for NMR based screening. It demands an extremely low protein concentration without isotopically label, with no previous knowledge of its structure and above all, there is no constrained for protein size. It is likewise plausible to allocate the binding orientation of the ligands in a huge library; purported as group epitope mapping (MAYER, M.; MEYER, B., 2001). Very recently, A. G. Ferreira and co-workers (TANOLI, S. A. K., *et al*., 2013) distinguished the strong ligands from the natural extract with the application of saturation transfer difference NMR together with LC-SPE-NMR hyphenation framework; making NMR, an immediate device to be connected to drug discovery process. In this context, one may argue that if the compounds are known inhibitors and/ or potent then, can be used for specific targets, but by no means STD NMR infers the bio-activity of the compounds itself.

Moreover, Diffusion order-spectroscopy (DOSY) NMR has been accounted for to be viable in distinguishing weakly bound ligands in a mixture (LIN, M., *et al*., 1997); (DERRICK, T. S., *et al*., 2002). Since, DOSY phenomenon dependent upon molecular anisotropy; weight, size and shape of the given molecule, is otherwise called NMR chromatography by a few authors (JAYAWICKRAMA, D. A., *et al.*, 1998); (COLBOURNE, A. A., *et al.*, 2011), as it focused on the diffusion based spectroscopic partition (not physically). Imperatively, it is non-invasive technique (GOUNARIDEs, J. S., *et al.*, 1999) in nature and does not disturb the aggregation equilibrium. Despite that, the compound chemical shift sensitivity, a NMR specific characteristic that permits observing the diffusion changes by the addition of macromolecule in a mixture. Thus, it additionally makes simplicity in discrimination of the free ligands from the bound ones.

Usually, interaction studies have been performed on substantial molecular weight compounds or compounds with uncommon skeleton (novel). Despite that, there are still large number of bioactive compounds which are modest in size and basic in structure, however remain ignored may be due to the simplicity of accessibility or easily synthesized. We have chosen to tackle such a small biodynamic molecules, which are usually neglected for no reason. For our studies, we have chosen to track the interaction of 2-aminopyridine and isovanillin towards the Bovine Serum Albumin (BSA) protein. Both compounds posses a number of biological activities remained unable to make importance they merit for, as being a lead candidate. Isovanillin has recently been demonstrated antimicrobial (S. GOPALAKRISHNAN, E. V., 2011), antibacterial (FRIEDMAN, M., *et al.*, 2003) activities and as a competent d substrate of xanthenes oxidase (PANOUTSOPOULOS, G., *et al.*, 2004); (PANOUTSOPOULOS, G., *et al.*, 2004a), and powerful inhibitor of aldehyde oxidase (PANOUTSOPOULOS, G. I.; BEEDHAM, C., 2004). Aldehyde oxidase is a metabolizing compound whose activity has generally prevail in the liver (SASAKI, K., *et al.*, 1983; BEEDHAM, C., 1987; BEEDHAM, C., *et al.*, 1987 and plays an important role in metabolizing numerous N-heterocyclic and aldehydic drugs (PELSY, G, *et al.*, 1983;

BEEDHAM, C., 1885). On the other hand, 2-aminopyridine is utilized within the generation of numerous drugs, and a known inhibitor of Nitric Oxide Synthase (NOS) (BLUHM, U., *et al*., 2009; CONNOLLY, S., *et al*., 2004).

4.3 Results and discussion:

As a rule, scientists, organic chemists and biologists enraptured by a bigger and complicated structural compound like kibdelomycin, vancomycin etc., for the bioactivities or inhibition recognized useful for interaction studies. Conversely, the smaller size structure, those come to be starting point of most drugs are generally oversighted. Moreover, the modest size compound like sugars; amino acids are convincingly picking up much importance by playing additional imperative role in diverse disease curing and also in the infection aetiology by smoothly moving in the body fluid and straightforward approach to the desired destination after being ingested. Thus, keeping these favourable points of interest of the littler and simple structure and in addition the bioactivities (S. GOPALAKRISHNAN, E. V., 2011; FRIEDMAN, M., *et al*., 2003; PANOUTSOPOULOS, G., *et al*., 2004; SASAKI, K., *et al*., 1983; PELSY, G, *et al*., 1983; BEEDHAM, C., 1885; BLUHM, U., *et al*., 2009; CONNOLLY, S., *et al*., 2004) of the selected molecules (2-aminopyridine and isovanillin) pulled in us for focus on the interaction studies.

4.3.1 STD NMR studies

Nowadays, large numbers of high-throughput screening methods are, undoubtedly being utilized for the screening of huge libraries joining the spectroscopic, non-spectroscopic and computational methodologies simultaneously. Consequently, we got the advantage of STD NMR and diffusion ordered-spectroscopy (DOSY) to describe the binding of 2-aminopyridine and isovanillin to bovine serum albumin. STD-NMR analyses were done by utilizing 1:100 ratio protein to ligands concentration respectively, at 25°C in D_2O, buffer (having pH 7.4) and 5% DMSO-*d*$_6$. A freshly prepared solution of ligands and in addition the protein was utilized as a part of this characterization, and sample was irradiated at -0.5 ppm to accomplish the protein saturation (*on*-resonance) and at 30 ppm unbound ligands (*off*-resonance). The signal presence in the STD NMR spectrum is a sensible proof of the binding of compounds to protein as indicated in Figure 22. Conversely, the signal absence in STD NMR spectrum reveals the particular ligand is not suitable as inhibitor for this protein. In the present study, we displayed the technique for tying of ligands to serum albumin. The epitope mapping by STD-NMR portrayed here describes to the atomic binding level of ligands (2-aminopyridine and isovanillin) with BSA. The result of STD spectrum demonstrates that the signal shows clear interaction of both structures (2-aminopyridine and isovanillin) to serum albumin.

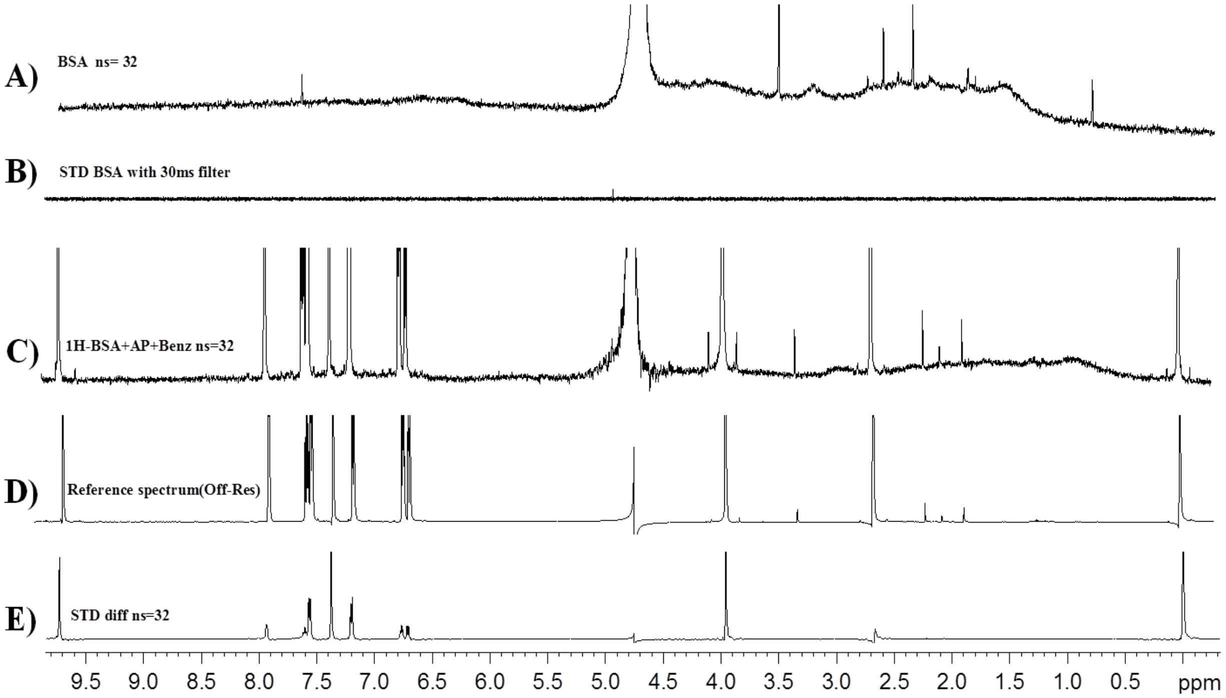


Figure 22. A) The top, ^{1}H-NMR spectrum of 50 µM BSA in phosphate buffer with the pH 7.4 (not corrected for D_2O)solution, showing a characteristic broad

signal obtained by using 3 mm NMR tube. B) STD Spectrum of protein with 30 ms of Spinlock filter, showed a complete removal of the protein signal from the spectrum. C) A ^{1}H-NMR spectrum of 50 µM BSA with 5mM of each, 2-aminopyridine and isovanillin. D) A reference spectrum (*off*-resonance) obtained with a cascade of 40 Gaussian shaped soft pulses, 2 seconds of saturation time and by applying 30 ms of spin lock filter (T_{1p}). For STD spectra, *on*-resonance saturation pulse was placed at −0.5 ppm, while the *off*-resonance at 30 ppm. E) A STD difference spectrum obtained by internally subtraction via phase cycling. All above spectra were produced on Bruker 600MHz AVANCE III spectrometer with a cryogenic TCI probe at 298K temperature.

4.3.2 Group Epitope Mapping (GEM) studies

In the STD NMR spectrum the signals from isovanillin (4-Hydroxy-3-methoxybenzaldehyde) were readily recognizable, which showed that this is an active ligand, whereas the little signals of 2-aminopyridine along these lines, is less dynamic compound. The methoxyl signal from isovanillin provided a larger integral value and was given a 100% STD effect, and whatever remains of the signals were given a relative STD effect to methoxyl signal as indicated in the Figure 23. Indeed, the methoxyl group occupies closest approach to the binding cavity of the protein, thus, accepted more saturation transfer from protein because of cross relaxation. Similarly, the second large integral values was discovered for an aldehydic proton of isovanillin with 73% STD effect, guaranteed a closer contact likewise. The ring protons (H-2 and H-5 of isovanillin) contributed almost equivalent STD Effect with 50 % and 51 % respectively, while the proton H-6 of isovanillin indicated a 63% STD effect.

Conversely, 2-aminopyridine provided significantly less intense signs, thus, less interaction with the protein. The greatest STD impact was seen from the proton number H-4 of the 2-aminopyridine with the 37 % effect, while the proton number H-5 and H-6 showed up with equivalent effect of 15 %, deliberately showing less contact to binding cavity of protein. Whereas, the proton number H-3 remained further a long way from the contact surface with the STD effect of 13 % and hence, more solvent exposed. In the example, which we displayed here, the STD saturation time of ca. 2 s discovered to be more efficient after evaluation of series of experiments with varying saturation times from 0.5 s to 5 s, and consequently further experiment including STD titration experiments were performed with this presaturation time. On the bases of screening outcome presented in Figure 22 and Figure 23 it is evident that the isovanillin bindings all the way tightly to protein while, the 2-aminopyridine is more towards the solvent side (a long way from protein's binding cavity).

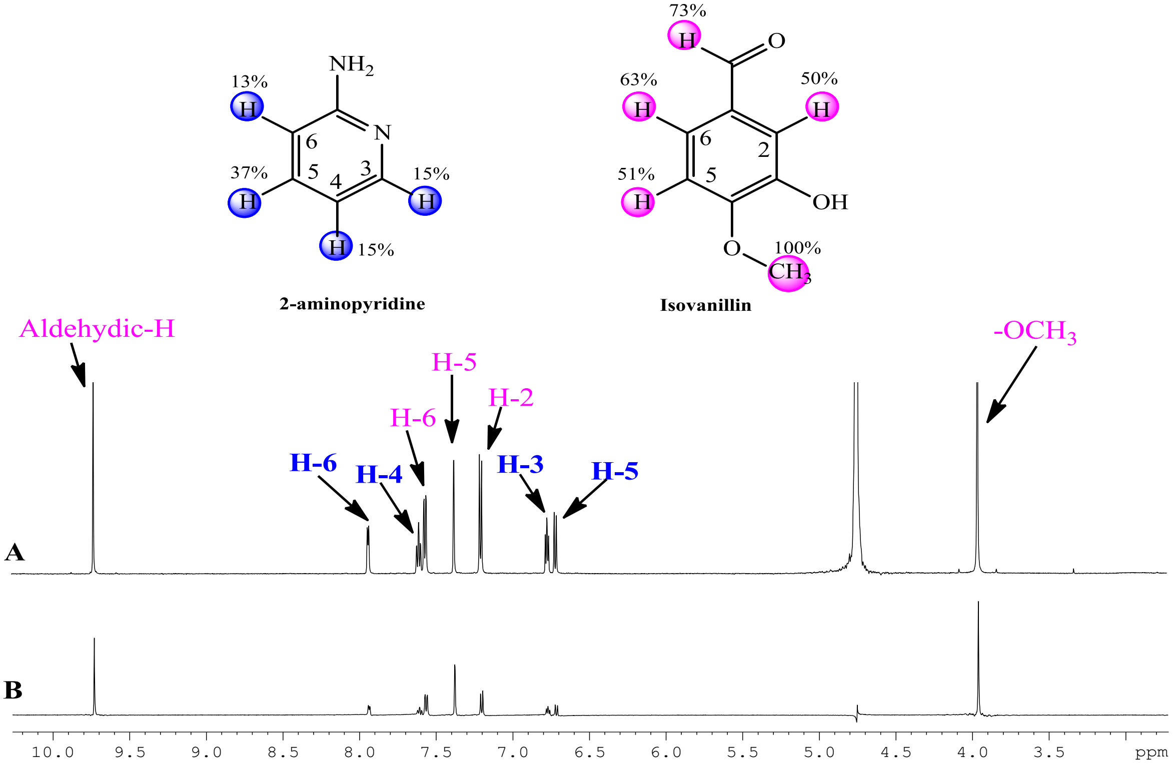


Figure 23. STD-NMR of 2-aminopyridine and isovanillin with BSA: (A) reference ^{1}H spectrum of 2-aminopyridine and isovanillin with BSA and (B) STD difference spectra of 2-aminopyridine (5mM) and isovanillin (5mM) with BSA (50 µM in

phosphate buffer with pH 7.4, not corrected for D_2O). The relative STD amplification factors as calculated from the STD difference spectrum are shown on the top of the spectra.

4.3.3 STD build-up studies

In the STD, build-up experiments (varying STD saturation time from 0.5 s-5 s) were performed for explicit evaluation of the STD amplification factor. By definition, the STD amplification factor is the fractional saturation received appropriated from the protein for a given proton multiplied by the molar excess of the ligand over the protein, as indicated below in a mathematical equation 1.

$$\boldsymbol{A_{STD} = \frac{I_{\circ}-I_{STD}}{I_{\circ}}\ X\ \frac{[L]}{[P]}} \qquad \textbf{(1)}$$

Where A_{STD} is an enhancement element of the given proton (I_{STD}/I^{o}), and [L] and [P] is given concentrations of particular ligands and protein separately. Indeed, the amplification factor is the power of the STD signal (average number of molecules) saturated for every atom of protein.

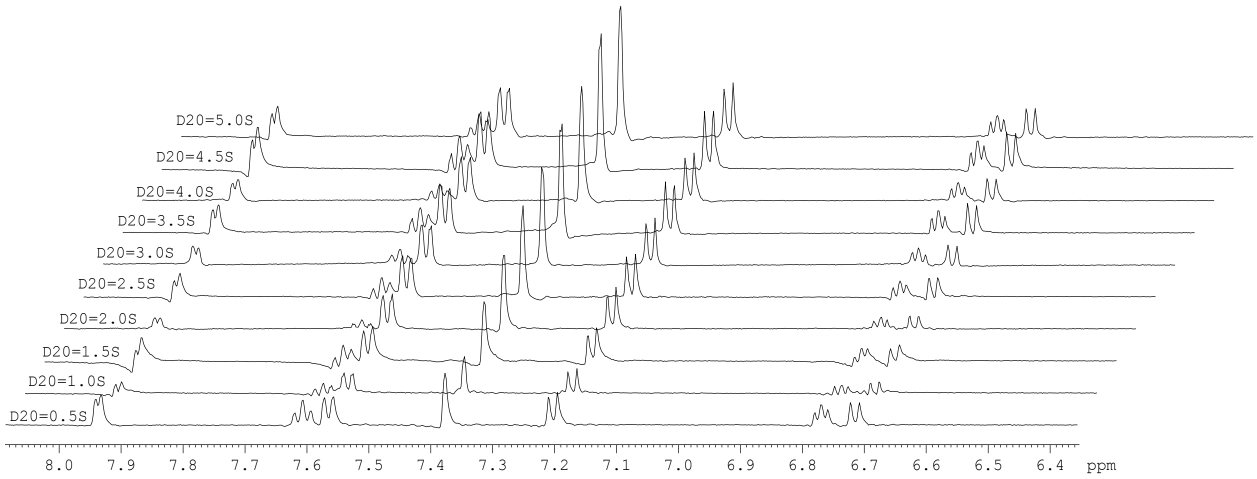
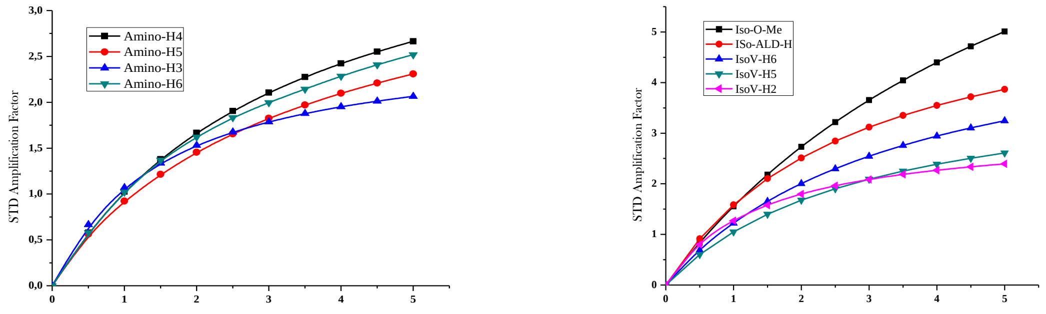


Figure 24. Stack plot of STD-NMR (2-aminopyridine and isovanillin with BSA) spectra at varying saturation times (0.5 s to 5.0 s). The STD-NMR spectra were performed with 50 µM BSA (~66.5-KDa) in phosphate buffer pH 7.4 (not corrected for D_2O) solution, with 5 mM solution of 2-aminopyridine and isovanillin each in 5% DMSO with 95% D_2O. Further details can be found in experimental section.

In principle, more extended saturation time or more ligand utilized, answerable for the biggest STD signal intensity (A_{STD}) and higher turnover rates. The confirmation about the saturation time came to be clear when different STD spectra were acquired by varying saturation times (0.5 s to 5.0 s) (Figure 24), and A_{STD} plotted against these saturation times (Figure 25a-b). A_{STD} came to be exceptionally imperative by quantitative expects at different saturation time, as the A_{STD} of giving hydrogen is very different from the other one, clarifying the proximity to the protein surface. From the saturation profile as demonstrated in figure 25 a-b, it is self-evident, at maximum saturation time; the receptor sites become saturated and no more ligands saturation take place results in a flattened intensity curve.



A) **B)**

Figure 25a-b. STD amplification factors plot of 2-aminopyridine (A) and isovanillin (B) as a function of saturation time, as a result of integral values obtained from different STD spectra (shown in Figure 24). Methoxyl signal with the largest intensity of isovanillin was given the 100% STD effect and rest signals were normalized with reference to this signal.

4.3.4 STD titration studies

Similarly, the increase in ligand concentration likewise is answerable for larger STD intensities; until the saturation of all binding sites of the protein. A_{STD} is also supportive for titration ponders. Dissociation constant value of these ligands (2-aminopyridine and isovanillin) might be computed with the assistance of the following mathematical equation 2.

$$A_{STD} = \frac{\alpha_{STD}[L]}{K_D+[L]} \tag{2}$$

Where K_D represents to dissociation constant, and α_{std} to maximum amplification factor, calculated at a given saturation time. For the K_D constant examination, we expected three things; 1) system (ligands and protein) is rapidly and reversibly exchanging, 2) all the binding interaction has same equilibrium and 3) binding surface (binding sites) of protein is independent. Unique STD examinations were performed with varying ligands concentration (see the experimental section) while, the saturation time for all experiments were kept same and was equivalent to 2 s. From the stack plot (see Figure 29 and 30 supporting information) it came to be clear that A_{STD} in addition got higher with an increase in ligands concentration. A parabolic curve obtained when the amplification factor (A_{STD}) was plotted against the different concentration of ligands, as demonstrated in Figure 26. The K_D values as estimated from equation 2 by taking help of Microsoft excel (see Figure 29 and 30 supporting information) were 584 µM (0.584mM) and 487 µM (0.487 mM) for the isovanillin and 2-aminopyridine respectively.

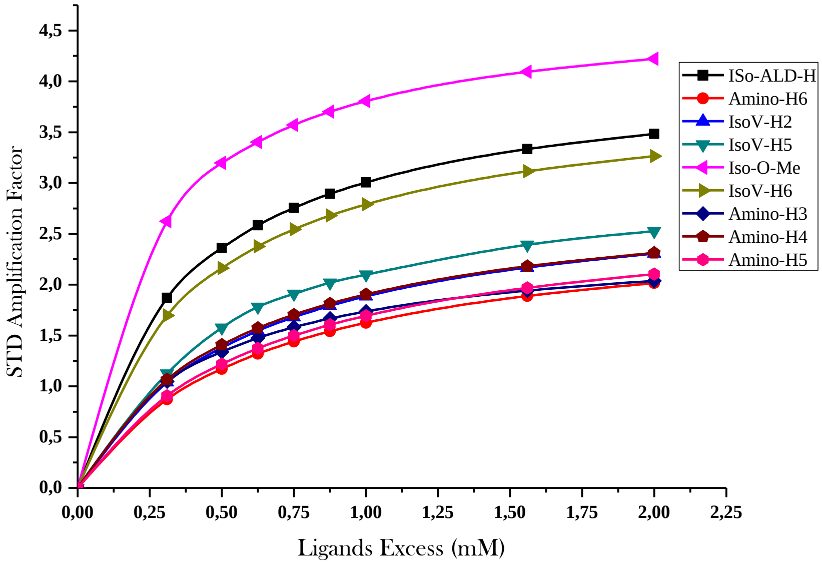


Figure 26. STD amplification factor curve plot of 2-aminopyridine and isovanillin as a function of varying concentrations ranging from 0.25 mM to 2.00 mM for each single parabolic line that determined from the STD titration spectra as shown in Figure 29 and 30 supporting information.

4.3.5 Diffusion ordered spectroscopy studies

Furthermore, we have also performed the diffusion-ordered spectroscopy of the selected molecules with (Figure 27) and without (see Figure 31 supporting information) the addition of protein. A quick inspection of 2D DOSY spectrum showed that both compounds diffusion coefficient was laid between -9.0 log m^2/s to -9.8 log m^2/s, as can be seen in the spectrum in figure 27. The calculation of the diffusion is very easy, as NMR chemical shifts (in ppm) stays in 2D DOSY spectrum along one axis and figured diffusion coefficients (in log m^2/s) along the other. By gathering, all straight-line signals indicate the diffusion coefficient of a one compound. Thus, the top five signals in a single line (Figure 27) represented to the diffusion coefficient of isovanillin on diffusion axis and chemical shift on separate axis, as shown from the 1D spectrum in Figure 22. Similarly, the second four signals in one line were acknowledged to 2-aminopyridine atom. Since, it is well understood that, the molecule having affinity towards the macromolecule, shows the change in diffusion coefficient with and without the addition of the macromolecule (LIN, M., *et al*., 1997); (DERRICK, T. S., *et al*., 2002). By comparing both spectra (with and without protein) it is quite clear that both molecules indicated a change in diffusion. Isovanillin in the presence of the protein showed a a -9.75 log (m^2/s) while in the absence -9.50 log (m^2/s) (see Figure 31 supporting information), moreover, 2-aminopyridine -9.643 log (m^2/s) (Figure 27) and -9.365 log (m^2/s) in the presence and absence of protein respectively. Adjusting HOD signal at -9.28 log (m^2/s) as reference for both spectra the diffusion coefficient of Isovanillin indicated 0.081 points while, the 2-aminopyridine showed 0.096 points change (see Figure 31 supporting information) in diffusion by addition of protein (Figure 27). Moreover, the average error, as calculated from DOSY Toolbox (Nilsson, M., 2009) for the mixture with and without protein was ± 0.029 and ± 0.031 respective, which is quite far from total change in diffusion in both cases. These effects prompted the conclusion that both molecules were involved in interaction, but the compound with little structure or less atomic weight indicated a bigger diffusion change with the addition of protein.

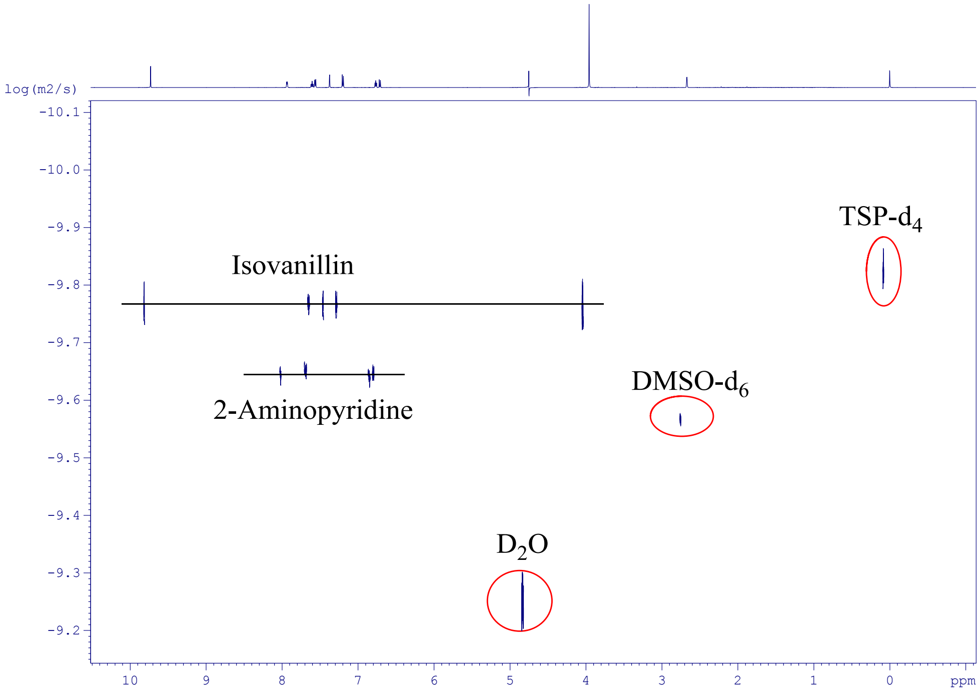


Figure 27: 2D-DOSY NMR spectrum of 2-aminopyridine (5 mM) and isovanillin (5 mM) with BSA (50 µM), where the x-axis contains the standard ^{1}H-NMR chemical shifts, while the y-axis contains the diffusion dimension. Spectrum produced as a result of total diffusion time (Δ) of 80 ms, 5 ms of eddy current, and 1000 µs of small delta (δ). The area in a circle represents the diffusion coefficient of reference solvents (D_2O, DMSO and TSPA-d_4) while, the solid straight lines represent the molecules of interests.

4.3.6 Docking studies

Further consolidation of our results, we got the help of docking simulation studies. In principle, docking is a strategy that ties the ligand through diverse interaction within the binding pocket of the target protein. To perform the docking studies, we have chosen the sub-space IIA of site 1 to see the different conformational adaptation with lowest possible ligand-protein complex energies for these compounds, as illustrated in the experimental section. Docking results uncovered the fact that, both compounds were small in size, and could only occupy a part of the active site and interacted with residues. Isovanillin, being a small molecule, interacted with only Trp213 through pi-pi interaction (Figure 32 supporting information) leaving all other residues vacant (Figure 28a). Correspondingly, 2-aminopyridine is likewise a small molecule cooperated progressively because of amino aggregation and upheld both hydrophilic and hydrophobic sorts of interactions. Arg217 donate its bond to NH_2 whereas Asp450 received hydrogen bond from NH_2 and Trp213 favored π-π interactions (Figure 33 supporting information). Molecular Operating Environment (MOE) 2011.1 stacked up these molecules (2-aminopyridine and isovanillin) as an active ligand, however being a smaller size they could not fill the cavity properly. Based on this study provided through MOE, we inferred that the small molecules interacted with BSA through hydrophobic (π-π interaction) and in addition hydrophilic (hydrogen bonding) interaction that stabilized this ligands-protein complex.

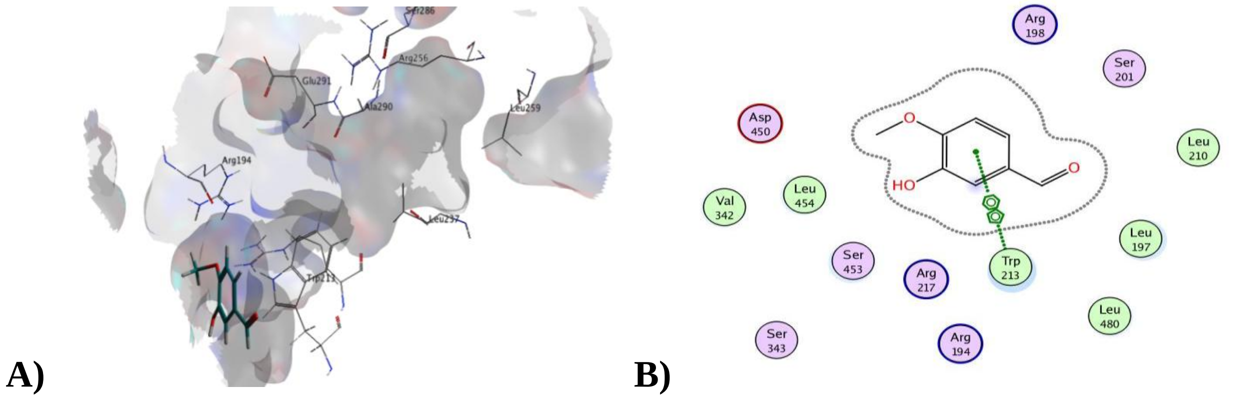


Figure 28a-b. Molecular Docking model of interactions between the Bovine Serum Albumin (BSA) and isovanillin, generated by using Molecular Operating Environment (MOE) 2011.10. The BSA PDB file was taken from protein data bank (http://www.rcsb.org/pdb/explore.do?structureId=4F5S) (see Figure 34 supporting information). The classical binding sites were marked in corresponding subdomain cations. See more details in the experimental section. Figure 28b representing a 2D scheme for interaction between the BSA binding sites and isovanillin molecules. Where, the π-π interactions between the isovanillin and Trp213 residues of BSA binding sites are shown with green dashed lines.

4.4 Conclusion

We have presented an STD NMR method that permitted the identification of high-affinity ligand based on fast exchange process (turnover rates) on the NMR time-scale. This study was represented with reference to the BSA (Bovine Serum Albumin: a model protein) –ligands complex framework. Affinity towards the receptor might be promptly understood from the ligands signal intensities in the STD difference spectrum. Larger the signal in STD spectrum, implies a greater affinity to the specific targets and thus, stronger will be the saturation gained from the receptor. In this study, we utilized Bovine Serum Albumin as a model, however this approach could be connected to different targets, and it used to uncover the approximate value of the dissociation constant of dynamic ligands. By utilizing the low protein to ligands concentration (50 µM: 5 mM respectively) we have successfully demonstrated that the isovanillin is a stronger ligand than 2-aminopyridine, as it is clear from the signal intensities (Figure 23). The estimated dissociation constant values as a result of STD NMR at different concentrations ranked isovanillin a potential ligand compared with 2-aminopyridine with the corresponding values 584 µM and 487 µM respectively. Moreover, diffusion ordered spectroscopy provided the same outcome about as acquired from larger changes in the diffusion coefficient value of isovanillin as well. Likewise, the docking simulation also discovered to be in magnificent concurrence with these NMR results; announced isovanillin as potential ligands that can binds through hydrogen bonding and additionally by means of π-π interaction. Effective STD NMR, DOSY NMR and docking comes about demonstrated above, depicted that, both compounds may be utilized as a transporter as a part of the circulatory system. Thus, on the premise of the combined results from the NMR techniques and docking simulation, there is no obvious reason for why that these smaller compounds to be ignored in the drug discovery.

4.5 Supporting information

The exploration of interaction studies of smaller size, mostly ignored yet intrinsically inestimable molecules towards BSA; an example of STD and DOSY NMR

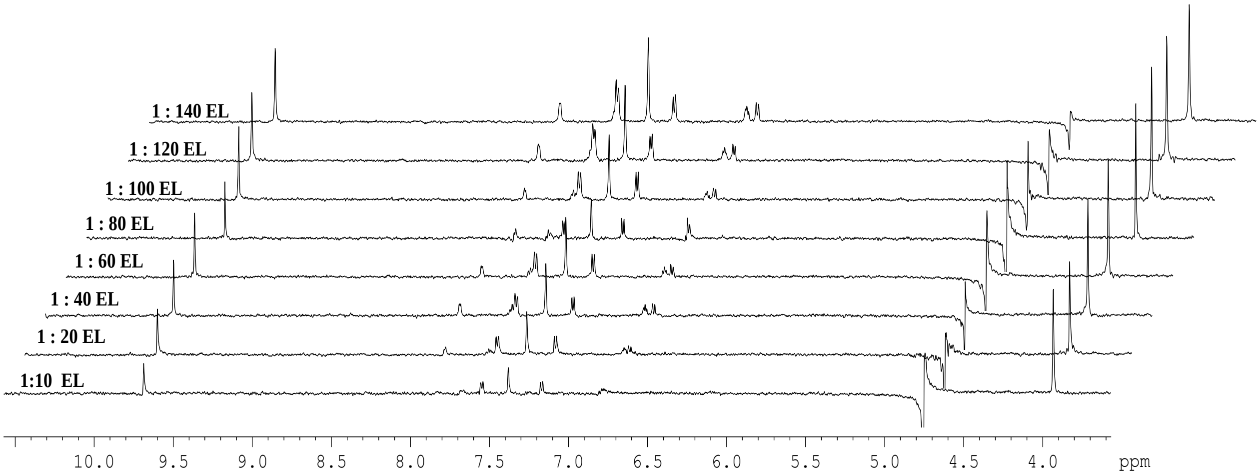


Figure 29. Stack plot of STD-NMR spectra (2-aminopyridine and isovanillin with BSA) at different protein vs ligands concentrations (1:10 - 1:140) of the ligands (2-aminopyridine and isovanillin). For all STD NMR experiments saturation time was kept constant (2.04 s) with cascade of 40 Gaussian shape soft pulses for the irradiation of samples, where BSA dissolved in phosphate buffer pH 7.4 solution while, 2-aminopyridine and isovanillin each with 5% DMSO in 95% D_2O. Further details can be found in the experimental section.

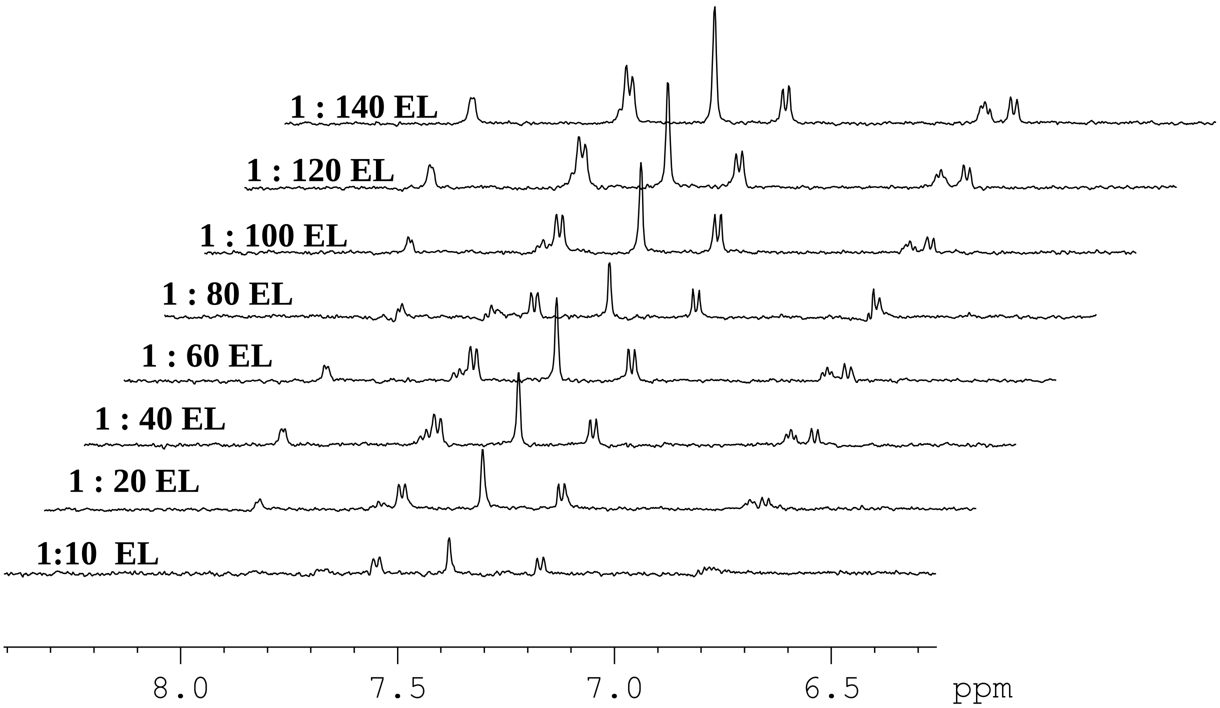


Figure 30. Magnified stack plot of STD-NMR spectra (2-aminopyridine and isovanillin with BSA) between the 6.3 - 8.4ppm regions. Where, all experiments were performed at different protein: ligands (2-aminopyridine and isovanillin) concentrations raging between 1:10 -1:140 folds. For all STD NMR experiments saturation time was kept constant (2.04 s) with cascade of 40 Gaussian shape soft pulses for the irradiation of samples, where BSA dissolved in phosphate buffer pH 7.4 solution while, 2-aminopyridine and isovanillin each with 5% DMSO in 95% D_2O. Further details can be found in the experimental section.

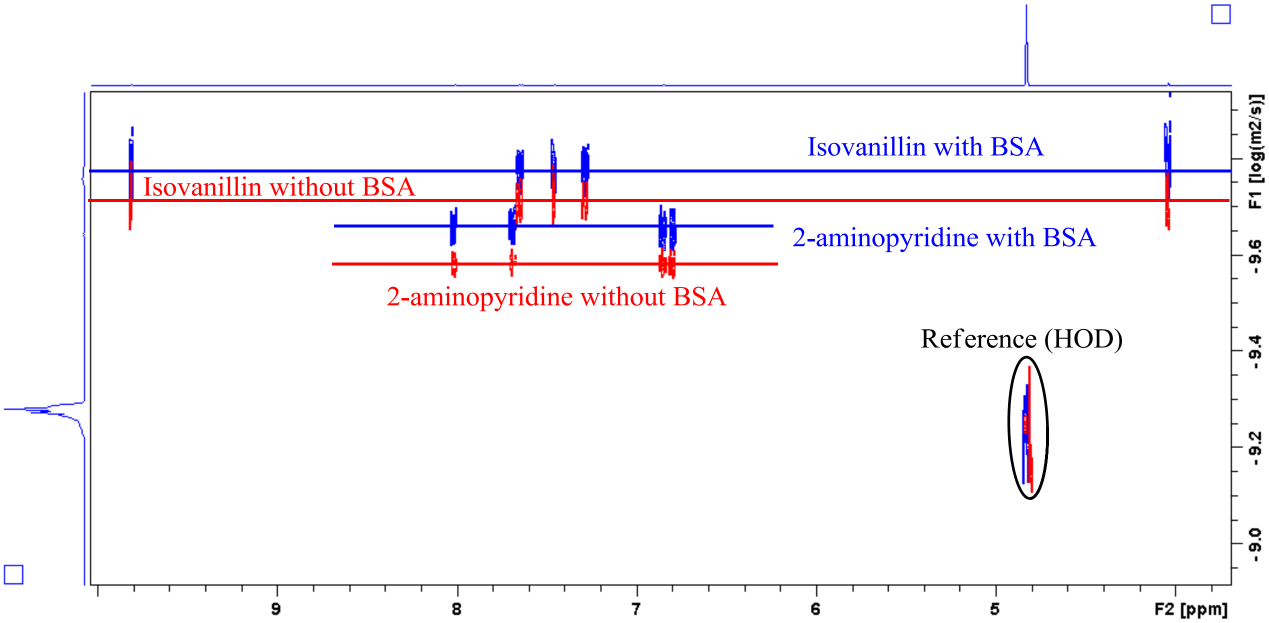


Figure 31. Comparison of 2D-DOSY NMR spectra with (Blue color) and without (red color) of protein and 2-aminopyridine and isovanillin mixture, where the x-axis contains the standard ^{1}H-NMR chemical shifts, while the y-axis contains the diffusion dimension. Spectrum produced as a result of total diffusion time (Δ) of 80 ms, 5 ms of eddy current, and 1000 µs of small delta (δ). The area represented the red solid lines straight as well as signals are the diffusion coefficients without protein, and the blue solid lines with respective blue signals, with the addition of protein for the molecules of interests.

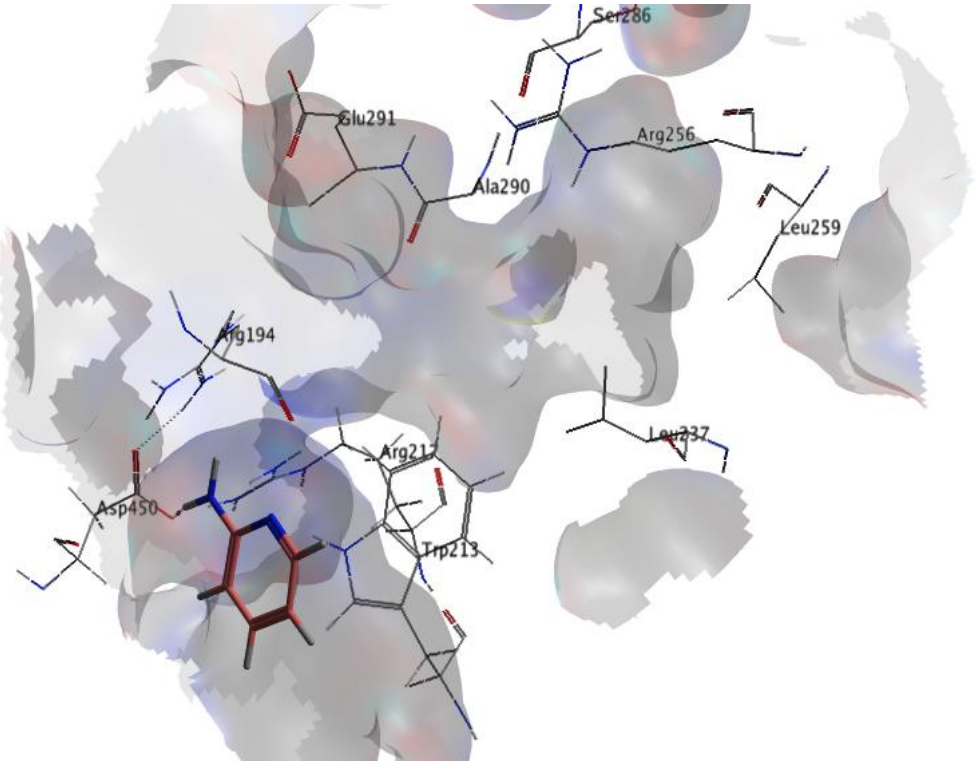


Figure 32. Molecular docking model of interactions between the Bovine Serum Albumin (BSA) and aminopyridine, generated by using Molecular Operating Environment (MOE) 2011.10. BSA PDB file was taken from protein data bank (http://www.rcsb.org/pdb/explore.do?structureId=4F5S) (4F5S). The classical binding sites were marked in corresponding sub domain cations. See more details in the experimental section.

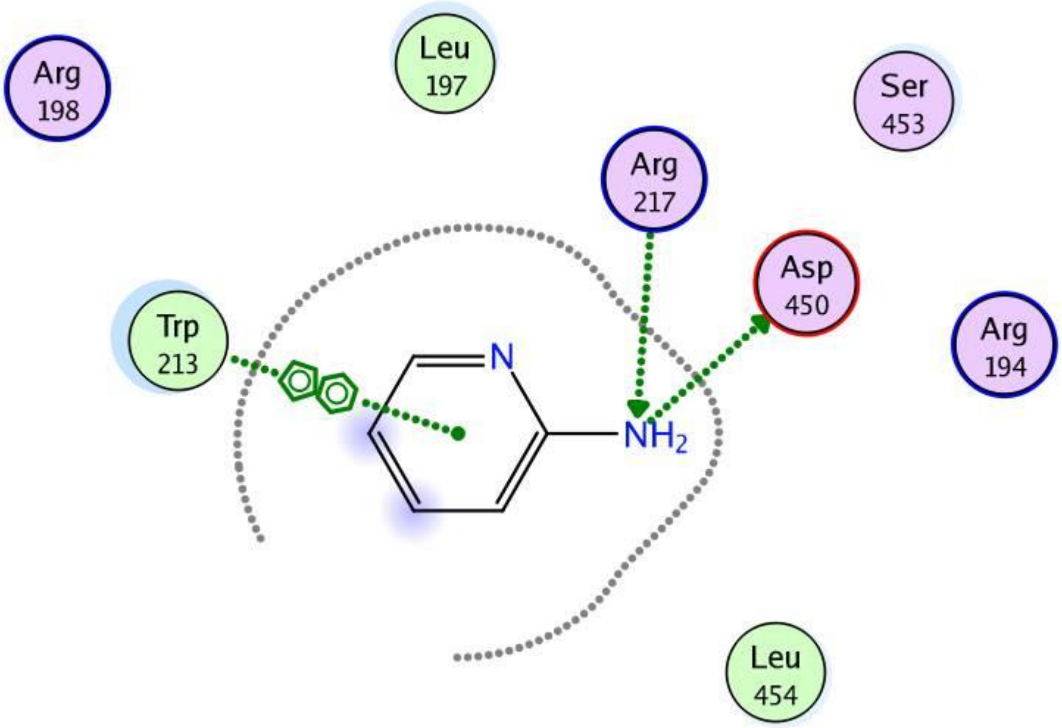


Figure 33. A 2D pictorial scheme for interaction between the BSA binding sites and Isovanillin molecules, generated by using Molecular Operating Environment (MOE) 2011.10. BSA PDB file was taken from protein data bank (http://www.rcsb.org/pdb/explore.do?structureId=4F5S) (4F5S). Where, the π-π interactions between the isovanillin and Trp213 residues of BSA binding sites are shown with green dashed lines between the aromatic ring and Trp213 protein residue. The donor accepter relationship was shown with green dashed lines with arrow heads among the Arg217, Asp450 of chain II of BSA and NH_2 group group of ligand molecule.

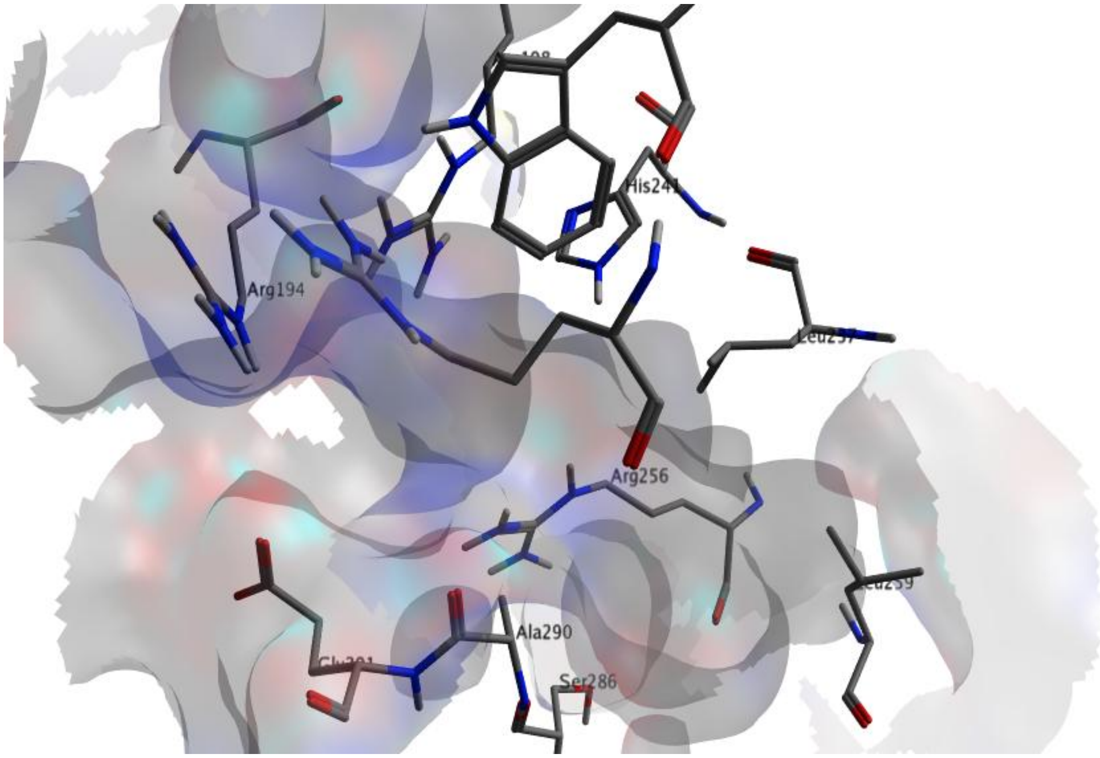


Figure 34: Selected active site of chain B of the Bovine Serum Albumin (BSA), generated by using Molecular Operating Environment (MOE) 2011.10. BSA PDB file was taken from protein data bank (http://www.rcsb.org/pdb/explore.do?structureId=4F5S) (4F5S). The classical binding sites were marked in corresponding subdomain cations.

5

Crude to leads: a triple-pronged direct NMR approach in coordination with docking simulation

5. Crude to leads: a triple-pronged direct NMR approach in coordination with docking simulation

5.1- Abstract

The screening of compounds that bind to the target of interest (specific proteins) plays a vital role in drug discovery. Usually, the identification of biologically active compounds is done from a library of structurally known compounds. However, we successfully illustrate here, that NMR techniques including saturation transfer difference (STD), transfer nuclear Overhauser spectroscopy (Tr-NOESY) and STD-TOCSY (total correlation spectroscopy) in combination with separation methods not only enable the rapid and comprehensive screening of active components, but also their unequivocal structural characterization. Furthermore, a time saving for the recognition of leads is also possible with this application. To probe the binding studies, a hydroethanolic fraction of crude extract (1 mg) from natural product (*Rauia resinous*) was used for the initial assessment with BSA protein. The docking simulation was performed with BSA in the region of Thr190, Arg198, Arg217, Trp213, Arg256, Ala290, and Tyr451 to further refine the active compound towards the leads. Docking results mimic binding as identified by STD, Tr-NOESY, and STD-TOCSY. Isovetexine-2-rhamnoside (2) was found to be most active through group epitope mapping results as well as the docking simulation with relative free energy of -7.2770 Kcal/mol. This experiment provided excellent results through the direct NMR screening method. Using Bovine Serum Albumin as a reference, we illustrate that this approach offers an excellent way for the first hand detection of the active constituents/inhibitors from natural remedies used in folk medicinal treatments.

5.2- Introduction

Natural products are an endless source of bioactive organic compounds with a broad spectrum of physical and chemical properties. Mankind has always been very curious to uncover the potential benefits of natural compounds since antiquity and this struggle continues today, because of the desire to find medicines for a novel and/or superior function. To date, plants are the major source of pharmaceuticals as most of current medicines have origins from a plant-derived product. Historically, use of plants for folk medicinal purposes was very common without having prior knowledge of the constituents responsible for the remedy. However, as science progressed, the availability of techniques such as chromatographic separation, spectroscopic identification, and structure elucidation in the early 1950s, made it possible to analyze natural products to identify the active constituents (TSWETT, M. S., 1906; MOORE S., *et al.,* 1958; HORVATH G. C., *et al.* 1967; WATANABE N., and NIKE E., 1978). Now drug discovery is a systematic process, which usually starts from the isolation of potent bioactive compounds using in vitro/in vivo experimentation followed by synthesis and structural modification to make them potent, efficacious, and non-toxic. This procedure is not a single-day's- or a weeks-work, but requires years of effort to design a drug and to bring it to the market. With so many potential drug discoveries, the time translation into money becomes a big issue whilst developing a drug and this cannot be accomplished without active industrial involvement because drug discovery is expensive, and it is very demanding and intensive work (CZARNIK, A. W., 1998). In order to reduce the time and labor, combinatorial chemistry has become attractive in the pharmaceutical industry which is expected to produce arrays of chemicals rapidly; however, screening a large library of compounds for a particular bioassay creates a problem of many false positive results (PELLECCHIA, M., *et al.,* 2008; BAUM, R., 1996). There is still a need for analytical tools in drug discovery that could reduce the aforesaid problems of screening large libraries efficiently and rapidly with few false positive results. NMR was introduced to provide solutions to these problems to yield data with high accuracy and swiftness (SHUKER, S. B., *et al.*, 1996; CALLE, L. P., *et al.*, 2011; DIERCKS, T., *et al*., 2001). However, slow turnover for the dereplications of compounds and structure identification in mixtures caused drawbacks, which hindered NMR for years from being the most routinely, used technique. More recently, the dereplication problem could be overcome by using hyphenation techniques (ALBERT, K., *et al.*, 1999; JAROSZEWSKI, J. W., 2005) and once again makes NMR an attractive technique in drug discovery.

A milestone came in the drug discovery process in 1999 when Bernd Meyer and co-workers (MAYER, M.; MEYER, B., 1999) used NMR as a tool for molecular recognition of a specific target without having any prior knowledge of protein function or any isotopic labelling unlike 3-FABS (DALVIT, C., *et al.*, 2003) (3 Fluorine Atoms for Biochemical Screening) and FAXS (DALVIT, C., *et al.*, 2003) (Fluorine chemical shift Anisotropy and Exchange for screening). Moreover, without any need for special instrumental setup. STD NMR is free from protein size constraint, an advantage over SAR by NMR (SHUKER, S. B., *et al.*, 1996). Similarly, STD has an advantage over X-ray analysis where requirement of single crystals acts as a bottleneck for structural characterization (FERNÁNDEZ, C.; JAHNKE, W., 2004). In addition, if the crystal quality is also not excellent it can miss numerous ligands in the ligand–receptor complex. Nowadays, the STD NMR method became a regular technique for ligand–protein interaction studies as well as for comprehensive epitope mapping and dissociation constant calculation (FIELDING, L., 2007; ANGULO, J., *et al*., 2010). In contrast to STD NMR, the fluorescence method for titration studies demands prior knowledge of the binding mechanism otherwise prone to errors (LILIANA B., 2007) and unable to resolve a complex mixture (YU, X., *et al*., 2011). Present studies employ the hydroethanolic extract of the *Rauia resinous* plant for the complete observation of molecular recognition and epitope mapping of small organic compounds to the protein by using different NMR techniques including STD, NOESY, Tr-NOESY, and 2D STD-TOCSY. A hyphenation system (LC-UV-SPE-NMR) was used for separation purposes. Earlier, NMR has already been used to analyse the crude extract and remained successful in evaluation (BILIA, A. R., *et al.*, 2001; TAGGI, A. E., *et al.*, 2004) but not for recognition of the active compounds. Recently, NMR has been used to find the target directly from the extract, but, structural information became a question because of the lack of separation methods (POLITI, M., *et al*., 2005). To our knowledge, the present work is the first time when the application of NMR methods together with a separation method is directly applied to the natural product extract for the recognition of active compounds through their binding properties, for drug discovery. The selection of a water-soluble extract was adopted for two reasons: first, the analyses of ligand–protein interaction studies in water are necessary for the dissolution of the protein and secondly, the polar components are usually not investigated. The genus Rauia belongs to the family Rutaceae; comprising 150 worldwide-distributed genera—a little is known about its potential for drug discovery (ALBARICI, T. R. 2006). Despite of the fact that a number of monoterpenes, sesquiterpenes and diterpenes; (VELOSO, E. D. S. 1995) coumarines (SILVA, G. F., *et al*., 1997) the flavone 7, 3, 4-trimethyl kaempferol, alkaloids lemobiline, *γ*-fagarin and esquimianine and *β*-sitosterol are already reported from this genus. To probe the binding studies we selected Bovine Serum Albumin (BSA), because of the structural homology of this protein with the Human Serum Albumin (HSA); a major constituent in circulatory systems that plays a pharmacological role in endogenous and exogenous protein transport in the body

(CARTER, D. C.; Ho, J. X., 1994). In the near future, this approach may open a way to new drug discoveries based on data from ligand–protein interactions.

5.3- Results and discussion

BSA is the most abundant protein in the body fluid; widely studied for the binding interactions through popular and extensively used techniques such as circular dichroism (CD), ultraviolet spectroscopy (UV), and mass spectrometry. In spite of having high sensitivity, these screening techniques suffer from a number of drawbacks such as demand for optical absorption/ activity (Johnson, W. C., 1990) lack of resolving overlapped signals (GARGALLO, J. J. a. R., 2012), lack of information regarding proton exchange (MAPLE, H. J., *et al*., 2011) with solvent, respectively. NMR could solve all these problems efficiently and in the last two decades, NMR spectroscopy has proven to be an excellent tool due to its ability to detect weak binding compounds. There are different NMR methods currently being used for structure based drug design. STD-NMR is the most robust technique to probe binding studies in large libraries as well as providing complete topology of the given ligands (Figure 35).

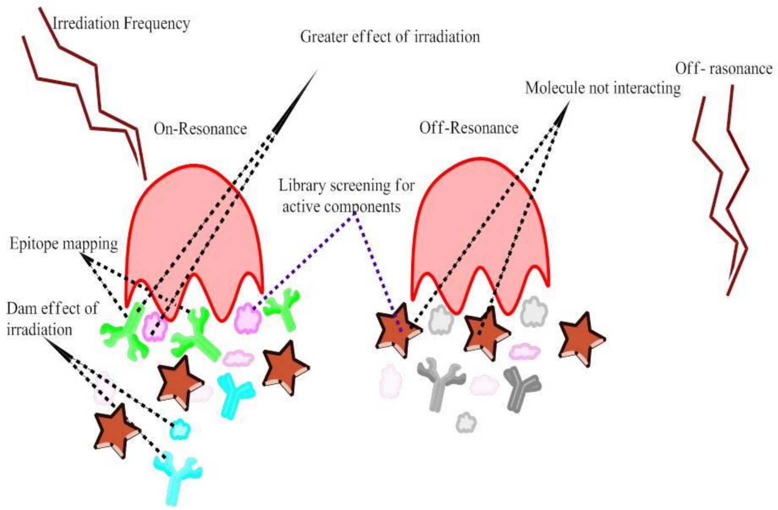


Figure 35. Diagrammatic presentation of STD-NMR. Two FIDs are obtained as a result of selective irradiation of the protein region (on-resonance) and nonprotein region (off-resonance). By the irradiation of the protein signal region; a cascade of protein hydrogens become saturated, and this saturation is transferred through cross relaxation to molecules having contact with protein. In off-resonance, irradiation has taken place in the region without the signals of protein and ligands. By subtraction of the on-resonance spectrum from the off-resonance by phase cycling a spectrum (STD difference) is obtained, having signals from molecules that bind to the protein only. The protein is saturated with a selective saturation pulse, usually a cascade of Gauss-shaped pulses with different saturation times ranging from 1–5 s. STD allows using lower protein concentration; generally, µM amount and ligands in mM amount. In a group of large molecular libraries, active molecules can be detected and the specific part of the ligand in contact with protein (epitope mapping) can also be recognized.

In the present study, we used the crude hydroethanolic extract of the *Rauia* species (isolation scheme Figure 41 supporting information) for the screening of active compounds. The binding studies of compounds of rauia for BSA (Bovine Serum Albumin) protein were carried out on approximately 1 mg of the extract G (eluted from the LH-20 column at 80: 20% v/v methanol–water, respectively) dissolved in D_2O–CD_3OD (95 : 5% v/v) and mixed with 50 mM protein (BSA) solution in the sodium phosphate buffer (pH 7.4). The STD spectrum, as shown in Figure. 36, clearly indicates the aromatic moieties of rauia's constituents have the closest approach to the binding cavity of the protein. The line broadening caused by ligand resonances was observed in the close vicinity with the protein's binding cavities. Likewise, the relaxation rate was increased causing a slow tumbling rate of the protein–ligand complex.

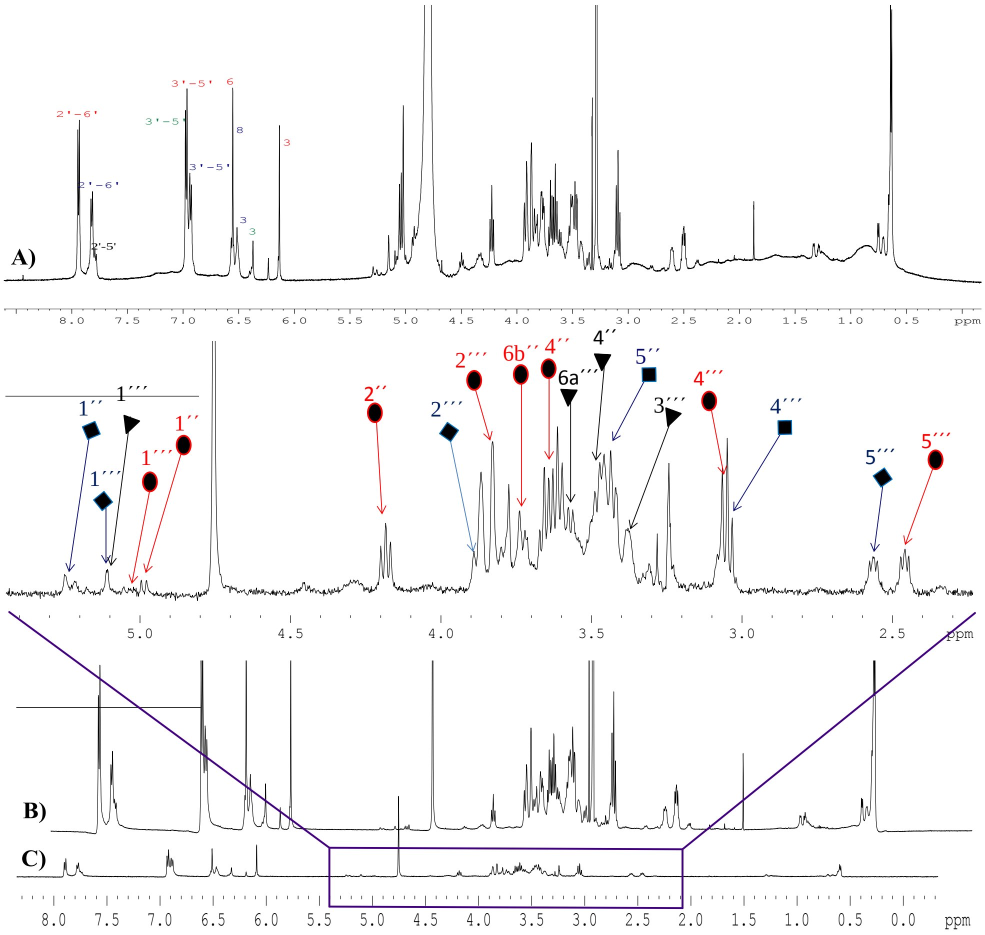


Figure 36. The top ^{1}H NMR spectrum (A) is of 1 mg extract with 50 mM protein solution taken in a 3 mm NMR tube. The spectrum was produced at 298 K on a 600 MHz Bruker AVANCE III spectrometer. The aromatic region is markedwith different colored numbers. Red colour indicates aromatic protons of vetexine-2-rhamnoside, blue color isovetexine-2-O-rhamnoside and green color neoschaftoside, respectively. In the lower two stack plots of STD spectra, B, represents the reference spectrum while, C, is the STD spectrum in the presence of BSA, respectively. The magnified region (2.2–5.4 ppm) of STD spectrum C in the box is shown up to spectrum B, the arrows starting from a circle represent the vetexine-2-O-rhamnoside protons, the arrows starting from a square represent isovetexine-2-O-rhamnoside while the arrows starting from a triangle represent the

neoschaftoside protons showing STD effects, respectively. See more details in the experimental section.

The magnified STD spectrum in Figure 36 C is shown with arrows and distinctive colors for the different starting points. The unequivocal representation of the STD spectrum suggests that almost all signals from this extract showed some level of enhancement, which establishes that all molecules in the extract show some interaction with protein. Further proof of this becomes apparent after the generation of 2D transfer NOESY and TOCSY spectra, which are the mirror reflection of this STD spectrum except having 2D dimensions.

The amplification factor in the STD spectrum depends upon two factors: (A) the greater the number of protons, the bigger the intensities and (B) the longer the saturation time, the larger the signal. The STD effect (amplification factor), therefore, reflects the relative amount of saturation transferred from the protein onto the ligands, and the number of ligands being saturated. The group epitope mapping calculation was effected by comparing the individual proton integrals, and are referenced to the strongest signal in the spectrum, which is assigned to 100% effect. Signals with high STD-effect appeared to be in a closer contact with the protein binding cavities, and lower STD-effect was more towards the solvent side (away from these binding cavities). The possible compounds that could be recognized via the complete 2D spectral information as accumulated from separation methods are shown in Figure 37 A, and the respective amplification factors are shown in different colors in Figure 37 B–D.

A saturation time of ca. 2. 5 s was found to be more efficient after evaluating a series of STD spectra with different saturation times ranging from 0.5 to 5 s. The standard Bruker pulse sequence with water suppression was used in the generation of STD spectra; background protein signals were eliminated by the application of a T_{1p} 30 ms spinlock filter. Intriguingly, the interaction footprint of the aromatic protons (2´/6´) of these compounds revealed almost in the same region as a doublet, and showed a strong STD effect, which provides evidence that the benzenoidal region remained closer to the protein surface more than the sugar moieties.

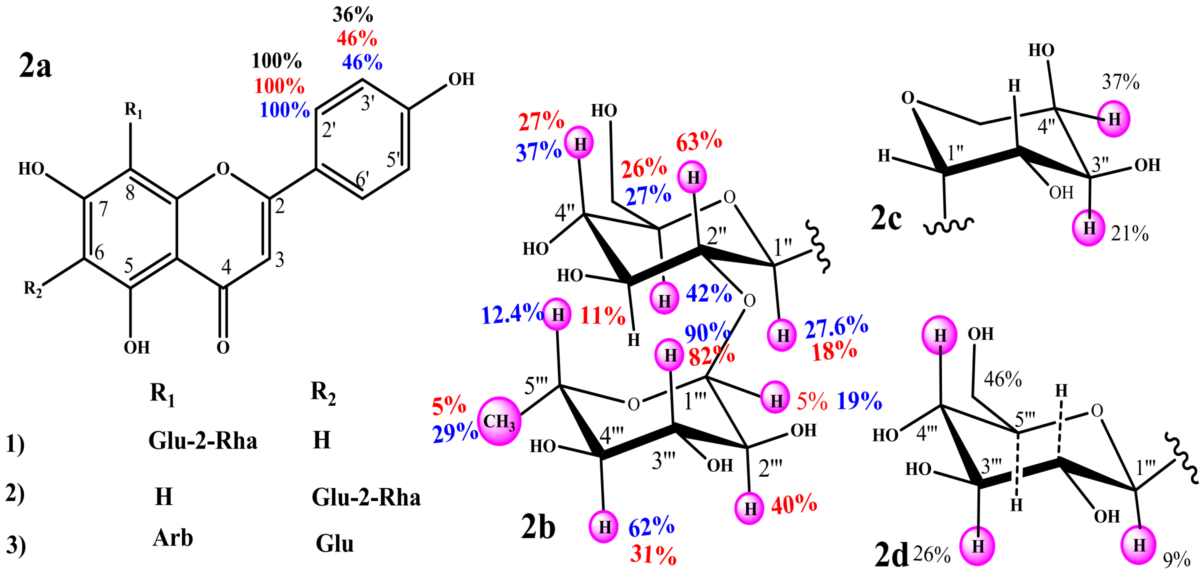


Figure 37. (A) The relative amplification factors (AF) with absolute numbering for the recognizable compounds: where red and blue colours represent AF for compound **1** and **2** respectively (B), whereas the black colour represents the nomenclature numberings of these compounds. (C) and (D) represent the AF for positions 8 (L-arabinose) and 6 (D-glucose) of compound **3** respectively.

The 100% STD effect was shown by the 2´/6´ of these compounds indicating a major contact to protein, thereby, receiving more saturation transfer via cross relaxation. Whereas, the hydrogens (3´/5´) of both compounds vetexine- 2-O-rhamnoside (**1**) and isovetexine-2-O-rhamnoside (**2**) contributed 46% amplification, on the other hand, the same hydrogens of neoschaftoside (**3**) were more distant and showed 36% STD effect. The spectral data remarkably showed that all axial hydrogens of the sugar moieties had some interactions towards serum albumin. However, from equatorial hydrogens only the anomeric protons of the rhamnoside skeleton of **2** and number 4´´ protons of the arabinose moiety of **3** showed STD signals and thereby, interactions. Second highest interaction from the STD spectrum corresponds to compound **1** and **2** protons 3´´´ from the rhamnose sugar, showing amplification between 80 and 90% as is clear from Figure 37 B–D. Similarly, hydrogen number 4´´´ from rhamnose of compound **2** and 2´´ from compound **1** showed almost equal amplification of 62% and 63%, respectively. The amplification factor from the sugar region revealed the binding pattern for all compounds is similar. Interestingly, the sugar moieties of all these compounds remain unable to show the 100% epitope mapping, and thereby, showing less interaction towards the protein. In other words, these moieties mostly remained towards the solvent side, relative to the benzene region.

Epitope mapping (FIELDING, L., 2007; ANGULO, J., *et al*., 2010) for these compounds was a difficult process, but the application of LC-UV-SPE-NMR separation (see chromatogram Figure 42 supporting information) followed by structure characterization for this extract inferred that the major components were responsible for interaction (Figure 37). Observation from STD amplification indicated that the aromatic moieties of the ligands were found to be involved in stronger interactions with protein. Complete epitope mapping and structural characterization was done with the help of 2D COSY spectra obtained by LC-UV-SPE-NMR (data not shown here).

STD-NMR, despite having a number of advantages and being the most widely used technique for affinity screening in NMR labs, also has some limitations. For example, when the ligands bind very strongly (K_D equal to less than 1 nM) with the protein, false-negative amplification factors may occur as a result of slow turnover rates. In addition, when a number of potent ligands are present in the mixture, signals become overlapped and can complicate the correct identification (WAGSTAFF, j. L., *et al*., 2013; KRAMER, M.; KLEINPETER, E., 2010).

For clear identification of active components of this extract, additional experiments 2D-NOESY with and without protein and 2D STD-TOCSY with protein were also performed. In 2D NOESY phase sensitive measurements, all peaks showed positive NOEs with respect to the diagonal while some small signals were observed in Figure 38 A and B. The 6 ppm region was negative; this might be due to the higher molecular weight or uncorrected phase. The buildup rate was long enough with a mixing time of 400 ms. In transfer NOESY, the standard Bruker phase-sensitive pulse sequence was modified by putting the spinlock of 3500 Hz after the first 90º pulse to remove peaks of protein signals. In transfer NOEs, strong negative signals having the same significance to the diagonal were observed as shown in Figure S-43 supporting information, which indicates a clear agreement to the previously conducted STD-NMR experiment. In transfer NOESY, the buildup rate was much faster compared to that of 2D NOESY with the same mixing time. Furthermore, the cross peaks (Figure 38A and B) in the sugar region (3–4.6 ppm) became more intense compared to the NOESY without protein, strikingly illustrating that there exists a clear binding; showing short relaxation as in macromolecules, likewise, in the aromatic region (6–8 ppm) region (see Figure 43 supporting information). Notably, a flavonoidal skeleton exhibited interaction as shown by the spectral results of STD-NMR and 2D tr-NOESY; a fact further verified through the STD amplification factor as in Figure 37.

For further consolidation to the bound conformation of structures obtained from 1D STD-NMR and transfer NOESY, the same sample with protein was assessed for 2D TOCSY and STDTOCSY experiments. As is well known, the 2D STD-TOCSY experiment is an excellent method to measure the proximities of the protons to the protein (HASELHORST, T., *et al.*, 2007; MAYER, M.; MEYER, B., 1999). The MLEV-17 sequences were performed to obtain the standard TOCSY experiment with the receptor as shown in Figure 44 supporting information and the corresponding STD-TOCSY spectrum (Figure 45 supporting information), respectively.

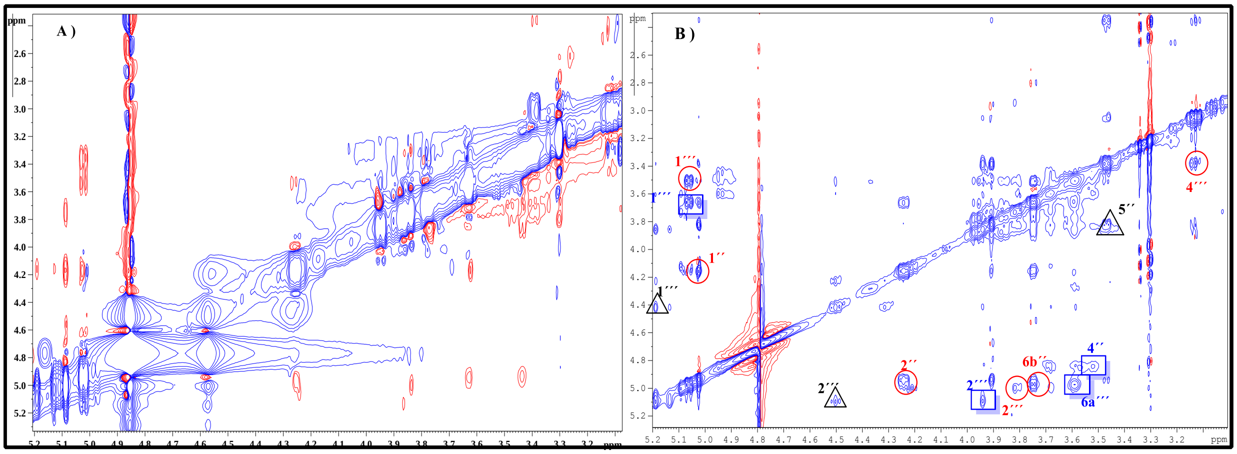


Figure 38. Two-dimensional NOESYspectra without BSA (A) and transfer NOESY (B), with Bruker standard phase sensitive pulse sequence for the extract in D_2O–methanol (95: 5% v/v) and BSA in buffer solution, respectively. In transfer NOESY the pulse sequence is modified by putting the 3500 Hz spinlock for the removal of unwanted protein signal before the first $\pi/2$ pulse. This spectrum is produced by using a 400 ms mixing time at the Bruker 600 MHz NMR spectrometer with a cryoprobe, at 298 K. In tr-NOE cross-peaks shown are of the same sign as of the diagonal signals (negative NOEs). The region from 2.2 to 4.5 ppm (red circles, blue squares, and black triangles represent protons of compounds **1**, **2** and **3**, respectively) showed for clearer recognition of the sugar signals. Other experimental conditions are given in the materials and methods section.

Observations of STD-TOCSY were established from the signal intensities; the more intense the signals, the more saturated the corresponding protons and hence, the closer to the receptor's binding surface. Sequential assignments of ^{1}H resonances were based on 1D STD-NMR signals, and NOE cross peak connectivities were observed. The cross peaks of the aromatic region of all compounds showed intense signals that could be attributed to the closest approach to protein in the ligand–protein complex. Due to the similarity in the parent structures, aromatic protons showed identical behaviour with different amplification factors as in Figure 37B–D. However, the magnified region 3.2– 4.5 ppm (Figure 39A and B) in standard TOCSY and STD-TOCSY spectra illustrated strong signs corresponding to the protons 2´´, 6b´´ of the glucosyl moiety, and the axial protons 2´´´and 4´´´ of a rhamnosyl moiety of **1** (shown with a red circle) which provided clear validation of 1D-STD spectral data. Likewise, the protons 2´´´, 6a´´ belong to **2**, 5´´ of 3, shown as squares, and triangles (Figure 39A and B) respectively, revealed perfect agreement to the 1D-STD and the tr-NOESY data. Other signals that received weak saturation were also present in the spectrum, but could not be characterized in this mixture. Similarly signals in the 0.5–1.5 ppm region belonging to the methyl group of a rhamnosoyl moiety of these compounds remained prominent as shown in the full spectrum of TOCSY (Figure 44 and 45 supporting information). The enlarged spectra of these STD-TOCSY and TOCSY from 2.2–4.5 ppm regions inferred that there were also some signals that disappear in the STD-TOCSY spectrum due to their low degree of saturation. From the STD-NMR, 2D tr-NOESY and 2D STD-TOCSY spectral data provided the existence of clear flavonoidal skeletal information in this mixture, but still the exact structural confirmation of compounds involved in interaction required some explanation.

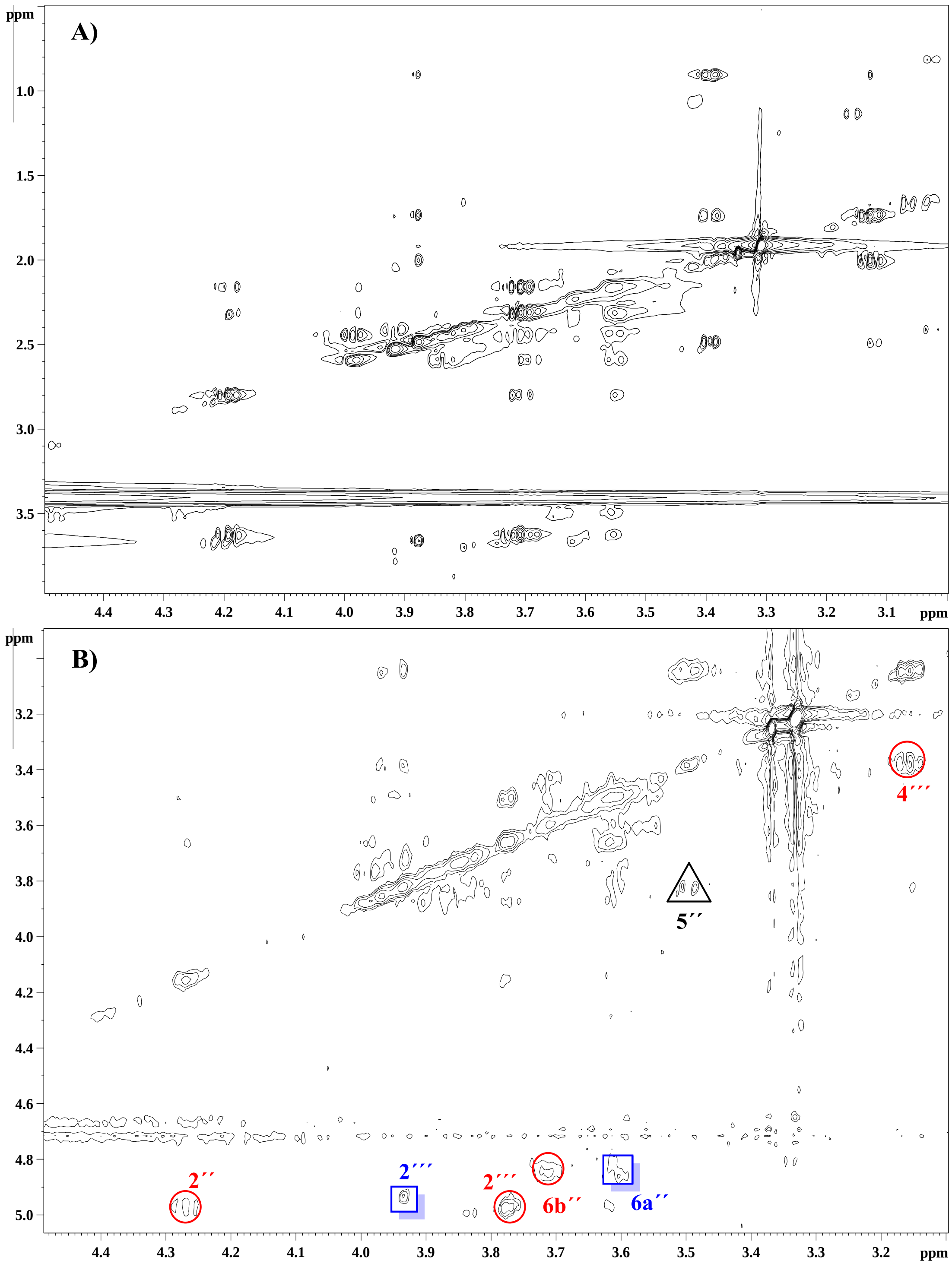


Figure 39. Two-dimensional TOCSY; A standard TOCSY (A) and STD TOCSY (B) NMR spectra of crude extract with 50 μM BSA complex is shown. Experiments were performed by using a Malcolm Levitt's composite-pulse

decoupling sequence (MLEV-17) with mixing time of 60 ms at 298 K with a 600 MHz spectrometer. In STD-TOCSY saturation of the protein was achieved with a cascade of 50 Gaussian pulses (50 ms duration each) with a delay of 200 ms in between each pulse, with a total saturation time of 2.5 s. while, the on-resonance saturation frequency was set to -0.5 ppm and off-resonance to 30 ppm. Both spectra were acquired in 1 mg of the extract in D_2O–methanol (95: 5% v/v) and BSA in buffer (7.4 pH, not corrected for D_2O) solution respectively. The magnified region (2.2–4.5 ppm) of both standard TOCSY and STD-TOCSY was compared to the signals of different sugars, where the red circles, blue squares, and black triangles represent protons of **1**, **2** and **3**, respectively. Other conditions are according to the experimental section.

With the reference to rauia unknown extract, to say anything about the confirmation of the structures was impossible only on the basis of STD, NOESY and TOCSY information. Therefore, to obtain clear information of the structure of the compounds involved in binding, still we require some other methods enabling separation and characterization. Hence, we used LC-UV-SPE-NMR for the separation of components responsible for binding as well as clarity in the characterization process. 1 mg of extract was dissolved in 1 ml of water–methanol (70: 30% v/v) respectively, and 10 ml (0.01 mg per injection) was injected to the Agilent 1200 HPLC connected to NMR and SPE prospect 2 system. A gradient mobil phase (methanol–water) was used at different concentrations (see Figure 46 supporting information) with a flow rate of 1 ml min^{-1}. The chromatogram showed six major components; concentrations of the two major compounds were sufficient to give 2D spectra (COSY, HSQC, and HMBC) for characterization, while the other four minor components (chromatogram Figure 42 supporting information) could provide ^{1}H NMR and COSY spectra. Spectral data were compared with already published data (Rayyan, S., *et al.*, 2005; Xie, C., *et al.*, 2003; Doyama, J.

T., *et al.*, 2005).Three compounds that were fully characterized from the chromatogram (Figure 42 supporting information) are vetexine-2-O-rhamnoside **1**, neoschaftoside **3** and isovetexine- 2-O-rhamnoside **2** corresponding to peaks 2, 4 and 5, respectively. Peak six is a dimer of vetexine-2-O-rhamnoside with molecular weight 1155.5 amu in electrospray ionization mass spectrometry (Figure S-47 supporting information), but we were unable to find the absolute connectivity of these two units.

In principle, affinity screening identifies molecules with physical affinity to the target protein. This, by no means, is unambiguously correlated with biological activity. The search for lead compound(s) from the crude extract becames rationalized after performing the molecular docking studies for conformational refinements. Direct NMR methods (STD, Tr-NOESY, and STD-TOCSY) could provide the evidence for the existence of binding between the ligands and protein. However, NMR remains unable to point out the position of attachment within the protein's cavities. The selected active site residues of chain B (Figure 48 supporting information) revealed that the compound **1** showed interaction by hydrogen bonding through Arg194 and Lys294 present within the active site as shown in the figure S-49a supporting information. The Arg194 acts as a hydrogen bond donor for oxygen of ligand, whereas Lys294 donated its hydrogen bond to the oxygen of the hydroxyl group that is attached to sugar part of the ligand (Figure S-49b supporting information). Different sites of ligand were exposed to the Thr190, Arg198, Arg217, Trp213, Arg256, Ala290 and Tyr451 residues to stabilize ligand residence in the pocket. Compound **2** fitted better than compound **1** in the binding pocket and showed more strong interactions with pocket residues (Figure 40A).

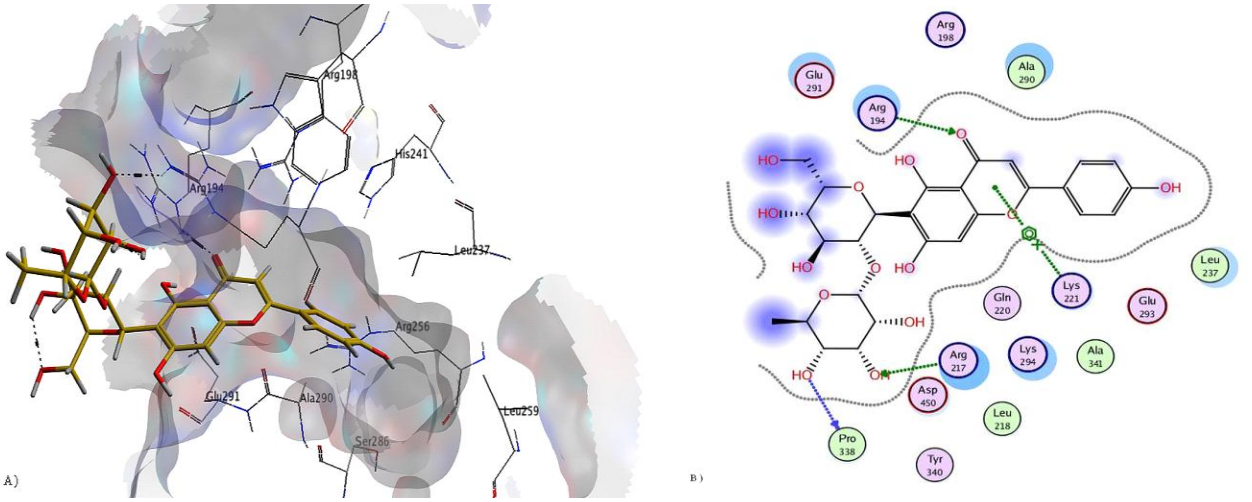


Figure 40(A) and (B): Molecular docking model of interactions between Bovine Serum Albumin (BSA) and isovetexine-2-O-rhamnoside **(2)** generated by using Molecular Operating Environment (MOE) 2011.10. BSA PDB file was taken from protein data bank (http://www.rcsb.org/pdb/explore.do?structureId=4F5S)(4F5S). The classical binding sites were marked in corresponding sub domain cations. See more details in Experimental section.

The dihydropyrine moiety of the ligand exhibited hydrophobic (π- π) interaction with Lys221, whereas hydrophilic interactions were maintained by Arg217 and Pro338. The Arg194 and 217 acted as a donor and Pro338 as acceptor for hydroxyl oxygen of the sugar moiety of compound **2** (Figure 40B). MOE ranked compound **2** on top (-7.2770 Kcal/mol) and was considered as the most active among the series, molecule **3** was second in terms of free energy (-7.0316 Kcal/mol) while molecule **1** showed lowest energy (-6.8345 Kcal/mol). Compound **3** is slightly different in structure from the others which docked in a quite displaced position, but maintained its interactions with some of the active site residues such as Arg194, Lys221, 294 and Glu291 (see Figure 50a and b, supporting information). The active residue sites Arg194, Lys221, 294 and Glu291 behaved as hydrogen-bond donors, whereas Glu291 also acted as acceptor. Other residues such as Arg435, Cys447, His287, Thr190 and Asp450 supported Glu291 binding. On the basis of this study, we can conclude that the phenolic compounds interacted with BSA through hydrophobic (π- π interactions) as well as hydrophilic (hydrogen bonding) interactions which stabilized the ligand–BSA complex. The docking result ranked compound **2** as the most active depending upon interactions to the selected active sites of bovine serum albumin. Indeed, the binding mode obtained by docking analysis was completely in agreement with that from the STD NMR, tr-NOESY and STD-TOCSY data, hence, validating the orientation of these compounds within the binding site.

5.4. Conclusions

Direct NMR methods (STD, Tr-NOESY, and STD-TOCSY) convincingly demonstrated that it is a fast method for the extract screenings towards the lead compounds annotation. A single and simple experiment (STD NMR) could provide significant ligand binding epitopes in a very short span of time by consuming less amount of the receptor. The use of hydroethanolic extract may allow a simple way to natural medicinal manufacture for further investigation. This could be a more economically favorable application for further investigation of natural products. To the best of our knowledge, this work represents the first example where the combination of results of screening techniques involving a separation method, and docking results delivered unambiguous confirmation of leads/ inhibitors, unequivocal structural characterization, and binding sites that could help in fast and facile drug design. In this study, we used Bovine Serum Albumin (BSA) as a model, but this approach can be applied to other target proteins and it may be an excellent way to obtain leads/inhibitors for a specific target in a short time. Herein, we used only 1 mg of crude material (extract) and 1 ml solution of target (50 mM) for a quick study in search of possible active ligands from the unknown natural product. In conclusion, we realistically demonstrated the efficient use of dynamic STD NMR followed by a chromatographic hyphenated system that may be a fruitful way of distinction of leads/inhibitors for future leads.

5.5 Supporting information

Crude to Leads: A triple-pronged direct NMR approach in coordination with docking simulation

.

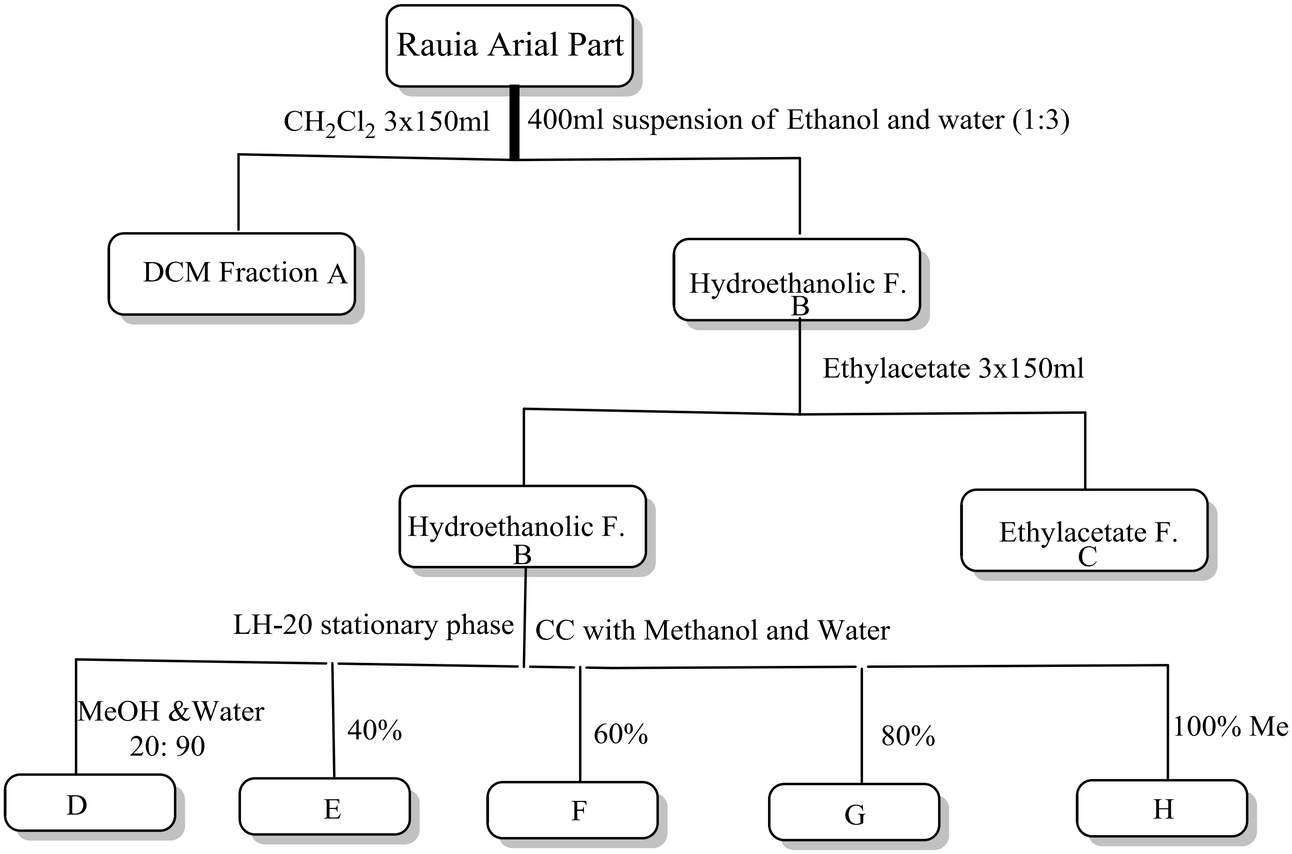


Figure 41. General isolation scheme for aerial parts of Rauia plant.

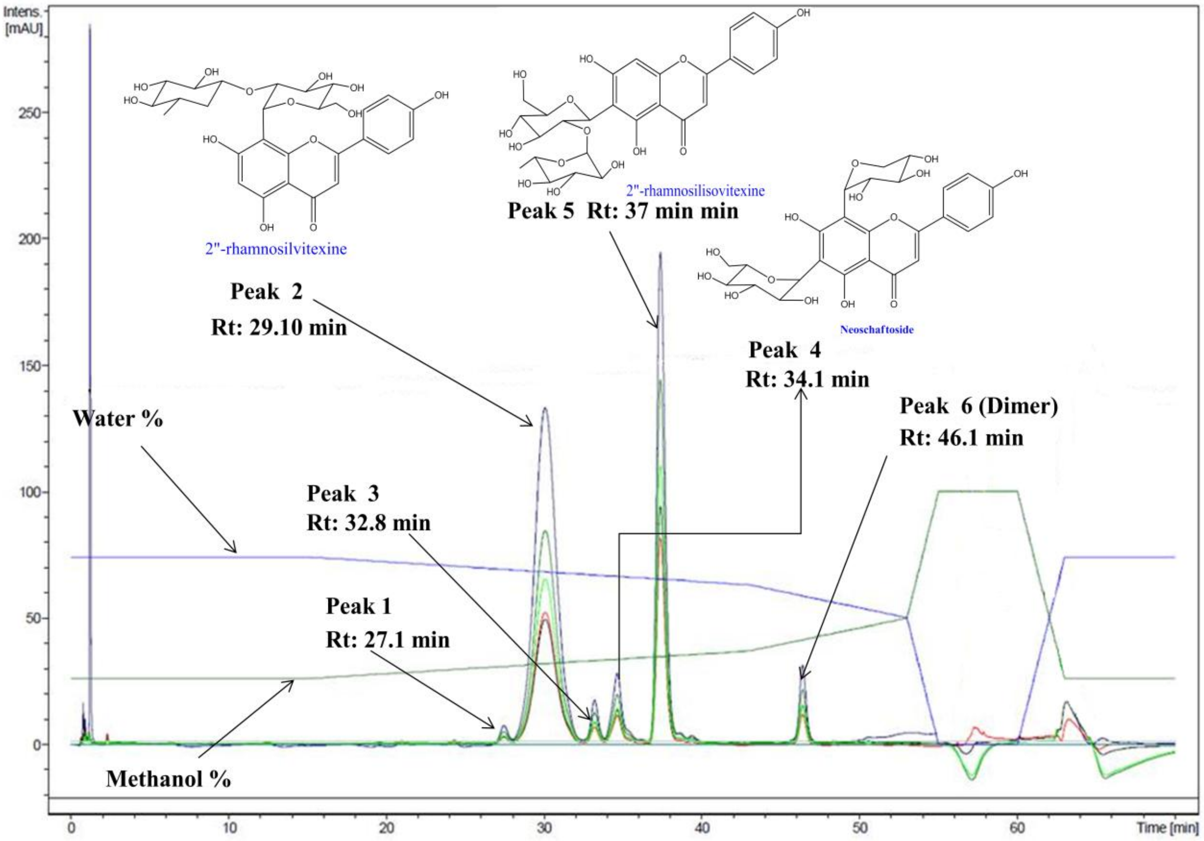


Figure 42. Chromatogram obtained as a result of separation with LC-method by using water and methanol as solvent at different gradients.

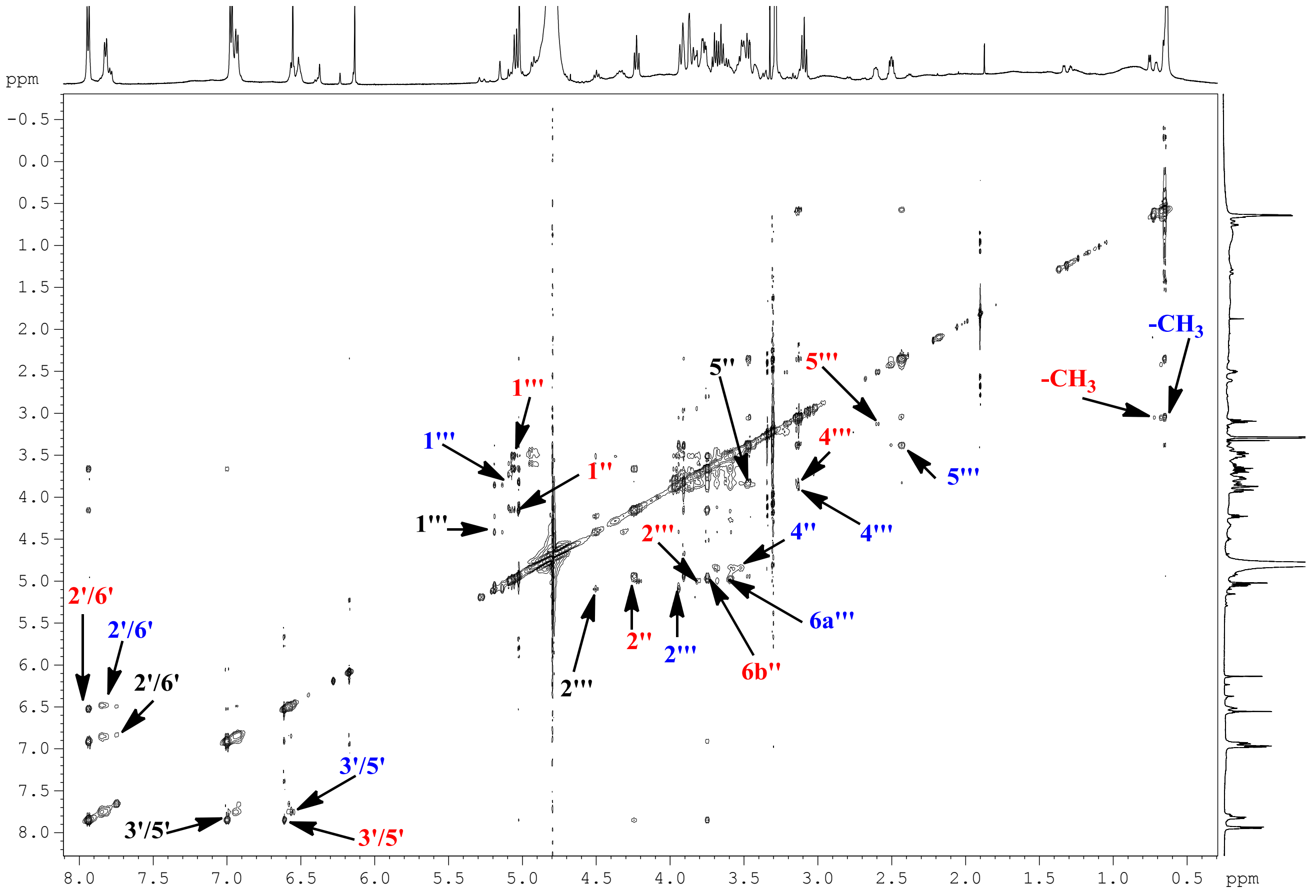


Figure 43. Two-dimensional transfer NOESY with Bruker standard phase sensitive pulse sequence for the extract in D_2O and methanol (95:5% v/v) and BSA in buffer solution . In transfer NOESY pulse sequence is modified my putting the 3500 Hz spin lock for the removal of unwanted protein signal before the first $\pi/2$ pulse. This spectrum is recorded by the use of 400 ms mixing time with at 600 MHz Bruker NMR spectrometer with a cryoprobe TCI at 298 K. NOE cross-peaks shown are of the same sign as of the diagonal signals (negative NOEs).

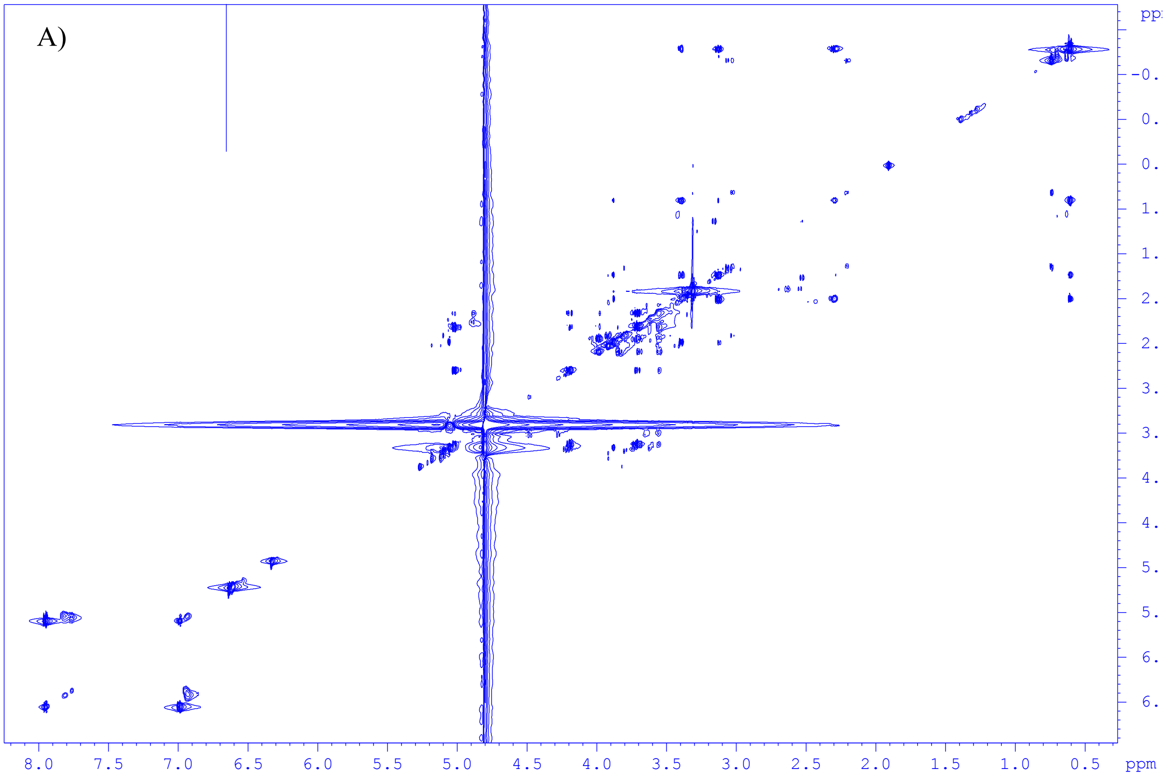


Figure 44. A standard 2D TOCSY NMR spectrum of crude extracts with 50μM BSA is shown. Experiment was done by using a Malcolm Levitt's composite-pulse decoupling sequence (MLEV-17) with mixing time of 60 ms at 298K with 600 MHz spectrometer. For standard TOCSY MLEV-17 sequence the experimented was recorded with 7211.539Hz spectral width in both dimensions (F_1 and F_2) with a total number of 56 scans, with 256 increments in t_1 were collected with 2 s relaxation delay for each scan; with the experimental time less than 10 hrs. Other conditions are according to the experimental section. TOCSY spectrum was acquired in 1mg of the extract in D_2O: methanol (95:5% v/v) and BSA in buffer (7.4 pH, not corrected for D_2O) solution respectively.

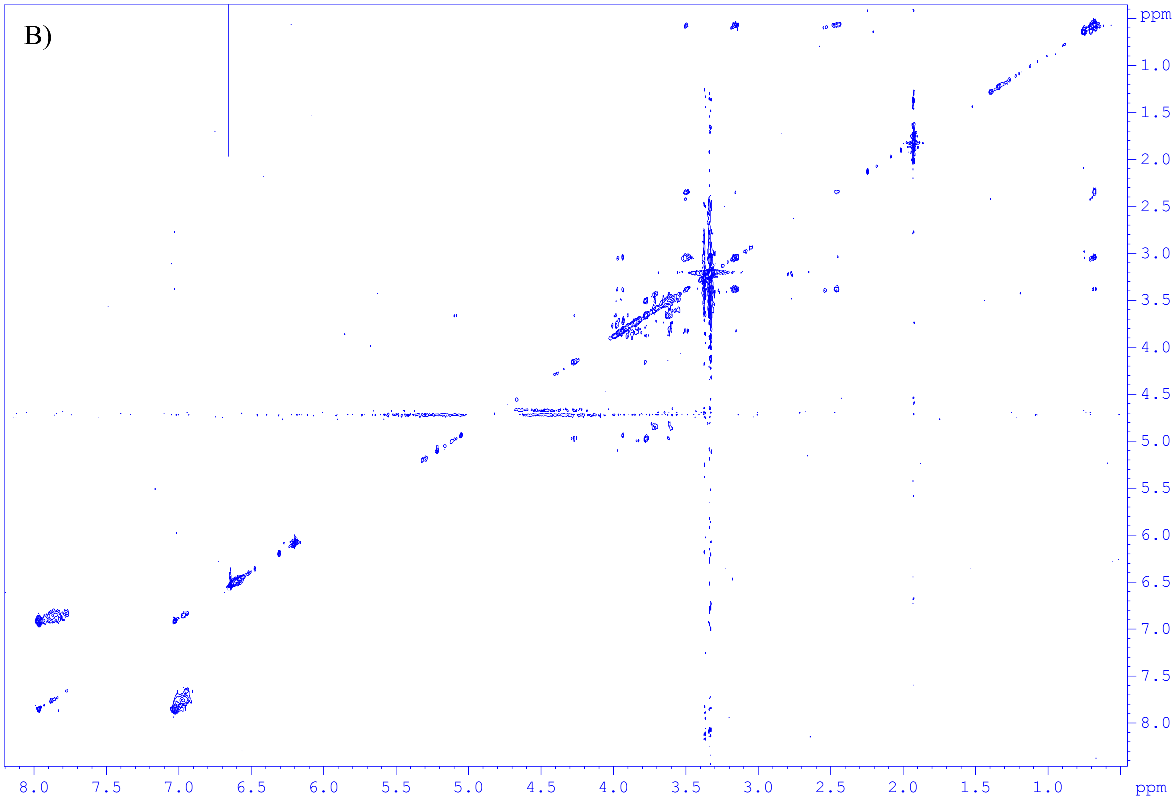


Figure 45. Two-dimensional STD-TOCSY NMR spectra of crude extracts with 50uM BSA by using a Malcolm Levitt's composite-pulse decoupling sequence (MLEV-17) with mixing time of 60 ms at 298K with 600 MHz spectrometer. In STD-TOCSY saturation of the protein was achieved with a cascade of 50 Gaussian pulse (50 ms duration each) with a delay of 200 ms in between each pulse, with a total saturation time of 2.5 s. The *on*-resonance saturation frequency was set to -0.5 ppm and *off*-resonance to 30 ppm. 2D STD-TOCSY Spectrum was recorded with 7211.5 Hz spectral width in both dimensions (F_1 and F_2) with 64 transients in t_2 and with 384 increments for the *on* and *off* resonance in t_1 with the total experimental time of more than 16 hrs. Both spectra (2D standard and STD TOCSY) were

acquired in 1 mg of the extract in D_2O: methanol (95:5% v/v) and BSA in buffer (7.4 pH, not corrected for D_2O) solution respectively.

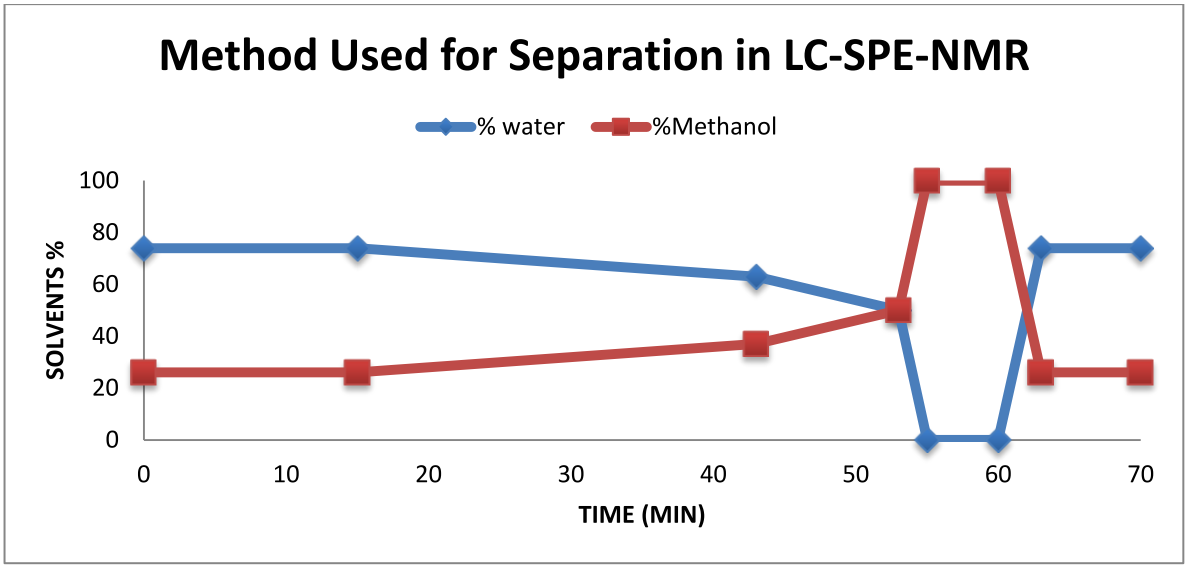


Figure 46. Separation method and the solvent gradient used in LC-method

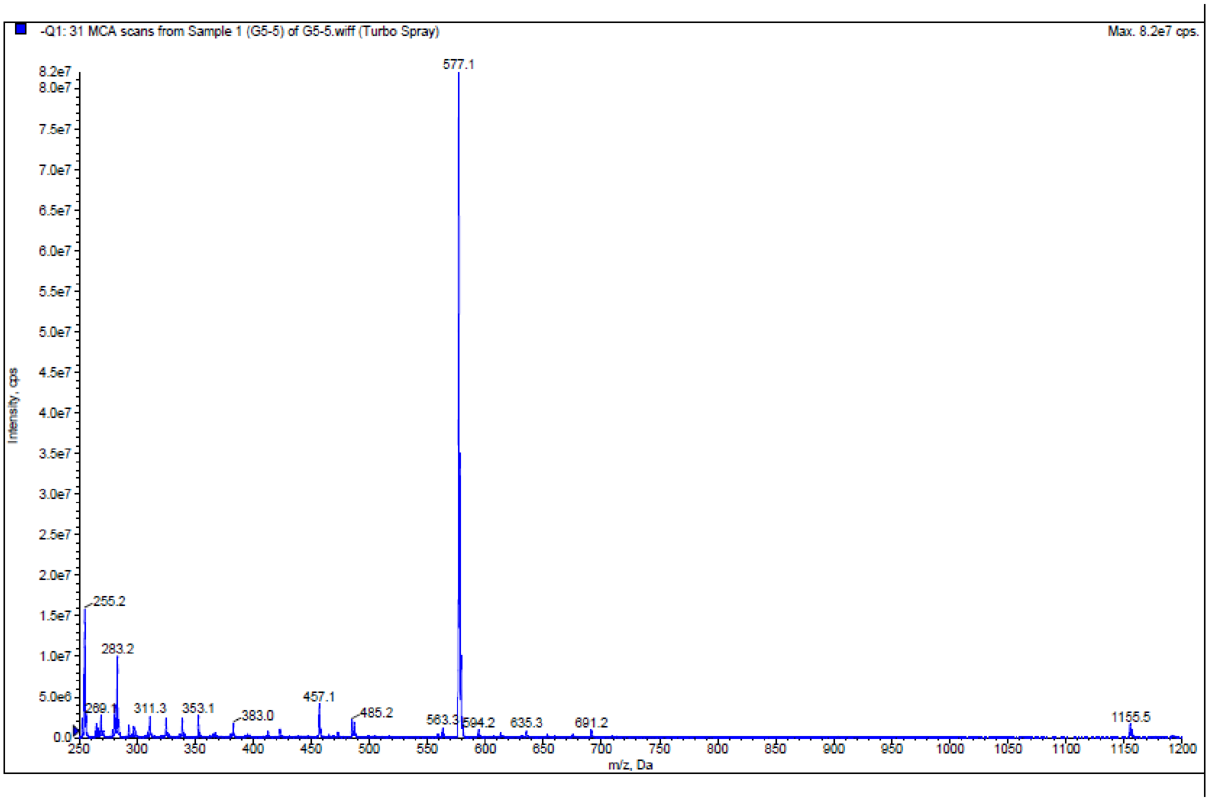


Figure 47. Electrospray Ionization (ESI) spectrum of peak 6 of an LC - chromatogram, showing the molecular weight 1155.5 a.m.u.

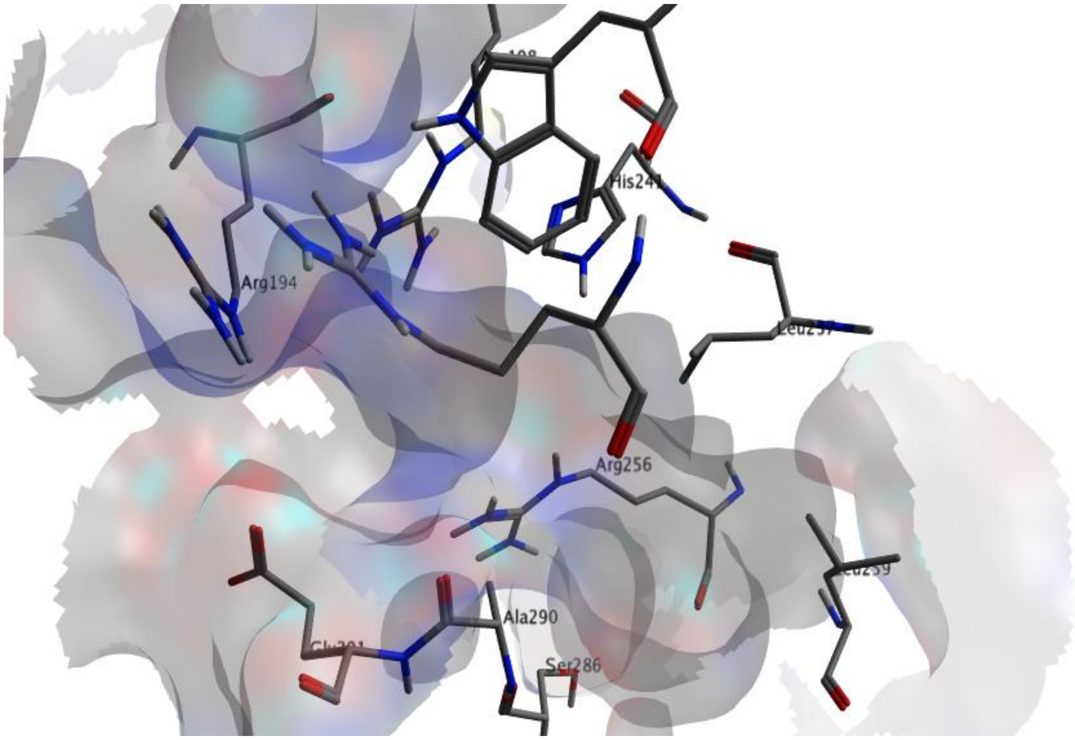


Figure 48: Selected active site of chain B of the Bovine Serum Albumin (BSA), generated by using Molecular Operating Environment (MOE) 2011.10. BSA PDB file was taken from protein data bank (http://www.rcsb.org/pdb/explore.do?structureId=4F5S) (4F5S). The classical binding sites were marked in corresponding subdomain cations.

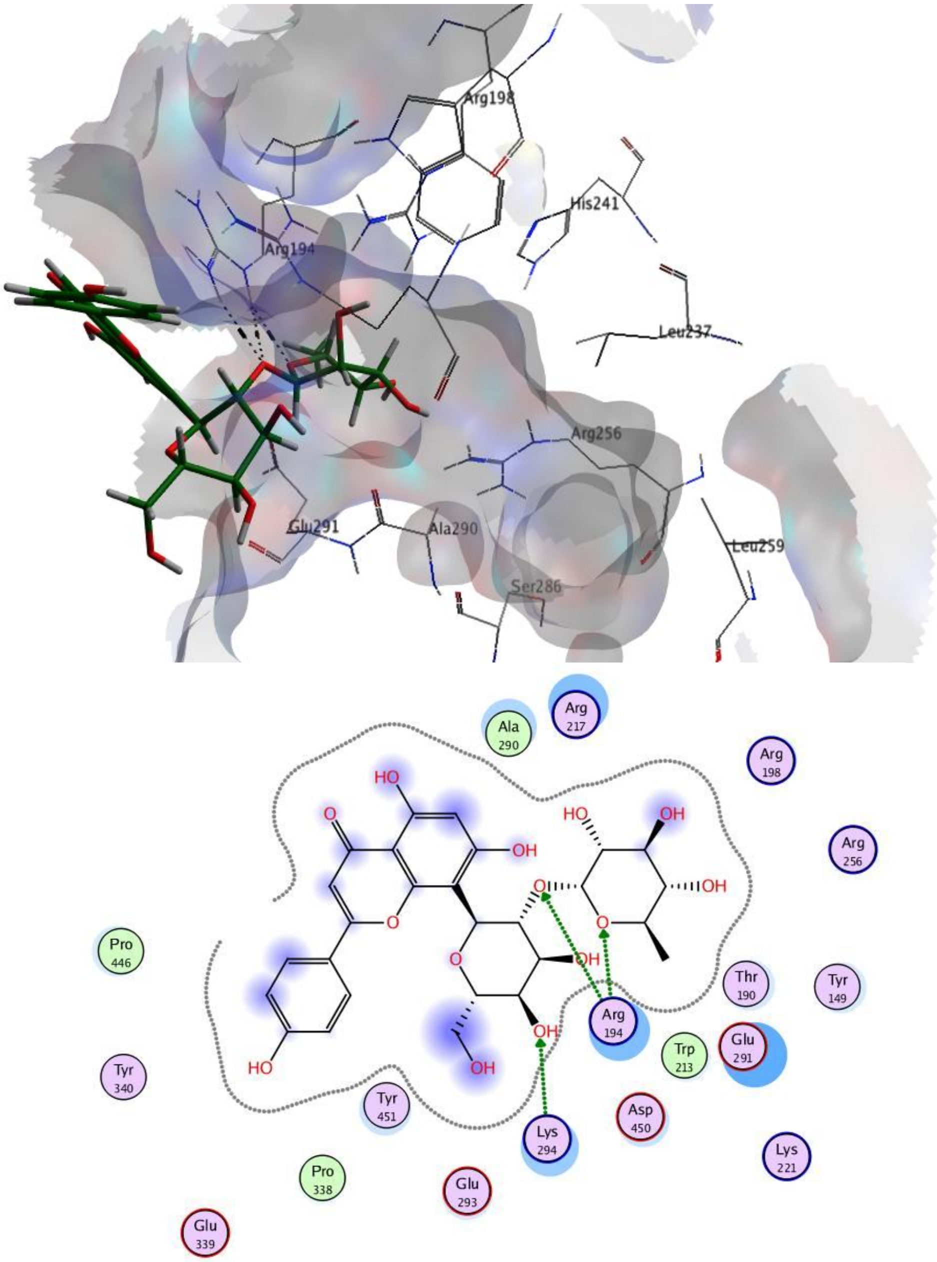


Figure 49a-b: Molecular docking model of interactions between the Bovine Serum Albumin (BSA) and vetexine-2-O-rhamnoside **(1)**, generated by using Molecular Operating Environment (MOE) 2011.10. BSA PDB file was taken from protein data bank (http://www.rcsb.org/pdb/explore.do?structureId=4F5S) (4F5S). The classical binding sites were marked in corresponding subdomain cations. See more details in experimental section.

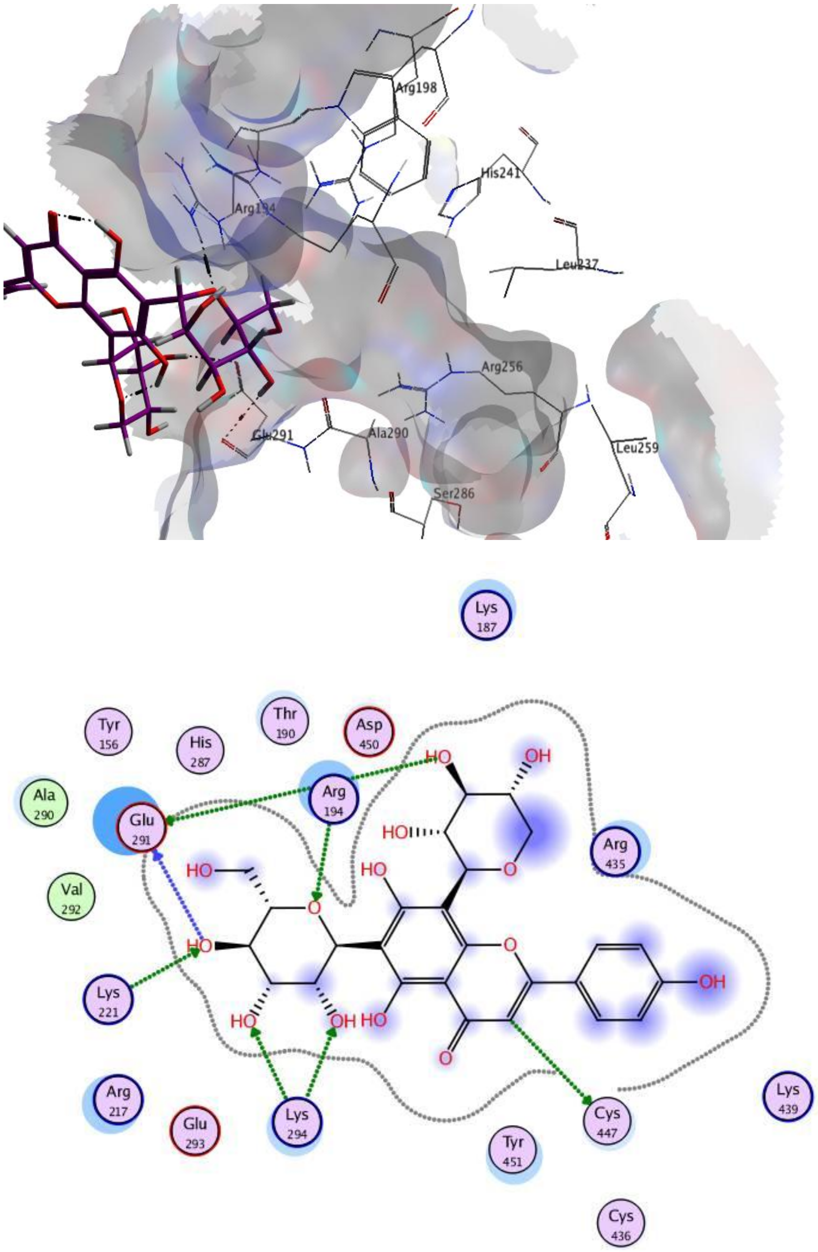


Figure 50a-b: Molecular docking model of interactions between the Bovine Serum Albumin (BSA) and neoschiftoside **(3)**, generated by using Molecular Operating Environment (MOE) 2011.10. BSA PDB file was taken from protein data bank (http://www.rcsb.org/pdb/explore.do?structureId=4F5S) (4F5S). The classical binding sites were marked in corresponding subdomain cations. See more details in experimental section.

❻

Direct NMR mediated ligands screening for a specific target (Human Serum Albumin): Exemplified by using a highly bioactive *Strypnodendron polyphyllum* fully bloomed flower extract.

6. Direct NMR mediated ligands screening for a specific target (Human Serum Albumin): Exemplified by using a highly bioactive *Strypnodendron polyphyllum* fully bloomed flower extract.

6.1 Abstract

Over the last two decades, new and more advanced strategies that help to rapidly screen and identify new ligands for a specific macromolecule become more important. In this viewpoint, the effectiveness of two-dimensional STD NMR, Tr-NOESY, and STD-TOCSY has been utilized in evaluating the binding potential of the natural extract, used as herbal medicine in Brazil, towards human serum albumin. Moreover, the 1D-DOSY experiment has also been performed with the aim to provide a distinction of different compounds based on molecular weight. STD NMR, Tr-NOESY, and TOCSY analysis revealed that the presence of sugar-attached flavonoidal skeleton is involved in binding however, the complete structural assignment was possible by LC-SPE-NMR utilization. The combined results from NMR spectroscopy and separation methods provided myricetin-3-O-rhamnopyranoside (**1**), quercetin-3-O-glucopyranoside (**2**), quercetin-3-O-arabinofuranoside (**3**), and quercetin-3-O-rhamnopyranoside (**4**) as an active site blocker. Moreover, the epitope results and additionally Tr-NOESY cross peaks illustrated the edge protons are involved in strong binding through flattened conformations. Similarly, STD NMR competition studies with the complex of ligand/HSA and additionally, by varying the concentration of spy molecule that selectively binds with Sudlow´s site I, was demonstrated. The competition results revealed that these recognizable compounds do not selectively bind with a site I of HSA. Therefore, molecular docking studies were utilized for both Sudlow sites (I and II) to track the bound conformations and their molecular interactions. Interestingly, similar results were obtained from docking simulation as well. Thus,

the sequence of techniques can be considered as a simple and fast analytical tool to screen natural extract to get better leads against one specific target.

6.2 Introduction

The recognition of ligands prior to separation from a folk medicinal extract is a great deal for new drug discovery, especially, where the components are sensitive towards the solvents, light or some isolation methods. Nowadays, herbal extracts are being produced by various pharmaceutical methods for fulfillment of this growing need in medication, nutritional supplements (SCHAUSS, A. G., *et al.*, 2006), in dietary purposes as well as in cosmetics industries (MASIH, N. G.; Singh, B. S.,2012; BOULDIN, A. S., *et al.*, 1999). In global market, there is a great demand of drugs derived from natural products (herbal medicine) regardless of country or region (JONES, W. P., *et al.*, 2006). Therefore, the recognition of ligand of interest from a bioactive mixture is a key to new drug discovery, and more significantly, this emerging industry (herbal medicine) demanding high standard of production in more economical and environmental friendly new methods.

Since, herbal extracts are sensitive towards the environmental changes; hence, NMR seems to be an ideal choice that offers economically and environmentally healthy conditions as required by the drug discovery process (MEYER, B.; PETERS, T., 2003). Unlike other biochemical assays, NMR neither demands the prior knowledge about the function of protein and/or protein labelling nor target specific instrumental set-up (POLITI, M., *et al.*, 2005). On the other hand, the screening techniques (Mass spectrometry, Circular dichorism, Ultraviolet spectroscopy etc.) used to analyze organic mixtures suffer with a number of drawbacks like, lack of information regarding exchange proton (MAPLE, H. J., *et al.*, 2011), unable to resolve overlaps in signals (GARGALLO, J. J. a. R., 2012) and demanding compounds with optical absorption (CARTER, D. C.; HO, J. X., 1994), respectively. Although the aforementioned techniques are widely used tools in drug discovery but it is often not easy to develop a simple and rapid ligands screening method by using these techniques because of differences in molecular polarities, solubility, presence of isomers as well as the lack of optical absorption etc. In addition, the structure assignments of unknown compounds and their contribution towards the biological activity are also uncertain.

Since, for the last two decades, NMR has been demonstrated for the fingerprinting content of the complex mixture (BILIA, A. R., *et al*., 2001; TAGGI, A. E., *et al*., 2004) and also for screening ligands but, structure characterization remains a bottleneck (POLITI, M., *et al*., 2005) unless, the involvement of some separation methods. However, the use of hyphenation methods like, LC-UV-SPE-NMR can beautifully eliminate this characterization bottleneck particularly, where the NMR suffers in making distinctions among the signals due to overlap (JAROSZEWSKI, J. W. 2005; JAROSZEWSKI, J. W. 2005a). Very recently, Antonio G. Ferreira, and coworkers (TANOLI, S. A. K., *et al*., 2013) could supply an excellent way to find leads from crude, through the significant results by using direct NMR techniques following the docking simulation. Thus, this development has, however made NMR arguably the most important tool for recognitions of ligands from the complex natural product.

In the present study, we have employed STD NMR and Tr-NOESY experiments to study the binding of ligands from *Strypnodendron polyphyllum* fully bloomed flower extract towards HSA protein. To date, to the best of our knowledge, no single study that has been reporting the interactions of this highly bioactive plant towards any specific target. The screening of ligands from natural product was conducted by using direct NMR applications along with hyphenation method; we have selected an ethyl-acetate extract of *Strypnodendron polyphyllum* (plant commonly known as Barbatimão). In folk medicine, the extract of barbatimão was frequently used for treatment of leucorrhea, diarrhea, sepsis, inflammatory disorders as well as for blood clotting and wound healing purposes in Brazil (LOPES, G. C., *et al.*, 2005; CORRÊA, M. P.; FLORESTAl, I. B. d. D., 1921). The antibacterial and antioxidant activity of barbatimão extract is also reported (LOPES, G. C., *et al.,* 2003). Therefore, in the near future, this approach may open a way to new drug discoveries based on data from ligand/protein interactions, and can provide a better evidential proof of understanding the usefulness of NMR.

6.3 Results and discussion

Drug discovery is foremost important and perhaps the most challenging task to pharmacist and chemist. In most of cases, the synthesis of a medicinal chemistry oriented bioactive compounds is expensive and time consuming especially, where a complex molecule having several equivalent reactive positions. Nevertheless, molecular recognition directly from the bioactive material may somehow compensate this tough task. Thus, over the past decades, the screening of natural-product mixtures with receptors by using NMR as a detector has been extensively reported (CALLE, L. P., *et al*., 2001). Since 1999, the STD NMR has become one of the most widely used methods for the ligands recognitions (KLEIN, J., *et al.*, 1999; MAYER, M.; MEYER, B. 1999), analyzing molecular kinetics and for determining the dissociation constant (MAYER, M.; MEYER, B. 2001; ANGULO J., *et al.*, 2010) of bound state ligands within the protein complexes.

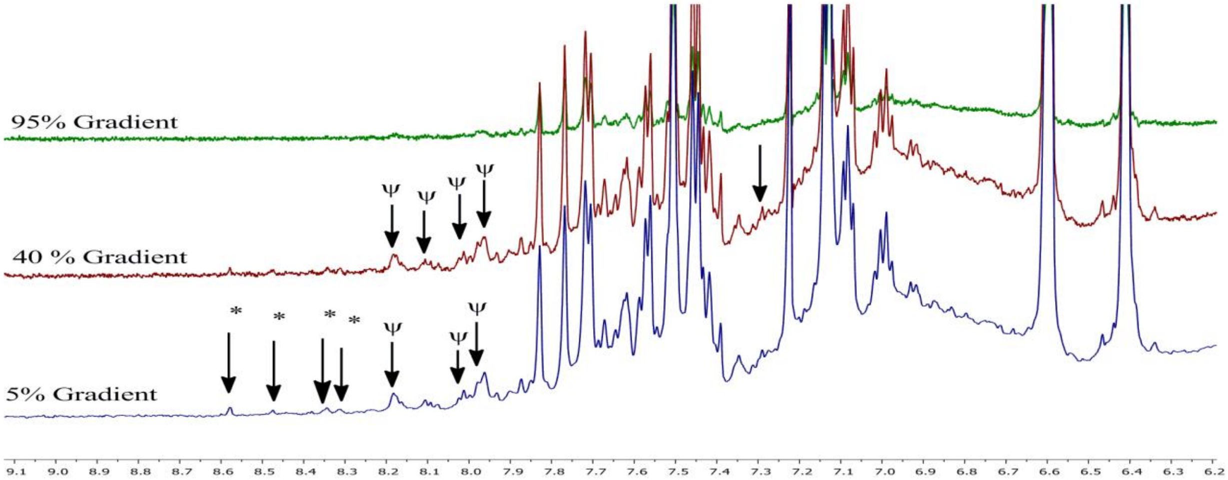


Figure 51. Comparison of the 1D DOSY spectra of extract, performed by using the big delta = 0.6 s, little delta = 1 ms, eddy current delays = 5 ms and with the following values of gradient strength: 95 % (a) 40 % (b) and 5 % (c). Where arrows with * and Ψ represent the small and medium size molecular weight compounds respectively.

Soon after the provision of screening strategies, and to have an idea regarding the size of different components within this fully bloomed flower extract, an alternate approach to bio-fractionation; a 1D-DOSY experiment was selected. It is well understood that the intensities of protons belonging to low molecular weight compounds are reduced or even completely wiped out by growing the gradient strength (POLITI, M., *et al*., 2006). Therefore, rather than utilizing bio-fractionation through the cut-off membrane, a series of 1D-DOSY analysis by varying gradient strength (5%, 20%, 40%, 80% and 95%) was employed to perform the fractionation based on molecular weight. Along these spectral results from the series of 1D-DOSY experiments, the assumption was made that, there at least two types of components are present as shown in figure 51. Change in signals is due to decrease in gradient strength from 40% to 5%, are indicated through arrows with asterisk (*) and psi (Ψ), providing an evidence of comprising almost similar component between δ = 7.9 to 12.4 ppm (Figure 51). The spectral results corresponding to gradient strength (20% and 80%) are not given because any perception of recognizable changes was not found in them. Therefore, after getting the realistic 1D-DOSY results in hand, we assumed that there are smaller to medium size components (full spectra Figure 60 in supporting information), and the decision about whether the interaction will originate from a smaller size or medium, the STD NMR conclusion will demonstrate later.

In the first step of this recognition process, screening was applied on a 1 mg of fully bloomed flower extract of *Strypnodendron polyphyllum* by taking help of STD NMR. First of all a ^{1}H-NMR in 5 µM of HSA (Figure 52 (A) was recorded that represents a ligands/protein complex system and act as a control, while, in Figure 51 (B) suppression of HOD signal was carried out to see the hidden anomeric protons under large HOD signal. Where, the given complete numbering was done by 2D NMR data (not shown here) thereby, separation methods. The STD NMR clearly showed (Figure 53) that aromatic as well as the sugar region of the spectra is involved in major bindings. Similarly, as an another evidence of binding the line broadening and increase in relaxation rate as a result of protein addition was also observed.

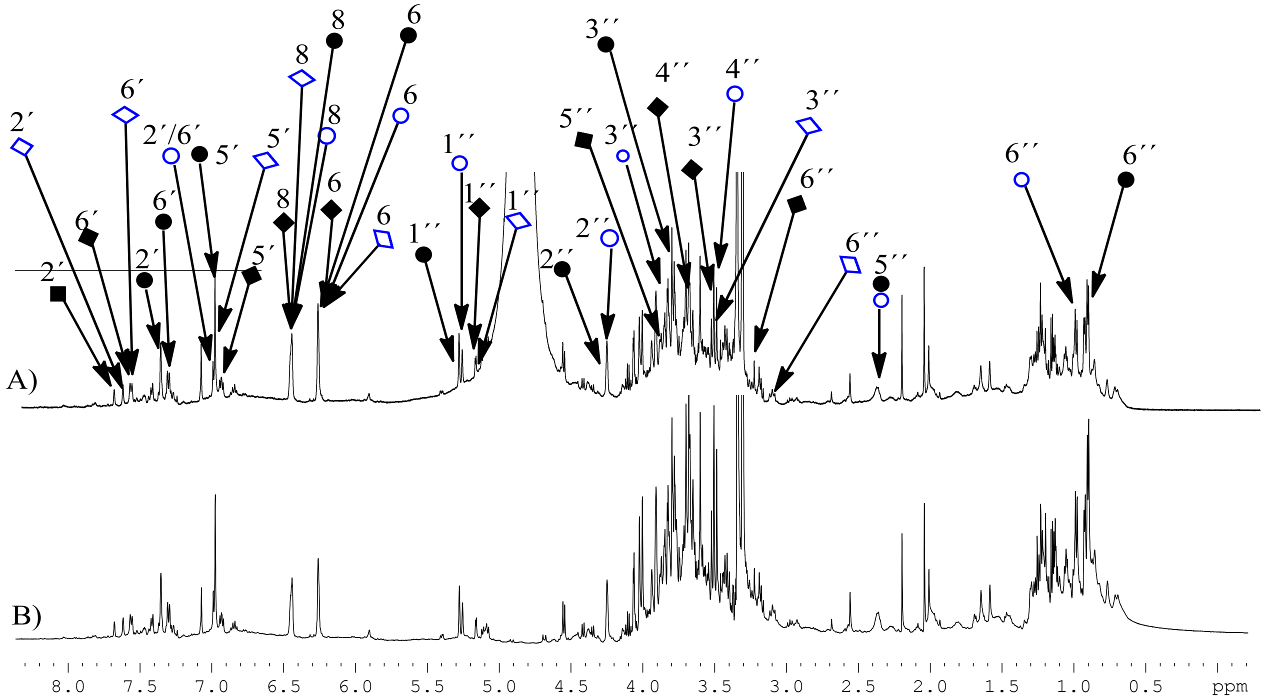


Figure 52. The ^{1}H-NMR spectra (A: Normal proton and B: with HOD signal suppression) of the ethyl acetate fraction of the fully bloomed flower extract of *Strypnodendron polyphyllum* plant in 50 µM solution of HSA protein acquired with Bruker 600 MHz AVANCE III spectrometer with a cryogenic TCI probe at 298 K . Where, signals were represented with the empty and filled square for compound quercetin-3-O-glucopyranoside (**2**) and quercetin-3-O-arabinofuranoside (**3**) and similarly, with an empty and filled circle for compound myricetin-3-rhmanopyranoside (**1**) and quercetin-3-O-rhamnopyranoside (**4**) respectively. Complete nomenclature numbering became possible after the evaluation of 1D and 2D spectra through separation method.

Figure 53, represented the proton NMR spectrum as a reference at the top, STD spectrum in the middle while the bottom spectrum is the magnified region (chemical shift δ = 3.0 to 8.2 ppm) of the STD spectrum in the dotted rectangle, for more clarity. The observations that come out from the figure are; some signals of impurities are present in the region δ = 2-2.5 ppm, background signals from protein, and almost all signals showed some level of enhancement, which indicates its good binding to HSA protein. Over the last decade, it has been known that the signals in STD spectrum represent the engagement in binding with some binding residues of receptor (MEYER, B.; PETERS, T. 2003; TANOLI, S. A. K., *et al.*, 2013; TANOLI, S. A. K., *et al.*, 2014). On the other hand, the absence of signals in STD spectrum represents no binding or no affinity towards respective receptor (MEYER, B.; PETERS, T. 2003; MAYER, M.; MEYER, B., 1999; MAYER, M.; MEYER, B., 2001).

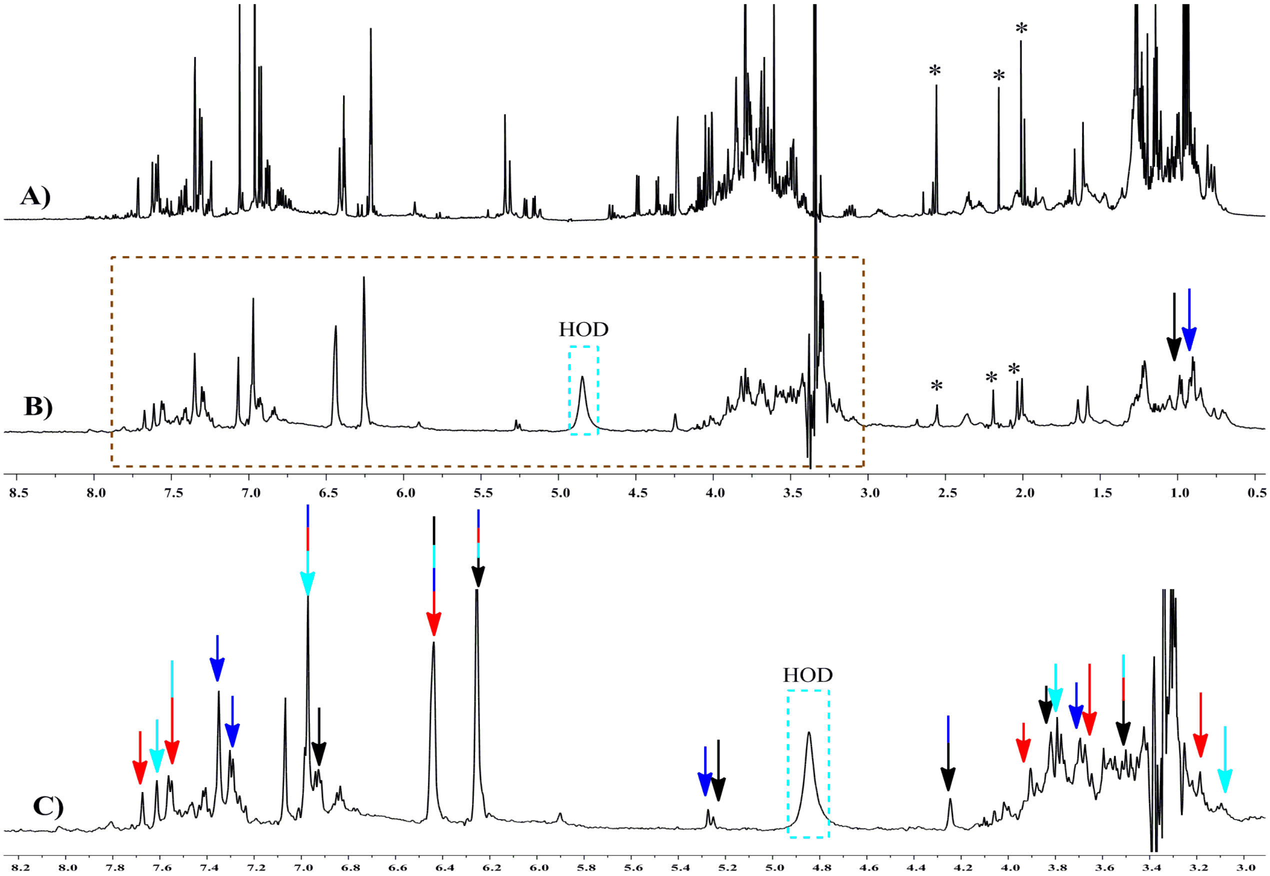
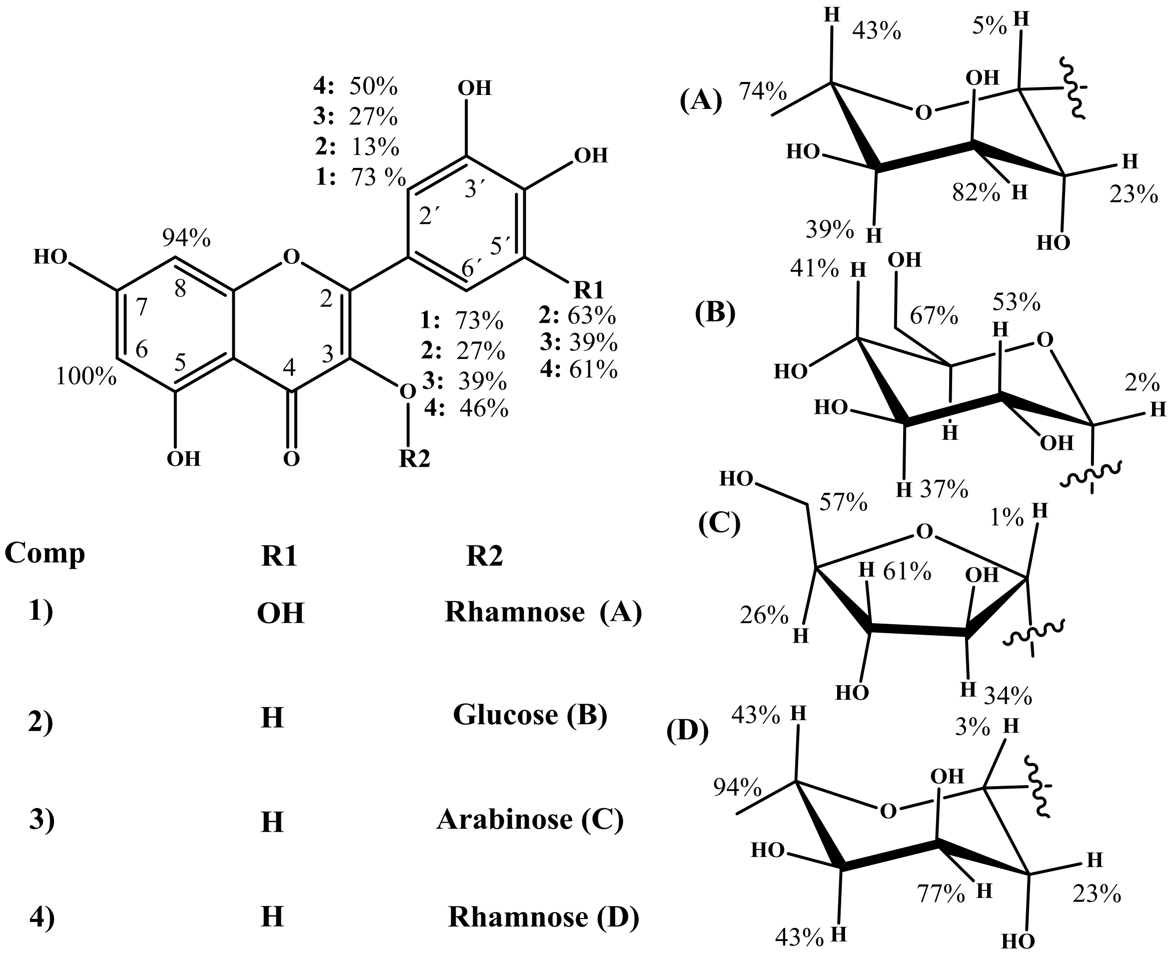


Figure 53. The ^{1}H-NMR with HOD signal suppression spectrum (A), STD NMR with water specrression spectrum (B) and the magnified region (δ = 3.0-8.2 ppm) of STD spectrum in a rectangle is shown in (C). All experiments were performed by using a 1 mg of extract in 50 µM protein solution and in a 3 mm NMR tube on Bruker 600 MHz spectrometer. To avoid the crowding as a result if numbering

sometime a single arrow with varying colors is shown, where red, sky blue, navy blue and black color represents the protons belonging to compound **3**, **2**, **1** and **4** respectively. Asterisk represents the signals from protein and some impurities. See more detail in experimental conditions.

The amplification gradient is the resultant of a saturation that received by the particular proton from the receptor, and it also determines the vicinity of ligand to the binding cavity. Thus, the intensity of the signal depends on the two factors, A) the length of saturation time, and B) the number of protons (TANOLI, S. A. K., *et al*., 2013; TANOLI, S. A. K., *et al*., 2014; MAYER, M.; MEYER, B., 1999). The STD build-up stack was obtained by performing experiments with varying saturation times (1s, 2s, 3s, 4s, and 5s), and the outcome stack model of spectra are shown in figure 61 supporting information. Group Epitope Mapping (GEM) calculations were made effective by comparing the individual proton integrals, and normalizing others with respect to this, which was assigned to 100% effect. After evaluation of series of experiment performed at a different saturation time, a saturation time of ca. 3 s was discovered more efficient where all signals were having excellent resolution and intensity by using Bruker stander pulse sequence with HOD suppression. The possible recognizable compounds with respective amplification factor after comparing the data from STD NMR, Tr-NOESY, TOCSY and the separation method are shown in Figure 54, where the sugar moieties of all compounds are represented separately. Intriguingly, the interaction footprints from these molecules revealed that the protons at position H-6 (6.26 ppm) and H-8 (6.44 ppm) showed singlet. Following these, the second large signals were obtained from sugar region. From these signals, it is obvious that the major stakeholders in interactions were from aromatic regional and then sugars moieties.

A 100% STD effect was obtained from the H-6 protons of all recognizable compounds, showing a major involvement in interaction that might be, due to large concentrations of compounds in this mixture and/or receiving larger saturation from the protein´s cavity through cross relaxation. However, the H-8 protons of these compounds contributed 94% STD effect. On the other hand, among other aromatic signals, position H-2´/H-6´ of myricetin-3-O-rhamnopyranoside (**1**) revealed taller with 73% STD effect, while the same protons from the Quercetin-3-O-glucopyranoside (**2)** provided between the 46% to 50% STD effect. So, the protons H-2´/H-6´ from compound **1** engaged in stronger binding as compared to same protons from compound **4**. However, the protons H-6´ and H-2´ from compound **2** and **3** revealed different amplification factors with 13% and 27 %; 27% and 39% respectively. Conversely, the proton position H-5´ remained prominent with 63%, 39% and 61% STD effect from the compound **2**, **3** and **4** respectively. From these aromatic epitope results, it is clear that the only sides of all recognizable compounds were towards the binding cavities of protein thereby receiving more saturation transfer, and hence the larger interaction.



Comp	R1	R2
1)	OH	Rhamnose (A)
2)	H	Glucose (B)
3)	H	Arabinose (C)
4)	H	Rhamnose (D)

Figure 54. The structures and the relative amplification factors (AF) in percentage with absolute numbering for the all-recognizable compounds: where the structure **1)** with substituent **A** representing myricetin-3-rhmanopyranoside, **2)** with **B** to quercetin-3-O-glucopyranoside, **3)** with **C** to quercetin-3-O-arabinofuranoside and **4)** with **D** to quercetin-3-O-rhamnopyranoside respectively. For more clarity of the amplification factors in attachments, all sugars were represented separately.

Similarly, the epitope results remarkably showed that both rhamnose units could provide better interaction compared to glucose and arabinose unit. In addition, all the axial protons from these molecules provided greater amplification and had good interaction towards HSA protein, while among the equatorial protons; H-3´´ from compound **1** and **4** could provide the noticeable STD effect with 82% and 77% respectively. Second higher STD effect was found from sugar moiety of compound **1**, H-4´´ and H-5´´ provided well-neigh effect of 39% and 43% represent its moderate interactions. While the similar protons from compound **4**, could provide almost same effect of 43%. By combining results of compound **1** and **4** for position H-4´´ and H-5´´, interacting behavior with HSA was found to be similar. However, the signal for H-4´´ from compound **2** and **3** also provided the STD effect of 41% and 26 % respectively. On the other hand, H-5´´ for both compounds showed a bit higher binding with 57% and 41% STD effect. Interestingly, the anomeric protons from all compounds showed a negligible STD effect between 1-5%. From these epitope mapping it can be concluded that the part of sugars that are towards flavones (aglycone) skeleton showed less binding compared to the part far from aglycone. Thus, these epitope results revealed by the evidence of interaction from both edges are in intimate contact with the binding sites.

Furthermore, Tr-NOESY experiments were performed to identify the bound conformations of these STD-exposed interacting molecules within the HSA-ligands complex. As, Tr-NOESY experiment is best-known experiments for the evaluations of bound conformation for the ligand-receptor complexes at equilibrium, where, the bound ligand's geometric information are transferred to unbound (MEYER, B., *et al*., 1997; ALBRAND, J. P., *et al*., 1979). In NOE, the major distinguishing parameter between the bound and unbound ligands is the correlation time ($\tau_{c)}$; small molecules posses short τ_c and longer relaxation rate while, in bound state this shorter correlation time (τ_c) converted to the longer, a characteristic of large mass molecules, results in a strong NOE. Similarly, the Tr-NOESY build-up rate in solution is also much faster if compared to NOESY (NI, F., 1994). In 2D NOESY without the addition of protein, all cross peaks were having positive signs comparative to diagonal (Figure 55 A), however, in the sugar region some signal also revealed negative signs that might be due to the larger size or uncorrected phase. Nevertheless, in Tr-NOESY all peaks represent same significance to the diagonal, and more importantly, the buildup rate was much faster when compared to NOESY without protein, an evidence of binding to protein (see in Figure 55 B). However, the mixing time for both NOE experiments were kept the same value of 400 ms. In the next step, the bound conformations of these compounds within the ligandreceptor complex were studied.

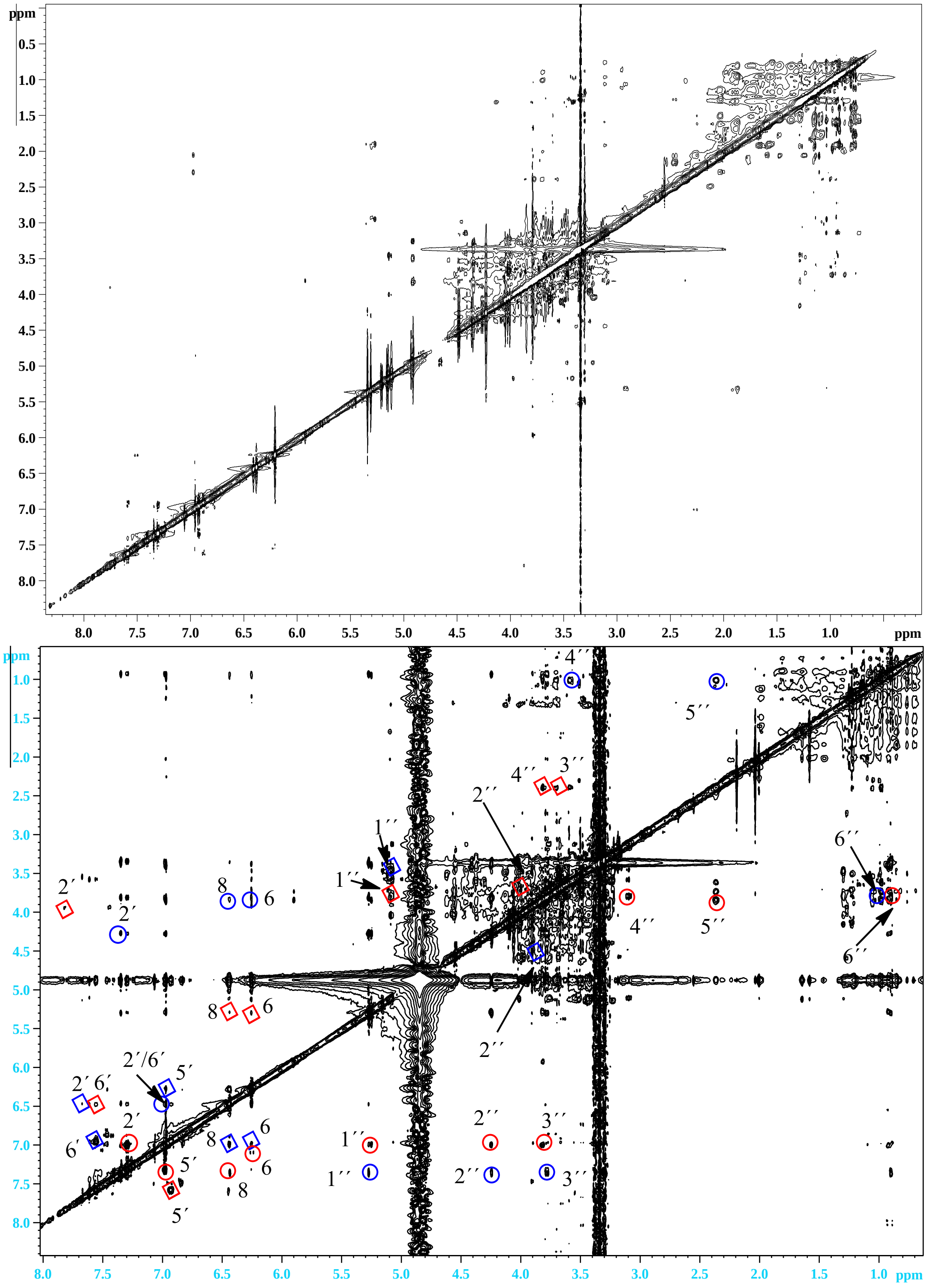


Figure 55. Two-dimensional NOESY spectrum without HSA protein (A) and Tr-NOESY (B), with Bruker standard phase sensitive pulse sequence for the extract in D_2O–methanol-d_4 (85 : 15% v/v) and protein in PBS in buffer solution, with pH 7.4 (not corrected for D_2O) respectively. For the Tr-NOESY, the pulse sequence was

modified by adding the 3500 Hz spinlock for the removal of background protein signals before the first π/2 pulse. Both spectra were acquired by applying a 400 ms mixing time after evaluating a number of mixing time spectra at the Bruker 600 MHz NMR spectrometer with a cryoprobe, at 298 K. In NOESY, the cross peaks observed was with opposite to the diagonal while for Tr-NOESY spectrum both had the same sign. Again, the signals were represented with the blue and red square for compound **2** and **3** and similarly, blue and red circle for compound **1** and **4** respectively.

Since, the protons H-6´´, H-5´´ , H-4´´ and H-3´´ from compound **1** and **4** were not present in NOESY experiment however, in the Tr- NOESY experiment these were very prominent (Figure 55 B). Actually, the flexibility of these protons has been much reduced in the bound conformation. Thus, the short correlation time of these protons were no longer exist in the bound state thereby interacting with binding cavities and hence afforded the strong negative NOE results. Similarly, in the aromatic region, the strong signals from the protons H-6 and H-8 were also absent but they come into sight due to binding with protein in Tr-NOESY experiment. Interestingly, some protons from the glycone part of compound **2** and **3** like H-2´´ and H-3´´ also showed the strong negative NOEs in the Tr-NOESY spectrum. It is worth noting that, the signals from the anomeric protons were present in both spectra (Figure 55 A and B). The Tr-NOESY experiments mimic the epitope mapping STD results by providing the similar results. Therefore, by the collective evidence of both STD and Tr-NOESY experiments, we argue that these portions of small molecules (**1**, **2**, **3** and **4**) were engaged in binding and the bound conformations were assumed at edges of each molecule gave strong binding. Therefore, by taking compound **1** as a reference, the bound conformation has been suggested in figure 56 A, B and C.

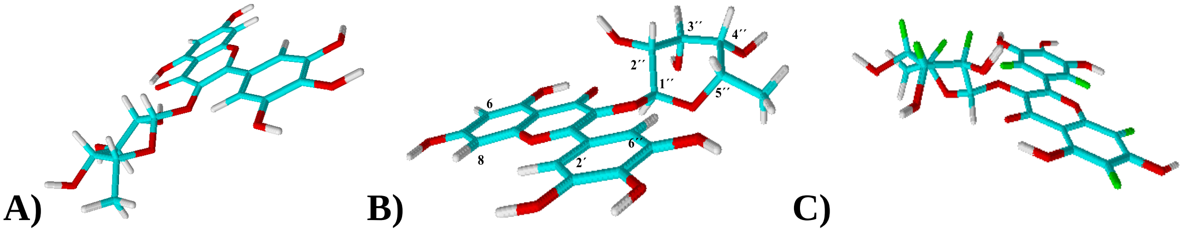


Figure 56. The suggested conformations obtained after the STD amplification factor calculations and Tr-NOESY cross peak information in the bound state within the solution of HSA and myricetin-3-O-rhamnopyranoside complex. In all conformations (A, B, and C), the structure is in extended form where, in B numbering and in C green color edge protons are shown representing the strong bindings.

Further confirmation about the structures and their bindings, the same sample (with protein) was considered for 2D TOCSY and STD-TOCSY experiments. The 2D TOCSY experiment is well known for complex structure identification and with the added advantages, ligands proton proximity to the neighboring binding sites of receptor (Haselhorst, T., *et al.*, 2007; Mayer, M.; Meyer, B., 1999). Signal intensities were the principle source for STD-TOCSY establishments; by comparing the *off*-resonance to the STD spectrum, more intense signals were considered a more intimate contact with protein binding sites, while, the less intense signals were more solvent exposed. The entire protons those were present in the STD NMR spectrum experienced the same saturation, resulting in strong cross peaks. Interestingly, the aromatic region (Figure 57 A and B) as well as the glycone part's (Figure 62 supporting information) protons, showed strong cross peaks that could be attributed to closest approach to binding sites in this ligandprotein complex.

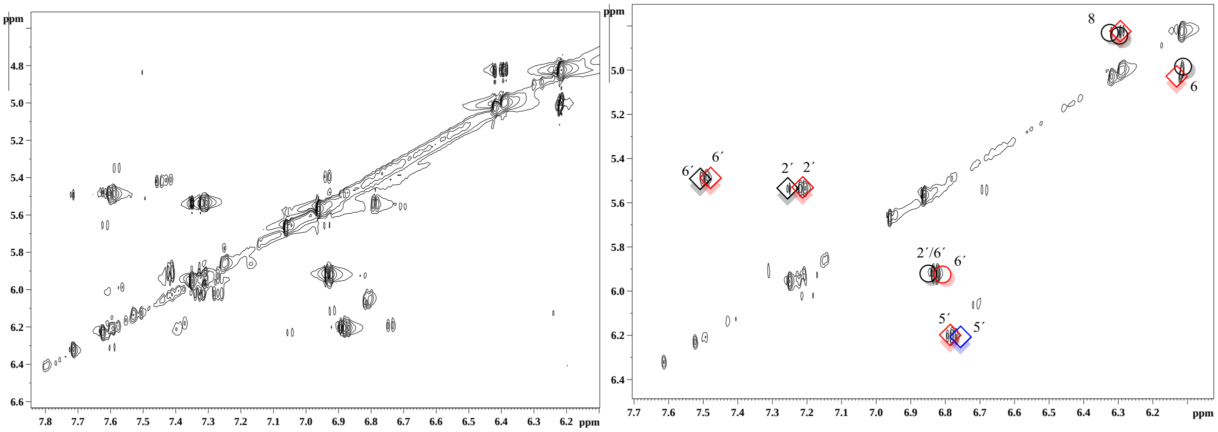


Figure 57. Two-dimensional standard TOCSY (A) and STD-TOCSY spectrum (B) acquired by using the 50µM solution of HSA in 1mg extract solution in methanol-d_4 and D_2O (15: 85 v/v) respectively. MLEV-17 sequences were employed to acquire these experiments by adding a mixing time of 60 ms at 298 K with a 600 MHz spectrometer. In STD-TOCSY, the parameters for the saturation of protein were kept same as used earlier for STD experiments. The magnified region (δ = 6–7.8 ppm) of both standard TOCSY and STD-TOCSY was shown and signals are represented with the blue and red square for compound **2** and **3** and similarly blue and red circle for compound **1** and **4** respectively.

Among aromatic protons of recognized compounds, H-6´ and H-5´´ of compound **2** and **3 (**small square having blue and red color), were prominent, and they could produce an evidence of binding in STD-TOCSY. Similarly, H-2´ from compound **1** and **4** (small circle having blue and red color respectively) and protons H-2´/6´ from compound **1,** provided an excellent agreement with STD spectrum and hence, interactions (Figure 56 B). Intriguingly, signals from H-6 and H-8 positions from these compounds (**1**, **2**, **3** and **4**) exactly mimic the STD and Tr-NOESY spectral results of having strong cross peaks, thereby, strong contact with binding sites. Thus, 2D STD-TOCSY again provided the strongest evidence of edge contact of recognizable ligands. In addition, the methyl signals of rhamnosyl sugar from compound **1** and **4** provided the same results (Figure 62 supporting information). In STD-TOCSY results, some signals from 2D TOCSY spectrum were missing, might be due to receiving a less degree of saturation from the protein because of the large distance from binding site or shallow site.

Up till now, on the premise of STD NMR, Tr-NOESY and STD-TOCSY experiment, the structural skeleton was almost clear and provided the evidence of the presence of a flaovonoidal type skeleton with attached sugars, but to guarantee about the absolute structural assignments was not clear. However, there was a substantial need of some separation method that could lead to fully characterization and a certainty about number of components involving in their binding. Thus, the obvious choice was the LC-SPE-NMR system which not only enables to characterize the structure but also helpful in providing a number of components involving as well (JAROSZEWSKI, J. W., 2005a; MARSTON, A.; HOSTETTMANN, K., 2009). To accomplish this purpose, 3 mg of the extract was dissolved in 1 ml of methanol–water (85: 15 % v/v) respectively, afterwards, a separation method having less run time and high resolution was observed (Figure 63 supporting information). By using the gradient solvent with varying concentrations and with the injection of 15 µl of sample with a flow rate of 1 ml/min, led to the chromatogram (see figure 64 supporting information) comprising 13 peaks (four major and 9 minor). After 30 times of trapping the major four compounds on the general phase cartridge of SPE system, each peak eluted to NMR probe by using the 230 µl of CD_3OD. All of trapped compounds provided very clear 2D NMR spectra (COSY, HSQC and HMBC), after match up to this spectral data with earlier reported results (KAWASHTY, S. A., *et al.*, 2000; TOM J., *et al.*, 1970; RODRIGUES, E. D., *et al.*, 2009), led the conclusion that the compounds myricetin-3-O-rhamnopyranoside, quecetin-3-O-glucopyranoside, quercetin-3-O-arabinofuranoside and quercetin-3-O-rhamnopyranoside as compound **1**, **2**, **3** and **4** respectively.

6.3.1 STD NMR titration studies

After getting the information related to the bound conformations of these interacting ligands through STD NMR and Tr-NOESY towards the HSA, the following step was conducted to determine the binding sites for each ligand. In this perspective, a competitive ligand with well-known and fully described binding site (so-called a spy molecule) was obviously necessary to make a comparison with the unknown molecule for the corresponding site. To assess this goal, we have been applying the STD NMR competition studies as it is well known, based on the signal intensities, the STD competition studies can validate whether this spy and unknown ligand compete for the same binding site of protein or not. The addition of a spy molecule in a solution containing other ligand either can reduce or completely eliminate the signals of that ligand, in case, if this spy is a better binder to this specific site. On the other hand, if this ligand binds itself tighter than this spy does, results the reduction or complete elimination of spy signal in STD experiment. Despite these possibilities, if both compounds having different binding sites, none of the signal going to be effected.

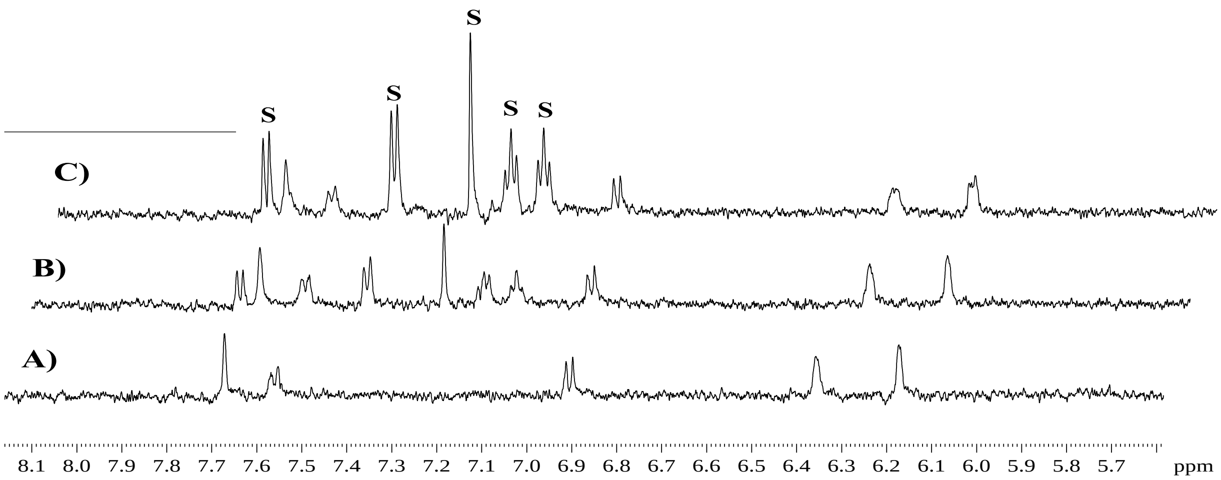


Figure 58. STD competition experiments showing only the aromatic region (δ =5. 7-8.1ppm) recorded with a Bruker 600 MHz spectrometer at 298K. A) STD spectrum of the HSA-quercetin-3-glucoside complex (HSA 50 µM), B) STD spectrum of the complex with HSA vs. spy molecule (1: 100 molar excess) and compound 2, C) STD spectrum of the complex system with HSA vs. spy molecule (1: 350 molar excess) and compound **2**. Protons from the spy molecule (L-tryptophan) are represented by S, whereas, quercetin-3-O-glucopyranoside is un-numbered.

Thus, by using STD NMR competition studies L-tryptophan was used as a reference (spy molecule) that preferably binds to the Sudlow's site II of human serum (CONRAD, M. L., *et al.,* 2009), and extensively applied for the site-selective competition studies (SUDLOW, G., *et al*., 1976; FEHSKE, K. J., *et al*., 1982). In this study, the unknown amount of trapped four natural products through SPE system were further utilized by making the sample of 1:100 and 1: 350 molar solutions of protein to L-tryptophan respectively, for each single STD experiment after having 1 and 2D spectra. All these competition experiments were performed by using the fixed concentration of HSA (50 µM) for every ligand sample. From the close observation of the spectra acquired from subjected ligands, it could be presumed that the STD intensities of these ligands protons remained same although by increasing the molar excess of the spy molecule (Figure 58 and Figure 65-67 supporting information). These outcomes of the competition studies strongly put forward that no ligand targets the selected site of HSA, and subsequently binding with some other site. Due to lack of sufficient quantities (from separation system), further STD competition experiments for the hunt of Sudlow site I was not performed with other spy molecules, selectively bind to the same locus. Thus, rather exploring the site selectivity of these ligands towards the Sudlow´s site I through STD competition experiments, further observation was made through the docking simulation and molecular dynamic studies.

6.3.2 Docking simulation

With a specific end goal to investigate the binding sites for the recognizable ligands, the next question was whether the ligands will prefer the Sudlow´s I or II for their best-fit accommodation. However, both sites (Sudlow´s I and II) were used here for docking simulation and compounds were analyzed individually within both binding sites for their binding potentials. Molecular docking simulation is a computational strategy used to tie up the ligand with targeted protein in order to identify their interaction patterns, binding affinities as well as their mechanism of action. Here, to uncover the binding and inhibitory capability of these sugars containing flavonoidal skeleton towards human serum albumin protein, the targeted amino acid residues in order to identify drug kinetics, for the present study are given in the experimental section. Each compound was individually analyzed for Sudlow’s site I and II. MOE showed that, compound **1** strongly interacted through its sugar hydroxyl group at 2.08 Å through NH of Arg257. Another quite weak hydrogen bond of 3.31 Å was also observed in between Arg222 NH and OH of ligands resorcinol shown in Figure 59. Similarly, compound **2** (Figure 68a supporting information), showed more intense interacting capability with three strong hydrogen bonds and highest binding free energy of -7.3211 Kcal/mol presented in table 11.

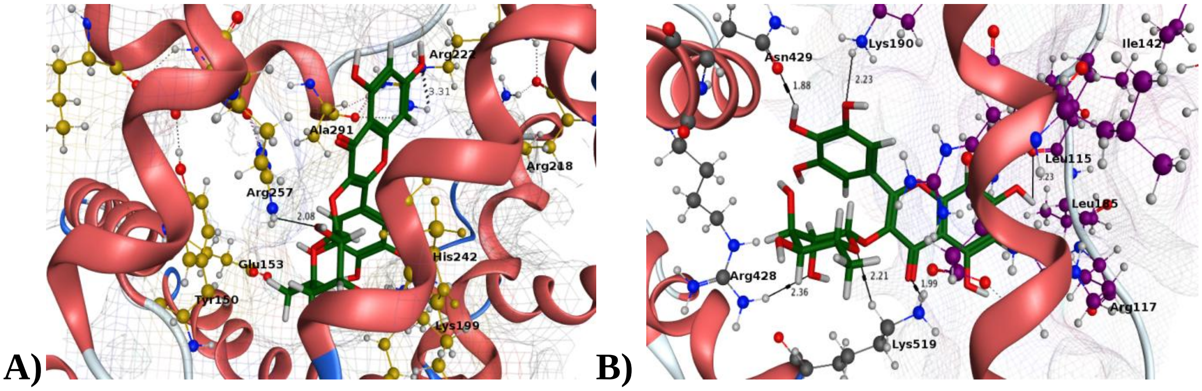


Figure 59. 3D bound conformations of the myricetin-3-O-rhamnopyranoside, generated by Molecular Operating Environment (MOE) 2011.10 within the HSA-myriscetin-3-O-rhamnopyranoside complex for the Sudlow´s site I (A) and site II (B). Where, the PDB file for HSA complex with amino camptothecin was taken from the protein data bank with ID: 4L8U. The best-fit accommodation of this structure after 30 different pose evaluations provided almost similar results as deduced earlier through STD-AF and Tr-NOESY studies.

These bonds mediated from the carboxylic oxygen of Glu153, guanidine NH of Arg257 and guanidine NH of Arg218 towards hydroxyl hydrogen and hydroxyl oxygen of sugar and hydroxyl oxygen of resorcinol moiety at two extremes of ligand molecule. However, compound **3** (Figure 69a supporting information) also accommodates itself tightly within the binding groove of the site I through the same number of hydrogen bonds with a binding free energy value of -7.2731 Kcal/mol. Conversely, compound **4** behaved quite differently and showed binding energy of -6.9405 Kcal/mol (table 11) via interacting with Asp451 and Ser192 (Figure 70a supporting information).

Site I				
Compounds	**S**	**E-score1**	**E-refine**	**E-score2**
Compound1	-6.969235	-15.22145	-17.91924	-6.969235
Compound2	-7.32114	-12.14727	-27.00223	-7.32114
Compound3	-7.27315	-12.33382	-25.02071	-7.27315
Compound4	-6.940517	-13.40245	-23.50109	-6.940517

Table-11: MOE docking and molecular dynamic results for the compounds (**1-4**) within the binding site1 of human serum albumin with their free binding energies (Kcal/mol).

Site II				
Compounds	**S**	**E-score1**	**E-refine**	**E-score2**
Compound1	-6.483576	-12.29182	-20.07261	-6.483576
Compound2	-6.516641	-11.81531	-8.509714	-6.516641
Compound3	-6.437744	-11.48641	-20.51512	-6.437744
Compound4	-6.61349	-11.05528	-17.81524	-6.61349

Table-12: MOE docking and molecular dynamic results for the compounds (**1-4**) within the binding site II of human serum albumin with their free binding energies (Kcal/mol).

On the other hand, when the same compounds were screened against the site II, quite different binding affinity pattern was revealed as shown in Table 12 with different conformations. Compound **1**(Figure 59b) placed itself at the edge of the defined binding site (purple) and interacted mostly with outside residues (gray). It interacted strongly through hydrogen bonds with 1.88, 1.99 and 2.36 Å originated from Asn429, Lys519 and Arg428 towards the OH of sugar. Furthermore, active site Lys190 and Leu115 are also interacted via hydrogen bond with a distance of 2.23 and 3.23 Å to the ligand to support its binding at the edge of protein's binding pocket. Orientation of compound **2** (Figure 68b supporting information) within the protein’s binding pocket, ensured the correct placement of molecules, surrounding its hot spot residues. Arg117 and Asp183, two important active site residues, formed hydrogen bonds of 1.92 and 2.22Å with hexose sugar oxygen and catecholic OH groups of compound, depicted in figure - 69b supporting information. Compound **4** presented in figure -70b supporting information also docked within binding groove catches protein's hot spot residues; Arg114 (2.00Å), Tyr161 (1.98Å), Val116 (1.95Å) and Arg114 (1.92Å) for hydrogen bonding, respectively.

6.4 Conclusions

In this study, direct NMR spectroscopic methods are suggested as a tool for the rapid and conclusive characterization of ligands recognition from the natural extract in the chase for new leads. Here, the combined use of STD-NMR, transfer NOESY experiments, and molecular docking furnished details on the molecular recognition of the sugar containing flavonoidal ligands, specifically, it delivered details of bound conformations of binding ligands in the complex. Moreover, similar STD NMR experiments were performed on ligand/HSA complex and the spy molecule, which were designed to observe competition for the Sudlow´s site II. It is worth noting that, the combined results from the STD competition experiments and docking protocol are in good agreements, and has empowered us to derive that these compounds do not compete for the same binding site as L-tryptophan, and have more affinity towards site I. Moreover, the number of hydrogen bonding with active site residues ensured their inhibitory potency and confirmed the stability of protein-ligand complexes. Compound **2** and **3** were found to be most active from analysis of their docking results. Similarly, the epitope mapping results were also in close agreement with the bound conformation results of molecular docking, and revealed a collective statement that these structures remain more flattened instead of twisting in the complex and showed a greater affinity from their edges. Therefore, the structural information gleaned from studying this ligand/protein interaction may facilitate new drug design development and better ligands/inhibitor for HSA protein. Moreover, this approach can also be extended to obtain a fingerprint for the quality promises and efficacy of crude extract.

6.5 Supplementary information

Direct NMR mediated ligands screening for a specific target (Human Serum Albumin): Exemplified by using a highly bioactive *Strypnodendron polyphyllum* fully bloomed flower extract.

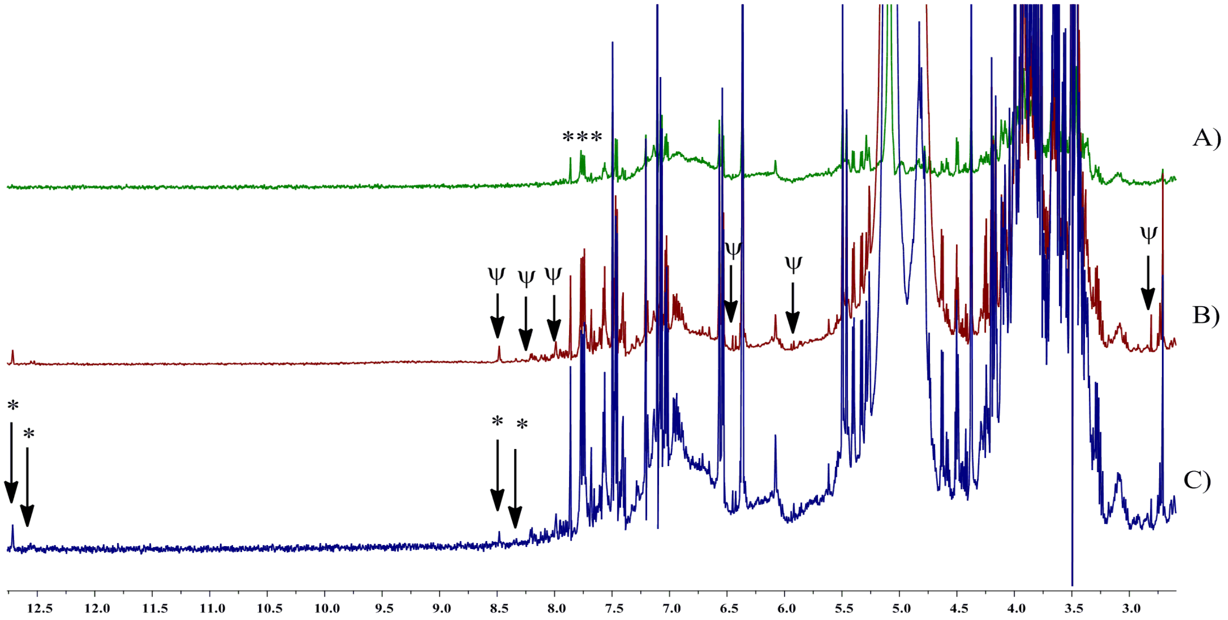


Figure 60. Comparison of the 1D DOSY spectra of extract, performed by using the big delta = 0.6 s, little delta = 1 ms, eddy current delays = 5 ms and with the following values of gradient strength: 95 % (a) 40 % (b) and 5 % (c). Where, arrows with * and Ψ represent the small and medium size molecular weight compounds respectively.

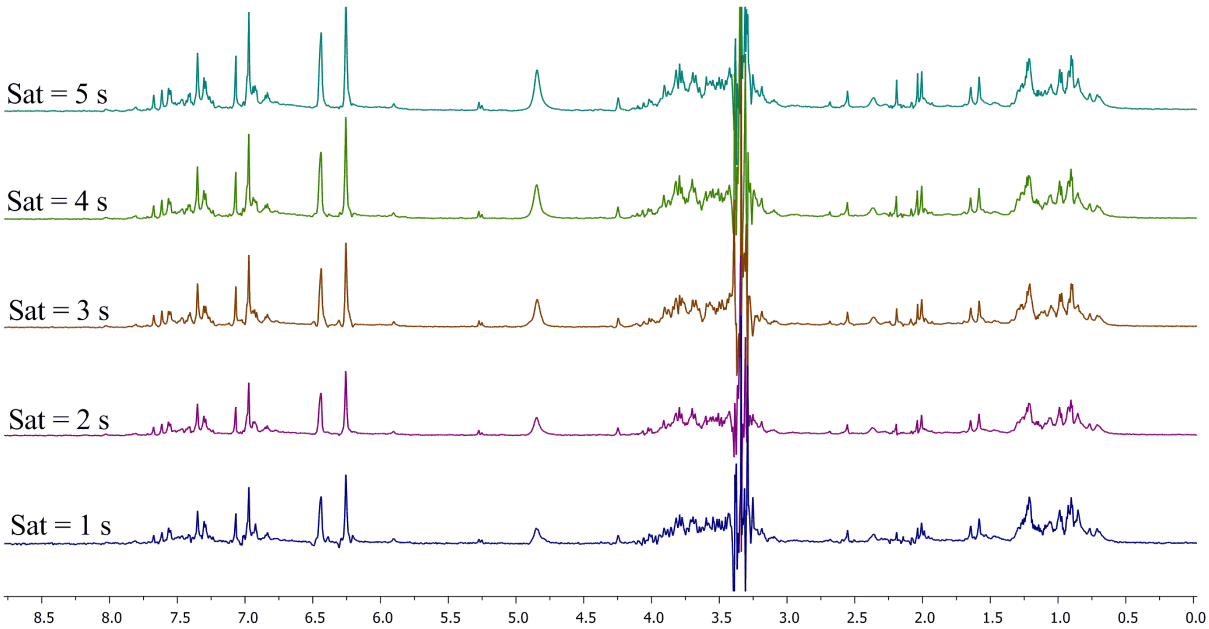


Figure 61. Stack plot of the STD NMR spectra, representing the five STD experiment with varying saturation time (1-5 s). All experiments were performed by using a 1 mg of extract in 50µM protein solution and in a 3 mm NMR tube on Bruker 600 MHz spectrometer at 98K. By increasing the saturation time of protein the intensity of the signal, increase gradually.

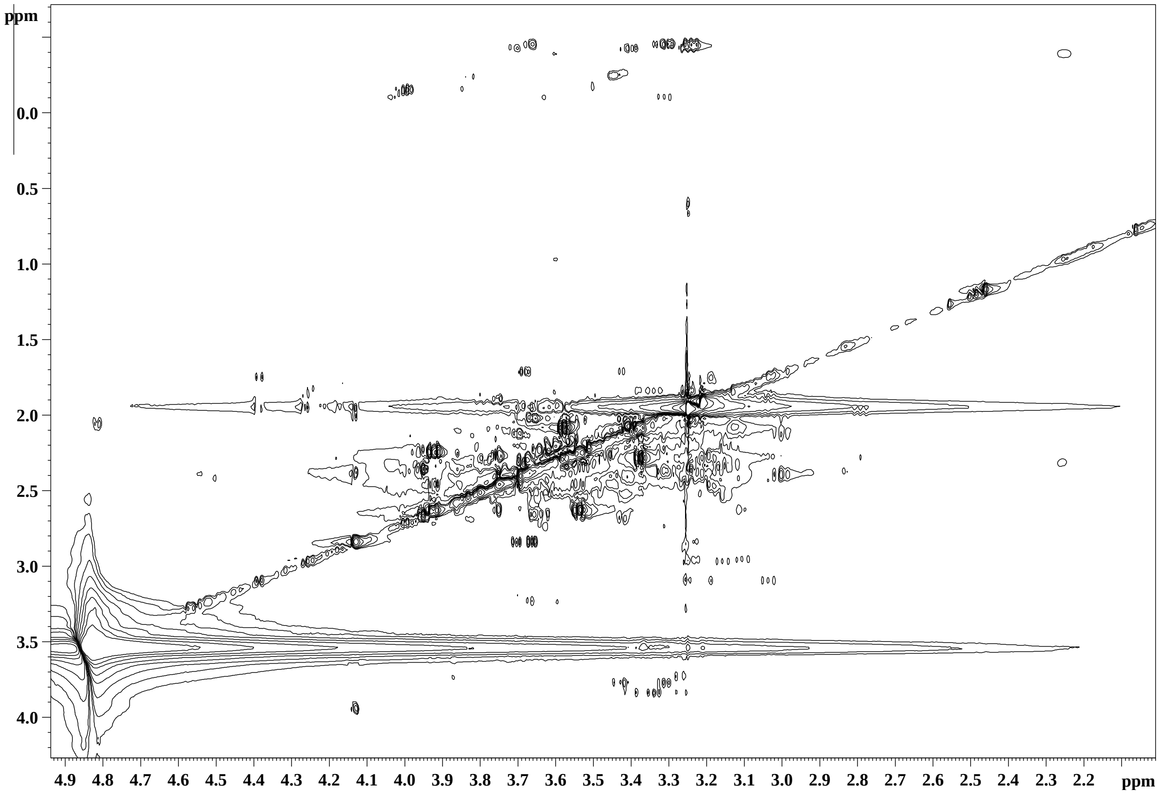


Figure 62. Two-dimensional STD-TOCSY acquired by using the 50µM solution of HSA in 1mg extract solution in methanol-*d*$_4$ and D_2O (15: 85 v/v) respectively. The MLEV-17 sequence was employed to acquire this experiment by adding a mixing time of 60 ms at 298 K with a 600 MHz spectrometer. However, the parameters for the saturation of protein were kept same as used earlier for STD experiments. The magnified region (δ = 2.2–4.9 ppm) has been shown here for more clarity of the binding outcomes provided by glycone moieties of recognizable compounds, where, the complete assignment can be seen in figure 5(B) in the main manuscript.

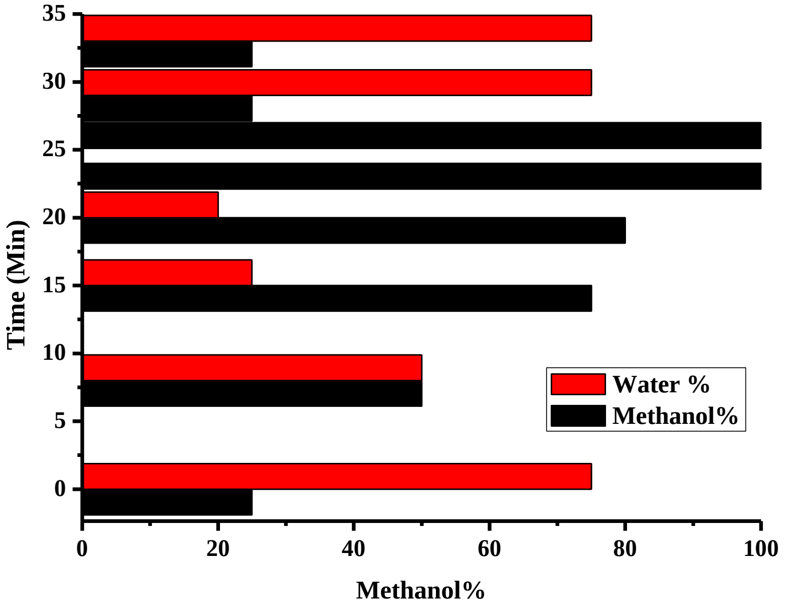


Figure 63. Separation method used to get the high resolution with maximum number of peaks by utilizing the solvent gradient system. The total run time added to 35 minutes with the gradient solvent system utilized (25:75, 50:50, 75:25, 80:20, 100:0, 100:0, 25:75 and 25:75 %) for the methanol and water respectively for the time 0, 8, 15, 20, 24, 27, 29, 35 minutes correspondingly. See the experimental section of the column specifications and other information.

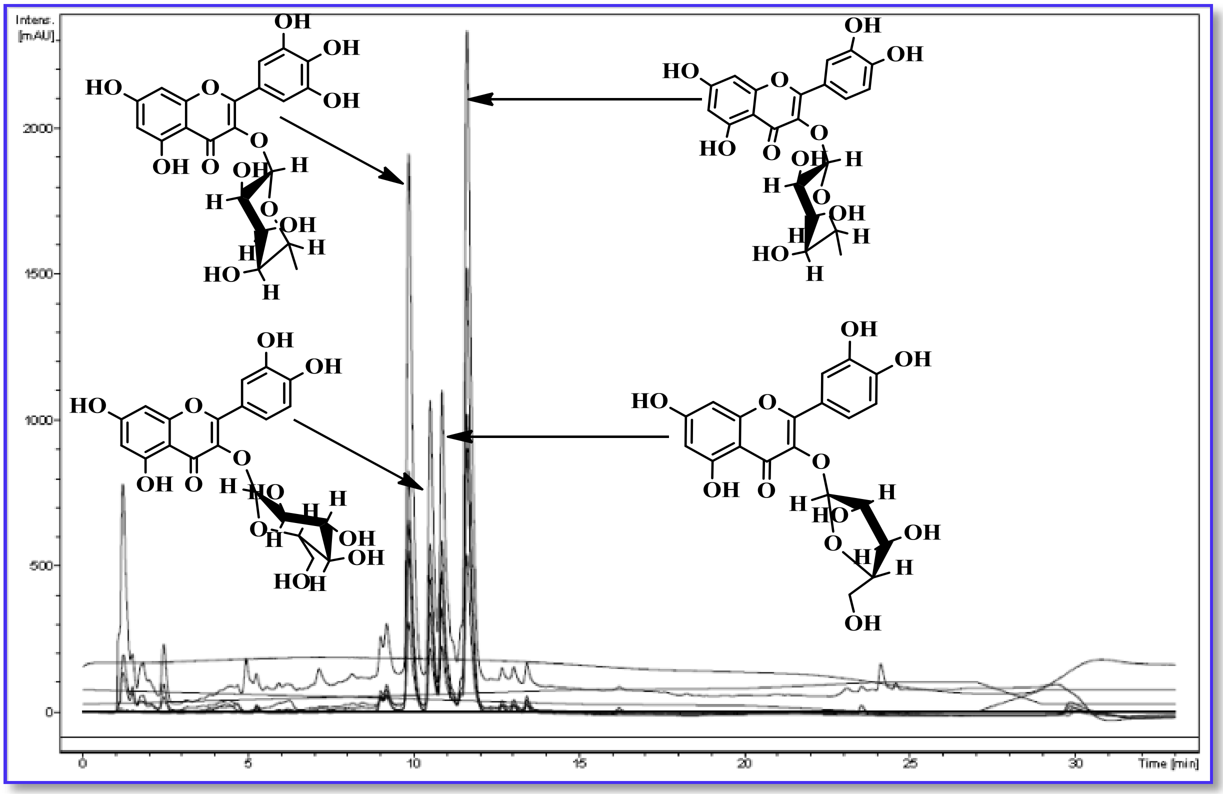


Figure 64. LC-chromatogram obtained as a result of injecting 15 µl of the extract sample with the flow rate of 1 ml/min whereas, the sample was prepared by using 3 mg extract in methanol and water (85: 15 v/v, respectively). Moreover, arrows also represent the compounds obtained from the respective chromatographic peaks.

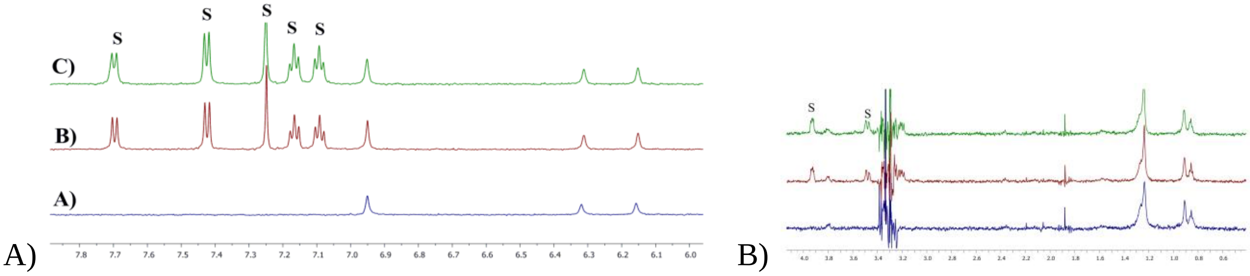


Figure 65. STD competition experiments showing only the aromatic region (**A)** and non-aromatic region (**B)** recorded with a Bruker 600 MHz spectrometer at 298 K. A) STD spectrum of the HSA- myricetin-3-O-rhamnopyranoside complex (HSA 50 μM), B) STD spectrum of the complex with HSA vs. spy molecule (1: 100 molar excess), C) STD spectrum of the complex with HSA vs. spy molecule (1: 350 molar excess) and compound **1**. Protons from the spy molecule (L-tryptophan) are represented by S, whereas, myricetin-3-O-rhamnopyranoside is un-numbered.

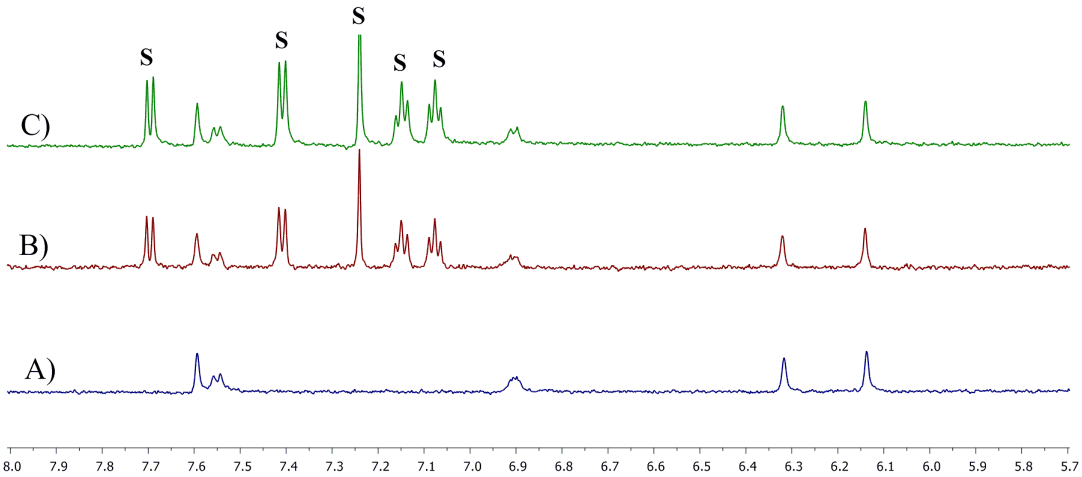


Figure 66. STD competition experiments showing only the aromatic region (A) and non-aromatic region (B) recorded with a Bruker 600 MHz spectrometer at 298 K. A) STD spectrum of the HSA- quercetin -3-O-arabinofuranoside complex (HSA 50 μM), B) STD spectrum of the complex with HSA vs. spy molecule (1: 100 molar excess) and compound **3**, C) STD spectrum of the complex with HSA vs. spy molecule (1: 350 molar excess) and compound **3**. Protons from the spy

molecule (L-tryptophan) are represented by S, whereas, quercetin -3-O-arabinofuranoside is un-numbered.

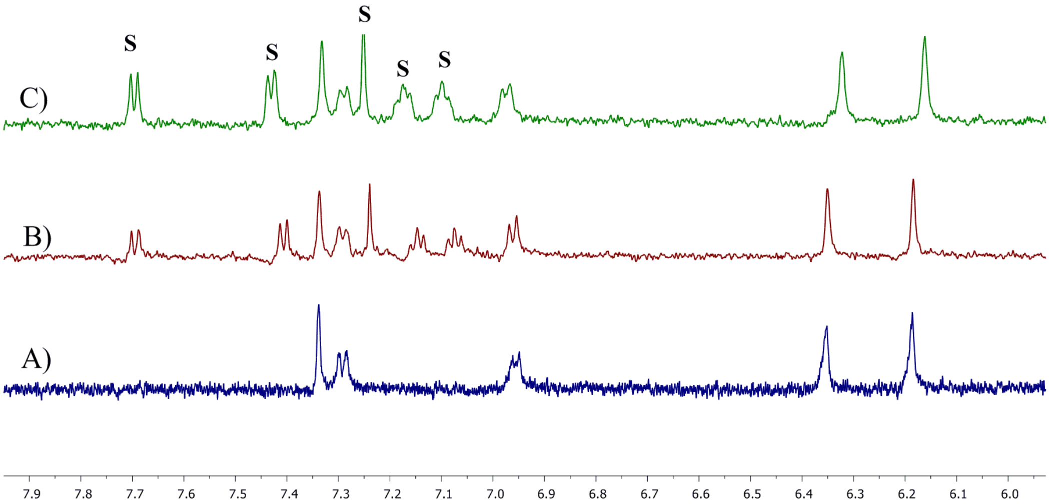


Figure 67. STD competition experiments showing only the aromatic region (A) and non-aromatic region (B) recorded with a Bruker 600 MHz spectrometer at 298 K. A) STD spectrum of the HSA- quercetin-3-O-rhamnopyranoside complex (HSA 50 µM), B) STD spectrum of the complex with HSA vs. spy molecule (1: 100 molar excess) and compound **4**, C) STD spectrum of the complex with HSA vs. spy molecule (1: 350 molar excess) and compound **4**. Protons from the spy molecule (L-tryptophan) are represented by S, whereas, quercetin-3-O-rhamnopyranoside is un-numbered.

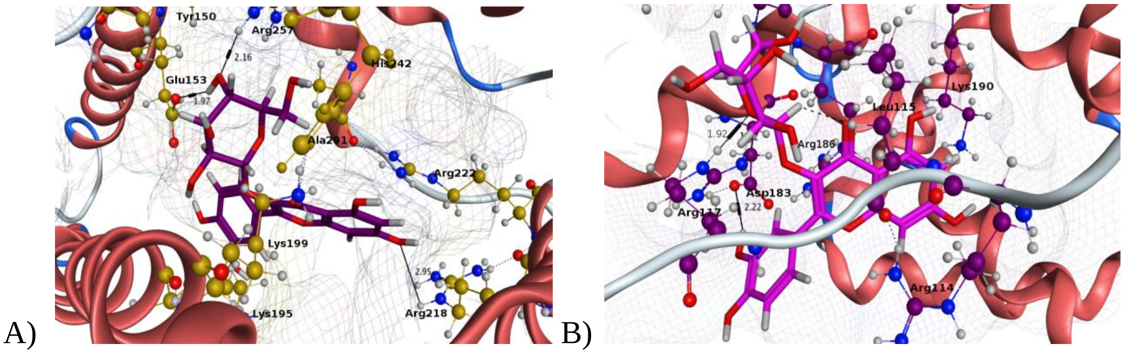


Figure 68. 3D bound conformations of the quercetin-3-O-glucopyranoside, generated by Molecular Operating Environment (MOE) 2011.10 within the HSA-quercetin-3-O-glucopyranoside complex for the Sudlow´s site I (A) and site II (B). Where, the PDB file for HSA complex with amino camptothecin was taken from the protein data bank with ID: 4L8U. The best-fit accommodation of this structure after 30 different pose evaluations provided almost similar results as deduced earlier through STD-AF and Tr-NOESY studies.

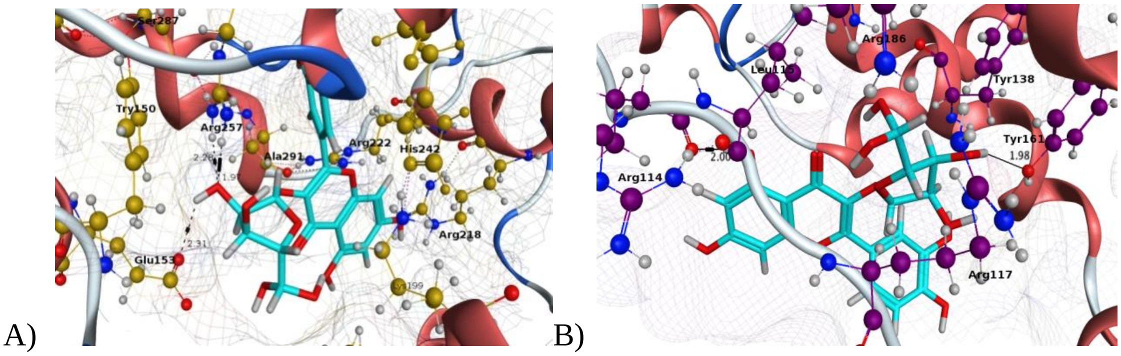


Figure 69. 3D bound conformations of the quercetin -3-O-arabinofuranoside, generated by Molecular Operating Environment (MOE) 2011.10 within the HSA-quercetin -3-O-arabinofuranoside complex for the Sudlow´s site I (A) and site II (B). Where, the PDB file for HSA complex with amino camptothecin was taken from the protein data bank with ID: 4L8U. The best-fit accommodation of this

structure after 30 different pose evaluations provided almost similar results as deduced earlier through STD-AF and Tr-NOESY studies.

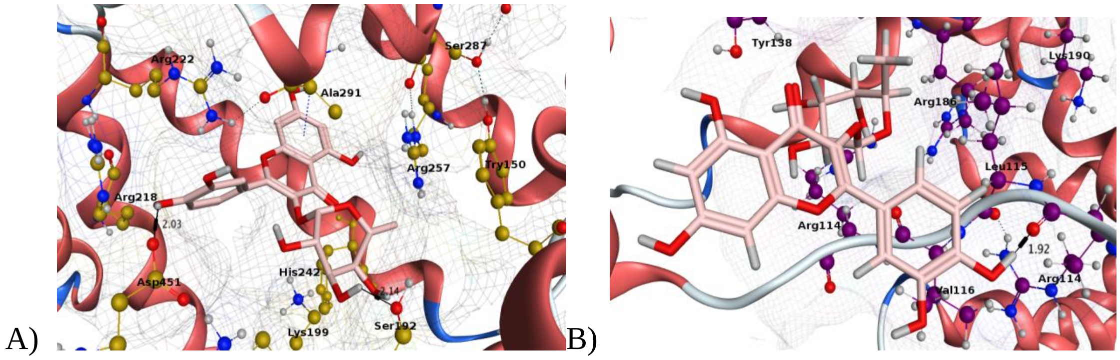


Figure 70. 3D bound conformations of the quercetin-3-O-rhamnopyranoside, generated by Molecular Operating Environment (MOE) 2011.10 within the HSA-quercetin-3-O-rhamnopyranoside complex for the Sudlow´s site I (A) and site II (B). Where, the PDB file for HSA complex with amino camptothecin was taken from the protein data bank with ID: 4L8U. The best-fit accommodation of this structure after 30 different pose evaluations provided almost similar results as deduced earlier through STD-AF and Tr-NOESY studies.

7

Mapping binding sites of Anti-Leishmanial and Anti-trypanosomiasis Synthetic library towards the Human Serum Albumin complex as probed by STD NMR, Tr-NOESY Experiments

7. Mapping binding sites of Anti-Leishmanial and Anti-trypanosomiasis Synthetic library towards the Human Serum Albumin complex as probed by STD NMR, Tr-NOESY Experiments.

7.1 Abstract

A library of seven compounds having moderate to potent activities against *Leishmania* and *Trypanossoma*, including 3 new moieties have been synthesized. Their interaction features, binding characterization, binding epitopes and titration with a spy molecule by using STD NMR, Tr-NOESY, and molecular docking simulation have been performed. The binding results and the NOESY conformational results were found to be in excellent agreement with the calculations obtained as a result of molecular docking.

7.2 Introduction

Leishmaniasis and Trypanosomiasis are the worldwide extend epidemic diseases; the dissemination of both diseases is a protozoan. According to the report of World Health Organization twelve millions of people have been infected in 88 countries, while, around 350 million people are at high risk level. In the perspective of the absence of safe prescription the serious secondary effects stimulated by the utilization of accessible chemotherapy (MCGREEVY, P. B., *et al.*, 1986), there is still need of new drugs for the treatment of leishmaniasis infections. Similarly, about 500,000 human race is suffering from Trypanosomiasis across the 36 African countries, and other lethal variant is Chagas disease, is a also a form of Trypanosomiasis which causes an average death toll of 21000 per annum, in Latin America (MAYA, J. D., *et al*. 2007). Thus, in a light this short introduction, our new synthetic library that possesses moderate to potent activity against these lethal diseases was an excellent choice to be treated against the Human Serum Albumin (HSA) for the binding studies. On the other hand, target protein (HSA) is well-know and most abundant constituents of blood (RODKEY, F. L., 1965). HSA has the large part in the regulation of different functions in the body such as osmotic pressure (PUTIGNANO, P. *et al.*, 2000), maintenance of pH (SHAKLAI, N. *et al.*, 1984), fatty acid transport (M, H.; AZZAZY, *et al.*, 1997), small molecule transport including the drugs (Koyama, H. *et al.*, 1997) etc.

NMR has a long history to work as a tool to probe binding studies (PELLECCHIA, M.; et al., 2008) and in characterization of macromolecular structures (BERMAN, H. M.; *et al.*, 2000). It can sense the perturbation involving both structures (protein and ligands) as a result of strong magnetic field. Thus, on the basis of changes, the screening may be ligandbased or the target-based. Since, the target based screening is a bit expensive in term of isotopic labelled protein a basic requirement for this strategy, conversely, ligand-based screening methods, neither require the prior knowledge of protein nor any special instrumental setup (PELLECCHIA, M.; *et al.*, 2002). A milestone came in the ligand based drug discovery when prof. B. Meyer introduced the saturation transfer transfer difference NMR technique (MAYER, M.; MEYER, B., 1999). Soon after its arrival it become the most robust and widely applicable techniques amongst other screening methods such that high throughput screening (RISHTON, G. M., 1997), SAR by NMR (SHUKER, S. B.; *et al.*, 1996) and fragment based drug design (SIEGAL, G.; *et al.*, 2007)etc. The wide applications of STD NMR are the consequences of the simple procedure, no necessary instrumentations, small protein use, no size limits on protein, can also work with the system with chemical exchange etc., (MEYER, B.; PETERS, T., 2003; WAGSTAFF, J. L., *et al.*, 2013). Therefore, a variety of reviews available illustrating the utility of STD in different areas likes, applications fragment-based drug design (BHUNIA, A., *et al.*, 2012), comparison with other methods (Ludwig C.; Guenther U. L., 2009; PENG, J. W., *et al.*, 2004) and alone as robust method (MEYER, B.; PETERS, T., 2003). In addition, the binding affinity of carbohydrates (Poveda, A.; Jimenez-Barbero, J., 2012), and extract studies (TANOLI, S. A. K., *et al*., 2013) as well.

7.3 Results and discussion

In a very recent report, Rahul B Aher *et al.* have shown that dibenzylideneacetone and their analogues and synthones have good potential to inhibit some parasite growth (AHER, R. B. *et al*., 2011). Therefore, keeping these favourable points of interest of these simple structure compounds, in this report we have selected a small library of compounds (shown in figure 71). Nowadays, large numbers of high-throughput screening methods are, available for the screening of large libraries joining the spectroscopic, non-spectroscopic and computational methodologies simultaneously. Consequently, we got the advantage of STD NMR and Tr-NOESY and docking simulation for to observe the interaction between these small molecules towards HSA. STD NMR experiments were performed by ligand: protein 200:1 molar excess, where protein was dissolved in PBS (Phosphate Buffer Saline) prepared in D_2O, while ligands solution were formed in 30: 70 % v/v CD_3OD in D_2O. Unfortunately, the parent structures of all compounds were similar, so to make a mixture of all compounds together was not possible. However, we have made a unique three mixtures to avoid extensive overlapping and in order to find the recognizable multiplicities. Thus, the mixture one had three and rest two mixtures were composed of each two compounds.

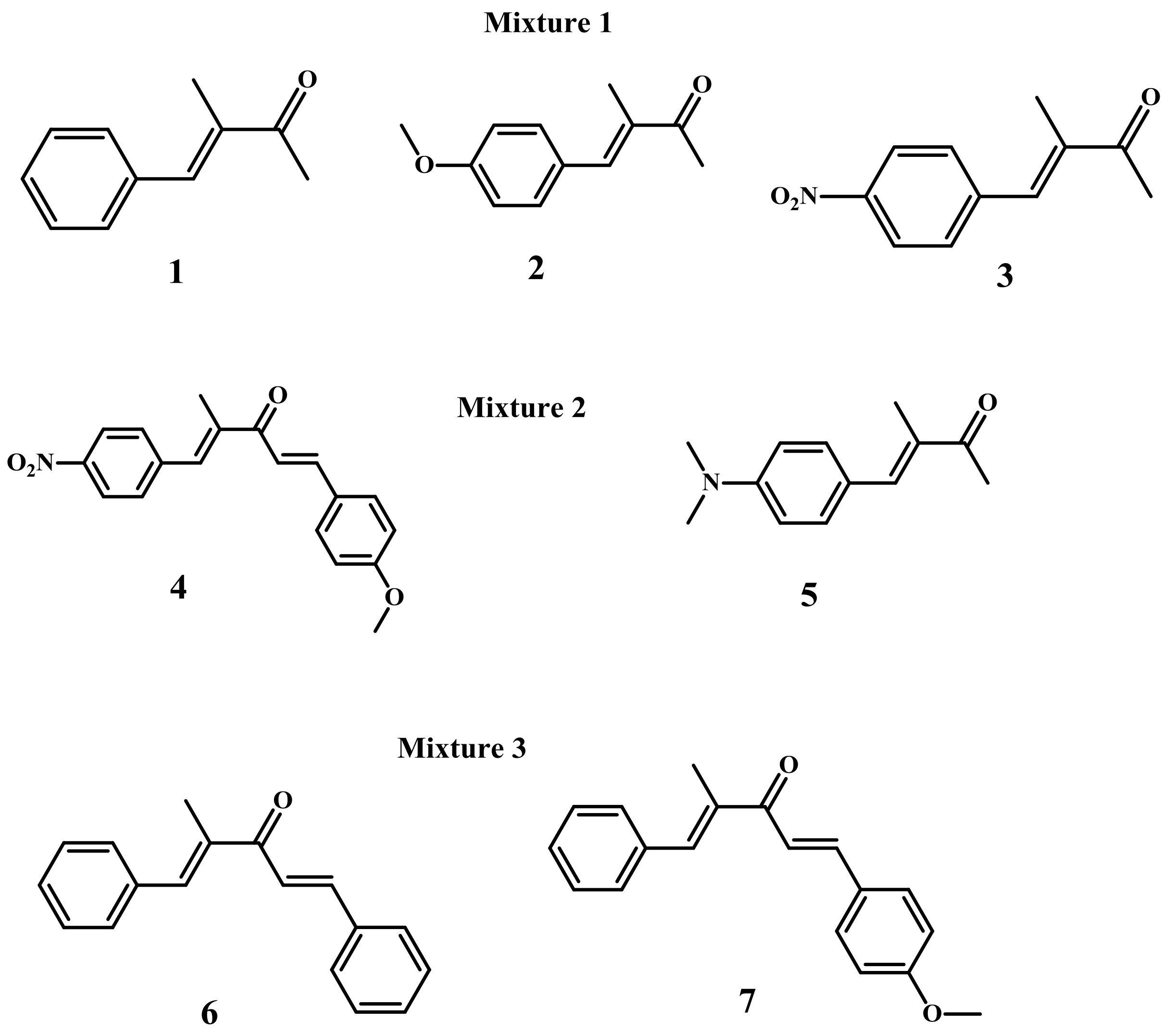


Figure 71. The synthetic library of selected compounds with their respective mixtures.

In the first experiment, a freshly prepared solution of mixture 1 was utilized and in addition, a small amount of protein was added as a part of this characterization. Therefore, the irradiation of protein was performed at -0.5 ppm to accomplish the protein saturation (*on*-resonance) with only the protein signals while for the unbound ligands (*off*-resonance) the irradiation frequency was placed at 30 ppm way from any signal. The saturation of protein was performed by the application of train of soft Guass shape pulses, with 50 ms of duration and where, each one was separated by 2 ms. The presence of signal in the in the STD NMR spectrum is a sensible proof of the binding of compounds to protein (MAYER, M.; MEYER, B., 2001) as indicated in Figure 72. For a clear understanding of the STD spectrum of mixture 1 (Figure 72) three different color arrow black, navy blue and sky blue was utilized for the compound **1**, **2** and **3** respectively.

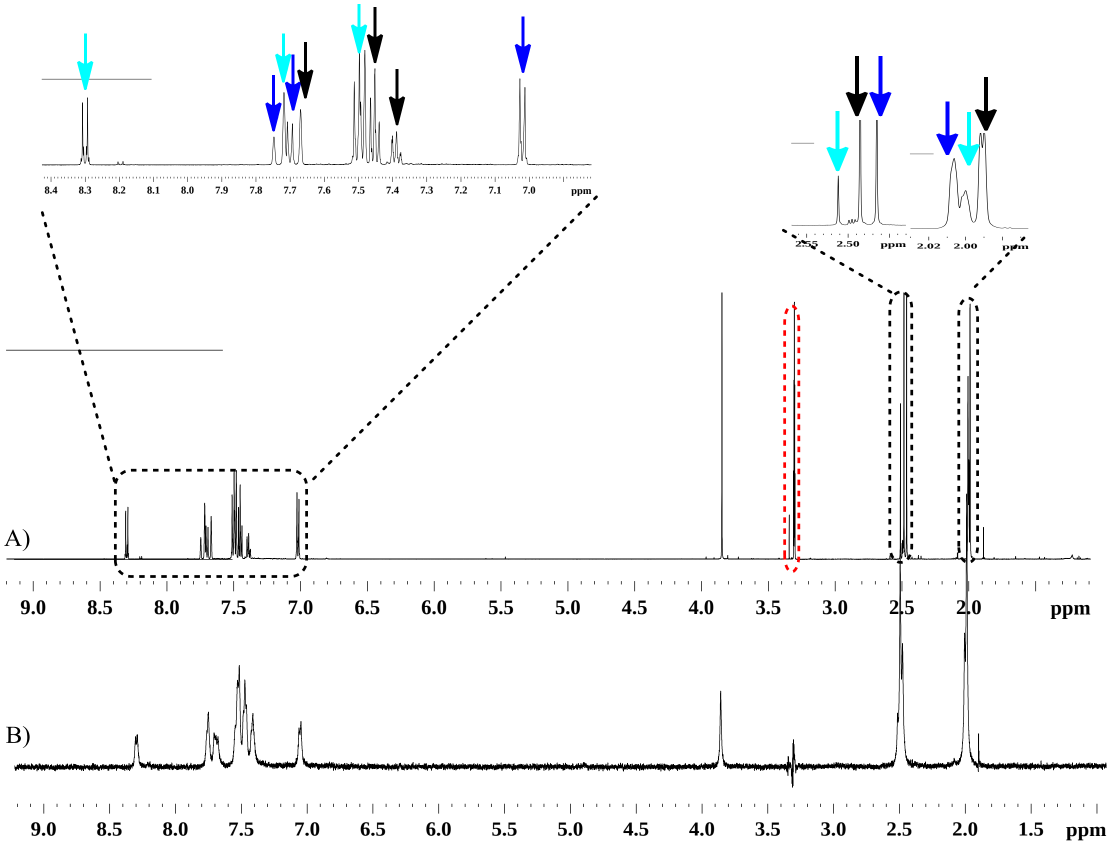


Figure 72. Saturation transfer difference NMR spectra of mixture 1, where, the top spectrum (*Off*-resonance) whiles the bottom (STD difference) NMR spectrum. The magnified regions were shown as out set at the top of spectra.

The binding epitopes were obtained after the STD build up step—a 5 experiments with varying saturation time between 1-5 s. The epitope mapping was made effective by the help of taking integration of all signals and then normalizing the all signals with respect to a large integral value; that was given the 100% STD amplification effect (MAYER, M.; MEYER, B., 2001). The binding epitopes showed the aromatic signals (3´, 4´ and 2´and 6´) from the compound **1** and **3** respectively, shares the equal contribution of 100% (Figure 73). However, the compound could show less than 50% effect. It is well documented that the STD NMR signal intensity depends upon the saturation and the number of protons being irradiated (TANOLI, S. A.K., *et al.*, 2013; TANOLI, S. A. K., *et al.*, 2014). Moreover, the vinyl protons from all compounds showed a prominent STD effect, hence interaction towards HSA with 65-80 for compound **1** and **2** while, the same protons of compound **3** could show up to 40 % effect.

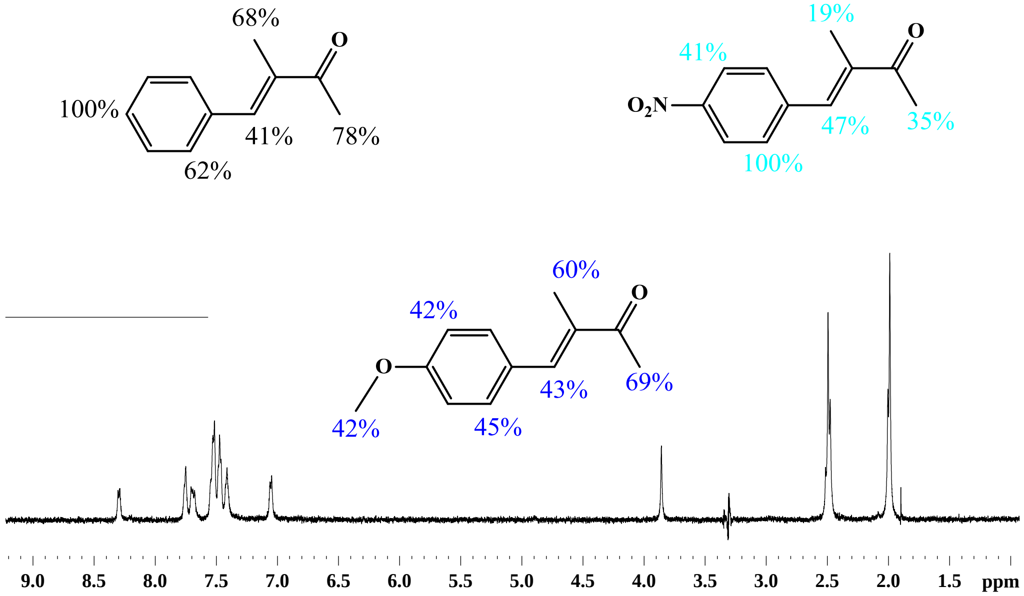


Figure 73. 1D STD difference NMR spectrum of mixture 1, with complete epitope mapping of three compounds. The 16 number of scans were applied to this spectrum, with 2 s of saturation time and 1s of relaxation delay.

The similar procedure was applied for mixture 2, ^{1}H-NMR spectrum of the mixture 2 can be seen in figure 74 below. Again, the same strategy was chosen to create a clear difference between the signals of two compound i.e., representations with different color arrows. The signals from compound **4** were presented with red, and conversely, the navy blue color was selected for compound **5** (Figure 74).

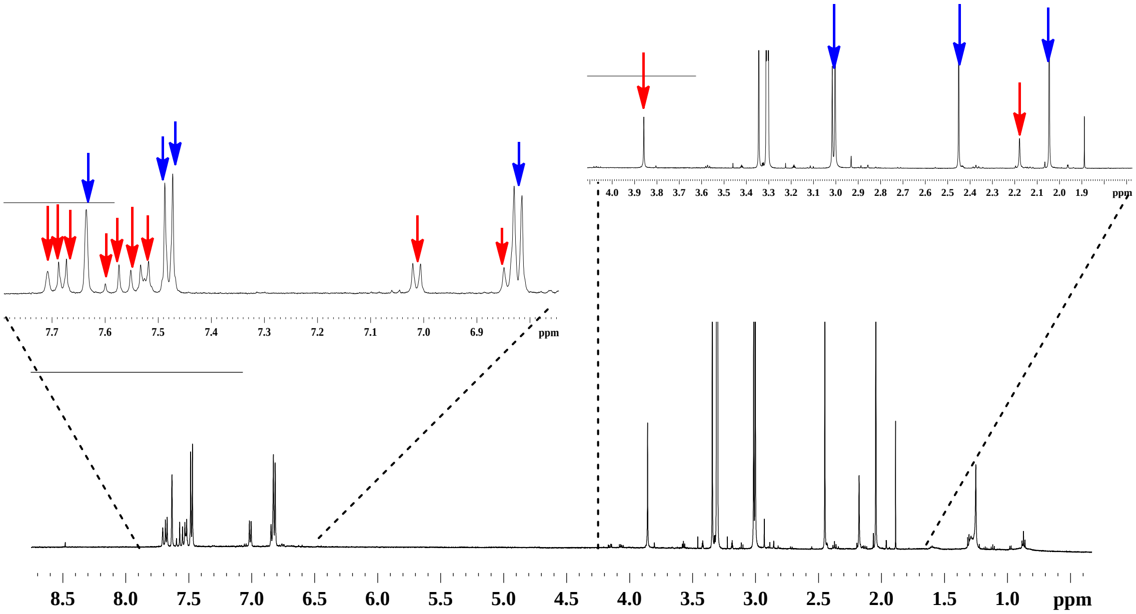


Figure 74. ^{1}H-NMR spectrum of mixture **2**, where navy blue and red color arrow representing the protons of **5** and **4** respectively. Magnification of the two regions was shown at the top with assignments through arrows.

The 1D hydrogen spectrum was quite clear so in order to see the clear multiplicity of each signals complete assignments of the both structures (**4** and **5**) are shown below in figure 75. In the figure below the impurity signal was shown by a red asterisk on the signal.

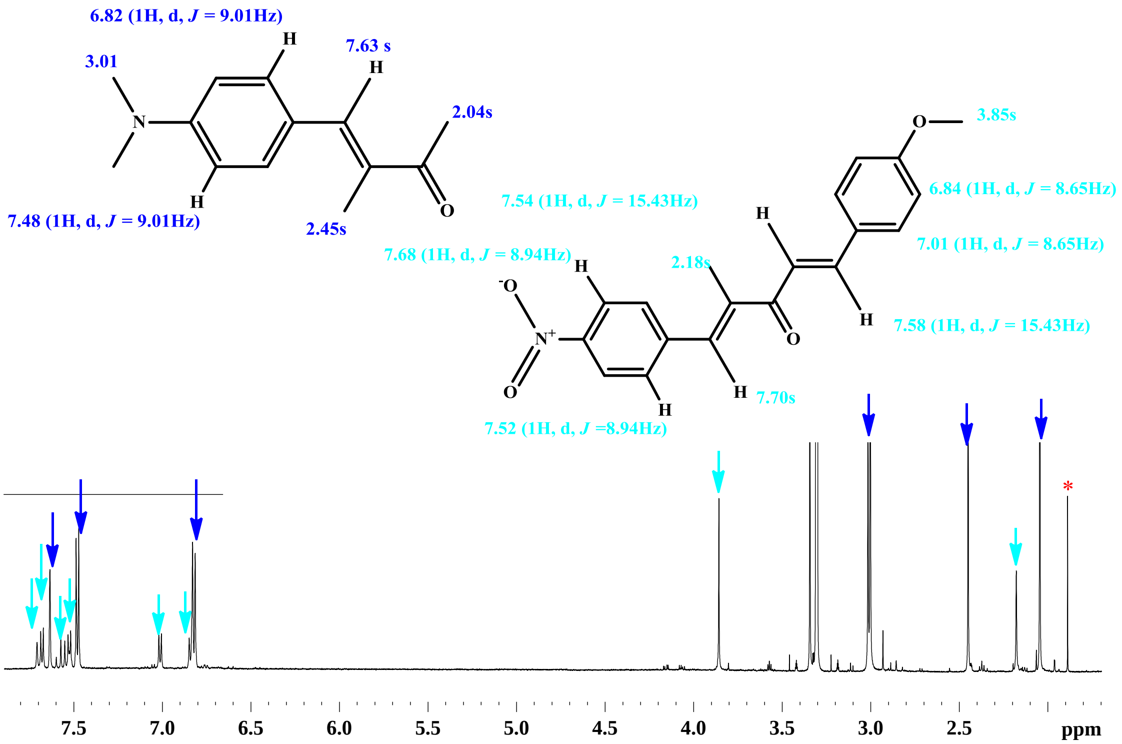


Figure 75. ^{1}H-NMR spectrum of mixture **2**, with the complete multiplicity of the signals are assigned on each compound.

The epitope results as an outcome of STD amplification factors from the Figure 76, clearly shows the smaller size ligand could reach the binding cavity easily and in turn showed 100 % STD effect through the outside methyl group of amine. The obvious reason comes to mind for the greater STD effect, is the side contact of compound **4** through its six protons. Similarly, the aromatic protons of the compound 4 again provided the second largest STD amplification (Figure 77). However, no prominent contact resulted from the compound **5**, except edges of aromatic protons and one vinylic proton. The region of contact which this molecule preferred is shallow or saturated with the compound **4** , and as the size is also a little expended that's why compound 5 remains unable to get in the vicinity of the binding pocket.

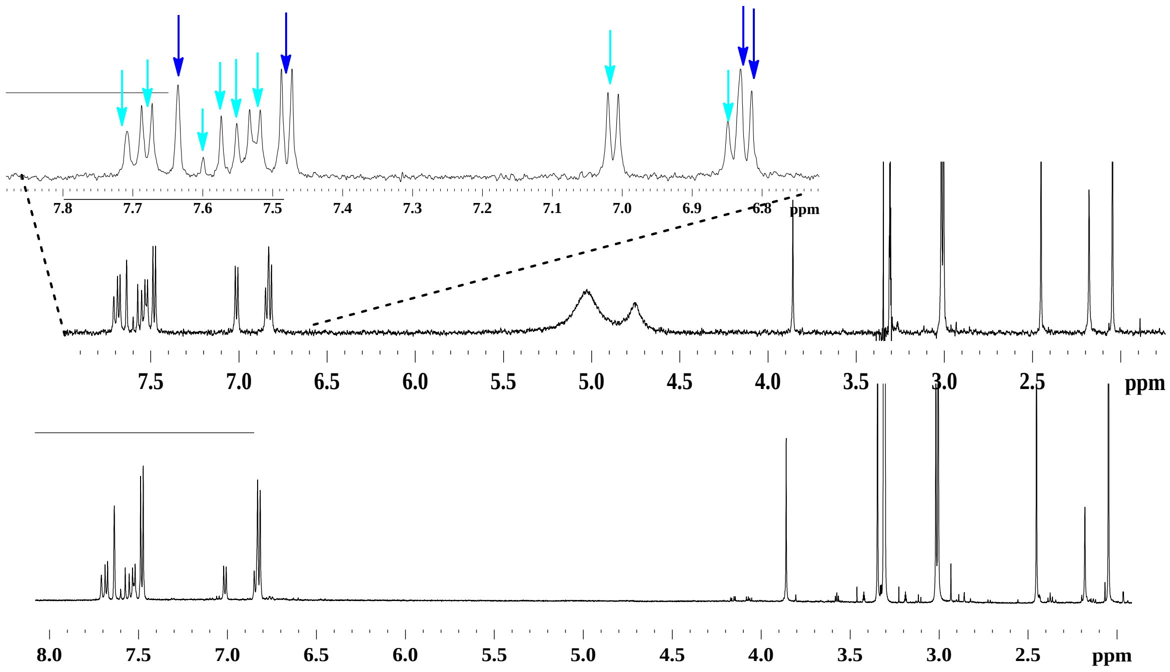


Figure 76. 1D STD NMR spectrum of mixture 2, where the region between the 6.8ppm to 7.8ppm was magnified and shown as outset at the top. Again, the same experimental procedure was applied as did for the mixture 1.

The next step to treat the last mixture with HSA and to find the binding epitopes as was performed for the earlier two mixtures.

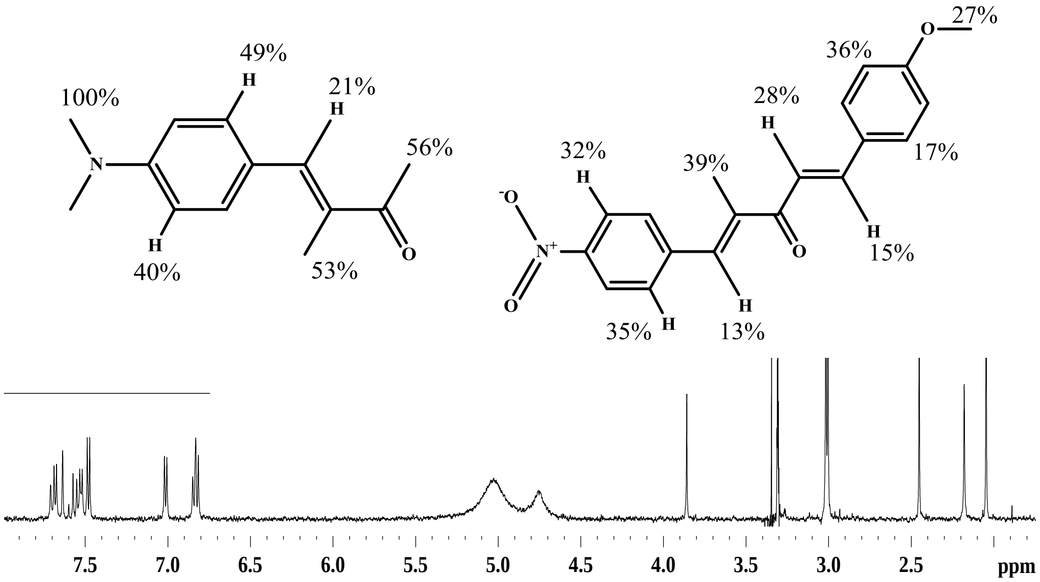


Figure 77. 1D STD spectrum of mixture 2 with complete epitopes as obtained from the STD build up procedure.

These two structures (**6** and **7**) had many similarities, and both had the two common aromatic rings, consequently, the resultant spectrum (Figure 78) becomes too crowded especially the aromatic region and even signal assignments become a tough task. Therefore for the completed assignments were made possible by the use of 2D COSY spectrum (data not shown here). The 16 number of scans were more than enough for the signal intensity in the proton spectrum (Figure 78). Therefore, by taking this spectrum together with COSY spectrum complete assignments of both compounds were made as shown in figure 79.

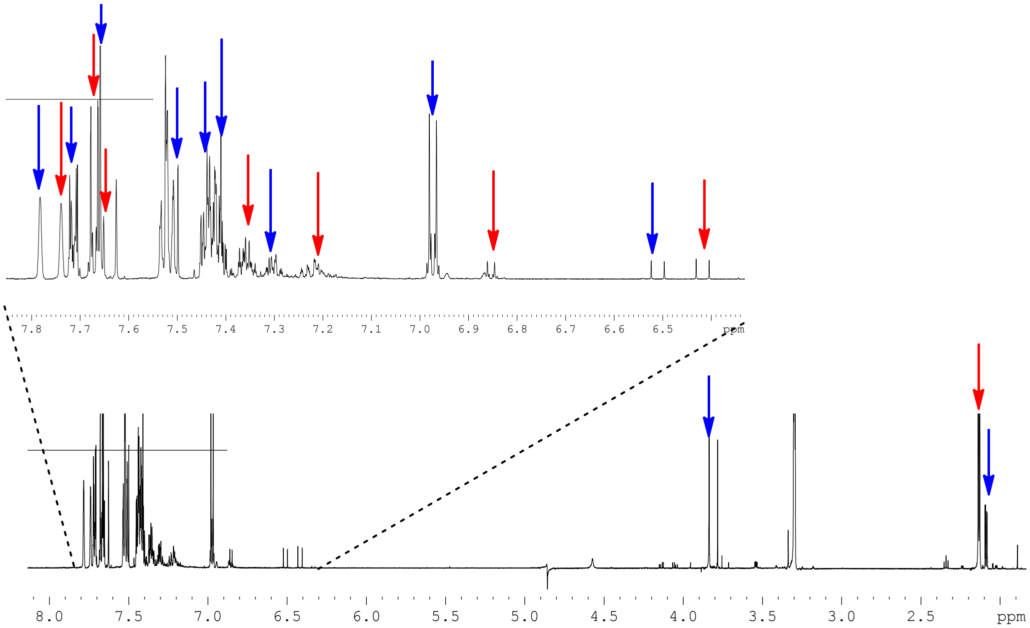
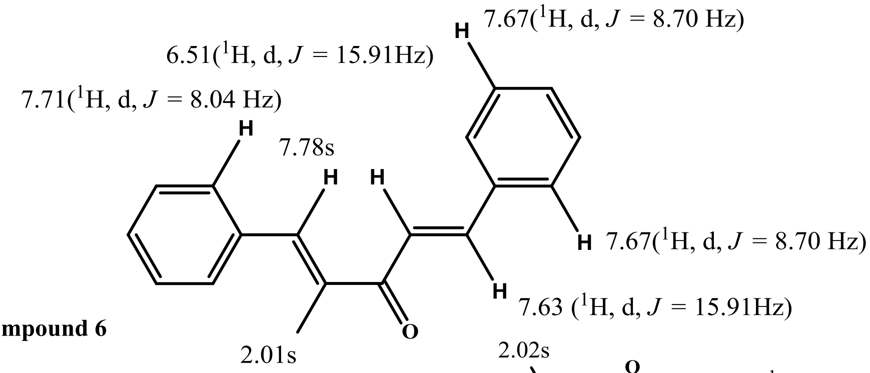
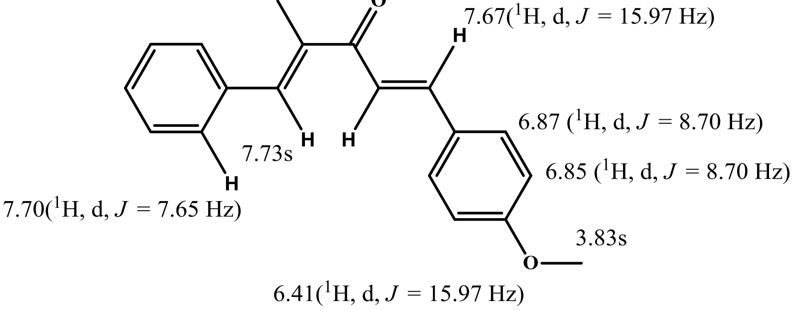


Figure 78. 1D ^{1}H-NMR spectrum of mixture **3**, the aromatic region was magnified to see the clear multiplicity and making assignments.

To see the clearance epitope results from the STD spectrum figure 80 was quite difficult, therefore, both ends of the aromatic without any substituents attached provided the maximum STD effect and hence bindings. It can be easily understood that these compounds **6** and **7**, repeating the same behavior as already seen in the case of compound **5**. They can´t find the proper binding cavity that can hold them that´s why they remains mostly towards the solvent exposed however, they provide the binding.



Compound 6



Compound 7

Figure 79. The structures of the compound **6** and **7** with complete multiplicity as measured from the ^{1}H NMR figure 78.

The next step was to perform the 2D NOESY and 2D Tr-NOESY spectra for these mixtures. Therefore, the ligand vs. protein ration was increased a little bit to get the saturation transfer easily and to see a better signal enhancement in the spectrum as shown figure 81 in the case of mixture **1**. The other two mixtures could not provide the spectrum very well, especially in the case of mixture, three may be due to the twisted structure the protons remains far apart from each other, and dipolar coupling is not efficient as it should be in the case when protons are in vicinal position (Figure 82).

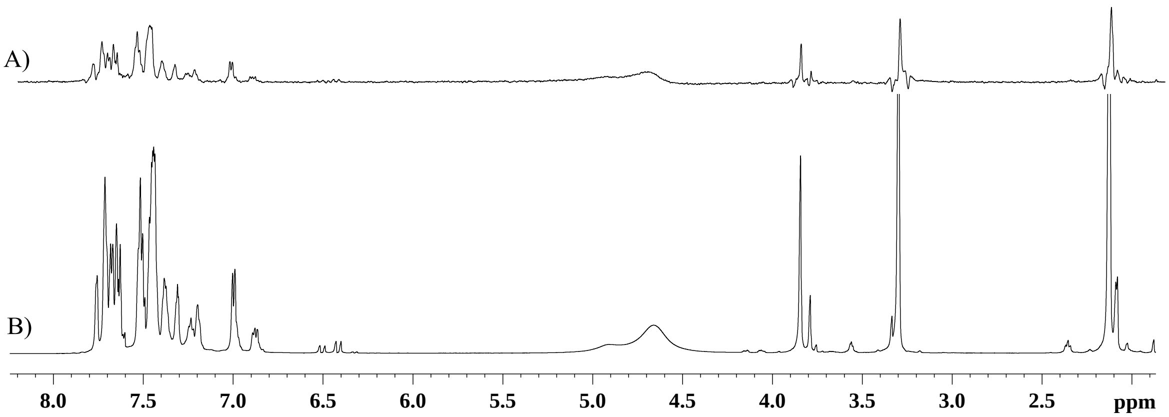


Figure 80. The ^{1}H NMR spectrum (bottom; B), the complete STD (top; A) of mixture **3** with 1:200 ratio of HSA and ligand in buffer solution at 7.2 pH (not corrected for D_2O).

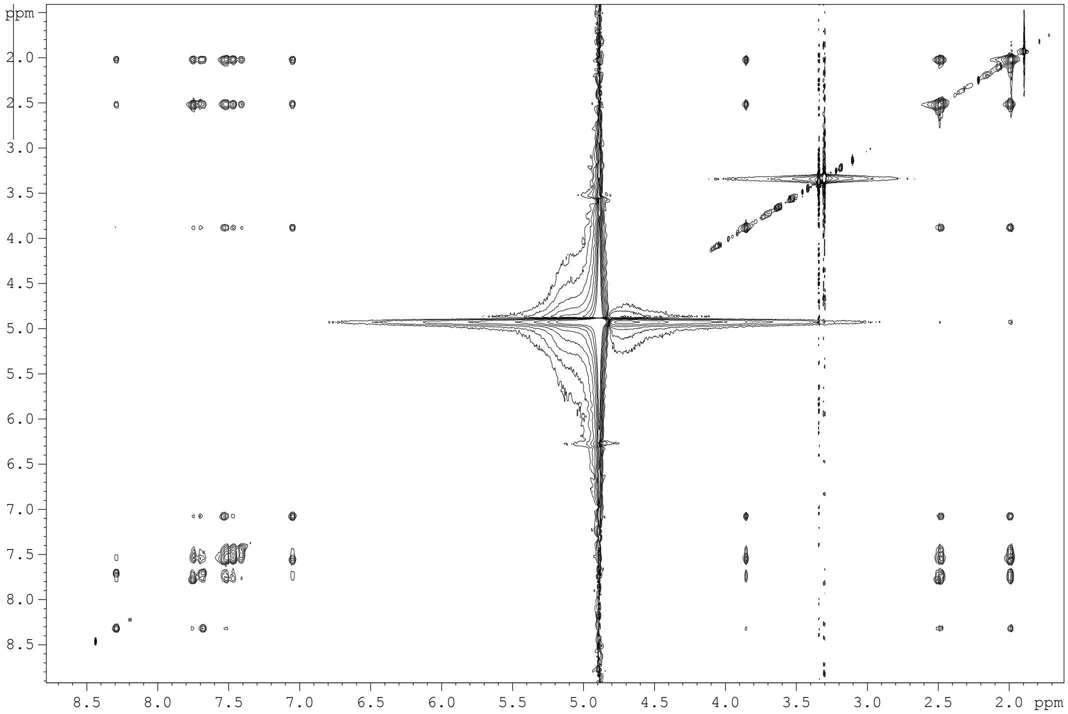


Figure 81. 2D Tr-NOESY of the mixture **1**, showing the excellent signal enhancements, the experiments were performed by 50: 1 ligand-protein ratio, at 600 MHz at 298K.

Then in the last step, a reporter molecule (L-tryptophan) was added into the mixtures, which preferably binds itself to Sudlow´s site II (CONRAD, M. L.; *et al.*, 2009). L tryptophan extensively used as a reporter molecule (SUDLOW, G.; 1976; FEHSKE, K. J., *et al*., 1982). It is well understood that the compounds that bind to the same site and show competition for one site, then by increasing the concentration of any one the other intensity is definitely going to be affected. Therefore, same conditions of STD NMR have been applied for all mixtures by changing the concentrations of spy, but we could not find any expected change within any spectra (Figure 83). However, from the STD NMR, Tr-NOESY NMR and this STD titration studies one can conclude that all of the compound (**1-7**) showed interaction, and as results of competitions studies none of them showed any changes in the signal intensity its mean, whether they are binding with some other side, or they are not competing for this site. (Sudlow site II).

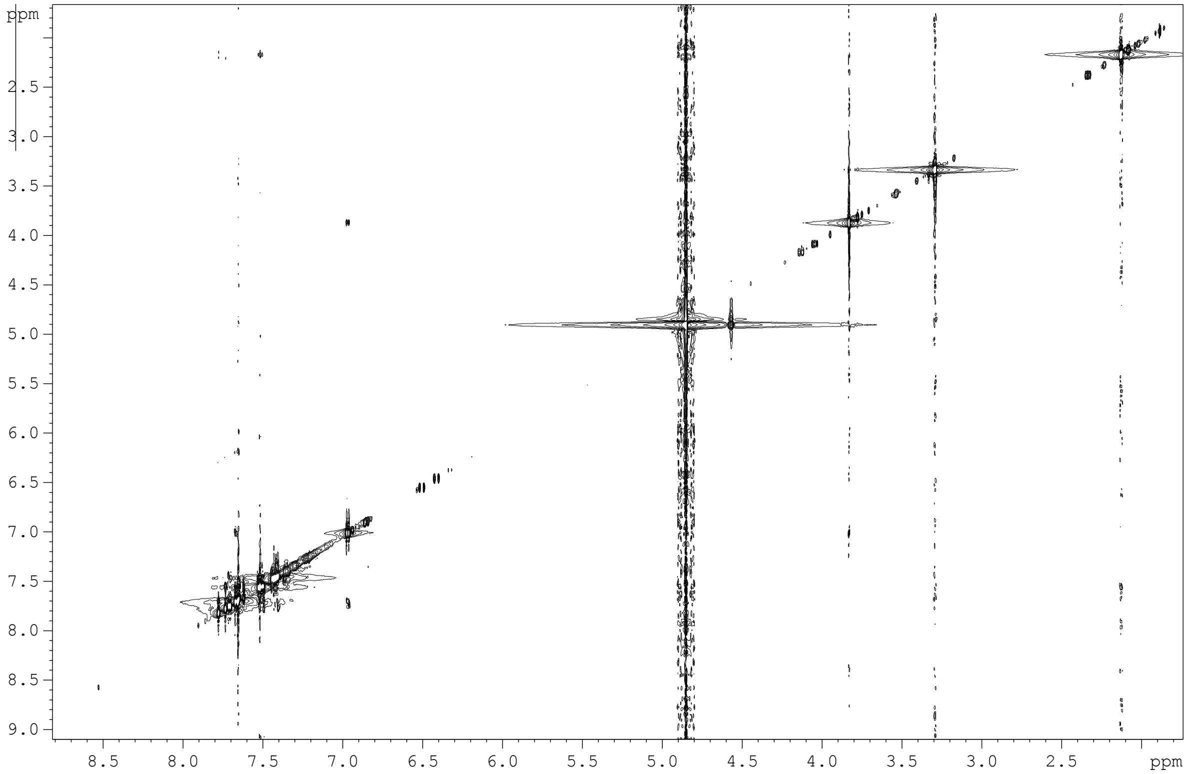


Figure 82. 2D Tr-NOESY of the mixture **3**, showing the excellent signal enhancements, the experiments were performed by 50: 1 ligand-protein ratio, at 600 MHz at 298K.

Then, in the last experiment we thought to take the help of docking studies to see verify the binding character of the NMR. Thus, to assess exactly what we could obtain from NMR, it was necessary to select both chains of HSA for the docking studies, and moreover, to verify whether our NMR results are correct or not suffering from some difficiencies. For this purpose, MOE 2012.10. (Molecular Operation Environment) software and on the other hand, human serum albumin protein was utilized with selected two distinct binding site residues with amino acids; Tyr150, Glu153, Lys195, Lys199, Arg218, Arg222, His242, Arg257, Leu284, Ser287, Ala291 and Arg114, Leu115, Val116, Arg117, Ile142, Tyr138, His146, Tyr161, Asp183, Leu185, Arg186, Lys190, Ser193 for site-1 and site-2, respectively. Based on high resolution a PDB file wit code “4L8U” was selected as a target protein from protein data bank.

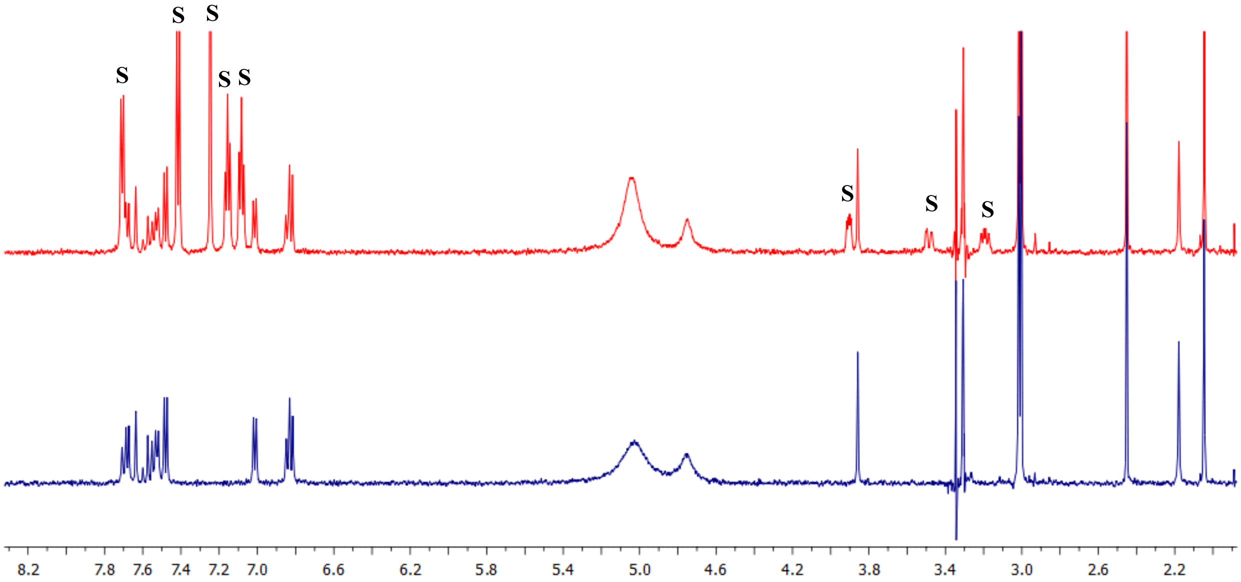


Figure 83. STD NMR spectrum of mixture **2** (Bottom) however, L-tryptophan a known reporter molecule was added with 1: 120 concentration, but the signal of the compound **4** and **5** remains same. Signals from spy are represented by S.

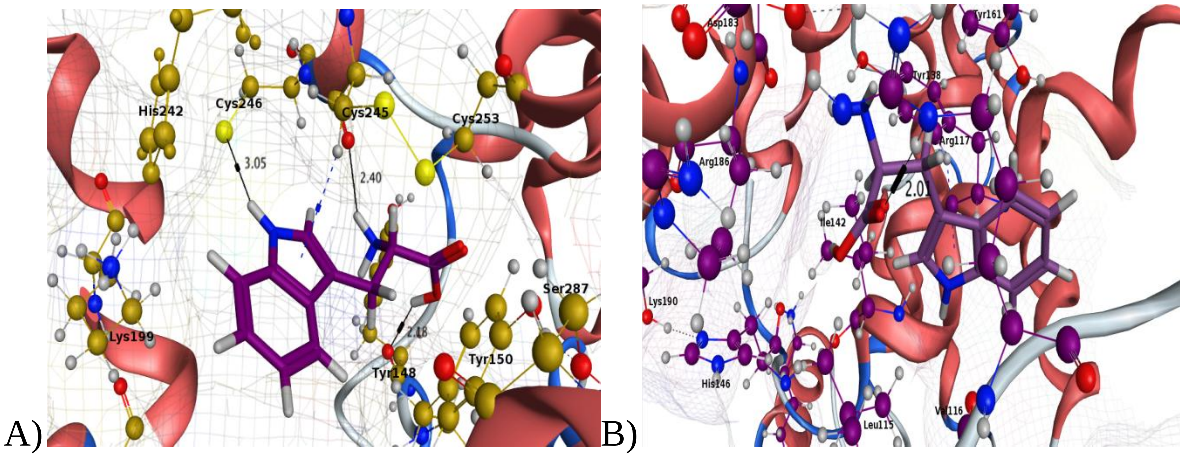


Figure 84. Binding site residues for the L-tryptophan for Sudlow site I and II respectively, obtained from MOE software.

Standard molecule upon docking fitted well within the binding cavity and interacted via hydrogen bonds as well as by hydrophobic interactions. Non-conserved active site residues Tyr148, Cys245 and Cys246 were involved in hydrogen bonding and metal ion interaction via ligand's carboxylic hydrogen, amine hydrogen and pyrrole NH, respectively. Beside these strong bonds, compound's binding was also supported by a hydrophobic interaction formed between CH of Cys245 and pyrrole ring of molecule, presented in figure-84. Unexpectedly, the binding energy of standard was found to be low that was not matched its orientation and interaction pattern that correlated well with its activity.

Idiosyncratic compounds analysis with reference to standard compound revealed that compound **1** when allowed to interact with binding site-1 of HSA, placed itself almost out of the pocket and displaying no interactions with hot spot residues. The binding free energy of such compound is least of all i.e; -5.0559 Kcal/mol ensuring its low potential towards this binding site of serum albumin. However, compound **3** with methoxy addition, localized in the same vicinity with quite favorable conformation and established a π-π interaction with Tyr150. A hot spot residue (Tyr150) interacted via its phenyl ring towards methoxy benzene CH of compound; moreover, single hydrophobic interaction could not be guarantee of its good activity. Similarly, compound **2** and **5** with nitro and dimethyl amine substitutions haven’t interacted with active site residues but interacted via strong hydrogen bond of 2.08Å and 2.13Å with Lys106, a residue out site the pocket. Low binding energies of -5.4263 and -5.5659 Kcal/mol were found to be well correlated with their conformations and interactions ability depicted in table-**1**, figure-85 and 86, respectively.

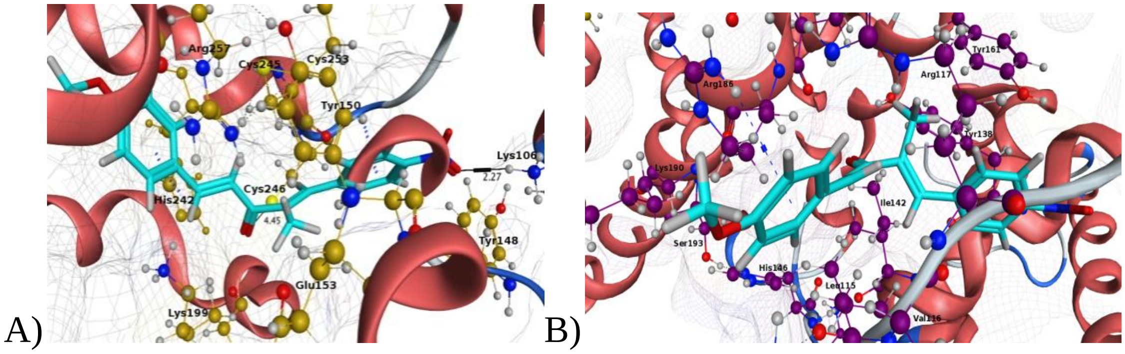


Figure 85. Binding site residues for Sudlow site I (A) and II (B) for compound **4** respectively, obtained from MOE software.

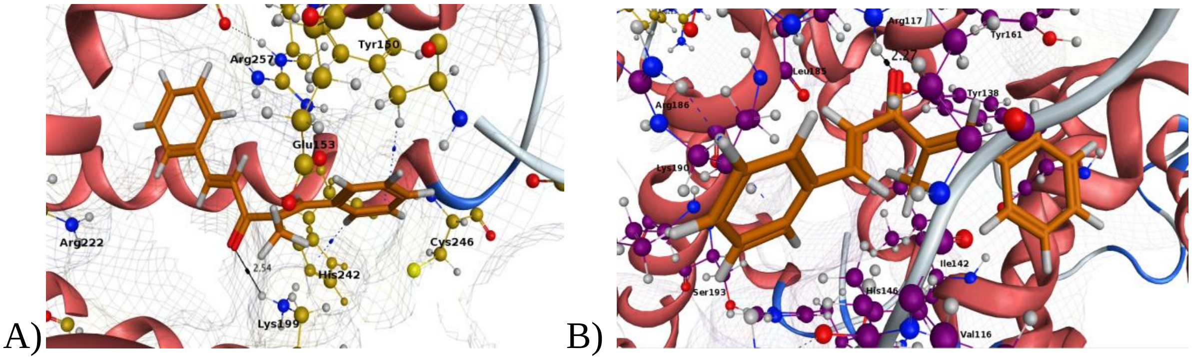


Figure 86. Binding site residues for Sudlow site I (A) and II (B) for compound **6** respectively, obtained from MOE software.

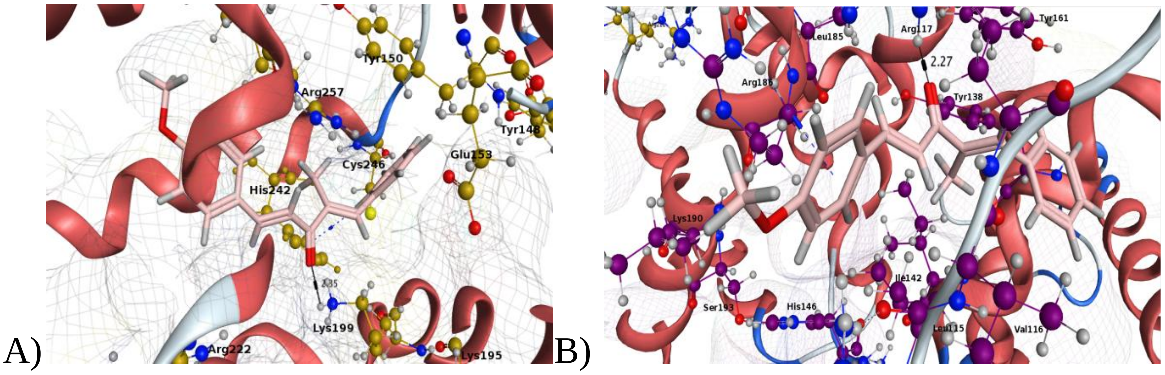


Figure 87. Binding site residues for Sudlow site I (A) and II (B) for compound **7** respectively, obtained from MOE software.

On the other hand, a compound with quite higher binding energy (-6.2895 Kcal/mol) was expected to interact more efficiently and was found to obtained perfect placement inside the cavity on visual inspection. Figure 87 showed the orientation of compound **7** within binding cavity and its ability to interact with protein's exposure site via hydrogen bond and two hydrophobic contacts i.e, with NH of Lys199, His242 imidazole ring and Tyr150 CH through its carbonyl oxygen and terminal phenyl ring. However, compound **5** with -7.8568 Kcal/mol interacted via hydrogen bonding, metal ion interaction as well as hydrophobic contacts with active site conserved residues and found to be most active of series. Oxygen of nitro-benzene mediated a strong hydrogen bond towards Lys106 NH with strength of 2.27Å, however, Cys246, His242 and Tyr150 were found to be involved in metal ion interaction with ketonic oxygen and π-π interactions with methoxy-benzene and nitro-benzene of compound **4**, respectively (figure-88).

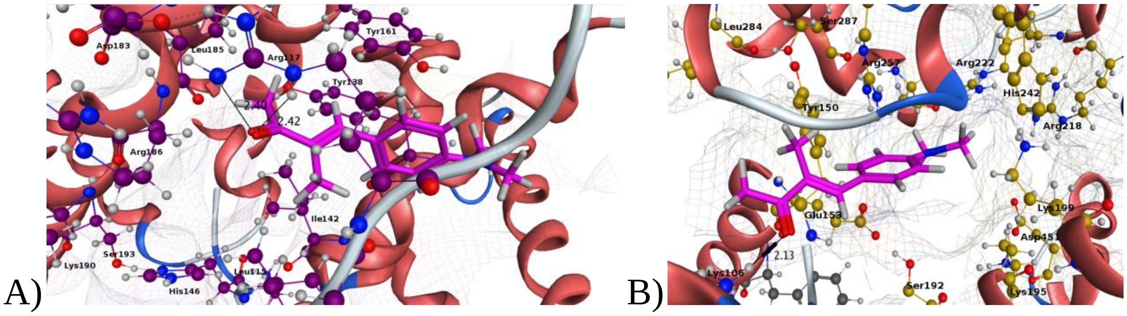


Figure 88. Binding site residues for Sudlow site I (A) and II (B) for compound **5** respectively, obtained from MOE software.

At the end, from overall analysis, it may be extracted that compounds (4, 7, with two phenyl rings and terminal methoxy group interacted more effectively within the human serum albumin binding site-1 to block it than others.

Docking results obtained from the other binding site of HSA were quite similar and indicating the worth of these compounds as a potential blocker of both sites. Standard compound, unexpectedly, bound with comparable binding free energy (-5.4272 Kcal/mol), in which Arg117 and Tyr138 via hydrogen bonding and π-π interaction played their role to hold compound within conserved site pocket residues.

Hydrophobic interactions were predominantly occurred in binding of compound **1** within site-2 of human serum albumin. Arg117 and Leu115 mediated hydrophilic-hydrophobic and hydrophobic-hydrophobic type of interactions to support ligand-protein binding. The binding energies for all compounds are shown in Table 13 and 14 for the Sudlow series I and II respectively.

Secreening of compounds (1-7) against site1

Compounds	S	E-score1	E-refine	E-score2
Compound1	-5,055872	-7,777197	-11,95248	-5,055872
Compound2	-5,532398	-7,251852	-12,84561	-5,532398
Compound3	-5,42634	-7,797433	-10,25321	-5,42634
Compound4	-7,856835	-10,08445	-22,24676	-7,856835
Compound5	-5,565903	-8,188156	-8,470411	-5,565903
Compound6	-6,28949	-8,482134	-17,01157	-6,28949
Compound7	-7,292704	-8,689496	-21,87032	-7,292704
Standard	-5,328514	-9,137387	-11,86802	-5,328514

Table 13. Free energy (Kcal/mol) of compounds with HSA in complex with reference to the binding site-1 of human serum albumin.

Secreeing of compounds (1-7) against site1I

Compounds	S	E-score1	E-refine	E-score2
Compound1	-4,945521	-7,441456	-12,59736	-4,945521
Compound2	-5,775546	-8,168342	-17,80937	-5,775546
Compound3	-5,282276	-8,021656	-17,32743	-5,282276
Compound4	-6,647553	-10,22077	-23,21007	-6,647553
Compound5	-5,84426	-8,185576	-15,31602	-5,84426
Compound6	-5,959717	-8,706781	-16,70379	-5,959717
Compound7	-6,274665	-9,581068	-19,48686	-6,274665
Standard	-5,427213	-9,264605	-18,40102	-5,427213

Table 14. Binding Free energy (Kcal/mol) calculated by MOE via scoring, refinement and rescoring of compound's scores to predict their binding affinity to binding site-II of human serum albumin.

8

Conclusions

8. Conclusions

Currently, the importance and significance of macromolecule dynamics for cooperative binding and molecular recognition became a matter of interest for biologist, pharmacist and chemist all over the world. However, the accuracy of the binding modes and ligands strengths still creates obstacles in the drug discovery process. Therefore, some computer-aided drug discovery programs are getting more importance in screening a large library of compounds within a very short time. Nevertheless, suffers with false results and considerable expertise, although screening arrays of ligands. On the other hand, conventional methods usually start from the isolation of some bioactive compound that becomes a reason for synthesis as results of some potential activity. However, that becomes a time intense task particularly when the structure posses multiple equivalent reactive sites.

In contrast to the aforementioned means, scientists are focusing for new strategies that would smooth the progress of drug discovery and additionally to scale back the time accustomed and economically friendly. In this viewpoint, we have adopted the most robust nuclear magnetic resonance methods that not only shown the time cutback but also the cost-effective in term of pharmaceutical industries. We have got devised most habitually used NMR techniques like, STD, Tr-NOESY and STD-TOCSY in coordination with a hyphenation system that would not exclusively recognized the ligands from natural extract but also aids in their unequivocal structure characterization. This recognition method followed by docking simulation that afforded leads directly from crude. During the course of these studies we have been used two targets (BSA and HSA) however, this studies can also be extended to the other drug targets for faster and wide screening strategy. Moreover, the scope of current method is also be employed for screening of natural herbal extract—used as a folk medicine without any knowledge of the effective constituents that are presents in the extract.

The overall binding results clearly showed that the compounds that are soluble in D_2O provided good STD NMR, Tr-NOESY, and STD-TOCSY results. The flavonoidal structure with sugars provided excellent results with these NMR screening methods, in contrast to the synthetic compounds that were used in chapter 7. Moreover, synthetic compounds utilized in chapter 7 were unable to show clear multiplicities as that were seen in the case of extract studies. These studies also providing a new idea to work on small molecules, although having some potent activates are supersede with no reason by the novel and big structures. In addition, the binding kinetics and epitopes of synthetic ligands is also helpful by using these methods. The NMR methods can help to screen the large library weakly binds ligands where most routinely and conventional screening methods fails.

Therefore, improvements in NMR-based screening methods can enhance the speed and the accuracy in structure determination within and without the bound conformation of receptor, will boost up the lead optimizations process and in other words the drug discovery process.

❾

References

9. References

ADAMS, C. P.; BRANTNER, V. V., Estimating The Cost Of New Drug Development: Is It Really $802 Million? *Health Affairs* **2006,** *25* (2), 420-428.

ADAMS, C. P.; BRANTNER, V. V., Spending on new drug development1. *Health Economics* **2010,** *19* (2), 130-141.

AHLQUIST, R. P., A STUDY OF THE ADRENOTROPIC RECEPTORS. *American Journal of Physiology - Legacy Content* **1948,** *153* (3), 586-600.

ALBARICI, T. R. Estudo fitoquímico de Rauia sp e Conchocarpus macrophyllu (Rutaceae) e avaliação da atividade antiparasitária de extratos e substâncias isoladas”. Federal university of São Carlos, Federal university of São Carlos, 2006.

ALBERT, K.; DACHTLER, M.; GLASER, T.; HÄNDEL, H.; LACKER, T.; SCHLOTTERBECK, G.; STROHSCHEIN, S.; TSENG, L.-H.; BRAUMANN, U., On-Line Coupling of Separation Techniques to NMR. *Journal of High Resolution Chromatography* **1999,** *22* (3), 135-143.

ALBRAND, J. P.; BIRDSALL, B.; FEENEY, J.; ROBERTS, G. C. K.; BURGEN, A. S. V. The use of transferred nuclear overhauser effects in the study of the conformations of small molecules bound to proteins. *International Journal of Biological Macromolecules* **1979**, *1*, 37-41.

ALLEN, T. M.; CULLIS, P. R., Drug Delivery Systems: Entering the Mainstream. *Science* **2004,** *303* (5665), 1818-1822.

ANGULO, J.; ENRÍQUEZ-NAVAS, P. M.; NIETO, P. M., Ligand–Receptor Binding Affinities from Saturation Transfer Difference (STD) NMR Spectroscopy: The Binding Isotherm of STD Initial Growth Rates. *Chemistry – A European Journal* **2010,** *16* (26), 7803-7812.

ANKERST, J.; MONTÁN, S.; SVANBERG, K.; SVANBERG, S., Laser-Induced Fluorescence Studies of Hematoporphyrin Derivative (HPD) in Normal and Tumor Tissue of Rat. *Applied Spectroscopy* **1984,** *38* (6), 890-896.

BAKHTIAR, R., Integrated strategies for drug discovery using mass spectrometry. *Journal of the American Society for Mass Spectrometry* **2006,** *17* (1), 108-110.

BALARAM, P.; BOTHNER-BY, A. A.; DADOK, J., Negative nuclear Overhuaser effects as probes of macromolecular structure. *Journal of the American Chemical Society* **1972,** *94* (11), 4015-4017.

BALAYSSAC, S.; TREFI, S.; GILARD, V.; MALET-MARTINO, M.; MARTINO, R.; DELSUC, M.-A., 2D and 3D DOSY 1H NMR, a useful tool for analysis of complex mixtures: Application to herbal drugs or dietary supplements for erectile dysfunction. *Journal of Pharmaceutical and Biomedical Analysis* **2009,** *50* (4), 602-612.

BAUER, C.; FREEMAN, R.; FRENKIEL, T.; KEELER, J.; SHAKA, A. J., Gaussian pulses. *Journal of Magnetic Resonance (1969)* **1984,** *58* (3), 442-457.

BAUM, R., COMBINATORIAL CHEMISTRY. *Chemical & Engineering News Archive* **1996,** *74* (7), 28-73.

BAX, A., A spatially selective composite 90° radiofrequency pulse. *Journal of Magnetic Resonance (1969)* **1985,** *65* (1), 142-145.

BECATTINI, B.; CULMSEE, C.; LEONE, M.; ZHAI, D.; ZHANG, X.; CROWELL, K. J.; REGA, M. F.; LANDSHAMER, S.; REED, J. C.; PLESNILA, N.; Pellecchia, M., Structure–activity relationships by interligand NOE-based design and synthesis of antiapoptotic compounds targeting Bid. *Proceedings of the National Academy of Sciences* **2006,** *103* (33), 12602-12606.

BECATTINI, B.; CULMSEE, C.; LEONE, M.; ZHAI, D.; ZHANG, X.; CROWELL, K. J.; REGA, M. F.; LANDSHAMER, S.; REED, J. C.; PLESNILA, N.; PELLECCHIA, M., Structure–activity relationships by interligand NOE-based design and synthesis of antiapoptotic compounds targeting Bid. *Proceedings of the National Academy of Sciences* **2006,** *103* (33), 12602-12606.

BECATTINI, B.; PELLECCHIA, M., SAR by ILOEs: An NMR-Based Approach to Reverse Chemical Genetics. *Chemistry – A European Journal* **2006,** *12* (10), 2658-2662.

BEEDHAM, C., 3 Molybdenum Hydroxylases: Biological Distribution and Substrate-Inhibitor Specificity. In *Progress in Medicinal Chemistry,* Ellis, G. P.; West, G. B., Eds. Elsevier: 1987; Vol. Volume 24, pp 85-127.

BEEDHAM, C., Molybdenum Hydroxylases as Drug-Metabolizing Enzymes. *Drug Metabolism Reviews* **1985,** *16* (1-2), 119-156.

BEEDHAM, C.; Bruce, S. E.; Critchley, D. J.; Al-Tayib, Y.; Rance, D. J., Species variation in hepatic aldehyde oxidase activity. *European Journal of Drug Metabolism and Pharmacokinetics* **1987,** *12* (4), 307-310.

BHUNIA, A.; BHATTACHARJYA, S.; CHATTERJEE, S., Applications of saturation transfer difference NMR in biological systems. *Drug Discovery Today* **2012,** *17* (9–10), 505-513.

BILIA, A. R.; BERGONZI, M. C.; MAZZI, G.; VINCIERI, F. F., Analysis of Plant Complex Matrices by Use of Nuclear Magnetic Resonance Spectroscopy: St. John's Wort Extract. *Journal of Agricultural and Food Chemistry* **2001,** *49* (5), 2115-2124.

BISMUTO, E.; GRATTON, E.; LAMB, D. C., Dynamics of ANS Binding to Tuna Apomyoglobin Measured with Fluorescence Correlation Spectroscopy. *Biophysical Journal* **2001,** *81* (6), 3510-3521.

BLUHM, U.; BOUCHER, J.-L.; BUSS, U.; CLEMENT, B.; FRIEDRICH, F.; GIRRESER, U.; HEBER, D.; LAM, T.; LEPOIVRE, M.; Rostaie-Gerylow, M.; Wolschendorf, U., Synthesis and evaluation of pyrido[1,2-a]pyrimidines as inhibitors of nitric oxide synthases. *European Journal of Medicinal Chemistry* **2009,** *44* (7), 2877-2887.

BOHACEK, R. S.; MCMARTIN, C.; GUIDA, W. C., The art and practice of structure-based drug design: A molecular modeling perspective. *Medicinal Research Reviews* **1996,** *16* (1), 3-50.

BOPPANA, K.; DUBEY, P. K.; JAGARLAPUDI, S. A. R. P.; VADIVELAN, S.; RAMBABU, G., Knowledge based identification of MAO-B selective inhibitors using pharmacophore and structure based virtual screening models. *European Journal of Medicinal Chemistry* **2009,** *44* (9), 3584-3590.

BOULDIN, A. S.; SMITH, M. C.; GARNER, D. D.; SZEINBACH, S. L.; FRATE, D. A.; CROOM, E. M., Pharmacy and herbal medicine in the US. *Social Science & Medicine* **1999,** *49* (2), 279-289.

BRAUN, W.; BÖSCH, C.; BROWN, L. R.; GŌ, N.; WÜTHRICH, K., Combined use of proton-proton overhauser enhancements and a distance geometry algorithm for determination of polypeptide conformations. Application to micelle-bound glucagon. *Biochimica et Biophysica Acta (BBA) - Protein Structure* **1981,** *667* (2), 377-396.

BRINGMANN, G.; RÜCKERT, M.; SAEB, W.; MUDOGO, V., Characterization of metabolites in plant extracts of Ancistrocladus likoko by high-performance liquid chromatography coupled on-line with 1H NMR spectroscopy. *Magnetic Resonance in Chemistry* **1999,** *37* (2), 98-102.

BROWN, J. Q.; VISHWANATH, K.; PALMER, G. M.; RAMANUJAM, N., Advances in quantitative UV–visible spectroscopy for clinical and pre-clinical application in cancer. *Current Opinion in Biotechnology* **2009,** *20* (1), 119-131.

Bruker Corporation/Spark Holland (Emmen, T. N. Bruker Corporation/Spark Holland (Emmen, The Netherlands).

BERMAN, H. M.; WESTBROOK, J.; FENG, Z.; GILLILAND, G.; BHAT, T. N.; WEISSIG, H.; SHINDYALOV, I. N.; BOURNE, P. E., The Protein Data Bank. *Nucleic Acids Research* **2000,** *28* (1), 235-242.

CALLE, L. P.; CANADA, F. J.; JIMENEZ-BARBERO; J. *Natural Product Reports* **2011**, *28*, 1118-1125.

CALLE, L. P.; CANADA, F. J.; JIMENEZ-BARBERO, J., Application of NMR methods to the study of the interaction of natural products with biomolecular receptors. *Natural Product Reports* **2011,** *28* (6), 1118-1125.

CALVERT, A. H.; PLUMMER, R., The Development of Phase I Cancer Trial Methodologies: the Use of Pharmacokinetic and Pharmacodynamic End Points Sets the Scene for Phase 0 Cancer Clinical Trials. *Clinical Cancer Research* **2008,** *14* (12), 3664-3669.

CARTER, D. C.; HO, J. X., Structure of Serum Albumin. In *Advances in Protein Chemistry*, C.B. Anfinsen, J. T. E. F. M. R.; David, S. E., Eds. Academic Press: 1994; Vol. Volume 45, pp 153-203.

CAVANAGH, J.; FAIRBROTHER, W. J.; PALMER III, A. G.; Rance, M.; Skelton, N. J., Chapter 3 - Experimental aspects of NMR spectroscopy. In *Protein NMR Spectroscopy (Second Edition)*, Eds. Academic Press: Burlington, 2007; pp 114-270.

CHEN, A.; SHAPIRO, M. J., Peer Reviewed: Affinity NMR. *Analytical Chemistry* **1999,** *71* (19), 669A-675A.

CHEN, A.; SHAPIRO, M., NOE Pumping: A ovel NMR technique for Identification of Compounds with Binding Affinity to Macromolecules. *Journal of the American Chemical Society* **1998,** *120*, 10258-10259.

CHENG, K. C.; KORFMACHER, W. A.; WHITE, R. E.; NJOROGE, F. G., Lead Optimization in Discovery Drug Metabolism and Pharmacokinetics/Case study:

The Hepatitis C Virus (HCV) Protease Inhibitor SCH 503034. *Perspectives in Medicinal Chemistry* **2007,** *1* (PMC-1-Cheng-et-al-(Kuo-Chi)), 1-9.

CLAASEN, B.; AXMANN, M.; MEINECKE, R.; MEYER, B., Direct Observation of Ligand Binding to Membrane Proteins in Living Cells by a Saturation Transfer Double Difference (STDD) NMR Spectroscopy Method Shows a Significantly Higher Affinity of Integrin αIIbβ3 in Native Platelets than in Liposomes. *Journal of the American Chemical Society* **2004,** *127* (3), 916-919.

CLARK, D. E., What has computer-aided molecular design ever done for drug discovery? *Expert Opinion on Drug Discovery* **2006,** *1* (2), 103-110.

COLBOURNE, A. A.; MORRIS, G. A.; NILSSON; M., Local Covariance Order Diffusion-Ordered Spectroscopy: A Powerful Tool for Mixture Analysis. *Journal of the American Chemical Society* **2011,** *133* (20), 7640-7643.

COLE, J. C.; MURRAY, C. W.; NISSINK, J. W. M.; TAYLOR, R. D.; TAYLOR, R., Comparing protein–ligand docking programs is difficult. *Proteins: Structure, Function, and Bioinformatics* **2005,** *60* (3), 325-332.

CONNOLLY, S.; ABERG, A.; ARVAI, A.; BEATON, H. G.; CHESHIRE, D. R.; COOK, A. R.; COOPER, S.; COX, D.; HAMLEY, P.; MALLINDER, P.; MILLICHIP, I.; NICHOLLS, D. J.; ROSENFELD, R. J.; ST-GALLAY, S. A.; TAINER, J.; TINKER, A. C.; WALLACE, A. V., 2-Aminopyridines as Highly Selective Inducible Nitric Oxide Synthase Inhibitors. Differential Binding Modes Dependent on Nitrogen Substitution. *Journal of Medicinal Chemistry* **2004,** *47* (12), 3320-3323.

CONRAD, M. L.; MOSER, A. C.; HAGE, D. S; Evaluation of indole-based probes for high-throughput screening of drug binding two human serum albumin: Analysis by high performance affinity chromatography. *Journal of Separation Science* **2009**, *32*, 1145-1155.

CORRÊA, M. P.; Florestal, I. B. d. D. *Dicionário das plantas úteis do Brasil e das exóticas cultivadas*; Rio de Janeiro: Imprensa Nacional, VI., **1921**; Vol. 1, p 747.

CZARNIK, A. W., Peer Reviewed: Combinatorial Chemistry. *Analytical Chemistry* **1998,** *70* (11), 378A-386A.

DALVIT, C.; ARDINI, E.; FLOCCO, M.; FOGLIATTO, G. P.; MONGELLI, N.; VERONESI, M., A General NMR Method for Rapid, Efficient, and Reliable Biochemical Screening. *Journal of the American Chemical Society* **2003,** *125* (47), 14620-14625.

DALVIT, C.; FAGERNESS, P. E.; HADDEN, D. T. A.; SARVER, R. W.; STOCKMAN, B. J., Fluorine-NMR Experiments for High-Throughput Screening: Theoretical Aspects, Practical Considerations, and Range of Applicability. *Journal of the American Chemical Society* **2003,** *125* (25), 7696-7703.

DALVIT, C.; FOGLIATTO, G.; STEWART, A.; VERONESI, M.; STOCKMAN, B., WaterLOGSY as a method for primary NMR screening: Practical aspects and range of applicability. *Journal of Biomolecular NMR* **2001,** *21* (4), 349-359.

DALVIT, C.; PEVARELLO, P.; TATÒ, M.; VERONESI, M.; VULPETTI, A.; SUNDSTRÖM, M., Identification of compounds with binding affinity to proteins via magnetization transfer from bulk water*. *Journal of Biomolecular NMR* **2000,** *18* (1), 65-68.

DAVIS, A. M.; TEAGUE, S. J.; KLEYWEGT, G. J., Application and Limitations of X-ray Crystallographic Data in Structure-Based Ligand and Drug Design. *Angewandte Chemie International Edition* **2003,** *42* (24), 2718-2736.

DEAR, G. J.; BEAUMONT, C.; ROBERTS, A.; SQUILLACI, B.; THOMAS, S.; NASH, M.; FRASER, D., Approaches for the rapid identification of drug metabolites in early clinical studies. *Bioanalysis* **2011,** *3* (2), 197-213.

DERRICK, T. S.; MCCORD, E. F.; LARIVE, C. K., Analysis of Protein/Ligand Interactions with NMR Diffusion Measurements: The Importance of Eliminating the Protein Background. *Journal of Magnetic Resonance* **2002,** *155* (2), 217-225.

DIDENKO, T.; BOELENS, R.; RÜDIGER, S. G. D., 3D DOSY-TROSY to determine the translational diffusion coefficient of large protein complexes. *Protein Engineering Design and Selection* **2012,** *25* (6), 319.

DIERCKS, T.; COLES, M.; KESSLER, H., Applications of NMR in drug discovery. *Current Opinion in Chemical Biology* **2001,** *5* (3), 285-291.

DIMASI, J. A.; HANSEN, R. W.; GRABOWSKI, H. G., The price of innovation: new estimates of drug development costs. *Journal of Health Economics* **2003,** *22* (2), 151-185.

DINGLEY, A.; MACKAY, J.; CHAPMAN, B.; MORRIS, M.; KUCHEL, P.; HAMBLY, B.; KING, G., Measuring protein self-association using pulsed-field-gradient NMR spectroscopy: Application to myosin light chain 2. *Journal of Biomolecular NMR* **1995,** *6* (3), 321-328.

DOYAMA, J. T.; RODRIGUES, H. G.; NOVELLI, E. L. B.; CEREDA, E.; VILEGAS, W., Chemical investigation and effects of the tea of Passiflora alata on biochemical parameters in rats. *Journal of Ethnopharmacology* **2005,** *96* (3), 371-374.

DREWS, J., Drug Discovery: A Historical Perspective. *Science* **2000,** *287* (5460), 1960-1964.

DUNNE, A.; JOWETT, M.; REES, S., Use of Primary Human Cells in High-Throughput Screens. In *High Throughput Screening*, Janzen, W. P.; Bernasconi, P., Eds. Humana Press: 2009; Vol. 565, pp 239-257.

EDRADA, R. A.; HEUBES, M.; BRAUERS, G.; WRAY, V.; BERG, A.; GRÄFE, U.; WOHLFARTH, M.; MÜHLBACHER, J.; SCHAUMANN, K.; SUDARSONO; BRINGMANN, G.; PROKSCH, P., Online Analysis of Xestodecalactones A−C, Novel Bioactive Metabolites from the Fungus Penicillium cf. montanense and Their Subsequent Isolation from the Sponge Xestospongia exigua#. *Journal of Natural Products* **2002,** *65* (11), 1598-1604.

EHRLICH P., Die Behandlung der Syphilis mit dem Ehrlichschen Präparat 606. *Deutsche medizinische Wochenschrift*: **1910**, 1893-1896.

EWING, T. A.; Makino, S.; Skillman, A. G.; Kuntz, I., DOCK 4.0: Search strategies for automated molecular docking of flexible molecule databases. *Journal of Computer-Aided Molecular Design* **2001,** *15* (5), 411-428.

FEHSKE, K. J.; SCHLÄFER, U.; WOLLERT, U.; MÜLLER, W. E.; Characterization of an important drug binding area on human serum albumin including the high-affinity binding sites of warfarin and azapropazone. *Molecular Pharmacology* **1982**, *21*, 387-393.

FERNÁNDEZ, C.; JAHNKE, W., New approaches for NMR screening in drug discovery. *Drug Discovery Today: Technologies* **2004,** *1* (3), 277-283.

FIELDING, L., NMR methods for the determination of protein–ligand dissociation constants. *Progress in Nuclear Magnetic Resonance Spectroscopy* **2007,** *51* (4), 219-242.

FLEMING, A., On the Antibacterial Action of Cultures of a Penicillium, with Special Reference to Their Use in the Isolation of B. influenzae. *British Journal of Experimental Pathology* **1929,** *10*, 226-236.

FOX, S.; FARR-JONES, S.; SOPCHAK, L.; BOGGS, A.; NICELY, H. W.; KHOURY, R.; BIROS, M., High-Throughput Screening: Update on Practices and Success. *Journal of Biomolecular Screening* **2006,** *11* (7), 864-869.

FRIEDMAN, M.; HENIKA, P. R.; MANDRELL, R. E., Antibacterial Activities of Phenolic Benzaldehydes and Benzoic Acids against Campylobacter jejuni, Escherichia coli, Listeria monocytogenes, and Salmonella enterica. *Journal of Food Protection* **2003,** *66* (10), 1811-1821.

GALE, D. C.; GOODLETT, D. R.; LIGHT-WAHL, K. J.; SMITH, R. D., Observation of Duplex DNA-Drug Noncovalent Complexes by Electrospray Ionization Mass Spectrometry. *Journal of the American Chemical Society* **1994,** *116* (13), 6027-6028.

GARGALLO, J. J. A. R., Experimental Methods for Studing the Interactions between G-Quadruplex Structures and Ligands. *Current Pharmaceutical Design* **2012,** *18* (4), 1900-1916.

GIBBS, S. J.; JOHNSON JR, C. S., A PFG NMR experiment for accurate diffusion and flow studies in the presence of eddy currents. *Journal of Magnetic Resonance (1969)* **1991,** *93* (2), 395-402.

GOODWIN, C. R.; FENN, L. S.; DEREWACZ, D. K.; BACHMANN, B. O.; MCLEAN, J. A., Structural Mass Spectrometry: Rapid Methods for Separation and Analysis of Peptide Natural Products. *Journal of Natural Products* **2012,** *75* (1), 48-53.

GOTTARELLI, G.; LENA, S.; MASIERO, S.; PIERACCINI, S.; SPADA, G. P., The use of circular dichroism spectroscopy for studying the chiral molecular self-assembly: An overview. *Chirality* **2008,** *20* (3-4), 471-485.

GOUNARIDES, J. S.; CHEN, A.; SHAPIRO, M. J., Nuclear magnetic resonance chromatography: applications of pulse field gradient diffusion NMR to mixture analysis and ligand–receptor interactions. *Journal of Chromatography B: Biomedical Sciences and Applications* **1999,** *725* (1), 79-90.

GRAUL, A. I.; REVEL, L.; TELL, M.; ROSA, E.; CRUCES, E., Overcoming the obstacles in the pharma/biotech industry: 2009 update. *Drug News Perspect* **2010,** *23* (1), 48-63.

GRIFFITHS, L., Optimization of Liquid Chromatography–NMR Spectroscopy: II — Saturation and Flow in On-Flow Liquid Chromatography – NMR Spectroscopy. *Magnetic Resonance in Chemistry* **1997,** *35* (4), 257-261.

GRIFFITHS, L., Optimization of NMR and HPLC Conditions for LC-NMR. *Analytical Chemistry* **1995,** *67* (22), 4091-4095.

HAJDUK, P. J.; OLEJNICZAK, E. T.; FESIK, S. W., One-Dimensional Relaxation- and Diffusion-Edited NMR Methods for Screening Compounds That Bind to Macromolecules. *Journal of the American Chemical Society* **1997,** *119* (50), 12257-12261.

HALPERIN, I.; MA, B.; WOLFSON, H.; Nussinov, R., Principles of docking: An overview of search algorithms and a guide to scoring functions. *Proteins: Structure, Function, and Bioinformatics* **2002,** *47* (4), 409-443.

HARTMANN, T.; SCHMITT, J.; RÖHRING, C.; NIMPTSCH, D.; NÖLLER, J.; MOHR, C., ADME related profiling in 96 and 384 well plate format--a novel and robust HT-assay for the determination of lipophilicity and serum albumin binding. *Current drug delivery* **2006,** *3* (2), 181-192.

HARTSHORN, M. J.; VERDONK, M. L.; CHESSARI, G.; BREWERTON, S. C.; MOOIJ, W. T. M.; MORTENSON, P. N.; MURRAY, C. W., Diverse, High-Quality Test Set for the Validation of Protein−Ligand Docking Performance. *Journal of Medicinal Chemistry* **2007,** *50* (4), 726-741.

HASELHORST, T.; BLANCHARD, H.; FRANK, M.; KRASCHNEFSKI, M. J.; KIEFEL, M. J.; SZYCZEW, A. J.; DYASON, J. C.; FLEMING, F.; HOLLOWAY, G.; COULSON, B. S.; VON ITZSTEIN, M.; STD NMR spectroscopy and molecular modeling investigation of the binding of N-acetylneuraminic acid derivatives to rhesus rotavirus VP8* core. *Glycobiology* **2007**, *17*, 68-81.

HEGDE, A.; SEETHARAMAPPA, J., Fluorescence and circular dichroism studies on binding and conformational aspects of an anti-leukemic drug with DNA. *Molecular Biology Reports* **2013**, 1-5.

HEGENER, O.; JORDAN, R.; HÄBERLEIN, H., Benzodiazepine Binding Studies on Living Cells: Application of Small Ligands for Fluorescence Correlation Spectroscopy. In *Biological Chemistry*, 2002; Vol. 383, p 1801.

HENRICHSEN, D.; ERNST, B.; MAGNANI, J. L.; WANG, W.-T.; MEYER, B.; PETERS, T., Bioaffinity NMR Spectroscopy: Identification of an E-Selectin Antagonist in a Substance Mixture by Transfer NOE. *Angewandte Chemie International Edition* **1999,** *38* (1-2), 98-102.

HILLISCH, A.; PINEDA, L. F.; HILGENFELD, R., Utility of homology models in the drug discovery process. *Drug Discovery Today* **2004,** *9* (15), 659-669.

HODGE, P.; MONVISADE, P.; MORRIS, G. A.; PREECE, I., A novel NMR method for screening soluble compound libraries. *Chemical Communications* **2001,** (3), 239-240.

HOPKINS, A. L.; GROOM, C. R., The druggable genome. *Nature Reviews Drug Discovery* **2002,** *1* (9), 727-730.

HOWE, I.; JARMAN, M., New Techniques for the Mass Spectrometry of Natural Products. In *Fortschritte der Chemie organischer Naturstoffe / Progress in the Chemistry of Organic Natural Products*, Herz, W.; Grisebach, H.; Kirby, G. W.; Tamm, C., Eds. Springer Vienna: 1985; Vol. 47, pp 107-152.

HSIEH, Y.; CHENG, K. C.; WANG, Y.; CHACKALAMANNIL, S.; XIA, Y.; KORFMACHER, W. A.; WHITE, R. E., The Role of Exploratory Drug Metabolism and Pharmacokinetics in New Drug Research: Case Study-Selection of a Thrombin Receptor Antagonist for Development. *Current Pharmaceutical Design* **2009,** *15* (19), 2262-2269.

HUGHES, J. P.; REES, S.; KALINDJIAN, S. B.; PHILPOTT, K. L., Principles of early drug discovery. *British Journal of Pharmacology* **2011,** *162* (6), 1239-1249.

HUNKELER, W.; MOHLER, H.; PIERI, L.; POLC, P.; BONETTI, E. P.; CUMIN, R.; SCHAFFNER, R.; Haefely, W., Selective antagonists of benzodiazepines. *Nature* **1981,** *290* (5806), 514-516.

HUO, R.; WEHRENS, R.; DUYNHOVEN, J. V.; BUYDENS, L. M. C., Assessment of techniques for DOSY NMR data processing. *Analytica Chimica Acta* **2003,** *490* (1–2), 231-251.

IMMING, P.; SINNING, C.; MEYER, A., Drugs, their targets and the nature and number of drug targets. *Nature Reviews Drug Discovery* **2006,** *5* (10), 821-834.

JAHNKE, W.; RÜDISSER, S.; ZURINI, M., Spin label enhanced NMR screening. *Journal of the American Chemical Society* **2001,** *123* (13), 3149-3150.

JAMESON, D. M.; CRONEY, J. C.; MOENS, P. D. J., [1] Fluorescence: Basic concepts, practical aspects, and some anecdotes. In *Methods in Enzymology*, Gerard Marriott, I. P., Ed. Academic Press: 2003; Vol. Volume 360, pp 1-43.

JAMESON, D. M.; SEIFRIED, S. E., Quantification of Protein–Protein Interactions Using Fluorescence Polarization. *Methods* **1999,** *19* (2), 222-233.

JAROSZEWSKI, J. W., Hyphenated NMR Methods in Natural Products Research, Part 2: HPLC-SPE-NMR and Other New Trends in NMR Hyphenation. *Planta Medica* **2005,** *71* (09), 795-802.

JAYAWICKRAMA, D. A.; LARIVE, C. K.; MCCORD, E. F.; ROE, D. C., Polymer additives mixture analysis using pulsed-field gradient NMR spectroscopy. *Magnetic Resonance in Chemistry* **1998,** *36* (10), 755-760.

JENKINS, K. M.; ANGELES, R.; QUINTOS, M. T.; XU, R.; KASSEL, D. B.; ROURICK, R. A., Automated high throughput ADME assays for metabolic stability and cytochrome P450 inhibition profiling of combinatorial libraries. *Journal of Pharmaceutical and Biomedical Analysis* **2004,** *34* (5), 989-1004.

JI, Z.; YAO, Z.; LIU, M., Saturation transfer difference nuclear magnetic resonance study on the specific binding of ligand to protein. *Analytical Biochemistry* **2009,** *385* (2), 380-382.

JIN, X.-L.; WEI, X.; QI, F.-M.; YU, S.-S.; ZHOU, B.; BAI, S., Characterization of hydroxycinnamic acid derivatives binding to bovine serum albumin. *Organic & Biomolecular Chemistry* **2012,** *10* (17), 3424-3431.

JOHNSON, J. I.; DECKER, S.; ZAHAREVITZ, D.; RUBINSTEIN, L. V.; VENDITTI, J. M.; SCHEPARTZ, S.; KALYANDRUG, S.; CHRISTIAN, M.; ARBUCK, S.; HOLLINGSHEAD, M.; SAUSVILLE, E. A., Relationships between drug activity in NCI preclinical in vitro and in vivo models and early clinical trials. *British Journal of Cancer* **2001,** *84* (10), 1424-1431.

JOHNSON, W. C., Protein secondary structure and circular dichroism: A practical guide. *Proteins: Structure, Function, and Bioinformatics* **1990,** *7* (3), 205-214.

JONES, G.; WILLETT, P.; GLEN, R. C., Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *Journal of Molecular Biology* **1995,** *245* (1), 43-53.

JONES, W. P.; YOUNG-WON, C.; KINGHORN, A. D., The Role of Pharmacognosy in Modern Medicine and Pharmacy. *Current Drug Targets* **2006,** *7* (3), 247-264.

JORISSEN, R. N.; GILSON, M. K., Virtual Screening of Molecular Databases Using a Support Vector Machine. *Journal of Chemical Information and Modeling* **2005,** *45* (3), 549-561.

KAWASHTY, S. A.; MOSHARRAFA, S. A. M.; El-Gibali, M.; Saleh, N. A. M. *Biochemical Systematics and Ecology* **2000**, *28*, 915-917.

KITCHEN, D. B.; DECORNEZ, H.; FURR, J. R.; BAJORATH, J., Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature Reveiws Drug Discovery* **2004,** *3* (11), 935-949.

KLEBE, G., Virtual ligand screening: strategies, perspectives and limitations. *Drug Discovery Today* **2006,** *11* (13–14), 580-594.

KLEIN, J.; MEINECKE, M. R.; MEYER, B.; *Journal of the American Chemical Society* **1999**, *121*, 5336-5337.

KNOWLES, J.; GROMO, G., Target selection in drug discovery. *Nature Reviews Drug Discovery* **2003,** *2* (1), 63-69.

KOLA, I.; LANDIS, J., Can the pharmaceutical industry reduce attrition rates? *Nature Reviews Drug Discovery* **2004,** *3* (8), 711-716.

KOLKMAN, A. J.; Application and Development of NMR Based Methodologies for P450BM3 Drug Metabolism, Metabolite Identification and Mixture Analysis. Radboud University Nijmegen, Printed by Ipskamp Drukkers, 2013.

KORFMACHER, W. A., Lead optimization strategies as part of a drug metabolism environment. *Current opinion in drug discovery & development* **2003,** *6* (4), 481-485.

KOYAMA, H.; SUGIOKA, N.; UNO, A.; MORI, S.; NAKAJIMA, K., Effects of glycosylation of hypoglycaemic drug binding to serum albumin. *Biopharmaceutics & Drug Disposition* **1997,** *18* (9), 791-801.

KRAMER, M.; KLEINPETER, E., STD-DOSY: A new NMR method to analyze multi-component enzyme/substrate systems. *Journal of Magnetic Resonance* **2010,** *202* (2), 245-249.

KUFAREVA, I.; RUEDA, M.; KATRITCH, V.; STEVENS, RAYMOND C.; ABAGYAN, R., Status of GPCR Modeling and Docking as Reflected by Community-wide GPCR Dock 2010 Assessment. *Structure (London, England : 1993)* **2011,** *19* (8), 1108-1126.

KUNTZ, I. D.; BLANEY, J. M.; OATLEY, S. J.; LANGRIDGE, R.; FERRIN, T. E., A geometric approach to macromolecule-ligand interactions. *Journal of Molecular Biology* **1982,** *161* (2), 269-288.

KUNTZ, I. D.; MENG, E. C.; SHOICHET, B. K., Structure-Based Molecular Design. *Accounts of Chemical Research* **1994,** *27* (5), 117-123.

LANGLEY, J. N., On the reaction of cells and of nerve-endings to certain poisons, chiefly as regards the reaction of striated muscle to nicotine and to curari. *The Journal of Physiology* **1905,** *33* (4-5), 374-413.

LEE, H., Pharmaceutical Applications of Liquid Chromatography Coupled with Mass Spectrometry (LC/MS). *Journal of Liquid Chromatography & Related Technologies* **2005,** *28* (7-8), 1161-1202.

LENGAUER, T.; LEMMEN, C.; RAREY, M.; ZIMMERMANN, M., Novel technologies for virtual screening. *Drug Discovery Today* **2004,** *9* (1), 27-34.

LI, D.; DEROSE, E.; LONDON, R., The inter-ligand Overhauser effect: A powerful new NMR approach for mapping structural relationships of macromolecular ligands. *Journal of Biomolecular NMR* **1999,** *15* (1), 71-76.

LI, D.; KERESZTES, I.; HOPSON, R.; WILLIARD, P. G., Characterization of Reactive Intermediates by Multinuclear Diffusion-Ordered NMR Spectroscopy (DOSY). *Accounts of Chemical Research* **2008,** *42* (2), 270-280.

LI, W.-Y.; XU, J.-G.; GUO, X.-Q.; ZHU, Q.-Z.; ZHAO, Y.-B., Study on the interaction between rivanol and DNA and its application to DNA assay. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **1997,** *53* (5), 781-787.

LILIANA BÎRLĂ, B. P., THIERRY NOGUEL, JEAN VIGO AND ANNE-CÉCILE RIBOU.; Comparative study of the interaction between BSA and three fluorescent probes: investigation of some critical parameters. *Revue Roumaine de Chimie* **2007,** *52* (7), 639-646.

LIN, M.; SHAPIRO, M. J.; WAREING, J. R., Diffusion-Edited NMR−Affinity NMR for Direct Observation of Molecular Interactions. *Journal of the American Chemical Society* **1997,** *119* (22), 5249-5250.

LIPINSKI, C. A.; LOMBARDO, F.; DOMINY, B. W.; FEENEY, P. J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews* **2001,** *46* (1–3), 3-26.

LIU, M.; MAO, X.-A.; YE, C.; Huang, H.; Nicholson, J. K.; Lindon, J. C., Improved WATERGATE Pulse Sequences for Solvent Suppression in NMR Spectroscopy. *Journal of Magnetic Resonance* **1998,** *132* (1), 125-129.

LOMBARDINO, J. G.; LOWE, J. A., The role of the medicinal chemist in drug discovery-then and now. *Nature Reviews Drug Discovery* **2004,** *3* (10), 853-862.

LONDON, R. E., Theoretical Analysis of the Inter-Ligand Overhauser Effect: A New Approach for Mapping Structural Relationships of Macromolecular Ligands. *Journal of Magnetic Resonance* **1999,** *141* (2), 301-311.

LOPES, G. C.; NAKAMURA, C. V.; DIAS FILHO, B. P.; MELLO, J. C. P.; Estudo físico-químico, químico e biológico de extrato das cascas de

Stryphnodendron polyphyllum Mart. (Leguminosae). *Revista Brasileira de Farmacognosia* **2003**, *13*, 24-27.

LOPES, G. C.; SANCHES, A. C. C.; NAKAMURA, C. V.; DIAS FILHO, B. P.; HERNANDES, L.; MELLO, J. C. P. D.; Influence of extracts of Stryphnodendron polyphyllum Mart. and Stryphnodendron obovatum Benth. on the cicatrisation of cutaneous wounds in rats. *Journal of Ethnopharmacology* **2005**, *99*, 265-272.

LORUSSO, P. M., Phase 0 Clinical Trials: An Answer to Drug Development Stagnation? *Journal of Clinical Oncology* **2009,** *27* (16), 2586-2588.

LUCH, A., *Molecular, Clinical and Environmental Toxicology: Volume 1: Molecular Toxicology*. Springer London, Limited: 2009.

LUDWIG C.; GUENTHER U. L.; Ligand based NMR methods for drug discovery. *Frontiers in Bioscience* **2009**, 1(14), 4565-4574.

LUDWIG, C.; MICHIELS, P. J. A.; LODI, A.; RIDE, J.; BUNCE, C.; GÜNTHER, U. L., Evaluation of Solvent Accessibility Epitopes for Different Dehydrogenase Inhibitors. *ChemMedChem* **2008,** *3* (9), 1371-1376.

LUDWIG, C.; MICHIELS, P. J. A.; WU, X.; KAVANAGH, K. L.; PILKA, E.; JANSSON, A.; OPPERMANN, U.; GÜNTHER, U. L., SALMON: Solvent Accessibility, Ligand binding, and Mapping of ligand Orientation by NMR Spectroscopy. *Journal of Medicinal Chemistry* **2007,** *51* (1), 1-3.

M, H.; AZZAZY, E.; CHRISTENSON, R. H., All About Albumin: Biochemistry, Genetics, and Medical Applications. Theodore Peters, Jr. San Diego, CA: Academic Press, 1996, 432 pp, $85.00. ISBN 0-12-552110-3. *Clinical Chemistry* **1997,** *43* (10), 2014a-2015.

MACARRON, R.; BANKS, M. N.; BOJANIC, D.; BURNS, D. J.; CIROVIC, D. A.; GARYANTES, T.; GREEN, D. V. S.; HERTZBERG, R. P.; JANZEN, W. P.; PASLAY, J. W.; SCHOPFER, U.; SITTAMPALAM, G. S., Impact of high-throughput screening in biomedical research. *Nature Reviews Drug Discovery* **2011,** *10* (3), 188-195.

MAGDE, D.; WEBB, W. W.; ELSON, E. L., Fluorescence correlation spectroscopy. III. Uniform translation and laminar flow. *Biopolymers* **1978,** *17* (2), 361-376.

MAPLE, H. J.; GARLISH, R. A.; RIGAU-ROCA, L.; PORTER, J.; WHITCOMBE, I.; PROSSER, C. E.; KENNEDY, J.; HENRY, A. J.; TAYLOR, R.

J.; CRUMP, M. P.; CROSBY, J., Automated Protein–Ligand Interaction Screening by Mass Spectrometry. *Journal of Medicinal Chemistry* **2011,** *55* (2), 837-851.

Marston, A.; Hostettmann, K.; Natural product analysis over the last decades. *Planta Med* **2009**, *75*, 672-682.

MARTIN, S. F.; WOOD, A. D.; MCROBBIE, M. M.; MAZILU, M.; MCDONALD, M. P.; SAMUEL, I. D. W.; HERRINGTON, C. S., Fluorescence spectroscopy of an in vitro model of human cervical precancer identifies neoplastic phenotype. *International Journal of Cancer* **2007,** *120* (9), 1964-1970.

MASIH, N. G.; SINGH, B. S., Phytochemical Screening of Some Plants Used in Herbal Based Cosmetic Preparations. In *Chemistry of Phytopotentials: Health, Energy and Environmental Perspectives*, Khemani, L. D.; Srivastava, M. M.; Srivastava, S., Eds. Springer Berlin Heidelberg: 2012; pp 111-112.

MASILAMANI, V.; VIJMASI, T.; AL SALHI, M.; GOVINDARAJ, K.; VIJAYA-RAGHAVAN, A. P.; Antonisamy, B., Cancer detection by native fluorescence of urine. *Journal of Biomedical Optics* **2010,** *15* (5), 057003-057003-9.

MAURER, H., Current role of liquid chromatography–mass spectrometry in clinical and forensic toxicology. *Analytical and Bioanalytical Chemistry* **2007,** *388* (7), 1315-1325.

MAYA, J. D.; CASSELS, B. K.; ITURRIAGA-VÁSQUEZ, P.; FERREIRA, J.; FAÚNDEZ, M.; GALANTI, N.; FERREIRA, A.; MORELLO, A., Mode of action of natural and synthetic drugs against Trypanosoma cruzi and their interaction with the mammalian host. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology* **2007,** *146* (4), 601-620.

MAYER, M.; MEYER, B., Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy. *Angewandte Chemie International Edition* **1999,** *38* (12), 1784-1788.

MAYER, M.; MEYER, B., Group Epitope Mapping by Saturation Transfer Difference NMR To Identify Segments of a Ligand in Direct Contact with a Protein Receptor. *Journal of the American Chemical Society* **2001,** *123* (25), 6108-6117.

MCGREEVY PB & MARSDEN PD (1986). CAMPBELL WC & REW RS (Editors), Chemotheraphy of Parasitic Diseases. Vol. 1, Plenum Press, New York, 115-127.

MCINNES, C., Virtual-screening strategies in drug discovery. *Current Opinion in Chemical Biology* **2007,** *11* (5), 494-502.

MERCK, G., Vorläufige Notiz über eine neue organische Base im Opium. *Justus Liebigs Annalen der Chemie* **1848,** *66* (1), 125-128.

MEYER, B.; PETERS, T., NMR Spectroscopy Techniques for Screening and Identifying Ligand Binding to Protein Receptors. *Angewandte Chemie International Edition* **2003,** *42* (8), 864-890.

MEYER, B.; WEIMAR, T.; PETERS, T.; Screening Mixtures for Biological Activity by NMR. *European Journal of Biochemistry* **1997**, *246 (3)*, 705-709.

MICHELINI, E.; CEVENINI, L.; MEZZANOTTE, L.; COPPA, A.; RODA, A., Cell-based assays: fuelling drug discovery. *Analytical and Bioanalytical Chemistry* **2010,** *398* (1), 227-238.

MILLER, J. N., Recent developments in fluorescence and chemiluminescence analysis. Plenary lecture. *Analyst* **1984,** *109* (3), 191-198.

MOORE, K.; REES, S., Cell-Based Versus Isolated Target Screening: How Lucky Do You Feel? *Journal of Biomolecular Screening* **2001,** *6* (2), 69-74.

MORRIS, K. F.; JOHNSON, C. S., Diffusion-ordered two-dimensional nuclear magnetic resonance spectroscopy. *Journal of the American Chemical Society* **1992,** *114* (8), 3139-3141.

MORRIS, K. F.; JOHNSON, C. S., Resolution of discrete and continuous molecular size distributions by means of diffusion-ordered 2D NMR spectroscopy. *Journal of the American Chemical Society* **1993,** *115* (10), 4291-4299.

MÜHLEBACH, A.; ADAM, J.; SCHÖN, U., Streamlined approach to high-quality purification and identification of compound series using high-resolution MS and NMR. *Journal of Separation Science* **2011,** *34* (21), 2983-2988.

NI, F., Recent developments in transferred NOE methods. *Progress in Nuclear Magnetic Resonance Spectroscopy* **1994,** *26, Part 6* (0), 517-606.

NIENHAUS, K.; NIENHAUS, G. U., Probing Heme Protein-Ligand Interactions by UV/Visible Absorption Spectroscopy. In *Protein-Ligand Interactions*, Ulrich Nienhaus, G., Ed. Humana Press: 2005; Vol. 305, pp 215-241.

NILSSON, M., The DOSY Toolbox: A new tool for processing PFG NMR diffusion data. *Journal of Magnetic Resonance* **2009,** *200* (2), 296-302.

On-line LC-NMR and Related Techniques. Albert, K., Ed. Wiley & Blackwell: august 2002; p. 306. http://www.wiley.com/WileyCDA/WileyTitle/productCd-0471496499.html

ORLOFF, J.; DOUGLAS, F.; PINHEIRO, J.; LEVINSON, S.; BRANSON, M.; CHATURVEDI, P.; ETTE, E.; GALLO, P.; HIRSCH, G.; MEHTA, C.; PATEL, N.; SABIR, S.; SPRINGS, S.; STANSKI, D.; EVERS, M. R.; FLEMING, E.; SINGH, N.; TRAMONTIN, T.; GOLUB, H., The future of drug development: advancing clinical trial design. *Nature Reviews Drug Discovery* **2009,** *8* (12), 949-957.

PALAMAR, J., How ephedrine escaped regulation in the United States: A historical review of misuse and associated policy. *Health Policy* **2011,** *99* (1), 1-9.

PANOUTSOPOULOS, G. I.; BEEDHAM, C.; Enzymatic oxidation of phthalazine with guinea pig liver aldehyde oxidase and liver slices: inhibition by isovanillin. *Acta Biochimica Polonica* **2004,** *51* (4), 943-951.

PANOUTSOPOULOS, G. I.; KOURETAS, D.; BEEDHAM, C., Contribution of Aldehyde Oxidase, Xanthine Oxidase, and Aldehyde Dehydrogenase on the Oxidation of Aromatic Aldehydes. *Chemical Research in Toxicology* **2004,** *17* (10), 1368-1376.

PANOUTSOPOULOS, G.; KOURETAS, D.; GOUNARIS, E.; BEEDHAM, C., Metabolism of 2-phenylethylamine and phenylacetaldehyde by precision-cut guinea pig fresh liver slices. *European Journal of Drug Metabolism and Pharmacokinetics* **2004,** *29* (2), 111-118.

PARELLA, T.; Pulse Program Catalogue, NMRGuide4.0, Bruker BioSpin, 2004.

PARK, J.-H.; PARK, Y.-S.; LEE, M.-H.; RHIM, S.-Y.; SONG, J.-C.; LEE, S.-J.; KIM, J.-M.; SHAW, L. M.; KANG, J.-S., Determination of plasma topiramate concentration using LC-MS/MS for pharmacokinetic and bioequivalence studies in healthy Korean volunteers. *Biomedical Chromatography* **2008,** *22* (8), 822-829.

PEARCE, B. C.; LANGLEY, D. R.; KANG, J.; HUANG, H.; KULKARNI, A., E-Novo: An Automated Workflow for Efficient Structure-Based Lead Optimization. *Journal of Chemical Information and Modeling* **2009,** *49* (7), 1797-1809.

PELLECCHIA, M.; BERTINI, I.; COWBURN, D.; DALVIT, C.; GIRALT, E.; JAHNKE, W.; JAMES, T. L.; HOMANS, S. W.; KESSLER, H.; LUCHINAT, C.; MEYER, B.; OSCHKINAT, H.; PENG, J.; SCHWALBE, H.; SIEGAL, G., Perspectives on NMR in drug discovery: a technique comes of age. *Nature Reviews Drug Discovery* **2008,** *7* (9), 738-745.

PELSY, G.; KLIBANOV, A. M., Remarkable positional (Regio)specificity of xanthine oxidase and some dehydrogenases in the reactions with substituted benzaldehydes. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* **1983,** *742* (2), 352-357.

PENG, J. W.; MOORE, J.; ABDUL-MANAN, N., NMR experiments for lead generation in drug discovery. *Progress in Nuclear Magnetic Resonance Spectroscopy* **2004,** *44* (3–4), 225-256.

PEREIRA, A.; PFEIFER, T. A.; GRIGLIATTI, T. A.; ANDERSEN, R. J., Functional Cell-Based Screening and Saturation Transfer Double-Difference NMR Have Identified Haplosamate A as a Cannabinoid Receptor Agonist. *ACS Chemical Biology* **2009,** *4* (2), 139-144.

PIOTTO, M.; SAUDEK, V.; SKLENÁŘ, V., Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *Journal of Biomolecular NMR* **1992,** *2* (6), 661-665.

POCARD, M.; SORIA, J.-C.; ALDAZ-CARROLL, L.; BELLET, D., Phase 0 Clinical Trials in Oncology: An Exploratory Methodology for Constructing a Study With Patients Undergoing Surgery for Metastatic Disease. *Journal of Clinical Oncology* **2010,** *28* (30), 4551-4553.

POLITI, M.; ALVARO-BLANCO, J.; GROVES, P.; PRIETO, A.; LEAL, J. A.; CAÑADA, F. J.; JIMÉNEZ-BARBERO,J.; *European Journal of Organic Chemistry* **2006**, *2006,* 2067-2073.

POLITI, M.; CHÁVEZ, M. I.; CAÑADA, F. J.; JIMÉNEZ-BARBERO, J., Screening by NMR: A New Approach for the Study of Bioactive Natural Products? The Example of Pleurotus ostreatus Hot Water Extract. *European Journal of Organic Chemistry* **2005,** (7), 1392-1396.

POST, C. B., Exchange-transferred NOE spectroscopy and bound ligand structure determination. *Current Opinion in Structural Biology* **2003,** *13* (5), 581-588.

POVEDA, A.; JIMENEZ-BARBERO, J., NMR studies of carbohydrate-protein interactions in solution. *Chemical Society Reviews* **1998,** *27* (2), 133-144.

PRAKASH, C.; SHAFFER, C. L.; NEDDERMAN, A., Analytical strategies for identifying drug metabolites. *Mass Spectrometry Reviews* **2007,** *26* (3), 340-369.

PUIG, O.; CASPARY, F.; RIGAUT, G.; RUTZ, B.; BOUVERET, E.; BRAGADO-NILSSON, E.; WILM, M.; SÉRAPHIN, B., The Tandem Affinity Purification (TAP) Method: A General Procedure of Protein Complex Purification. *Methods* **2001,** *24* (3), 218-229.

PUTIGNANO, P.; KALTSAS, G. A.; KORBONITS, M.; JENKINS, P. J.; MONSON, J. P.; BESSER, G. M.; GROSSMAN, A. B., Alterations in Serum Protein Levels in Patients with Cushing's Syndrome before and after Successful Treatment. *Journal of Clinical Endocrinology & Metabolism* **2000,** *85* (9), 3309-3312.

RANG, H. P., *Drug discovery and development: technology in transition*. Churchill Livingstone / Elsevier: 2006.

RAREY, M.; KRAMER, B.; LENGAUER, T.; KLEBE, G., A Fast Flexible Docking Method using an Incremental Construction Algorithm. *Journal of Molecular Biology* **1996,** *261* (3), 470-489.

RAYYAN, S.; FOSSEN, T.; SOLHEIM NATELAND, H.; ANDERSEN, O. M., Isolation and identification of flavonoids, including flavone rotamers, from the herbal drug 'Crataegi folium cum flore' (hawthorn). *Phytochemical Analysis* **2005,** *16* (5), 334-341.

RISHTON, G. M., Reactive compounds and in vitro false positives in HTS. *Drug Discovery Today* **1997,** *2* (9), 382-384.

RODA, A., Analytical chemistry for drug discovery and development. *Analytical Bioanalytical Chemistry* **2010,** *398* (1), 27-28.

RODRIGUES, E. D.; DA SILVA, D. B.; DE OLIVEIRA, D. C. R.; DA SILVA, G. V. J.; *Magnetic Resonance in Chemistry* **2009**, *47*, 1095-1100.

RODKEY, F. L., Direct Spectrophotometric Determination of Albumin in Human Serum. *Clinical Chemistry* **1965,** *11* (4), 478-487.

RUBIO, S.; GOMEZ-HENS, A.; VALCARCEL, M., Analytical applications of synchronous fluorescence spectroscopy. *Talanta* **1986,** *33* (8), 633-640.

RUSS, A. P.; LAMPEL, S., The druggable genome: an update. *Drug Discovery Today* **2005,** *10* (23–24), 1607-1610.

S. GOPALAKRISHNAN, E. V., Gc-ms analysis of some bioactive constituents of mussaenda frondosa linn. *International Journal of Pharma and Bio Sciences* **2011,** *2* (1), 313-320.

SÁNCHEZ-PEDREGAL, V. M.; REESE, M.; MEILER, J.; BLOMMERS, M. J. J.; GRIESINGER, C.; CARLOMAGNO, T., The INPHARMA Method: Protein-Mediated Interligand NOEs for Pharmacophore Mapping. *Angewandte Chemie* **2005,** *117* (27), 4244-4247.

SASAKI, K.; HOSOYA, R.; WANG, Y.-M.; RAULSTON, G. L., Formation and disposition of 7-hydroxymethorexate in rabbits. *Biochemical Pharmacology* **1983,** *32* (3), 503-507.

SCHAUSS, A. G.; WU, X.; PRIOR, R. L.; OU, B.; PATEL, D.; HUANG, D.; KABABICK, J. P., Phytochemical and Nutrient Composition of the Freeze-Dried Amazonian Palm Berry, Euterpe oleraceae Mart. (Acai). *Journal of Agricultural and Food Chemistry* **2006,** *54* (22), 8598-8603.

SCHLOSSER, J.; RAREY, M., Beyond the Virtual Screening Paradigm: Structure-Based Searching for New Lead Compounds. *Journal of Chemical Information and Modeling* **2009,** *49* (4), 800-809.

SCHLOTTERBECK, G.; CECCARELLI, S. M., LC–SPE–NMR–MS: a total analysis system for bioanalysis. *Bioanalysis* **2009,** *1* (3), 549-559.

SCHNEIDER, G.; BÖHM, H.-J., Virtual screening and fast automated docking methods. *Drug Discovery Today* **2002,** *7, Supplement 1* (0), 64-70.

SCHOONJANS, V.; QUESTIER, F.; BOROSY, A. P.; WALCZAK, B.; MASSART, D. L.; HUDSON, B. D., Use of mass spectrometry for assessing similarity/diversity of natural products with unknown chemical structures. *Journal of Pharmaceutical and Biomedical Analysis* **2000,** *21* (6), 1197-1214.

SCHWARTZ, W. B., The Effect of Sulfanilamide on Salt and Water Excretion in Congestive Heart Failure. *New England Journal of Medicine* **1949,** *240* (5), 173-177.

SEVRIOUKOVA, I. F.; LI, H.; Zhang, H.; Peterson, J. A.; Poulos, T. L., Structure of a cytochrome P450–redox partner electron-transfer complex. *Proceedings of the National Academy of Sciences* **1999,** *96* (5), 1863-1868.

SHAKLAI, N.; GARLICK, R. L.; BUNN, H. F., Nonenzymatic glycosylation of human serum albumin alters its conformation and function. *Journal of Biological Chemistry* **1984,** *259* (6), 3812-7.

SHOICHET, B. K., Virtual screening of chemical libraries. *Nature* **2004,** *432* (7019), 862-865.

SHUKER, S. B.; HAJDUK, P. J.; MEADOWS, R. P.; FESIK, S. W., Discovering High-Affinity Ligands for Proteins: SAR by NMR. *Science* **1996,** *274* (5292), 1531-1534.

SIEGAL, G.; AB, E.; SCHULTZ, J., Integration of fragment screening and library design. *Drug Discovery Today* **2007,** *12* (23–24), 1032-1039.

SIRAJUDDIN, M.; ALI, S.; BADSHAH, A., Drug–DNA interactions and their study by UV–Visible, fluorescence spectroscopies and cyclic voltametry. *Journal of Photochemistry and Photobiology B: Biology* **2013,** *124* (0), 1-19.

STARK, J.; POWERS, R., Application of NMR and Molecular Docking in Structure-Based Drug Discovery. In *NMR of Proteins and Small Biomolecules*, Zhu, G., Ed. Springer Berlin Heidelberg: **2012**; Vol. 326, pp 1-34.

STARK, J.; POWERS, R., Rapid Protein−Ligand Costructures Using Chemical Shift Perturbations. *Journal of the American Chemical Society* **2007,** *130* (2), 535-545.

STEJSKAL, E. O.; TANNER, J. E., Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient. *The Journal of Chemical Physics* **1965,** *42* (1), 288-292.

SUDLOW, G.; BIRKETT, D. J.; WADE, D. N., Further Characterization of Specific Drug Binding Sites on Human Serum Albumin. *Molecular Pharmacology* **1976,** *12* (6), 1052-1061.

SUH, D.; CHAIRES, J. B., Criteria for the mode of binding of DNA binding agents. *Bioorganic & Medicinal Chemistry* **1995,** *3* (6), 723-728.

SUN, H.; XIANG, J.; LIU, Y.; LI, L.; LI, Q.; XU, G.; TANG, Y., A stabilizing and denaturing dual-effect for natural polyamines interacting with G-quadruplexes depending on concentration. *Biochimie* **2011,** *93* (8), 1351-1356.

Symposium on Drug Responses in Man, W. G. E. W. C. F. In *Drug responses in man. Edited by Gordon Wolstenholme and Ruth Porter*, London, 1967; Churchill: London.

TAGGI, A. E.; MEINWALD, J.; SCHROEDER, F. C., A New Approach to Natural Products Discovery Exemplified by the Identification of Sulfated Nucleosides in Spider Venom. *Journal of the American Chemical Society* **2004,** *126* (33), 10364-10369.

TANOLI, S. A. K.; TANOLI, N. U.; BONDANCIA, T. M.; USMANI, S.; KERSSEBAUM, R.; FERREIRA, A. G.; FERNANDES, J. B.; UL-HAQ, Z.; Crude to leads: a triple-pronged direct NMR approach in coordination with docking simulation. *Analyst* **2013**, *138*, 5137-5145.

TANOLI, S. A. K.; TANOLI, N. U.; USMANI, S.; UL-HAQ, Z.; FERREIRA, A. G., The exploration of interaction studies of smaller size, mostly ignored yet intrinsically inestimable molecules towards BSA; An example of STD and DOSY NMR. *Central Eurpean Journal of Chemistry.* **2014,** *12* (3), 332-340.

TAUFER, M.; CROWLEY, M.; PRICE, D. J.; CHIEN, A. A.; BROOKS, C. L., Study of a highly accurate and fast protein–ligand docking method based on molecular dynamics. *Concurrency and Computation: Practice and Experience* **2005,** *17* (14), 1627-1641.

TENNANT, J. M.; BARKHUIZEN, D. A.; RAGHOO, M.; MAGUIRE, G. E. M. APTA: A Simple Ligand for Lanthanide Energy Transfer Luminescence in Aqueous Solution. *South African Journal of Chemistry* **2006**, 59, 86-89.

TOM J. MABRY KRM, M. B. THOMAS.; The Systematic Identification of Flavonoids. In: *Angewandte Chemie International Edition.* Berlin Heidelberg: Springer Berlin Heidelberg; 1970: 355.

TREFI, S.; GILARD, V.; BALAYSSAC, S.; MALET-MARTINO, M.; MARTINO, R., Quality assessment of fluoxetine and fluvoxamine pharmaceutical formulations purchased in different countries or via the Internet by 19F and 2D DOSY 1H NMR. *Journal of Pharmaceutical and Biomedical Analysis* **2008,** *46* (4), 707-722.

TREFI, S.; ROUTABOUL, C.; HAMIEH, S.; GILARD, V.; MALET-MARTINO, M.; MARTINO, R., Analysis of illegally manufactured formulations of tadalafil (Cialis®) by 1H NMR, 2D DOSY 1H NMR and Raman spectroscopy. *Journal of Pharmaceutical and Biomedical Analysis* **2008,** *47* (1), 103-113.

TSENG, L.-H.; BRAUMANN, U.; GODEJOHANN, M.; LEE, S.-S.; ALBERT, K., Structure Identification of Aporphine Alkaloids by On-Line Coupling of HPLC-NMR with Loop Storage. *Journal of the Chinese Chemical Society* **2000,** *47* (6), 1231-1236.

TWOMBLY, R., Slow Start to Phase 0 as Researchers Debate Value. *Journal of the National Cancer Institute* **2006,** *98* (12), 804-806.

VAN DE WATERBEEMD, H.; SMITH, D. A.; BEAUMONT, K.; WALKER, D. K., Property-Based Design: Optimization of Drug Absorption and Pharmacokinetics. *Journal of Medicinal Chemistry* **2001,** *44* (9), 1313-1333.

VELOSO, E. D. S.; Fitoquímica comparada dos Gêneros Angostura, Almeidea e Rauia (Rutaceae). . Universidade Federal de São Carlos, UFSCar, 1995.

VELOZO, E. S.; OLIVEIRA, D. J.; ARRUDA, A. C.; VIEIRA, P. C.; FERNANDES, J. B.; DAS, M. F.; DA SILVA, G. F.; CARACELLI, I.; ZUKERMAN-SCHPECTOR, J., RAUIANIN.; A New Coumarin from Rauia Resinosa. *Natural Product Letters* **1997,** *9* (4), 237-244.

VIDAL, D.; THORMANN, M.; PONS, M., A Novel Search Engine for Virtual Screening of Very Large Databases. *Journal of Chemical Information and Modeling* **2006,** *46* (2), 836-843.

VIGNESH, G.; NEHRU, S.; MANOJKUMAR, Y.; ARUNACHALAM, S., Spectroscopic investigation on the interaction of some surfactant-cobalt(III) complexes with serum albumins. *Journal of Luminescence* **2014,** *145* (0), 269-277.

WAGSTAFF, J. L.; TAYLOR, S. L.; HOWARD, M. J., Recent developments and applications of saturation transfer difference nuclear magnetic resonance (STD NMR) spectroscopy. *Molecular BioSystems* **2013,** *9* (4), 571-577.

WAGSTAFF, J. L.; VALLATH, S.; MARSHALL, J. F.; WILLIAMSON, R. A.; HOWARD, M. J., Two-dimensional heteronuclear saturation transfer difference NMR reveals detailed integrin [small alpha]v[small beta]6 protein-peptide interactions. *Chemical Communications* **2010,** *46* (40), 7533-7535.

WAN, K. X.; GROSS, M. L.; SHIBUE, T., Gaphase stability of double-stranded oligodeoxynucleotides and their noncovalent complexes with DNA-binding drugs as revealed by collisional activation in an ion trap. *Journal of the American Society for Mass Spectrometry* **2000,** *11* (5), 450-457.

WARIDEL, P.; WOLFENDER, J.-L.; LACHAVANNE, J.-B.; HOSTETTMANN, K., ent-Labdane glycosides from the aquatic plant Potamogeton lucens and analytical evaluation of the lipophilic extract constituents of various Potamogeton species. *Phytochemistry* **2004,** *65* (7), 945-954.

WARREN, G. L.; ANDREWS, C. W.; CAPELLI, A.-M.; CLARKE, B.; LALONDE, J.; LAMBERT, M. H.; LINDVALL, M.; NEVINS, N.; SEMUS, S. F.; SENGER, S.; TEDESCO, G.; WALL, I. D.; WOOLVEN, J. M.; PEISHOFF, C. E.; HEAD, M. S., A Critical Assessment of Docking Programs and Scoring Functions. *Journal of Medicinal Chemistry* **2005,** *49* (20), 5912-5931.

WHITE, R. H., The Difficult Road from Sequence to Function. *Journal of Bacteriology* **2006,** *188* (10), 3431-3432.

WILLIAMSON, M. P., Using chemical shift perturbation to characterise ligand binding. *Progress in Nuclear Magnetic Resonance Spectroscopy* **2013,** *73* (0), 1-16.

WOODS, A.; KOOMEN, J.; RUOTOLO, B.; GILLIG, K.; RUSSEL, D.; FUHRER, K.; GONIN, M.; EGAN, T.; SCHULTZ, J. A., A study of peptide—Peptide interactions using MALDI ion mobility o-TOF and ESI mass spectrometry. *Journal of the American Society for Mass Spectrometry* **2002,** *13* (2), 166-169.

WRIGHT, P., Metabolite identification by mass spectrometry: forty years of evolution. *Xenobiotica* **2011,** *41* (8), 670-686.

XIA, Y.; ZHU, Q.; JUN, K.-Y.; WANG, J.; GAO, X., Clean STD-NMR spectrum for improved detection of ligand-protein interactions at low concentration of protein. *Magnetic Resonance in Chemistry* **2010,** *48* (12), 918-924.

XIE, C.; VEITCH, N. C.; HOUGHTON, P. J.; Simmonds, M. S. J., Flavone C-Glycosides from Viola yedoensis MAKINO. *Chemical and Pharmaceutical Bulletin* **2003,** *51* (10), 1204-1207.

YANG, Y.; ADELSTEIN, S. J.; KASSIS, A. I., Target discovery from data mining approaches. *Drug Discovery Today* **2009,** *14* (3–4), 147-154.

YEKKALA, R.; ADAMS, E.; VAN SCHEPDAEL, A.; HOOGMARTENS, J., Quality control of protease inhibitors. *Journal of Pharmaceutical Sciences* **2008,** *97* (6), 2012-2021.

YU, X.; YANG, Y.; SHIYU, L.; YAO, Q.; HETING, L.; XIAOFANG, L.; PINGGUI, Y., The fluorescence spectroscopic study on the interaction between imidazo[2,1-b]thiazole analogues and bovine serum albumin. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **2011,** *83* (1), 322-328.

YU, X.; YAO, Q.; TAO, H.; YANG, Y.; LI, L.; ZHENG, B.; ZHU, S., A study on the interaction between 3-spiro-piperidones and bovine serum albumin using spectroscopic approaches. *Luminescence* **2013,** *28* (5), 705-712.

YUANLONG, Y.; YANMING, Y.; FUMING, L.; YUFEN, L.; PAOZHONG, M., Characteristic autofluorescence for cancer diagnosis and its origin. *Lasers in Surgery and Medicine* **1987,** *7* (6), 528-532.

ZHENG, J.; SUGRUE, R. J.; TANG, K., Mass spectrometry based proteomic studies on viruses and hosts – A review. *Analytica Chimica Acta* **2011,** *702* (2), 149-159.

ZOLGHADRI, S.; SABOURY, A. A.; Amin, E.; Moosavi-Movahedi, A. A., A spectroscopic study on the interaction between ferric oxide nanoparticles and human hemoglobin. *Journal of the Iranian Chemical Society* **2010,** *7* (2), S145-S153.