1
Introduction

Intermolecular recognition, including physiological and pathological mechanisms, is a vital phenomenon in a variety of biological systems (like ligand-protein, protein-protein, and protein-DNA interactions). It is, therefore, significant to study the principles of these interactions, particularly the binding efficacy and the underlying physical forces. This interaction analysis offers better qualitative and quantitative information in terms of kinetic parameters such as binding-affinity, the stoichiometry of interaction, and the relevant thermodynamic parameters, namely enthalpy and entropy of binding, in both *in vitro* and *in vivo* fashion. Further, the analysis of the ligand-protein interaction is crucial for designing the novel bioactive molecules and also to understand the biological functions. The present thesis aims to describe the implementation of ligand-based solution-state ^{1}H Nuclear Magnetic Resonance (NMR) spectroscopy as the major technique to understand the ligand-protein interaction besides employing computational method and biophysical experiments as supporting analytical tools. The main focus of the thesis is to provide a quantitative description of ligand-protein, specifically the organophosphate pesticides (OP)-protein interaction in *in vitro* fashion.

1.1 LIGAND-PROTEIN INTERACTION

"The secret of life is molecular recognition; the ability of one molecule to "recognize" another through weak bonding interactions." Linus Pauling

Ligand-protein interactions are the molecular basis of many ubiquitous and essential biological functions. These interactions have a strong influence on their dynamic and kinetic behavior, which affects the distribution, absorption, and metabolism (ADME) of the ligand as well as the structure and physiological action of interacting proteins [Cala et al, 2014b; Copeland, 2003; Du et al, 2016; Homans, 2007; Hansen et al, 2002; Ludwig and Guenther, 2009; Mittag et al, 2003], as shown in Figure 1.1. In addition, the number of dynamic events such as chemical processes involving the ligand affecting the biological systems may also alter the ligand-protein interaction. Proteins play a vital function in cellular activity. Proteins have a vast range of functions in the body, including cell signaling, structural, mechanical, and biochemical processes, by interaction with other exogenous and endogenous molecules. In the biological system, the ligands are usually small molecules that exhibit a huge range of physicochemical functions. Ligands bind with macromolecules (protein) with high affinity and specificity in the reversible, non-covalent manner. The importance of analyzing ligand-protein binding interactions is that it provides valuable information about binding like kinetics, thermodynamics, conformations of targets, and ligand efficiency [Copeland, 2003]. To comprehend these intermolecular interactions, understanding of structural and functional properties of the target protein and ligand is essential besides deciphering the physicochemical mechanism governing such complexation.

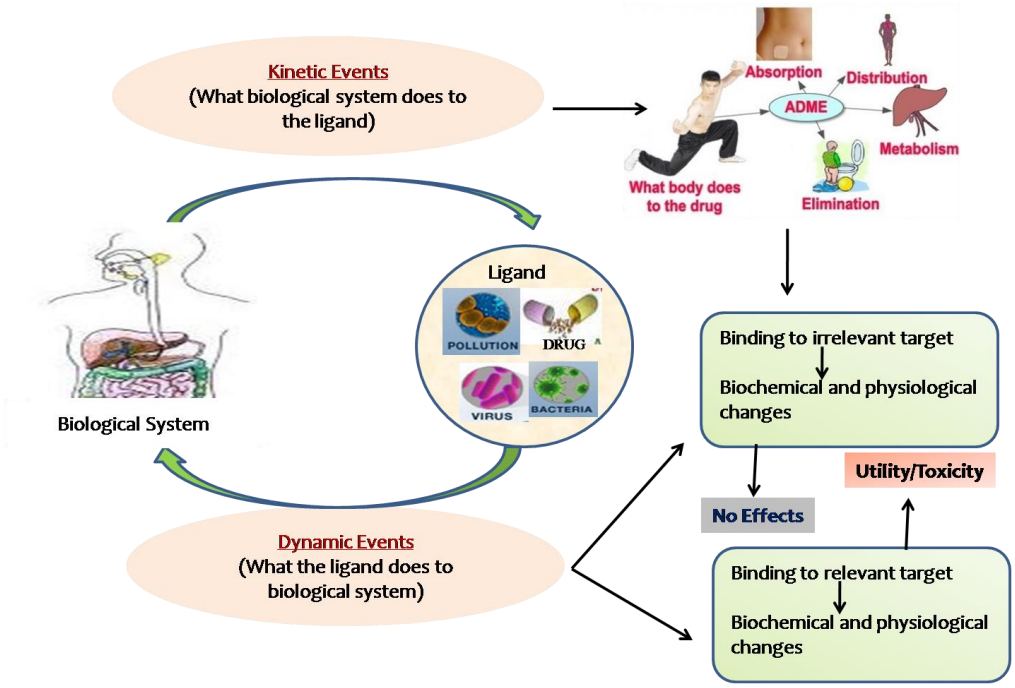

Figure 1.1 : Kinetic and dynamic events of ligand-protein interaction.

Figure 1.2 gives a schematic representation of ligand-protein complex formation and subsequent binding parameters involved during the interaction. A comprehensive review of the literature is presented for the experimental and theoretical methods used to assess these interactions, especially focusing on solution-state Nuclear Magnetic Resonance (NMR) [Angulo and Nieto, 2011; Cala et al, 2014b; Fielding, 2003, 2007; Ludwig and Guenther, 2009; Martini et al, 2006; Meyer and Peters, 2003c, 2003a] along with Fluorescence spectroscopy [Albani, 2004; Dewey, 1991a; Du et al, 2016; Eftink, 1991; Lakowicz and Weber, 1973], UV-vis spectroscopy Wang et al , 2009], Circular Dichroism (CD) [Amaraneni et al, 2014; Dahiya et al, 2017; Greenfield, 2006; Wang et al, 2017], Isothermal Titration Calorimetry (ITC) [Du et al, 2016; Gasymov and Glasgow, 2007; Homans, 2007; Olsson et al, 2008]), Surface Plasmon Resonance (SPR) [Gordon & Perugini, 2016,Manager & Specifications, 2000,Patching, 2014,Pattnaik, 2005] and Molecular Docking (MD) [Liu et al, 2018; Wang et al, 2016; Wu et al, 2015; Wu et al, 2003]. These experimental methods usually provide information regarding binding kinetics and thermodynamics. The computational techniques involve the ligand-protein docking and binding free energy calculations.

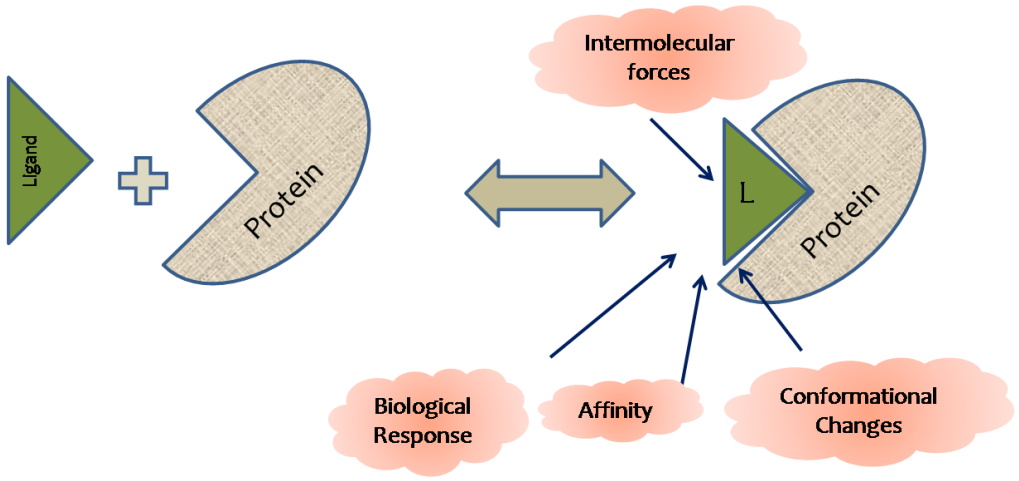

Figure 1.2: Schematic representation of ligand-protein interaction and binding parameters involved.

1.2 PHYSICOCHEMICAL MECHANISM OF LIGAND-PROTEIN INTERACTION

This section describes the binding kinetics relevant to the ligand-protein association governed by these intermolecular interactions.

1.2.1 The Binding Kinetics of Ligand-Protein Interaction

A binding event between a ligand and a protein has been modeled as a dynamic equilibrium among free ligand (L), free protein (M) and the ligand-protein complex ML [Ross & Subramanian, 1981]:

$$M + L \rightleftharpoons ML \tag{1.1}$$

where ML is the ligand-protein complex. k_1: the rate constants for the forward reaction and k_2: the rate constants for the reverse reaction. Eq.(1.2) represents the relation between the rate of forward reaction and the rate of reverse reaction at equilibrium:

$$k_1[P][L] = k_2[PL] \tag{1.2}$$

where at equilibrium, the square brackets specify the molar concentration in solution. Further, the association constant K_A is given by Eq.(1.3) [Copelansd, 2003; Fielding, 2003; Manager and Specifications, 2000; Mittag et al., 2003]:

$$K_A = \frac{k_1}{k_2} = \frac{[PL]}{[P][L]} = \frac{1}{K_D} \tag{1.3}$$

where K_D is the dissociation constant (in units of M) and determines the binding strength of the ligand-protein complex.

1.2.2 The Thermodynamic Parameters Relevant to the Ligand-Protein Interaction

The ligand-protein system is composed of solute (ligand and protein) and the solvent (water and buffer ions) molecules. During intermolecular interactions among these molecules in solution, heat exchange takes place. The heat transfer and subsequent complex formation (if any) is dictated by the laws of thermodynamics [Bronowska, 2011; Homans, 2007; Olsson et al, 2008; Ross and Subramanian, 1981]. At constant temperature and pressure, the thermodynamic potential, *i.e.*, Gibbs free energy is the most critical parameter that measures the capacity of the thermodynamic system to undergo a certain change. In the case of ligand-protein interaction, Gibbs free energy determines the interaction forces involved in the formation of the complex. In a spontaneous process, the ligand-protein interaction occurs with a change in Gibbs free energy (ΔG) of the system being negative. Further, the magnitude of ΔG determines the stability of the complex that defines the binding affinity between ligand and protein as depicted in Eq.(1.4)

$$\Delta G = -RTlnK_A \tag{1.4}$$

where K_A is the association constant; R: universal gas constant and T: the absolute temperature at which binding occurs. ΔG can also be defined in terms of enthalpy (ΔH) and entropy (ΔS) *i.e.*:

$$\Delta G = \Delta H - T\Delta S \quad (1.5)$$

The thermodynamic parameters for ligand-protein binding help in understanding the change in free energy, entropy, and enthalpy on binding and provide the information regarding the main driving forces involved in the binding process.

1.3 METHODS USED TO STUDY LIGAND-PROTEIN INTERACTION

Several experimental as well as theoretical methods based on spectrophotometry, Calorimetry, Nuclear Magnetic Resonance (NMR) and molecular docking have been established in literature to analyze ligand-protein interactions [Amaraneni et al, 2014; Dewey, 1991b; Ding et al, 2011; Du et al, 2016; Fielding, 2007]. The experimental methods like NMR, X-Ray diffraction, ITC, and fluorescence can be used to study the ligand-protein interaction, which provides complementary information regarding the change in structure and dynamics between the free and bound molecules and also about the binding events relevant to these interactions. Table 1.1 summarized the application of various techniques used to detect ligand-protein interactions. The detailed description is given only for the methods used in the present thesis. The rest of the methods are beyond the scope of the present thesis.

Table 1.1: Ligand-protein interaction analysis using various methods.

METHOD	***Applications***	***Advantages***	***References***
Circular Dichroism	Information regarding the secondary structure of the proteins, folding and unfolding of proteins.	Requires a low concentration of the sample, and any macromolecule can be investigated.	[Dodero et al, 2011; Greenfield, 2006; Martin and Schilstra, 2008]
Differential Scanning Calorimetry	Identification of structure, stabilization of the ligand-proteins.	Very high temperature can be used and reaction temperature can be determined accurately.	[Du et al, 2016]
Fluorescence Spectroscopy	Information regarding the binding stoichiometry, binding kinetics, and thermodynamic parameters.	High sensitivity and specificity.	[Albani, 2004; Du et al, 2016; Gordon and Perugini, 2016; Lakowicz, 2006]
FTIR Spectroscopy	Determination of the structural changes of macromolecules after ligand binding.	Provide information on multiple parameters simultaneously.	[Glassford et al, 2013]

Isothermal Titration Calorimetry (ITC)	Determination of thermodynamic parameters of the ligand-protein interaction, stoichiometry of complex and dissociation constants.	A single ITC experiment can provide complete thermodynamic and kinetic profiling.	[Di Trani et al., 2018, Du et al., 2016, Homans, 2007, Su & Xu, 2018]
Mass Spectrometry	Identification of the binding sites and conformational change in protein structure.	High throughput capability for compound library screening.	[Maple et al, 2012]
Molecular Docking	Binding parameters determination and binding site identification.	Provide every possible conformation based on the ligand and macromolecule complex formation.	[Du et al, 2016; Huang and Zou, 2010]
NMR Spectroscopy with detection of ligand resonances	Information on ligand-target binding kinetics.	Don't require protein-labeling, cost-effective.	[Cala et al, 2014a; Cala et al, 2014b; Fielding, 2003, 2007; Ludwig and Guenther, 2009; Meyer and Peters, 2003b]
NMR Spectroscopy with detection of target resonances	Information on the structure of the protein and binding site determination.	Information about the 3D structure of the protein.	[Cala et al, 2014b; Fielding, 2007; Unione et al, 2014]
Surface Plasmon Resonance (SPR)	Binding and dissociation kinetics determination and is independent of protein concentration.	SPR can be used to monitor real-time analysis of binding events.	[Gordon & Perugini, 2016,Patching, 2014,Pattnaik, 2005]
UV-vis Spectroscopy	Detection of the functional groups and chemical kinetics.	Solid, liquid, semi-solid, and powder samples can be analyzed.	[Nienhaus & Nienhaus, 2005]
X-Ray Crystallography	3D structure determination.	3D structure can be obtained.	[Gordon & Perugini, 2016,Homans, 2007]

Figure 1.3 represents the literature survey of published scientific data conducted for a ten years period ranging from 2009-2019, using the web of science database. It is clearly seen that Fluorescence, ITC, and NMR are the prominent analytical methods used besides molecular docking. Figure 1.4 shows the implementation of NMR spectroscopy for ligand protein interaction is increased from 2000-2019.

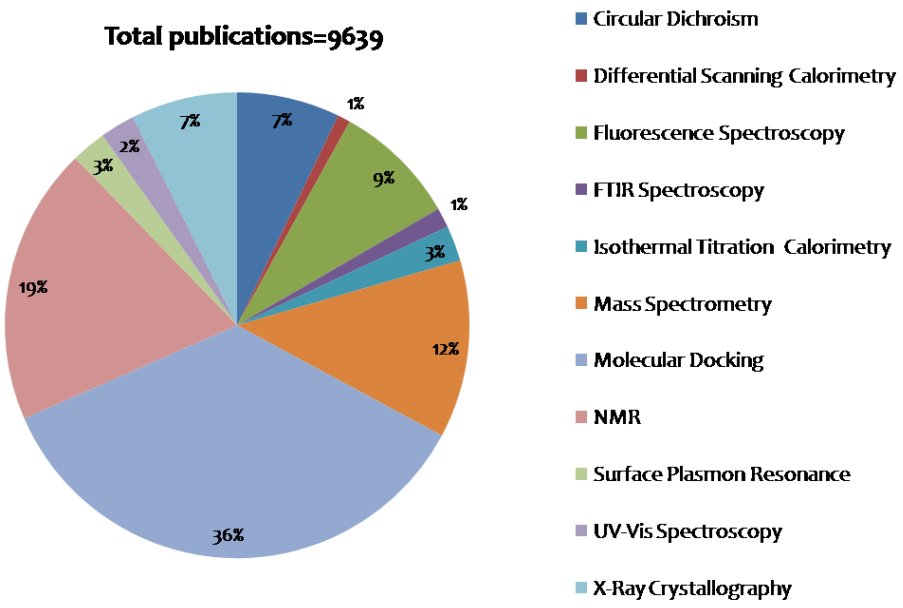

Figure 1.3 : Segmentation of Different biophysical methods available for ligand-protein interaction (Source: web of Science database, generic keyword used for different biophysical techniques is: technique name for ligand-protein interaction.

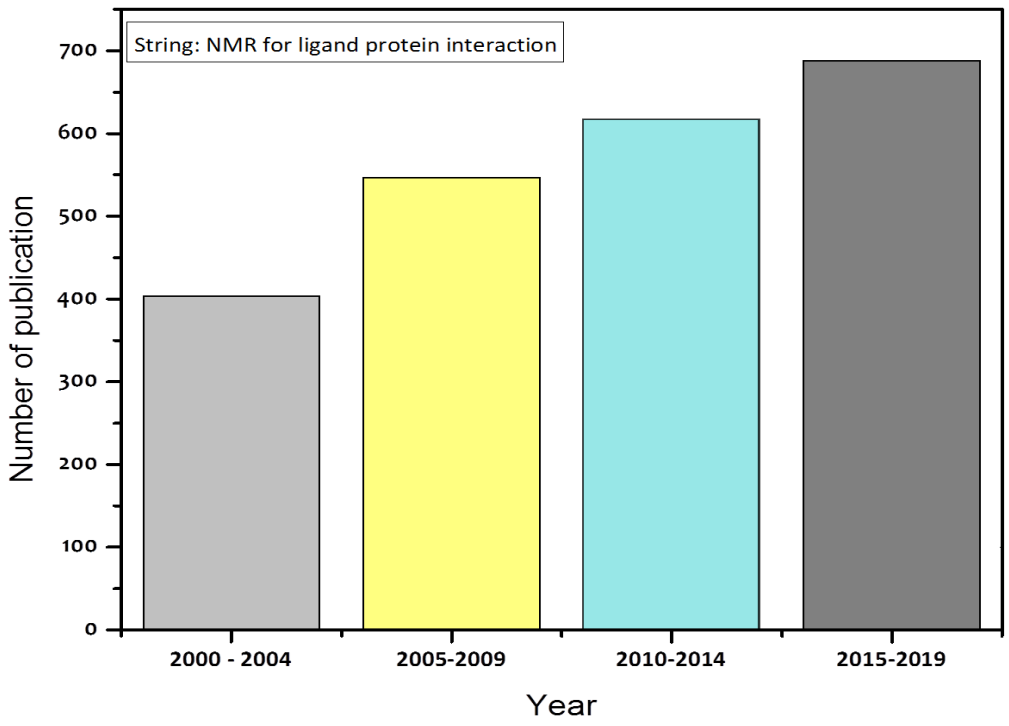

Figure 1.4 : The number of NMR data published for ligand-protein interaction (Source: Pubmed).

1.3.1 Ligand-protein interaction by NMR

NMR is one of the most popular and widely used methods to monitor and characterize ligand-protein interaction, which provides structural and dynamic information. Over the last two decades, a range of novel NMR methods is introduced and have prevalent applications in both pharmaceutical and research industry. The ligand-macromolecule interaction causes modifications in the properties of both the ligand and protein, such as changes in conformation, structural constituents, chemical environment, and diffusion properties to name a few. The interaction between ligand-protein can be probed and quantified by determining the changes in the various NMR parameters of either the ligand or the protein or both. Theoretically, all NMR parameters are essential and offer information about the ligand-protein binding; however, quantitative evaluation of these interactions is possible with NMR methods based on magnetization transfer and relaxation. Methods based on spin-lattice relaxation and Nuclear Overhauser Enhancement (NOE) are not only highly sensitive but also allows interpretation of the ligand-protein interaction in terms of binding strength and complex stability. NMR offers both protein-based and ligand-based methods by comparing relevant NMR parameters of the free and the complex state of the molecules, as shown in Figure 1.5 [Cala et al, 2014a].

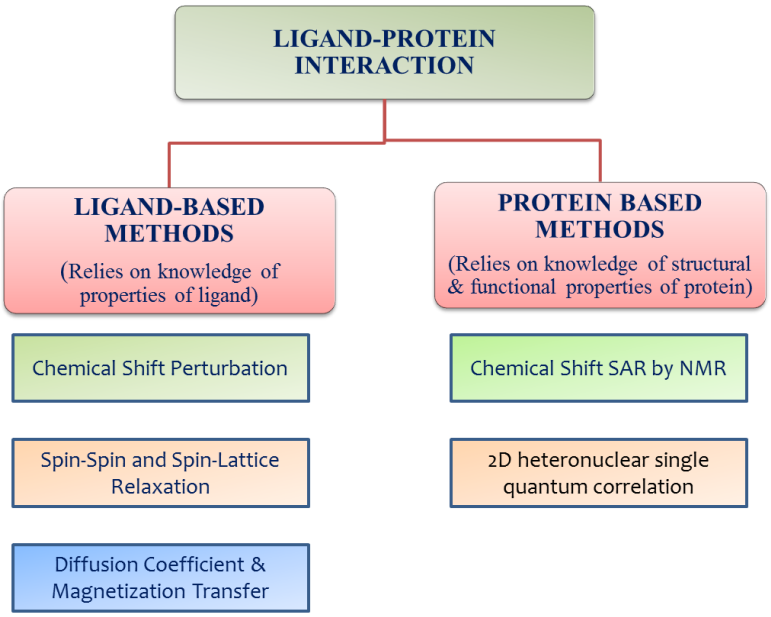

Figure 1.5: Different NMR approaches to study the ligand-protein interaction.

In the last decade, a vast amount of work has been published on NMR investigation of the ligand-protein interaction. In the current thesis work, only ligand-based solution-state ^{1}H NMR is employed to study the ligand-protein interaction. The binding process is considered as a dynamic equilibrium condition between the association and dissociation processes resulting in measurement of NMR parameters that are the weighted average of the free and bound state of the ligand, or the protein Figure 1.6 provides a cartoonic representation of a specific ligand binding to a protein from a mixture of ligands. The binding phenomenon is modeled as a chemical exchange process between the free and complex state of the ligand and protein. Such chemical exchange alters the NMR spectral appearance for the ligand/protein. In the case of ligand observed NMR spectra, two separate peaks are exhibited for the free and bound state of the ligand in a slow exchange regime while a single broad peak appears in the case of an intermediate exchange regime. In case of fast exchange between the free and complex state of the ligand, a single sharp peak is observed in the NMR spectrum of the ligand.

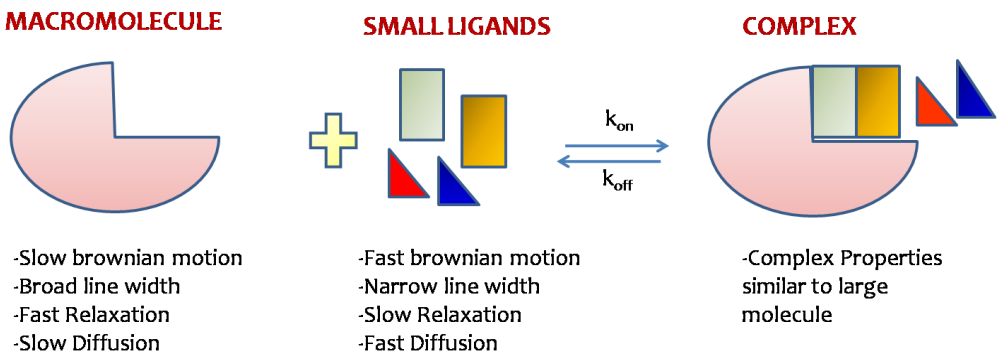

Figure 1.6 : NMR parameters of ligand-protein interaction in the free and bound state. During ligand-protein binding, the ligand presumes the properties of the macromolecule.

The association constant of binding reactions of ligand-protein complex range from 10^2 to 10^{12} M^{-1}. The NMR properties for the free state and bound state are different, as shown in Figure 1.6 [Fielding, 2003; Han et al, 2012; U Hansen et al, 2002; Ludwig and Guenther, 2009]. Hence, changes observed and quantified for different NMR parameters of either ligand or the protein are sufficient to comment on the binding process. For ligands in a non-viscous solution, the rotational correlation time (τ_c) is short due to fast Brownian motion, and positive NOE is observed. Whereas, for macromolecules, rotational correlation time (τ_c) is large due to slow molecular motion, and reduction in intensity occurs, giving rise to negative NOE. Hence complex formation by a small ligand will definitely manifest its effect by altering the motional properties of the ligand and will lead to the measurement of population weighted-average NMR parameters of the ligand in solution. There are two different approaches to measure these changes in NMR parameters.

(a) The Protein-detected Methods

The easiest and typical parameter for protein-based detection is chemical shift mapping. After the addition of ligand to the protein system, changes in the chemical shift of the protein backbone and side-chain resonances observed to localize the ligand-binding site and further distinguish between specific from non-specific binding. The 3D structure of the ligand-protein complex can be resolved using heteronuclear experiments on isotopically labeled protein samples ^{13}C, ^{15}N, ^{2}H. The main limitations of these methods are (a) long experimental time and (b) requirement of highly stable and soluble protein [Cala et al, 2014b; Fielding, 2003] as well as (c) necessity of isotope labeling of the protein under study and (d) measurement at preferably high magnetic field.

(b) The Ligand-detected Methods

Ligand-based methods are inexpensive and can be performed at a lower magnetic field with a minimal concentration of proteins, while protein-based methods are limited due to the requirement of a higher magnetic field, longer data collection, a large quantity of isotopically labeled proteins, *etc*. Several Ligand-based experiments are proposed in the literature that includes chemicals shift parameter, relaxation analysis, saturation transfer difference, diffusion experiments, Nuclear Overhauser Effect (NOE), water LOGSY, SALMON, INPHARMA [Cala et al, 2014a; Fielding, 2003, 2007; Fisher and Bain, 2014; Krishnan, 2005; Ludwig and Guenther, 2009; Unione et al, 2014]. The comprehensive literature report for these techniques is provided in Table 1.2. The present thesis employs two of the most popular methods, namely Saturation transfer difference NMR and spin-lattice relaxation analysis besides the preliminary confirmation of ligand binding by analyzing chemical shift and line width modifications. Detailed information is only provided for the aforementioned methods in the subsequent Chapters.

Table 1.2: The literature review of ligand-based NMR methods for ligand-protein interaction.

Methods	Applications	References
Chemical Shift and Line width change	Chemical structure and binding study determination	[Dorai and Kumar, 2001; Fielding, 2007; Liu et al, 1997]
INPHARMA	Active site identification and bound conformation of the ligand	[Carlomagno, 2012, Sanchez-pedregal et al., 2005, Cala et al, 2014a; Ishima and Torchia, 2000]]
Proton Relaxation Time T_1 and T_2	Correlation time, binding affinity and thermodynamic parameters study	[Bloembergen et al, Fielding, 2003; Rossi et al., 2001; Unione et al., 2014]
Saturation Transfer Difference (STD) NMR	Screening in mixtures, dissociation constant determination, epitope mapping, and multiple binding modes	[Angulo et al, 2010; Cala and Krimm, 2015; Krishnan, 2005; Mayer and Meyer, 2001]
Transferred NOESY	Screening in mixtures for FBDD and bound conformation of ligands	[Anglister et al, 2016; Ludwig, 2009; Plesniak et al, 2008]
WATERLOGSY	Screening in mixtures, K_D determination, epitope mapping	[Dalvit et al, 2001; Huang et al, 2017; Ludwig, 2009]

1.4 AGROCHEMICAL-PROTEIN INTERACTION

Agrochemicals are the substances or the mixture of substances used to alleviate or prevent pests like insects, bacteria, rodents, and other harmful organisms that harm the crops and diminish crop production. Agrochemicals are broadly classified into two categories, namely biopesticides and chemical pesticides. Biopesticides are developed from natural sources like animals or plants, *e.g.,* microbial pesticides. On the other hand, chemical pesticides are synthetic molecules and can be categorized by target organisms, e.g., insecticides, fungicides, herbicides, nematicides, plant growth regulators, and rodenticides. Furthermore, insecticides are grouped into Organochlorines (*e.g.*, endosulfan, hexachlorobenzene), Organophosphates (e.g., diazinon, chlorpyrifos, parathion), Carbamic and thiocarbamide derivatives (e.g., aldicarb, carbofuran), Urea derivatives (e.g., fenuro, monuron), Fluorine-containing compounds (e.g., cryolite, acetoprole, dichlofluanid) and synthetic pyrethroids (e.g., allethrin) [Bernardes et al, 2015; Gallo and Lawryk, 1991] based on the chemical structure. The widespread use of these chemicals can lead to environmental contamination that can affect the non-target organisms, including humans and animals, through water, air, polluted soil, and food [Mostafalou & Abdollahi, 2013]. Therefore, it is essential to understand the mechanism of action of these pesticides on biomacromolecule to analyze their toxic effects. In literature, various biophysical methods are available to detect the impact of pesticides on animals and humans like the computational approach [Ramalho et al, 2016; Moreira et al, 2016; Sharma et al, 2011], equilibrium dialysis [Lv et al., 2014,Mourik & de Jong, 1978], biosensor methods [Vakurov et al, 2004; Wang et al, 2011], MALDITOFMS [Amaraneni et al, 2014], Mass Spectrometry [Thompson et al, 2010], Multi-spectroscopic methods *viz.* Fluorescence, UV-vis, and Circular Dichroism (CD) [Han et al, 2012; Jafari et al, 2017; Hansen, 2013; Sogorb et al, 2008; Mourik and Jong, 1978; Østergaard and Larsen, 2007; Silva et al, 2010b; Silva, et al, 2010; Sogorb et al, 2007; Suganthi and Elango, 2017], Thin liquid chromatography [Sharma and Kocher, 2013], ^{1}H NMR Metabolomics [Yoon et al, 2016], ^{1}H NMR spectroscopy [Martini et al, 2010]. These methods help in understanding the potential effect of various pesticides on animals and humans. A literature survey using Web of Science conducted for the past decade from 2009-2019, as represented in Figure 1.7, reveals that a considerable amount of research effort has been made to understand agrochemicals-protein interaction. Insecticide-protein interaction has received major attention in specific indicating constant growth in commercial availability and field applications of these molecules. The current thesis revolves

around analyzing OP-protein interaction in different solvent conditions targeting quantification of binding efficacy and structural modification of both OP and protein on interaction.

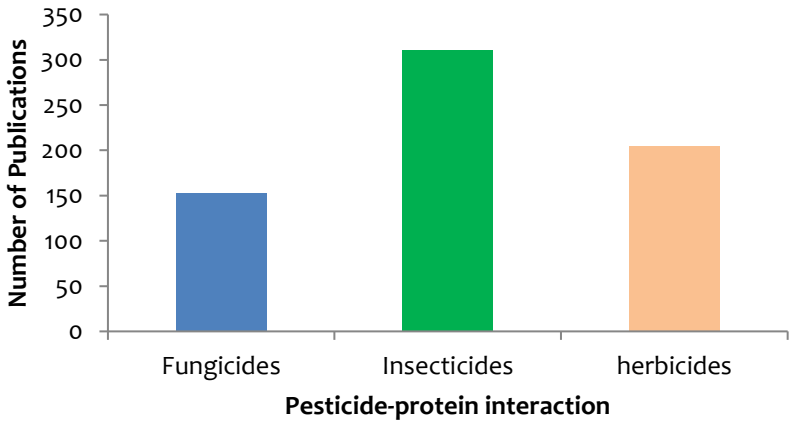

Figure 1.7 : Number of published data on pesticide-protein interaction from 2009 onwards. (Source: web of science).

1.4.1 ORGANOPHOSPHATE-PROTEIN INTERACTION STUDY

Amongst the various pesticide families organophosphate (OP) pesticides are the most abundantly used chemicals in agriculture, household and industry worldwide due to their better biodegradability and cost-effective production [Bravo et al, 2004; Gwinn et al, 2005; Kavvalakis and Tsatsakis, 2012; Magnarelli and Fonovich, 2013]. The general structure for OP is given in Figure 1.8. The structure typically contains a pentavalent phosphorus atom forming a double bond with oxygen or sulfur atom.

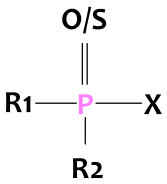

Figure 1.8 : General structure of OP

Due to the specified molecular structural properties *viz.* presence of phosphorous and carbonyl moieties [Reffstrup et al, 2010]. OP can be easily absorbed through lungs, skin, and gastrointestinal tract, which in turn creates severe health complications in younger animals and humans [Eskenazi et al, 2007; Kavikarunya and Reetha, 2012; Pailan and Saha, 2015]. OP impart their toxicity by attaching to red blood cells (RBC) that allow them to interact with acetylcholinesterase (AChE). This interaction causes inhibition of AChE activities at synaptic junctions which further results in accumulation of AChE at nerve endings, leading to hyperactivation of receptors [Čolović et al, 2013]. Consequently, the impact of organophosphate pesticides on living systems has become a significant issue these days [Damalas & Eleftherohorinos, 2011]. The hydrolysis of OP yields a dialkyl phosphate and a leaving group

[Sogorb & Vilanova, 2002]. OP act primarily by disrupting the protein phosphorylation pathways that are associated with metabolic regulation, hormone signaling, neuronal functions, cell survival, and death [Hargreaves, 2012,Magnarelli & Fonovich, 2013]. Their tendency to inhibit many metabolic and physiological enzymes like acetylcholinesterase (AchE), cytochrome P450, protein kinase C, and glutathione S-transferases (GSTs), makes them toxic and leads to neurotoxicity in humans [Fukuto, 1990; Sharma et al, 2011]. Although it might be expected that this reaction results in a decreased toxicity, as the leaving group and dialkyl phosphate do not inhibit cholinesterase enzymes, there are several reports that confirm the potential toxic effects of these metabolites due to their higher water solubility and mobility [Aktar et al, 2009; Bernardes et al, 2015; Sinclair and Boxall, 2003]. Further, besides the OP, their degradation products also play crucial roles in polluting the non-target environment [Aktar et al, 2009]. Therefore, it will be equally significant to study the long term impact of these metabolites in tissues of humans and animals, as shown in Figure 1.9 [Benedetti et al, 2014].

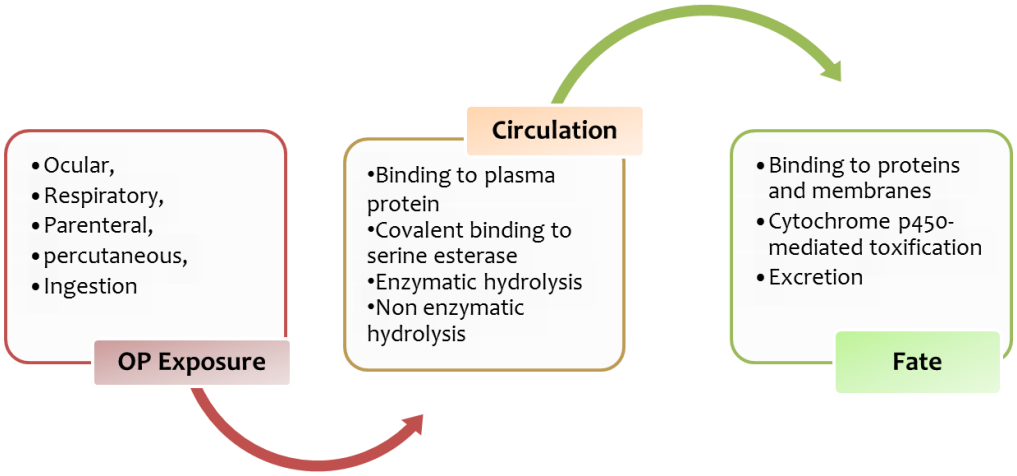

Figure 1.9 : Importance of OP pesticides-protein interaction study. This figure is modified from [Worek et al, 2016].

A quick review of the literature reveals that there are number of instances where researchers have discussed both *in vivo* rat model-based experiments as well as *in vitro* spectroscopic methods to analyze OP pesticides toxicity in terms of the interaction of OP pesticides and their metabolites with proteins [Davies and Holub, 1980; Haque et al, 1973; Osman, 2011; Kim and Ahn, 2009; Li et al, 2010; Costa, and Furlong, 2013; Mourik and Jong, 1978; Narang et al,, 2015; Sharma et al, 2011; Silva et al, 2010a]. Table 1.3 illustrated the application of NMR methods reported in the literature to study OP-protein interaction. In most of the cases, researchers have employed solution-state ^{1}H and ^{31}P NMR experiments monitoring chemical shift changes to confirm the OP-protein interaction.

Table 1.3 : NMR methods used for OP-proteins interaction and their findings.

System	Findings	References
Chymotrypsin, α-Lytic proteinasechymotrypsin, chymotrypsinogen, trypsin, trypsinogen, atropinesterase, subtilin, AChE, BChE, chymotrypsin Diisopropyl fluorophosphate -Using ^{31}P NMR Spectroscopy	OP conjugate stereochemistry and OP conjugate aging, chemical shift correlation, binding site determination -These findings help in understanding the enzyme inhibition mechanism	[Gokalp et al, 2005; Fataneh Jafari et al, 2018, 2018; Saadati and Mirzaei, 2016]
Fenitrothion, Several OP pesticides, Isoparathion methyl, Profenofos, Phosphorothiolates, Phosphonofluoridates, Acephate, methamidophos, $POCl_3$ -^{31}P NMR spectroscopy	Metabolite identification, Purity check of chemicals and Identification of sulfur-containing oxidation products -These findings help in understanding the toxicity caused by OP	[Gokalp et al, 2005; Saadati and Mirzaei, 2016]
Malathion, Chlorpyrifos, Vamidothion, Phosphorodithioates, FP-biotin, Chlorpyrifos, Glyphosate -^{31}P and ^{1}H NMR spectroscopy	Identification of metabolites -These findings help in bio-monitoring of OP	[Koskela, 2010]
Bovine serum albumin -^{1}H NMR spectroscopy	Identified the TCPy and PM interaction with BSA and found that halogen-containing OP has a much stronger affinity to protein than others. Further, BSA increases the hydrolysis for OP-oxons. -These findings help in designing OP biomarker	[Dahiya et al, 2019]
Humic acid -using ^{31}P NMR	Identified the interaction of humic acid with OP	[Šmejkalová and Piccolo, 2008; Šmejkalová et al, 2009]

1.5 PURPOSE OF THE STUDY

The main aim of the current thesis is to implement solution-state ^{1}H NMR methods in combination with molecular docking, ITC, and fluorescence quenching analysis as complementary biophysical techniques to analyze and quantify the organophosphate-protein and metabolite-protein interaction by *in vitro* methods. The major goals of the thesis can be broadly classified as identification of OP-protein/OP metabolite-protein system, qualitative confirmation of OP-protein/OP metabolite-protein binding interaction in solution, quantification of OP-protein/OP metabolite-protein binding efficacy and evaluation of thermodynamic parameters using biophysical techniques with a major focus on the implementation of one-dimensional ligand-based solution-state NMR methods. Two major NMR techniques, namely Saturation transfer difference NMR and selective spin-lattice relaxation measurements are identified as the ligand-based NMR experiments for quantification of association constants and group epitope mapping of the OP-protein complexes. Two major complementary biophysical techniques *viz.* Fluorescence quenching analysis and ITC are selected not only to support the NMR experimental outcome but also to shed light on the thermodynamics of the binding process. Figure 1.10 highlights the different biophysical techniques employed in the current thesis to reveal various binding parameters. The work has been carried out with the following molecular systems: OP and OP metabolite, namely CPF (chlorpyrifos), DZN (diazinon) and PA (parathion) as parent OP and TCPy, IMP (2-isopropyl-6- methyl-pyrimidin-4-ol), and PM as the corresponding metabolites; relevant biologically active proteins, *i.e.,* major carrier protein serum albumin and gut enzyme trypsin. The forthcoming sections provide the brief introduction to these molecular species.

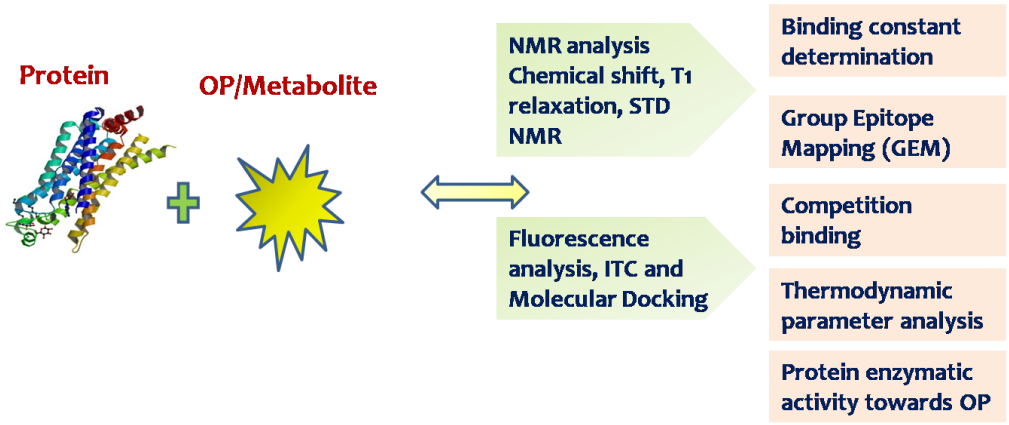

Figure 1.10 : Schematic representation of the general overview OP/OP metabolites-protein interaction under *in vitro* conditions.

1.5.1 Organophosphate Pesticides Used in the Present Thesis

This section provides detailed information regarding the test OP and their metabolites used to study the molecular interaction with proteins of choice. Chlorpyrifos (CPF), Diazinon (DZN), Parathion (PA), and their metabolites were chosen as the ligands used for interaction with the proteins.

(a) Chlorpyrifos (CPF)

Chlorpyrifos (O, O-diethyl O-3,5,6-trichloropyridin-2-yl phosphorothioate, chlorpyrifos-ethyl) is a chlorinated organophosphate pesticide used on crops to prevent the pests and insects [Anwar et al, 2009]. It has been registered over a hundred countries for agriculture purposes. Commercially it is registered under brand name mainly Dursban, Lorsban, Agromil, Dhanwan, Dorson, and Omexan. CPF inhibits the enzyme acetylcholinesterase (AChE), in the brain and peripheral nervous system leading to decreased degradation of a neurotransmitter, acetylcholine that results in overstimulation

of the associated synaptic systems. CPF was first introduced in 1965 and is now the most widely used insecticide worldwide to control pests in agriculture and households. When ingested accidentally or inhaled with air or absorbed through the skin, cytochrome P450 (CYP-450) enzyme metabolizes the CPF [He and Li, 2007]. These enzymes degrade CPF into 3,5,6-trichloro-2-pyridinol (TCPy) and diethyl thiophosphate through oxidative ester cleavage (dearylation). Alternatively, CPF can degrade in CPF-oxon through oxidative desulfation[Smith et al, 2010]. CPF has low water solubility (2 mg/L) but exhibits a high soil-absorption tendency with a varying range of half-life of 10 to 120 days in the soil[Xu et al, 2008]. It exhibits a dose-dependent effect on the plasma cholinesterase depression [Han et al, 2012b] leading to acute toxicity on aquatic organisms especially fish while excessive exposure makes humans prone to acute phosphorus poisoning due to phosphorylation of AcHE [Gollapudi et al, 1995]. The degradation of CPF is shown in Figure 1.11. 3,5,6-Trichloro-2-pyridinol (TCPy) is considered as one of the major degradation product of CPF [Yang et al, 2005]. CPF-methyl [Kim & Ahn, 2009] as well as another systemic herbicide, triclopyr [Li et al, 2010] are also identified as a degradation products. The enzymes those are present in the human liver cause oxidative desulfuration of CPF, which ultimately gets hydrolyzed into TCPy and diethyl phosphate or diethyl thiophosphate as depicted in Figure 1.11 representing a generalized degradation pathway of CPF. At neutral pH, TCPy is a charged species and has more water solubility compared to CPF [Anwar et al, 2009]. TCPy is excreted in the urine, with an average half-life of about 27 hrs [Morgan et al, 2005]. Toxicity of TCPy is still debatable as it does not inhibit cholinesterase enzymes, unlike its parent [NPIC, 2010]. Conversely, on various other occasions, it is considered as more toxic due to its higher water solubility causing contamination of both soil and aquatic environments [John & Shaike, 2015a].

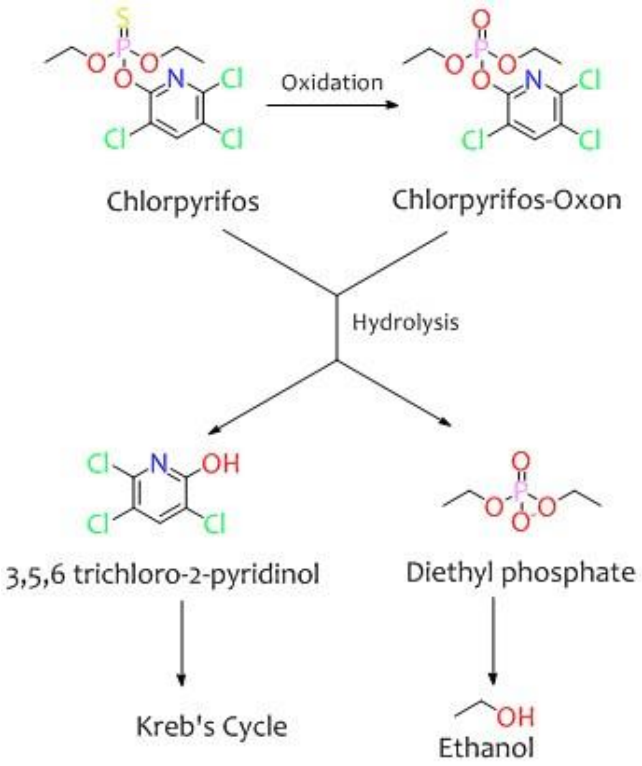

Figure 1.11: The degradation pathway of Chlorpyrifos.

(b) Diazinon (DZN)

Diazinon (O, O-diethyl-O-(2-isopropyl-4-methyl-6-pyrimidinyl phosphorothionate) is a class I organophosphate pesticide used worldwide for agriculture purpose. Due to its highly toxic effect on water and food resources, the use of DZN is restricted in some countries. The degradation of DZN occurs through direct oxidation degradation hydrolysis, which results in two metabolites 2-isopropyl-6- methyl-pyrimidin-4-ol and diazoxon [Kouloumbos et al, 2003] . The degradation pathway for DZN is shown in Figure 1.12. *In vivo* metabolism of DZN occurs through cytochrome P450. Due to oxidative desulphuration (P=S) of DZN, the diazoxon is produced (P=S). The impact of exposure to diazinon results in acute health problems in animals and humans. A quick literature review unveils that DZN is very toxic to proteins and can lead to a change in protein structure [Jafari et al, 2018; Saadati and Mirzaei, 2016]. Hence, it is important to study the interaction of DZN and its metabolites with biologically relevant proteins [Davies & Holub, 1980].

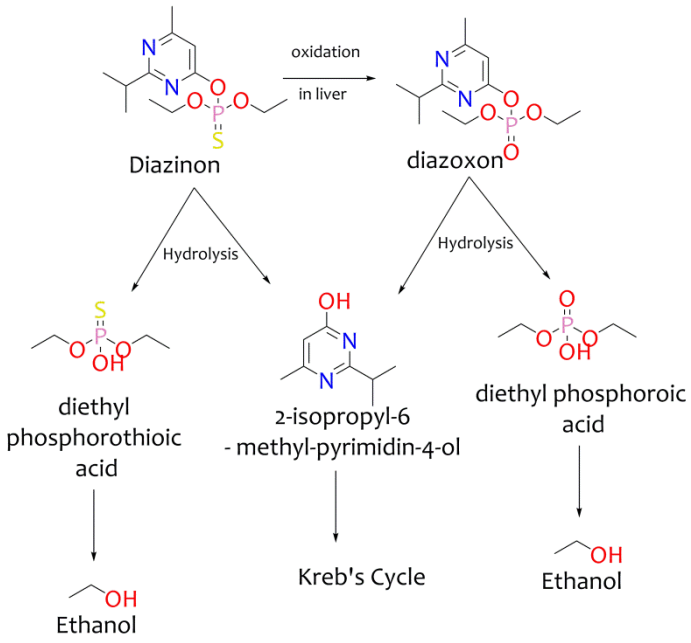

Figure 1.12 : The degradation pathway of Diazinon.

(c) Parathion (PA)

Parathion (*O*, *O*-Diethyl *O*-(4-nitrophenyl) phosphorothioate) is a highly active thiophosphorus ester organophosphate pesticide (OP). It is an aromatic nitro compound that is used worldwide in agriculture and fish industry. It causes inhibition of acetylcholinesterase (AChE), leading to the significant manifestation of organophosphate poisoning. United States Environmental Protection Agency (EPA) banned the use of PA due to its highly toxic nature. But still, in developing countries like India, it is used for pest control in crops [IPCS Inchem, 1992; Wu and Linden, 2008]. The biotransformation of PA results in paraoxon methyl, which also acts as an acetylcholinesterase inhibitor and can harm animals and humans. The schematic degradation of Parathion to paraoxon methyl and p-nitrophenol is given in Figure 1.13.

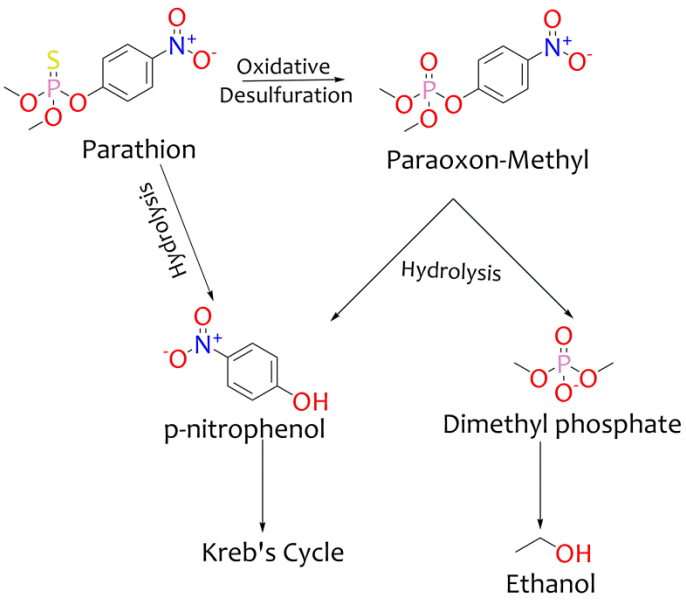

Figure 1.13: The degradation pathway of Parathion

1.5.2 Proteins used in the present thesis

Two proteins are used to decipher the interaction of OP with proteins, one is bovine serum albumin, and the other is trypsin. Bovine serum albumin is a widely studied model protein that can be used for the assessments of the effects of OP on the protein as well as benchmarking the various biophysical methods employed for such interaction analysis. The following brief details about the structure and function of the two proteins are provided.

(a) Serum Albumin

For *in vitro* ligand-protein interaction analysis, serum albumin is an ideal model protein. Among various possible pesticide-protein interactions, it is essential to study OP-serum albumin interaction that controls the free concentration of OP inside the body and therefore defines their effects on various proteins [Silva et al., 2010a, 2010b; Silva et al, 2004; Tamura et al, 1990]. Further, the strong binding affinity of pesticides with plasma protein indicates their lesser diffusion to body tissues leading to lesser toxicity; however, a stronger complexation with plasma protein may also hinder the excretion process of the metabolites. Serum albumin is a globular protein present in blood plasma with a molecular weight ca. 66 kDa with 583 amino acid residues. It executes different biological and physiological functions in the body, *such as* transport of various endogenous and exogenous compounds, maintaining blood osmotic pressure. Serum albumins have three domains (I-III), with each domain having subdomains (A and B). The most commonly used model albumin proteins are HSA and BSA to access the ligand-protein interactions, and both the serum proteins exhibit 75% sequence homology and 76% tertiary structures similarity. The ADMET (absorption, distribution, metabolism, excretion, and toxicity) of the ligands are strongly affected by their interaction with serum protein. BSA is commonly used for investigating biological interaction studies due to its commercial availability and homology with human serum albumin [Lee and Lee, 1995; Zhang, and Kokot, 2012]. BSA is a multifunctional protein that possesses catalytic properties against a variety of xenobiotic substrate, and known for its pseudo-enzymatic activity [Goncharov et al, 2015]. The hydrolysis study of OP in the presence of BSA can provide useful information regarding the catalytic degradation of these molecules [Hansen et al, 2002].

(b) **Trypsin**

Proteases control various pathological and physiological processes. The interaction of ligands-proteases is essential to reveal the binding mechanism features and the proteases conformation [Gonçalves et al, 2011; Wang et al , 2016]. Trypsin is a water-soluble protease. It is the most important digestive enzyme extracted from the pancreatic acinar cells, which has been used as a typical digestive protein [Wang and Zhang, 2014; Wu et al, 2017]. Trypsin's molecular weight is 23,300 Dalton and contains 223 amino acid residues with six disulfide bridges that hold the individual chains together. Trypsin protein has two domains, with nearly the same size with six antiparallel β-sheet. It contains ten tyrosine residues (Tyr), four tryptophan residues (Trp), and six phenylalanine residues (Phe). Being a digestive protein, it helps in digestion of food proteins, apoptosis, immune response, hemostasis, and signal transduction. Therefore reduced activity of trypsin is harmful as it results in reduced absorption of the nutrients [Liu et al, 2017; Liu et al, 2015; Wang et al., 2016]. Investigation of the interaction of OP with trypsin is important since such interactional analysis will shed light on the effect of accidental OP ingestion on the digestive system. The literature review presented in the thesis provides limited information on the binding of OP-trypsin and the consecutive effect of such binding trypsin activity. Hence it is worthwhile to invest time in understanding OP-trypsin interaction.

1.6 SCOPE OF THE THESIS

The present research work has primarily focused on unraveling the molecular interaction between organophosphate-protein under *in vitro* solution conditions. An attempt has been made to employ the ligand-based solution-state ^{1}H NMR methods as a major technique with molecular docking, ITC, and fluorescence quenching analysis as complementary techniques to study the kinetic and the thermodynamic properties of OP-protein association by monitoring various binding parameters. Analysis and quantification of molecular interaction of the three most widely used OP *viz.* chlorpyrifos, diazinon, parathion, and their metabolites with two different proteins, namely bovine serum albumin and trypsin are considered in details. The study reveals the subtle advantages of *in vitro* measurements that require a low volume of materials which generates a limited amount of toxic waste material, limited or no need for animals, low cost, high level of standardization, controlled testing conditions, absence of systematic effects [Jang et al, 2014]. The findings of this thesis indicate that the presence of OP in food or environment could be harmful to animals and humans since the interaction of these toxic substances with biologically relevant proteins may lead to structural and functional changes in these proteins. However thorough in-vivo analysis of OP-protein interaction will only enable commenting on the impact of these molecules on living systems. The present work also investigates the use of BSA as a biomarker for OP oxons under *in vitro* conditions.

Chapter 1 ***Introduction*** of the present thesis reviews the current literature to compare the various methods available to analyze the ligand-protein interaction and further the effect of various agrochemicals like insecticides, herbicides, fungicides on the environment that provides a comprehensive introduction of the lethality of these agrochemicals. Chapter 2 ***Materials and Methods*** further gives the information regarding materials used for current thesis work and offers the background of Nuclear Magnetic Resonance (NMR) methods and other complementary techniques used in this thesis.

Chapter 3, ***Competitive binding of OP with BSA*** presents the key importance of the STD NMR technique to analyze the OP-BSA interaction. Three most commonly used OP has been chosen, and the respective OP-BSA systems are investigated in detail revealing the binding efficacy and binding epitope of the OP-BSA complexes. Extensive application of ^{1}H STD NMR has allowed an understanding of the competitive binding of the OP molecules with BSA. It is clearly demonstrated in this Chapter that structural niceties regulate the OP-protein interaction. In all the three cases, the aromatic ring moiety has been found to be in close proximity with BSA binding sites. The competition STD NMR is used to compare the binding affinity of the OP-BSA

complex. The binding affinity exhibits the order CPF>PA>DZN in case of OP-BSA interaction that is attributed to the fact that CPF contains three Cl atoms allowing halogen bonding with the protein binding sites. STD NMR assisted with ITC, and molecular docking has provided the structural model for OP-BSA interaction as well as the binding strength of CPF, DZN, and PA towards BSA.

Chapter 4 ***Comparative Interaction of TCPy and PM with BSA and relative esterase activity*** aim to demonstrate the employability of ^{1}H spin-lattice relaxation analysis to quantify ligand-protein interaction. Two stable OP metabolites TCPy and PM are chosen to unravel the binding interaction of these metabolites with BSA. The study provides an understanding of metabolite-BSA interaction in comparison to that of the parent-BSA interaction analyzed in Chapter 3. Selective and non-selective ^{1}H spin-lattice relaxation analysis has been highlighted as the major analytical technique quantifying the association constants and molecular correlation times of the bound ligands. Further, Fluorescence quenching analysis and molecular docking have been used as a complementary technique. The current study reveals that TCPy (due to halogen atoms present in its structure) forms a more stable complex with BSA than PM. Also, due to the presence of P=O in the case of PM, it undergoes hydrolysis in the presence of BSA. A thorough analysis reveals BSA esterase activity towards PM indicating the possible application of BSA as a biomarker for OP oxons.

Chapter 5 ***Interaction of OP and their metabolites with trypsin: An NMR case Study*** covers the investigation of CPF, DZN and their metabolite interaction with trypsin, and gives the idea about the OP-enzyme interaction. The current findings confirm that both OP and their metabolites affect the structure of trypsin. The binding affinity of OP and their metabolite to trypsin is less comparable to OP-BSA binding. Also, the test ligands inhibit the trypsin activity.

Chapter 6 ***Solvent dependent interaction of OP with BSA: A comparative fluorescence quenching analysis*** confirms the role of solvent polarity as well as pesticide concentration in dictating OP-protein interaction. Both CPF and TCPy exhibit static quenching of BSA fluorescence emission in solution, indicating the formation of the ground-state complex. Based on the binding interaction analysis, it might be speculated that free concentration of CPF in the plasma will be higher compared to TCPy, resulting in greater diffusion of CPF to target tissues while TCPy will be retained in plasma as TCPy-BSA complex.

Chapter 7 ***Summary*** summarizes the research work of whole thesis and further discuss about its future scope.

…