Structural dynamics of acetylcholinesterase and its implications in reactivators design

Gianluca Santoni

UNIVERSITÉ DE GRENOBLE

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Physique pour les sciences du vivant**

Arrêté ministériel : 7 Aout 2006

Présentée par

Gianluca SANTONI

Thèse dirigée par **Martin WEIK**
et codirigée par **Florian NACHON**

préparée au sein **de l'Institut de Biologie Structurale de Grenoble**
et de **l'école doctorale de physique**

Structural dynamics of acetylcholinesterase and its implications in reactivator design

Thèse soutenue publiquement le **30/01/2015**,
devant le jury composé de :

Dr. Yves Bourne
Directeur de recherche CNRS, AFMB Marseille, Rapporteur
Dr. Etienne Derat
Maitre de conference, Université Pierre et Marie Curie, Paris, Rapporteur
Prof. Pierre-Yves Renard
Professeur, Université de Normandie, Rouen, Examinateur
Prof. Israel Silman
Professeur, Weizmann Institute of Science,Rehovot, Examinateur
Dr. Yvain Nicolet
Chargé de recherche CNRS, IBS, Grenoble, Examinateur
Dr. Jacques-Philippe Colletier
Chargé de recherche CNRS, IBS, Grenoble, Examinateur
Dr. Martin Weik
Directeur de recherche CEA, IBS, Grenoble, Directeur de thèse
Dr. Florian Nachon
IDEF, IRBA, Bretigny sur Orge, Co-Directeur de thèse

UNIVERSITÉ DE GRENOBLE

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Physique pour les sciences du vivant**

Arrêté ministériel : 7 Aout 2006

Présentée par

Gianluca SANTONI

Thèse dirigée par **Martin WEIK**
et codirigée par **Florian NACHON**

préparée au sein **institut de biologie structurale de Grenoble**
et de **l'école doctorale de physique**

Dynamique structurale de l'acetylcholinesterase et ses implications dans la conception de reactivateurs.

Thèse soutenue publiquement le **30/01/2015**,
devant le jury composé de :

Dr. Yves Bourne
Directeur de recherche CNRS, AFMB Marseille, Rapporteur
Dr. Etienne Derat
Maitre de conference, Université Pierre et Marie Curie, Paris, Rapporteur
Prof. Pierre-Yves Renard
Professeur, Université de Normandie, Rouen, Examinateur
Prof. Israel Silman
Professeur, Weizmann institute,Rehovot, Examinateur
Dr. Yvain Nicolet
Chargé de recherche CNRS, IBS, Grenoble, Examinateur
Dr. Jacques-Philippe Colletier
Chargé de recherche CNRS, IBS, Grenoble, Examinateur
Dr. Martin Weik
Directeur de recherche CEA, IBS, Grenoble, Directeur de thèse
Dr. Florian Nachon
IDEF, IRBA, Bretigny sur Orge, Co-Directeur de thèse

Contents

List of Figures

Preface

This thesis work is devoted to the study of the structural dynamics of acetylcholinesterase and its interaction with organophosphate nerve agents and reactivators. Intoxication by organophsophates is a major health problem, with about 2 million cases per year in the world, expecially in developing countries, and the currently available countermeasures are not efficient in protecting against these potent inhibitors.

In the design of a new family of reactivators, we applied the principles of rational drug design. The aim was to capitalize the knowledge of the enzyme and its structure to improve the currently available reactivators. In particular, we looked for neutral molecules, capable of passing the blood-brain barrier and restore AChE activity in the central nervous system, differently from what is currently on the market. This process led us to develop a first lead molecule, named KM297, and to improve it through rational optimization to its derivative, JDS207, that conserves the reactivation properties of the lead molecule, but improves its binding mode in the active site of the enzyme.

This process has revealed the need for improved knowledge of the dynamics of the enzyme, in particular in the computational modeling of binding modes. For this reason, after using traditional docking methods and x-ray crystallography, a large effort has been devoted to the improvement of the computational characterization of sidechain rotamers, that have been included in the rational design process.

The manuscript is composed of two main parts. In the first one, chapters 1, 2 and 3, I introduce the main thematics of the work and the experimental methodologies. Chapter 1 is devoted to acetylcholinesterase structure and function, with a

particular focus on its interaction with organophosphates and reactivators. Chapter 2 introduces the thematic of rational drug design, and the implications of protein structural dynamics in the process. Chapter 3 introduces the main methods used in the thesis: x-ray crystallography, molecular dynamics simulations and molecular docking.

The second part, chapters 4 and 5, is focused on the experimental results. Chapter 4 presents all the steps required to go from the design of KM297 to its optimization. Various crystal structures, along with simulation studies, are involved in the process leading to the design and evaluation of the derivative JDS207. Chapter 5 describes the efforts to develop a new method to include sidechain flexibility in molecular docking. I will present the main results that range from a qualitative observation of the sidechain movements in molecular dynamics simulations to the development of a computer program that generates receptors for docking reproducing the simulation results.

Finally, the appendix section includes the docking code and the manuscripts published during the thesis.

Chapter 1

Acetylcholinesterase

1.1 Biological function

Acetylcholinesterase (AChE) is an essential enzyme notably present in the nervous system of superior eukariotes. In the central nervous system, it is found at cholinergic synapses, where its role is to hydrolyze the neurotransmitter acetylcholine (ACh) (Dale, 1914) restoring in this way the excitability of the synapse and allowing the transmission of the subsequent nervous signal (Silman and Sussman, 2005). Due to its function, its turnover must be faster than a nerve signal transmission process, which takes a few milliseconds. For this reason, AChE is one of nature's fastest enzymes. Depending on the species considered, AChE can hydrolyze between 1000 and 20000 substrate molecules per second. Its catalytic efficiency has been estimated to be in the order of $10^8 M^{-1}s^{-1}$, in the range of the theoretical limit of diffusion controlled reaction (between $10^8 M^{-1}s^{-1}$ and $10^9 M^{-1}s^{-1}$) (Bar-Even et al., 2011). Therefore it has been suggested that the trafic of substrates and products to and from the active site should be the limiting factor for the efficiency of the enzyme (Quinn, 1987).

At neuromuscular junctions (figure 1.1), AChE is present in anchored tetrameric form. The C-terminal helices of the four monomers interact by binding to a Proline-Rich Attaching Domain (PRAD) (Massoulié et al., 2005). PRAD peptides are found in two main proteins. One is a collagenic protein ColQ (Cousin et al., 1996) and the other is the neuron membranes PRiMA (Proline-Rich Membrane Anchor) (Perrier et al., 2002).

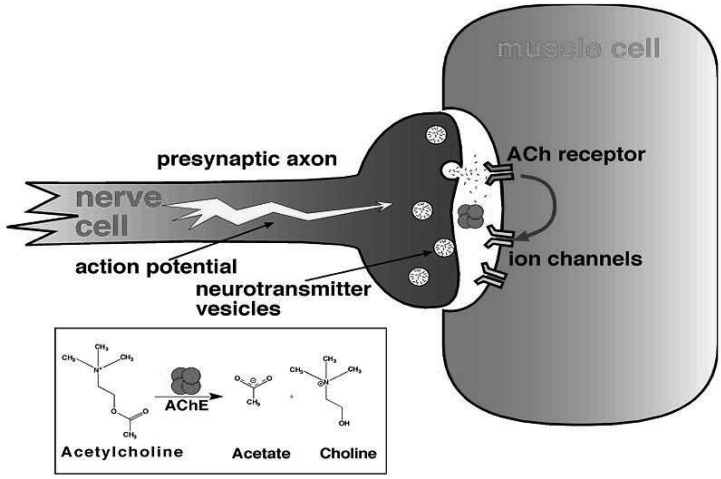


Figure 1.1: Schematic representation of a neuromuscular junction. AChE is found in the synaptic cleft and restores transmission by hydrolyzing neurotransmitter acetylcholine (ACh).

Another type of cholinesterase, the butyrylcholinesterase (BChE), is widely present in eukariotes. It has been characterized by structural and functional studies (Masson et al., 2009), but its role is still not completely understood.

1.2 Structure and dynamics of AChE

1.2.1 Three dimensional structure of AChE

The first crystals of an AChE were obtained from the electric fish *Torpedo californica* (*Tc*AChE) in the late 1980s (Sussman et al., 1988) and the crystal structure was solved soon after (Sussman et al., 1991). The choice of this organism was motivated by the fact that its electric organ is formed by a modified muscle containing multiple neuromuscular junctions, in which a dimeric anchored form of AChE is expressed in large quantities.

Each of the monomers is folded in an α/β hydrolase fold (Ollis et al., 1992), and contains 15 α-helices and 11 β-strands, connected by flexible loops. The contact region between two monomers, made by four helices, is called the 4-helix bundle. This first AChE structure (Sussman et al., 1991) revealed that the catalytic triad of

the enzyme, formed by residues Ser200, His440, Glu327 [1], is located at the heart of the enzyme (figure 1.2). The substrate can access the active site through a 20 Å deep gorge that is less then 5 Å wide at its narrowest point, the so-called bottleneck. A cluster of aromatic residues, Trp279, Tyr70 and Tyr124, forms a peripheral binding site at the entrance of the gorge.

Crystallographic structures of complexes between AChE and substrate analogues revealed how the catalytic machinery is optimized to fit the transition state of the substrate (ACh). In particular, the structure of *Tc*AChE in complex with a non-hydrolysable substrate analogue 4-oxo-N,N,N-trimethylpentanaminium (OTMA, figure 1.3) (Colletier et al., 2006a; Bourne et al., 2006) provides us with the conformation of the active site in the intermediate state of the hydrolysis reaction.

We can observe (figure 1.2 B) how the OTMA molecule bound to the catalytic serine interacts with the active site environment, to adopt an optimal position for hydrolysis. The negative charge of the carbonyl oxygen of the acyl group is stabilized by a series of hydrogen bonds within the oxy-anion hole, formed by the main chain NH groups of Gly118, Gly119 and Ala201. The methyl group of the OTMA carbonyl interacts with the so-called acyl-binding pocket, formed by residues Phe288, Phe290 and Trp233. The limited size of this pocket has an essential role in the substrate selectivity (Harel et al., 1992).

At the other end of the substrate analog, the most important interaction occurs between OTMA nitrogen quaternary ammonium and the choline-binding pocket. It is formed by residues Phe330, Glu199 and Trp84. While the role of the glutamic acid is uncertain due to controversy regarding its protonation state (Wahlgren et al., 2011), the main interaction is between the positively charged trimethylammonium of OTMA and the π electrons of Trp84 and Phe330 (Harel et al., 1993). This interaction has been calculated to account for about 50% of the interaction energy between the enzyme and its substrate (Harel et al., 1996). The substrate-binding mode, as observed from the structure, is optimal to position the ester function in front of the catalytic serine so is to permit hydrolysis.

The peripheral anionic site (PAS) at the entrance of the gorge (figure 1.2 A) is responsible for capturing incoming substrate molecules and for orienting them correctly. Its most important residue was found to be the highly conserved Trp279.

[1] Residue numbering is from *Torpedo californica* enzyme in the whole manuscript, unless stated otherwise.

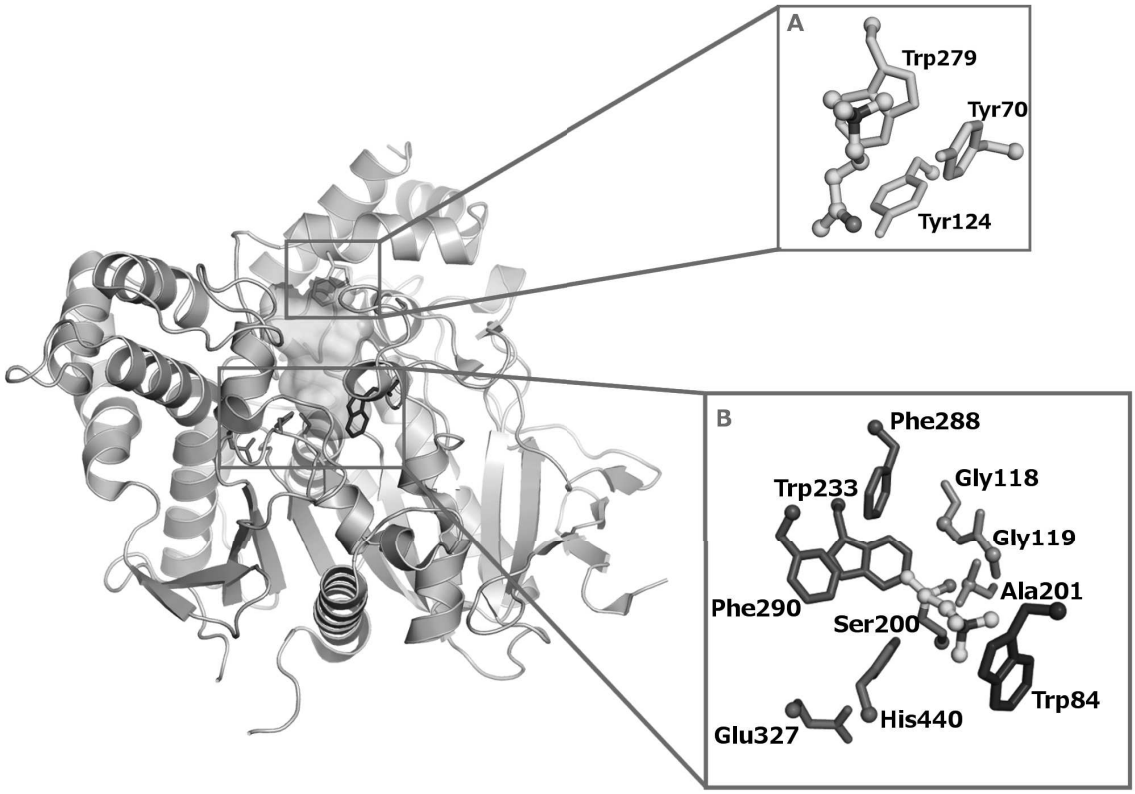
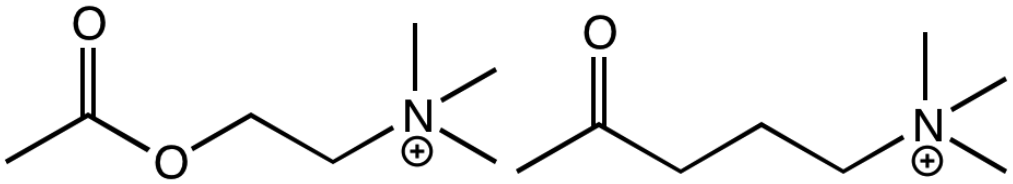


Figure 1.2: The 3-dimensional structure of *Tc*AChE (Sussman et al., 1991). The catalytic triad is shown in red sticks at the core of the enzyme. In this orientation, the entrance of the active site gorge, whose surface is shown in green, is at the top of the figure. The red squares show the details of the interaction between the substrate analogue OTMA (carbons shown in cyan) and the residues in the gorge of *Tc*AChE (Colletier et al., 2006a). A) The peripheral aromatic site residues, shown in green at the top of the figure, with a substrate molecule bound in front of Trp279. B) The catalytic site. Bound to the triad, shown in red, we observe a second substrate analogue molecule, covalently bound to Ser200. The choline moiety is in cation-π interaction with Trp84 (blue). The oxygen of OTMA is stabilized in the oxyanion hole (orange), and the methyl of the acetyl group is nested in the acyl-binding pocket (purple).

Crystal structures of AChE in complex with high concentrations of substrate, where a molecule of ACh or acetylthiocholine was found in front of Trp279, confirmed the role of this residue in binding substrate (Bourne et al., 2006; Colletier et al., 2006a).



(a) Acetycholine (b) OTMA

Figure 1.3: Neurotransmitter acetylcholine (a) and its non-hydrolyzable analogue (b) 4-oxo-N,N,N-trimethylpentanaminium (OTMA)

1.2.2 Structural dynamics involved in substrate traffic and ligand binding

While the access of substrates through the gorge and the general catalytic mechanism are well understood, there is still a debate concerning the exit of reaction products from the active site. Either they can exit through the gorge, or through an alternative exit, the so called backdoor (Gilson et al., 1994). Crystallographic studies have shown the presence of the reaction product thiocholine, a choline analogue, just behind Trp84 (Bartolucci et al., 1999; Colletier et al., 2006b). Mutagenesis aimed at enlarging the putative backdoor has shown that products can exit through this route when the gorge entrance is blocked, though this enlargement did not increase the catalytic efficiency of the enzyme when the entrance was free (Kronman et al., 1994; Faerman et al., 1996; Nachon et al., 2008). It was observed by crystallographic studies for *Tc*AChE and for *Bf*AChE in complex with an antibody (Sanson et al., 2011; Bourne et al., 2014), that sidechains of Trp84 and of the related residue Tyr442 can move to open an exit channel at the bottom of the gorge. The reaction product thiocholine has been observed in MD simulations to leave the active site through this door (Xu et al., 2010). Molecular dynamics simulation studies have shown the possibilities for alternative routes to the active site to open (Gilson et al., 1994; Tai et al., 2001). Even though there is no other direct experimental evidence of these side doors openings has been observed, the binding of bulky ligands like galanthamine derivatives (Greenblatt et al., 2004), is possible only if these rearrangements, in particular of the Trp279-Ser291 loop, can occur.

Conformational changes in the gorge environment have been shown to occur upon ligand binding both by MD simulations and crystallographic structures. The

access of the inhibitor huperzine-A to the active site can only occur through a rearrangement of the bottleneck region, where the distance between residues Phe330 and Tyr121 goes from 2 to 6 Å during the passage of Hup-A (Xu et al., 2003).

An interesting case is the one concerning the bifunctional inhibitor NF595. This molecule has been shown to bind to an alternative conformation of the peripheral site (Colletier et al., 2006b). Molecular dynamics simulations have helped to conclude that the molecule binds to a pre existing conformation, rather than through a mechanism of induced fit (Xu et al., 2008).

Covalent binding of the potent organophosphate cresyl saligenin phosphate induces a conformational change of the acyl-binding pocket (Carletti et al., 2013). In this case, the conformation assumed by Phe297[2] has not been observed in molecular dynamics simulations of the native form of the enzyme, suggesting that an induced fit mechanism regulates the interaction of this molecule with the enzyme.

1.2.3 Human AChE crystals

For two decades, the human enzyme was only crystallizable in complex with fasciculin-2 (FAS-2), a three-fingered toxin that inhibits AChE by binding at the peripheral site and blocking the access to the gorge. FAS-2 favorizes crystallization of the enzyme by participating in crystalline contacts (Harel et al., 1995; Bourne et al., 1995; Kryger et al., 2000). Structures of hAChE in complex with ligands have been solved with this method (Nachon et al., 2013). However, opportunities of application of these crystals have been limited by the occupation of the peripheral site and resulting difficulties in the access to the active site. Also, FAS-2 binding was shown to modify the dynamics of the enzyme (Bui et al., 2006; Bui and McCammon, 2006), meaning that observation of conformational changes in these crystals would not be conclusive. A first form of recombinant hAChE, truncated at residue 540, was solved without fasciculin (Dvir et al., 2010), but crystals diffracted poorly (3.2 Å resolution) and the access to the active site of one of the monomers was blocked by a symmetry related copy. Finally, a higher resolution structure (2.0 Å) of another recombinant form of hAChE truncated at residue 1 and 544 (Cheung et al., 2012), has allowed soaking of ligands into hAChE, revealing differences in the binding modes of some molecules, like the anti-Alzheimer molecule donepezil, previously solved in complex

[2]Mouse AChE numbering, Phe288 in *Tc*AChE

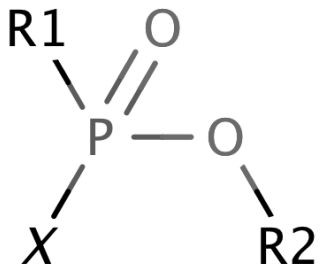


Figure 1.4: General structure of an organophosphate nerve agent.

with *Tc*AChE (Kryger et al., 2000).

1.3 Organophosphates and their inhibition of AChE

AChE is the target of various potent neurotoxic compounds, including organophosphate (OP) compounds. This family of molecules includes both pesticides and chemical warfare nerve agents. While the use of nerve agents is forbidden by international conventions, organophosphate pesticides are still largely used in the agriculture of most developing countries. Every year about 3 million intoxications by OP pesticides are recorded worldwide, including around 280000 fatalities.

The general structure of an OP nerve agent is shown in figure 1.4. A phosphorous atom forms the core of the molecule, and is bound to four substituents: an oxygen (or a sulfur) atom, a leaving group (X), an alkoxyle (R1) and another group (R2). Pesticides were the first molecules of this family to be developed, around the mid 19^{th} century. In the 1930s, Germans synthesized the G-agent series (G stands for 'german'). The first was tabun, or GA (1936) followed by sarin (GB, 1937) and soman (GD, 1944). The other main family of warfare nerve agents, the V agents (V stands for 'venomous') has been developed by the British, with the synthesis of VX, which has been followed by a few related molecules, mainly the Russian VX, or VR and the Chinese VX, or CVX. In general, G agents are more volatile then V agents, but the latter are more persistent and more toxic, especially through skin contact.

These powerful chemical weapons have only been used on rare occasions. Also, in 1993 , most countries have signed the Chemical Weapons Convention, forbidding the

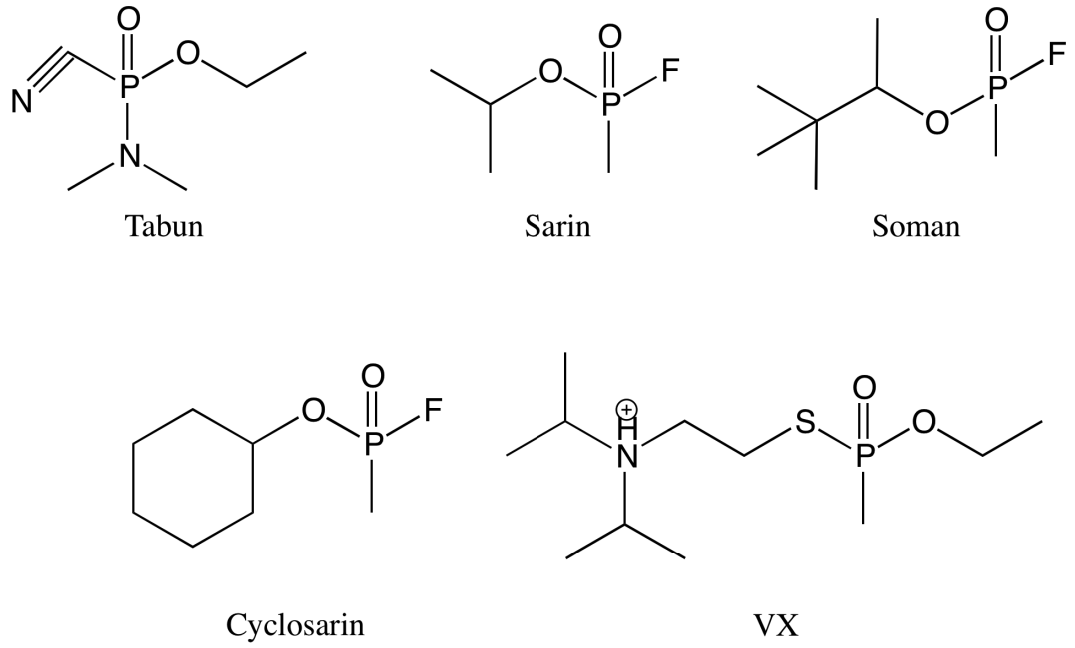


Figure 1.5: The main organophosphate agents. Molecules of the G-series: tabun (GA), sarin (GB), soman (GD), cyclosarin (GF); and lead molecule of the V-series: VX.

usage of chemical warfare agents. However, we can cite a few cases of terrorist/war acts using this kind of weapons in the recent years. A famous one is the 1994 Tokyo metro attack, where sarin was released in a metro station in the capital of Japan, killing 11 people and intoxicating around 3000. More recently, during the Syrian civil war, a sarin attack in August 2013 in the suburbs of Damascus has caused around 1000 deaths, according to American sources [3].

1.3.1 Inhibition mechanism

The catalytic serine of AChE is phosphylated by nucleophilic substitution. In this mechanism, represented in figure 1.6, the oxygen atom of the catalytic serine attacks the phosphorous of the OP, from the opposite side of the leaving group, forming a trigonal bipyramidal transition state. Then, the bond between the phosphorus and the leaving group breaks, leading to the formation of a tetrahedral serine adduct.

[3]See for exemple this article from the Washington Post: http://tinyurl.com/ksj2lc2

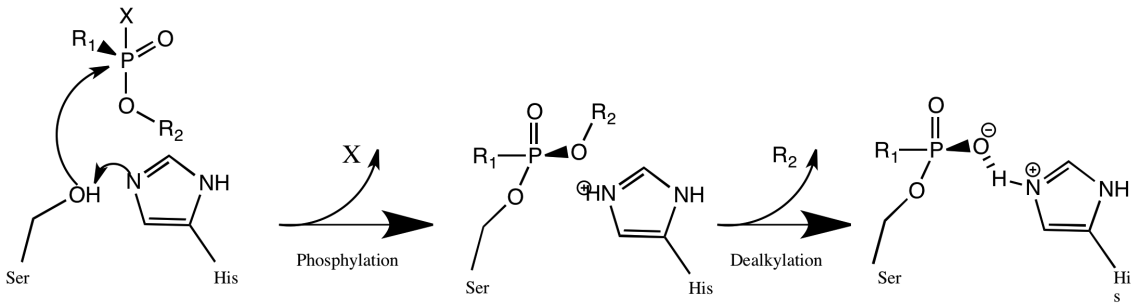


Figure 1.6: The inhibition mechanism of AChE by organophosphate nerve agents.

This adduct is a close structural analogue of the transition state of catalysis (figure 1.2).

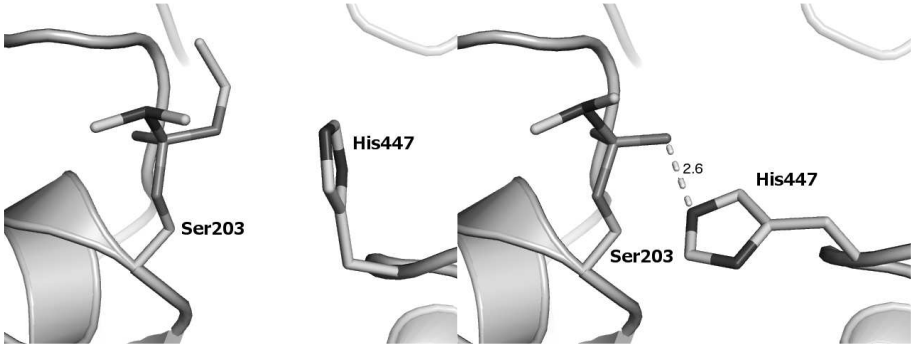
The catalytic triad is not, however designed to catalyze the hydrolysis of this adduct. Even if spontaneous dephosphylation by nucleophilic substitution with a water molecule is possible, it is extremely slow, with timescales spanning from hours to days.

After inhibition, the enzyme is able to catalyze a secondary reaction, referred to as 'aging' (Fleisher and Harris, 1965), wich involves the dealkylation of the alkoxyl group R2. The aged complex is even more stable than the 'non-aged', and dephosphylation of the catalytic serine after this stage is virtually impossible (Masson et al., 2010).

The most probable origin of the stability of the aged adduct, is the formation of a salt bridge between the imidazolium of the catalytic histidine and the freshly formed oxyanion group bound to the phosphorus atom.

1.3.2 Three dimensional structures of OP-inhibited AChE

The inhibition mechanism has already been clarified by mutagenesis studies, but its deeper understanding has been made possible by determining crystal structures of OPs in complex with AChE from various species (Millard et al., 1999b,a; Ekström et al., 2006a; Carletti et al., 2008; Sanson et al., 2009). It has been observed, e.g. that aging of tabun in *Tc*AChE is accompanied by a conformational change of the histidine of the catalytic triad, figure 1.7. This residue is pushed in a non-native conformation by the presence of the ethyl group of tabun, and goes back to its native position after aging, forming a salt bridge with the phosphate group of the



(a) Non-aged form (b) Aged form

Figure 1.7: Crystallographic snapshots of aging of tabun-inhibited mAChE (pdb acces code 3dl4 and 3dl7). We can observe that His447 moves from an alternative conformation forced by the ethoxy substituent of tabun back to its native one, forming a salt bridge with the oxyanion of the phosphoramidyl group corresponding to aged tabun (Millard et al., 1999a; Carletti et al., 2008).

inhibited catalytic serine (Millard et al., 1999a). The lack of non-aged structures of hAChE in complex with OPs, due to the difficulty in performing soaking experiments with hAChE-FAS crystals, will probably be addressed now that unliganded hAChE crystals of good quality begin to be available.

1.3.3 Treatment of OP intoxication

The current formulation for treatment of OP intoxication is composed of three types of drugs, with the goal of addressing the symptoms of AChE inhibition, while also liberating the enzyme from the presence of the OP: an antimuscarinic (atropine), and anticonvulsivant (diazepam) and an oxime reactivator of phosphylated AChE. These drugs are combined into an autoinjector for field treatment. Oximes (figure 1.8) are a class of highly nucleophilic molecules that can make a substitution, thus breaking the bond between the oxygen of the catalytic serine and the phosphorus of the OP (Mercey et al., 2012).

The first oxime to be synthetized, about 60 years ago, was the 2-pyridinium aldoxime methiodide (2-PAM) (Wilson and Ginsburg, 1955). Since then, hundreds of pyridinium aldoximes have been synthesized. HI-6 (Oldiges and Schoene, 1970) is considered as the best oxime to date, even if tabun-inhibited AChE is almost completely resistent. The mechanism of reactivation by an oxime is thought to happen

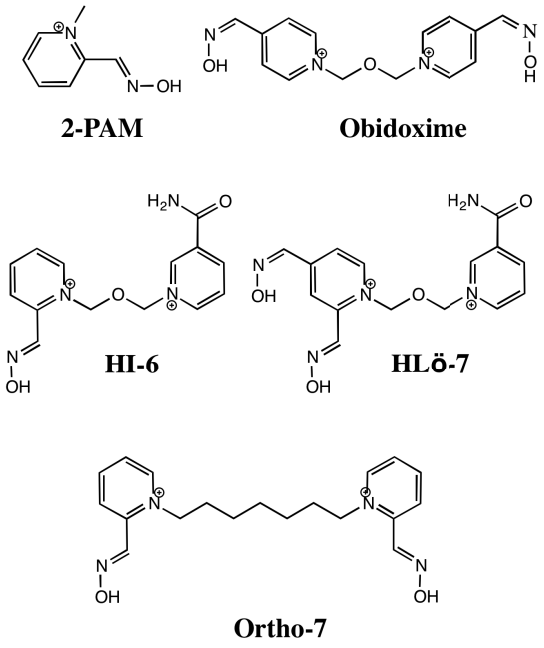


Figure 1.8: Structures of the most common oximes.

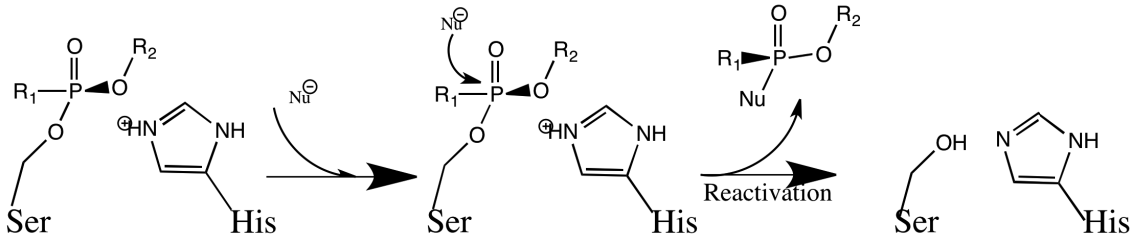


Figure 1.9: Reactivation mechanism of OP-inhibited AChE by a nucleophile Nu.

according to the following scheme (figure 1.9): the strong nucleophilic oximate attacks the phosphorus atom from the opposite side of the serine to form a reversible bipyramidal OP-oxime-AChE transition state. The leaving of the OP-oxime adduct liberates the catalytic serine restoring in this way the activity of the enzyme.

The oxime reaction mechanism has been studied by in vitro assays, simulations and X-ray crystallography. Crystal structures of AChE in complex with oximes have shown how they can bind to the enzyme. Most bisquaternary oximes have been found to bind at the PAS of the enzyme, through cation-π interactions. For example, one pyridinium group of HI-6 and HLö-7 is bound at the PAS of mAChE

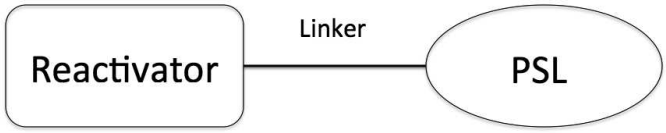


Figure 1.10: General structure of a bifunctional reactivator. It is formed by a group bearing an oxime function and a peripheral site ligand (PSL), connected by a flexible linker.

(Ekström et al., 2006b) forming an aromatic stacking involving Trp286 and Tyr124[4], while the stacking of Ortho-7 and obidoxime involves Trp286 and Tyr72[4]. The PAS tryptophan has to rearrange for stacking to oxime, supporting the hypothesis that the high level of flexibility of this residue (Xu et al., 2008) is necessary for the binding of these ligands to the enzyme.

These structural data helped in clarifying the mechanism of reactivation (Ekström et al., 2009) and in pointing out the drawbacks of some oximes. For example the oxime function of HI-6 is at the closest at more than 9 Å from the phosphorous atom, implying that the orientation of the oxime group is not optimal for the nucleophilic substitution. Also, 2-PAM has been found to be stacked in front of Trp84 in aged soman-*Tc*AChE with the oxime group forming an hydrogen bond with the mainchain carbonyl of the catalytic serine (pdb code 2vq6) (Sanson et al., 2009). In this conformation, reactivation is not possible, the oxime being also at more than 9 Å from the phosphorous atom of the OP.

The development of oximes has greatly taken advantage from the structure of ternary AChE-OP-oxime complexes. Observation of the binding of bispyridinium oximes at the peripheral site led to the concept of bifunctional reactivators (figure 1.10), for which one part of the molecule is a specific PAS binder providing affinity to the enzyme, connected through a carbon linker to an oxime function, that provides the reactivity of the molecule (Mercey et al., 2011).

Other strategies of protection against OP intoxications involve prophylactic treatments. During the first Gulf war, the US army used pyridostigmine, a carbamate that inhibits AChE by making a short-lived bond with the catalytic serine, thus rendering it unable to react with OPs. However, the suspected deleterious side effects pushed towards the developments of new prophylactic stategies, enzyme-based

[4]mAChE numbering.

scavengers being the most promising one. These enzymes can neutralize OPs before they can inhibit AChE at the synapses and neuromuscular junctions (Nachon et al., 2013). We distinguish between two kinds of OP scavengers, catalytic and stoichiometric. Stoichiometric scavengers are molecules that can trap a single OP molecule. Carboxylesterase (Maxwell and Brecht, 2001; Jackson et al., 2013) and butyrylcholinesterase (BChE) (Masson et al., 2009), belong to this type. Treatments based on injections of BChE are now under development, as production costs and purification efficiencies are improving (Brazzolotto et al., 2012).

Catalytic scavengers, like human paraoxonase (Rochu et al., 2007), phosphotriesterase (Griffiths and Tawfik, 2003) or self-reactivable mutants of BChE, have the ability of hydrolizing OPs and thus neutralizing multiple OP molecules. Therefore, a significantly lower dose would protect against the same amount of OP. Some mutants of paraoxonase and phosphotriesterase are now efficient enough to offer good protection and great efforts are devoted to develop safe formulations (Worek et al., 2014a,b).

Chapter 2

Drug design and protein flexibility

The process of designing new drugs is a long and tortuous road. At first there is the need to find a good ligand for a target protein in the organism. Then, this molecule has to be made into a drug, by improving its pharmacological properties without loosing affinity and specificity for the target. Finally, it has to pass through different levels of clinical tests, to assess its efficacity and characterize possible side effects (Mandal et al., 2009). With the term 'drug design' we denote the process of creation of a therapeutic molecule, based on the knowledge of a specific target in the organism, or in the modification of already known ligands. Definition of 'drug design' has been created to put in advance the rational procedure involved in this process, in opposition to the traditional 'trial and error' procedure. We should more precisely talk about ligand design, since this process usually stops at the first step of the development of a drug, when a good ligand for the target protein is found and its properties are optimized. Through the manuscript we will keep using the common expression 'drug design'.

There are two main approaches to drug design: ligand-based or receptor-based design. In a process of ligand-based design, the focus is made on how to use known molecules, eventually with modifications, to target receptors related to different diseases. In this thesis we used a receptor-based approach, in which we take advantage of the knowledge of the structure of a target enzyme to design ligands that specifically fit inside the active site.

In the past, knowledge of the receptor was limited to its crystallographic structure. However it is known that protein dynamics can have an impact on the mode

of interaction with substrates and ligands. After briefly discussing the type of movements involved in the dynamical behavior of proteins, we will describe some methods to include information about dynamics in the process of designing new ligands.

2.1 Structure based drug design (SBDD)

2.1.1 Definition of structure based drug design

Structure based drug design (SBDD) describes every methodology that uses a known 3-dimensional structure of a target to develop and optimize ligands that would bind to it. The first drug produced with this method is the anti-glaucoma molecule Dorzolamide (Greer et al., 1994), which became available on the market in 1995. Another famous exemple of a rationally designed drug is the tyrosine kinease inhibitor Imatinib, approved by the FDA in 2001 and commercialized by Novartis as Gleevec (Zimmermann, 1996). Even if some molecules created in a drug design process made it to the market, this method is still far from being completely efficient: return on investment is moderate with respect to the previsions made a couple of decades ago (Brown and Superti-Furga, 2003).

2.1.2 The search for new ligands

Once the target protein is chosen, the search for a ligand begins.

The most common techniques in the discovery of ligands are virtual screening (Jain, 2004), hight throughput screening (Hertzberg and Pope, 2000) and rational building (Hartenfeller and Schneider, 2011).

The first two methods consist in the screening, either *in vitro* or *in silico*, of large molecule libraries and of their activity against the chosen target. Virtual screening techniques use chemical databases, e.g. the ZINC database (Irwin et al., 2012), which are large libraries of chemical compounds in a standard digital format. All these compounds are then docked against the target, to find which of them present the higher affinity.

High throughput screening is conceptually similar but, in this case, activity of ligands on the target is directly evaluated *in vitro*, using screening plates. Technological advances have allowed to screen up to 100,000 ligands per day on a single

target (Hann and Oprea, 2004). Due to the huge quantity of data generated by both techniques, data processing has a key role in defining the good hits of the experiments (McInnes, 2007; Jacobsson et al., 2003). Statistical methods exist to combine *in silico* and *in vitro* and increase the accuracy of the screenings (Bajorath, 2002).

If multiple binding sites are present on the target, site-specific hits can eventually be combined together, to improve their affinity or specificity via a synergistic effect. This is the case for exemple of AChE and its distinct peripheral and catalytic sites.

Instead of searching for known molecules to fit the binding site of a protein target, another approach to SBDD consists in building a new ligand *de novo*, based on the conformation and the structure of the binding environment (Hartenfeller and Schneider, 2011). This design process is aimed to adapt the properties of the ligand to the receptor (Schneider and Fechner, 2005). This kind of design takes advantage of the structural knowledge of target-inhibitor complexes. These structures can be a starting point in the design of a ligand by screening for the optimal chemical functions to bind at a given position in the protein. An example of *de novo* design is the combinatory study of human cannabinoid (CB1) receptor inhibitors (Rogers-Evans et al., 2004). In this work, a series of previously known GPCR inhibitors has been used to generate fragments that were then combined to design inhibitors for the CB1.

In our project, we used a mixed approach to develop a family of bifunctional AChE reactivators. A new reactive oxime molecule has been coupled to a known ligand for the peripheral site of the enzyme. These results will be presented in chapter 4.

2.2 Protein structural dynamics

2.2.1 The conformational landscape of proteins

Proteins are in many cases erroneously represented as static, rigid objects. Yet, we know that many of their functional properties cannot be explained solely by their crystal structure, which is not a simple photography of a protein, but a temporal and spatial ensemble average image of all the molecules forming the crystal.

Protein dynamics can be understood in terms of an energy landscape of their conformations (Frauenfelder et al., 1991). If we consider that a protein can adopt

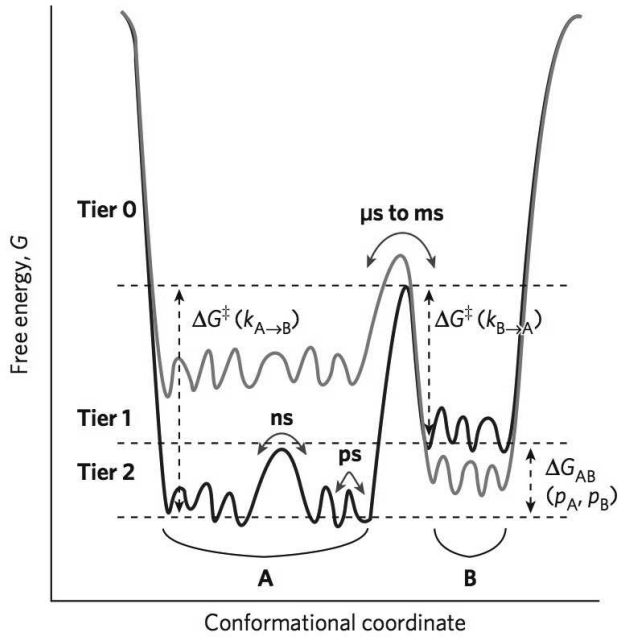


Figure 2.1: Simplified representation of the conformational landscape of a protein. The different heights of energy barriers are translated in the different timescales at which each of these changes occur. Extracted from Henzler-Wildman and Kern (2007).

a set of iso-energetic states, it is easier to imagine that it can evolve from one to another in a dynamic way.

Motions can occur on different time scales (Henzler-Wildman and Kern, 2007), each corresponding to a specific energetic level in the conformational landscape, as represented in figure 2.1. The most rapid movements are represented by the small energy barriers in the conformational landscape. They involve side chain motions and typically occur on the scale from the picosecond up to the nanosecond. The slowest movement, with timescales up to the second, correspond to all the large conformational changes occurring upon folding, interaction with ligands and rearrangements of subunits of large molecular assemblies.

All the possible movements have been proven to be involved in enzyme functioning (Henzler-Wildman et al., 2007), by for example modulating catalysis and access of substrates to the active site.

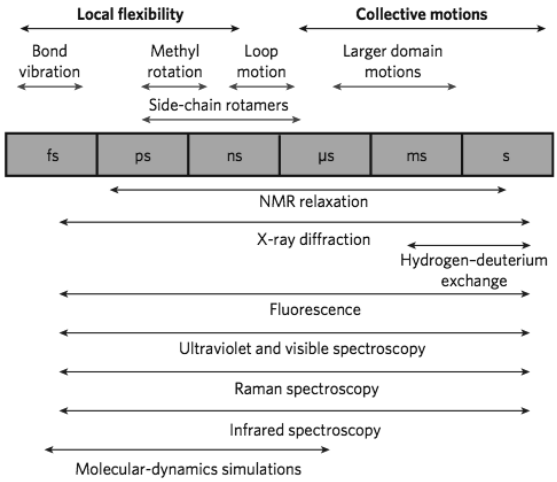


Figure 2.2: Schematic representation of the different techniques available to characterize the structural dynamcis of proteins. The timescales for all atoms MD simulations is constatly increasing, up to the millisecond. Extracted from Henzler-Wildman and Kern (2007).

2.2.2 Techniques to characterize protein structural dynamics

Different experimental and computational techniques are available to study spatial and temporal aspects of protein motions. Their respective temporal domains are represented in figure 2.2. With no intention of being exhaustive, we present here the most widespread in studying the dynamics at atomic resolution.

Nuclear magnetic resonance (NMR) is a technique that takes advantage of the resonance frequencies of certain nuclei to both characterize their motions and solve the structure of a protein. NMR structures are characterized by the fact that data are reproduced by an ensemble of structures, rather than a single one. This can provide insight into the dynamic processes regulating protein activity (Marion, 2013).

Kinetic crystallography is a general definition that includes multiple methods aimed at characterizing the intermediate states of protein 'at work' by solving their crystallographic structure (Bourgeois and Royant, 2005; Bourgeois and Weik, 2009). The common feature of all these methods is the need for synchronization between all the molecules in the crystals, to get the best snapshot possible for an intermediate of a reaction. To achieve this synchronization, we can for example recur to caged compounds, substrate analogs that can be activated through a signal, such as a laser pulse or an X-ray burn, to achieve a synchronized start of the enzymatic process in

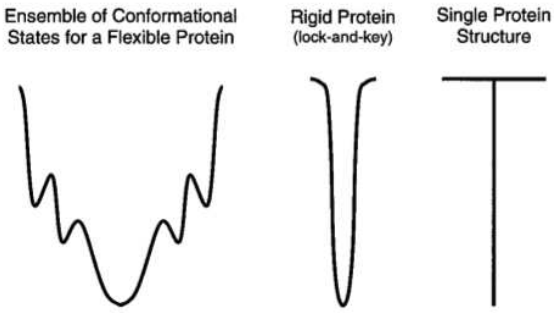


Figure 2.3: Representation of a single structure in the theory of conformation landscape. From Carlson and McCammon (2000).

all the molecules of the crystals (Colletier et al., 2007, 2008). Other techniques used in kinetic crystallography include substrate diffusion in crystals, variation of data collection temperature (Weik and Colletier, 2010), pH variation (Nar et al., 1991) and real time Laue crystallography (Schotte et al., 2003; Bourgeois et al., 2003).

Molecular dynamics simulation is widely used to extract kinetic information on the properties of a protein, for which the crystallographic 3D structure is known. Timescales of all atom simulations have constantly increased up to the millisecond, due to the always larger computational power and to the optimization of massively parallel integration algorithms (Shaw et al., 2009). More details on MD simulations can be found in chapter 3.

2.3 Including dynamical information in SBDD

As for enzymatic activity, structural protein dynamics is also involved in the interaction with ligands. The description of the conformational ensemble can provide a sharper understanding of this problem. A single structure, such as a crystallographic structure, represents only a single point in the entire conformational landscape of the protein, figure 2.3, and is therefore inadequate to represent the complexity of protein dynamics and enzymatic catalysis.

For this reason, large efforts have been devoted in the last 20 years to include dynamic information in SBDD. The increase of available computational power, with costs always lower, has opened a panoply of opportunities for the development of new techniques (Carlson and McCammon, 2000; Michel, 2014).

The most simple approach is the use of multiple structures as a framework for the design (Totrov and Abagyan, 2008). As explained above, an ensemble of structures can be obtained with NMR, but this approach is limited to relatively small proteins. To overcome this limitation, MD simulations have been used to produce an ensemble of structures, extracted from trajectory snapshots (see for example (Pang and Kozikowski, 1994)). Recently, room temperature crystallographic structures have also been successfully used to generate multiple receptors for docking and finding new inhibitors (Fischer et al., 2014).

More sophisticated methods can be included in the general category of soft docking, which will be described in practical details in the next chapter. We can distinguish between traditional flexible docking and more sophisticated techniques. We consider as a flexible docking experiment the case in which only the sidechains of chosen residues are free to move. A different case is when the mainchain of the protein is also given a certain degree of freedom. This is often done by using MD simulations and extracting multiple snapshots, to find rarely occupied protein conformations (Lin et al., 2002).

Including flexibility in drug design has been fundamental in the design of HIV-1 reverse transcriptase inhibitors (Das et al., 2005). In particular, they were designed to overcome the problem of drug resistance, by targeting multiple conformations of the target. Crystallographic structures have confirmed that they can bind to at least two different conformations of the active site (Das et al., 2004).

Another successful example is the HIV integrase inhibitor ISENTRESS, developed by Merck and approved by the FDA in 2007. This molecule has been designed following the discovery of a possible binding trench through molecular dynamics simulations (Schames et al., 2004), which have then been characterized as proper ligand binding sites by molecular docking techniques.

In chapter 5 of this manuscript we will describe a new method, based on MD simulations data, to characterize the dynamics of side chains of AChE and include it in a drug design process.

Chapter 3

Methods

During this work we recurred to multiple methods to characterize the dynamics of AChE and its interaction with OP nerve agents and reactivators. In this chapter we will briefly introduce the following methodologies: X-ray protein crystallography, molecular docking and molecular dynamics simulations. Materials and methods for each experiment will be described in subsequent chapters.

3.1 X-ray protein crystallography

X-ray protein crystallography is by far the most widely applied technique to solve protein structures. The Protein Data Bank (PDB) contains today 104125 structures [1], and almost 90% are obtained by means of X-ray crystallography. The accessibility of synchrotron radiation facilities around the world has been a great impulse to the growth of the number of solved structure, along with more user-friendly analysis software and the increase in computational power. The development of cryo-methods has also allowed to collect more higher quality data at synchrotron sources.

The bottleneck of protein structure determination is the generation of well diffracting crystals. Automated methods of screening with crystallization robots have made the process easier, but protein crystallization remains a trial and error procedure. While recent developments have allowed to solve structures with smaller crystals, it is usually easier to collect good quality data from large crystals, typically hundreds of microns of side.

[1] As of 16^{th} of October 2014

In the following sections we will introduce the principles of X-ray diffraction by crystals, then address the issue of obtaining protein crystals and, finally, illustrate the main methods to solve a protein structure from X-ray diffraction patterns.

3.1.1 Crystal symmetry

A crystal is defined as a solid which has a spatial periodicity. Spatial periodicity means that it is formed by a regular repetition of the same element over the three dimensional space. This defines the crystal symmetry, as a set of operations over the 3 dimensions that lead to an identical position. The minimal unit of the crystal is called the asymmetric unit. It is usually defined by three vectors (a, b, c), describing its dimensions, and the three angles (α, β, γ) between them. All the possible symmetries are classified in 7 lattice systems, and in 14 Bravais lattices. In each of the lattices, multiple space groups, the total number of possibilities is 230 for crystals of macro-molecules, can be defined. Each space group has a well defined and unique symmetry.

3.1.2 Diffraction of X-rays by a crystal

X-ray photons interact with matter, more specifically, with the electron shells of atoms. The regular arrangement of atoms into a crystal forms planes that, as mirrors, have the property to reflect X-rays. This reflection, due to the undulatory nature of light, can occur only in some well defined directions, where the interference between X-rays reflected by adjacent planes is constructive. This mechanism is described by Bragg's law (equation 3.1), which defines the condition to have a constructive interference between waves reflected by different planes of a crystalline solid. Its derivation is straightforward from a schematic representation of the diffraction event, figure 3.1.

$$n\lambda = 2d \sin\theta \tag{3.1}$$

Here n is an integer, λ is the incident wavelength and θ the angle between the incident wave and the crystal planes. At fixed wavelength, equation 3.1 gives us the value of the distances between crystalline planes, d, considering that the angle can be experimentally determined. For a macromolecular crystallography experiment, the

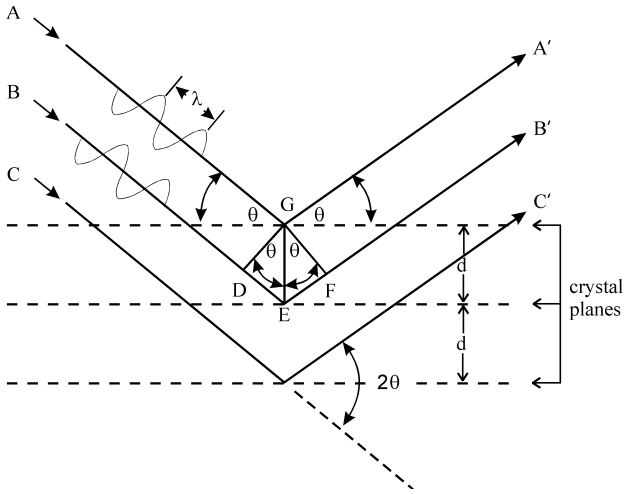


Figure 3.1: Schematic representation of Bragg's law.

wavelength is in the order of the Ångström[2]. As we can deduce from the expression of Bragg's law, incident X-rays are diffracted to discrete directions, related to the different distances between the crystalline planes.

Structure factor

A convenient way to write the Bragg's law in terms of scattering vectors and unit cell parameters is the Laue equations, equation 3.2. The Laue equations state that diffraction only occurs when the scalar product between the diffracted wave vector S and vectors (a,b,c) of the crystal is an integer

$$\begin{cases} a \cdot S = h \\ b \cdot S = k \\ c \cdot S = l \end{cases} \tag{3.2}$$

Diffraction by a crystal can be seen as the sum of the individual contributions of each atom of the system to the scattering of incident radiation. This means that every scattered amplitude $F(hkl)$ can be defined as the sum of the contributions of all the atoms in the volume, if we assume that the scattering by each atom is independent from that of its neighbors. Using the Laue equations, we can derive an

[2]The wavelength of a 12 keV photon is 1.03 Å

expression for the diffracted amplitude.

$$F(hkl) = \sum_{j=1}^{N} f_j e^{2\pi i(hx_j+ky_j+lz_j)} \tag{3.3}$$

Assuming the crystal to be ideally infinite, we can write the sum as an integral of the electron density over the volume of the crystal:

$$F(hkl) = V \int_{hkl} \rho(x,y,z) e^{2\pi i(hx+ky+lz)} dxdydz \tag{3.4}$$

We can recognize in equation 3.4 the Fourier transform of the electron density of the crystal, the structure factor F. This gives us a useful relation between measured quantities (diffracted intensities) and the electron density ρ of the atoms.

Reciprocal space

In order to solve the structure of a protein, we need to relate the (hkl) reflections observed with the spatial coordinates of the atoms of the protein. To do this, we have to define a reciprocal space, which can be defined as the wave space, in opposition to the direct space, which is the space defined by vector distances between atoms. Direct (a,b,c) and reciprocal (a*,b*,c*) space vectors are mutually related by equations 3.5. The denominator is the scalar triple product of the three vectors, which corresponds also to the volume of the unit cell in the direct space.

$$\begin{aligned} a^* &= \frac{b \times c}{a \cdot (b \times c)} & \qquad a &= \frac{b^* \times c^*}{a^* \cdot (b^* \times c^*)} \\ b^* &= \frac{c \times a}{b \cdot (c \times a)} & \qquad b &= \frac{c^* \times a^*}{b^* \cdot (c^* \times a^*)} \\ c^* &= \frac{a \times b}{c \cdot (a \times b)} & \qquad c &= \frac{a^* \times b^*}{c^* \cdot (a^* \times b^*)} \end{aligned} \tag{3.5}$$

From this definition we can deduce that the reciprocal space has the same origin then the direct space, and that if the direct lattice rotates, so does the reciprocal lattice.

Conditions to obtain diffraction are usefully represented by the Ewald construction (figure 3.2). This geometric construction arises from combining Bragg's law

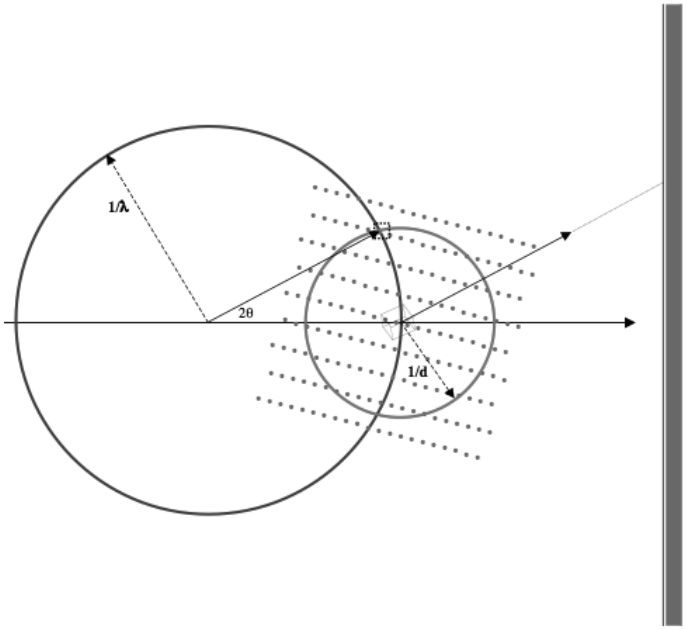
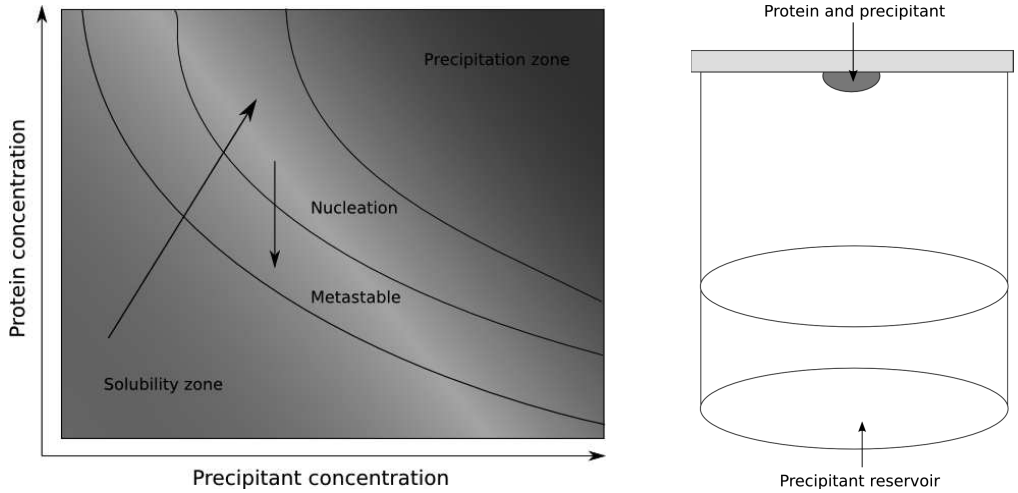


Figure 3.2: Representation of the Ewald construction. Every time there is an intersection between the Ewald sphere of radius $1/\lambda$ and a reciprocal lattice point contained into the resolution sphere of radius $1/d$, there is a constructive condition for diffraction. Image courtesy J.P. Colletier.

with the Laue conditions. If we build a sphere of radius $\frac{1}{\lambda}$ in the reciprocal space, we have diffraction every time a vector of the reciprocal space intersects the sphere. To optimize the number of diffraction spots, during an exposure to X-rays the crystal is rotated. In this way, more reflections are taken into account per image. Rotating the crystal also solves the problem that, due to the non ideal alignement of all the planes inside a crystal, not all the planes respect the Ewald condition at the same time. This effect, known as mosaicity, leads to the collection of partial reflections, that are difficult to handle. The reconstruction of the complete reciprocal space requires the collection of a series of images taken at different degrees of rotation. We need 360 degrees of rotation, which are reduced to 180 by the basic symmetry of the reciprocal space. High symmetry space groups allow to collect a complete dataset with a smaller value for the rotation angle.



(a) Phase diagram of a protein-precipitantsolution (b) Hanging drop method

Figure 3.3: a: Phase diagram for a protein-precipitant solution. We can observe how, during a crystallization experiment, the solution moves to the nucleation zone, where crystals start to appear. The system goes the towards rthe metastable zone, where crystals can grow in size. b: Schematic representation of a hanging drop crystallization experiment.

3.1.3 Protein crystals

Crystallogenesis of proteins by the vapor diffusion method

Growing good quality protein crystals is the *conditio sine qua non* to obtain the structure of a protein by means of X-ray diffraction. Crystals grow in a particular phase of a protein solution. If we look at the phase diagram represented in figure 3.3 we can observe a small nucleation region where crystals can start to grow.

Different methods exist to move the solution conditions towards the nucleation and the metastable region, the most used being the vapor diffusion method. In this technique, a protein solution is mixed with a buffered solution of precipitant to form a drop that is placed with a bigger reservoir of the same precipitant in an hermetic environment. Equilibration of the drop and reservoir through vapor diffusion leads to a slow concentration of precipitant and protein and eventually to the growth of small protein crystals in the drop, i.e. the nucleation process. By this phenomenon, the concentration of protein available in the solution decreases, resulting in the movement towards the metastable zone. This region of the phase

diagram is favorable forfurther crystal growth. A common setup to perform vapor diffusion is the hanging drop method,figure 3.3 b, where a drop of protein solution (typical volume is in the order of 1 μl) is deposed on a glass slide with the same ammount of precipitant solution. The slide is then used to close a crystallization pit, containing around 1 ml of the precipitant solution. The slide is sealed with grease. This creates an isolated atmosphere where vapor diffusion can take place.

3.1.4 Crystallographic data collection at synchrotron radiation sources

Once the crystals have been produced, they are exposed to an X-ray beam and diffraction patterns are recorded that can be integrated to determine the structure factors.

Synchrotrons

In a synchrotron, a beam of electrons is accelerated by a series of injectors and boosters up to energies of a few GeV and then stored on a circular accumulation ring. Electrons are kept into a pseudo-circular trajectory by a series of dipolar bending magnets (BM). For a phenomenon known as synchrotron radiation, electrons will lose energy by emitting X-rays in the direction of the tangent of the curve they are turning. Energy loss is then compensated by radiofrequency cavities, that are designed to keep the electron kinetic energy constant over a turn of the machine. The inconveniece is that a BM X-ray beam has a large energy spectrum and a large spatial dispersion.

Modern synchrotrons, usually referred to as third generation synchrotrons, use, along with bending magnets, another method of generation of X-rays called insertion device (ID). In an ID beamline, electrons are brought to oscillation by a series of consecutive permanent magnets of opposite polarization, called undulators. The spectrum of the emitted X-rays is narrower and the power is a couple of orders of magnitude higher than that of a BM.

X-ray detectors for protein crystallography

Electronic detectors of X-rays have been a major improvement in the history of crystallography. Before, we needed to collect diffraction patterns on photographic paper and then manually process the data.

In the last years, the most widely used detector was the charge-coupled device (CCD) detector. In a CCD detector, X-rays are initially converted into visible light by interaction with a scintillating material such for exemple Gd_2O_2S:Tb. Visible light is then brought to a CCD device through an optic fiber and is translated into an electric signal. The rise of high flux beamlines has made CCD too slow, so they are currently being replaced by faster photon counting detectors, such as the Pilatus^{tm} (Henrich et al., 2009). These detectors directly count X-ray photons, without the need for a scintillator, that increases enormously the noise into the signal by spreading a spot over multiple adjacent pixels. Contrary to CCD detectors, photon counting detectors have a short readout time, in the order of $10ms$ and can be read continuously.

Radiation damage and cryo-cooling

The incident beam interacts in multiple ways with the proteins in the crystal (Murray et al., 2005). To be more precise, only 2% of the incident beam actually interact with the crystal, while about 98% of the photons simply pass through it. At the energies used in a synchrotron crystallography experiment, in the order of 12 keV, only 8% of the interacting photons are diffracted. Most of the interaction, accounting for about 84% of the interacting photons, comes from photoelectric effect. The liberated photo-electron is highly energetic, and can lead to hundreds of secondary ionization events, leading to the formation of radical species (Garman, 2010). The effects of this series of events can be classified as specific and global radiation damage.

Global radiation damage, accounts for the general loss in diffraction quality of the crystal, which is reflected in the loss of intensity of diffraction spots and in the increase of Wilson B factors and mosaicity.

X-rays can also cause specific damage to the protein. It has been observed that some chemical groups, like disulfide bonds and carboxylic groups of glutamates and aspartates, are more sensitive than others to radiation damage (Weik et al., 2000; Ravelli and McSweeney, 2000; Burmeister, 2000).

The simplest way to mitigate the effects of radiation damage in protein crystals is cryo-cooling. This method, developed in the 90's, consists in collecting data at cryogenic temperature, usually 100 K, instead of room temperature. Crystals are cryo-cooled by mounting them in a loop (Teng, 1990) and then submerging them into liquid nitrogen. Data collection is then performed at the synchrotron beam line under a flux of nitrogen gas at a temperature of 100 K. At this low temperature, radiation damage is reduced by a factor of about 100 compared to room temperature, due to the reduced mobility of radical species. For example, the only radical that can still move in amorphous ice at 110 K are electrons, while OH radicals are immobilized (Garman, 2010).

During the cooling process, crystals have to be protected from the effects of this rapid drop of their temperature. In fact, the water of the crystallization solution forms crystalline ice that both dominates diffraction at high resolution and damages crystals, reducing their diffraction quality. For this reason, a cryo-protection is required. It consists in soaking the crystal in a liquor that includes some additive that prevents the formation of crystalline ice. The most used cryoprotectant is glycerol, but other substances, such as small PEG molecules, have also this characteristic. The choice of a good cryoprotectant can dramatically change the diffraction power of a crystal and is thus a key step in obtaining good data (Juers and Matthews, 2004).

3.1.5 Data processing

Diffraction data sets consist in a series of images, each of them characterized by its signal intensity and by two coordinates. By processing of the data we mean all the treatment required to find the diffraction spots in this images and then translate them into a 3 dimensional diffraction pattern, that can later be used to generate electron density maps with an operation of Fourier transform. We will describe here the three main steps during a data processing pipeline. Each of the most used software, typically XDS (Kabsch, 2010) and MOSFLM (Leslie, 2006), performs them in different ways.

Indexing

The first step in data processing is the indexing, meaning giving each single diffraction peak a Miller index (hkl). Each spacegroup is characterized by specific symmetry of the diffraction patterns, along with missing reflections. A few images are needed to properly index the reflections. From a practical point of view, indexing software usually rank possible space groups with a certain penalty score. The good solution is then chosen as the one with the lowest penalty score and the highest symmetry.

Integration

After assigning a unique Miller index to each spot, in the integration step the amplitude of each reflection is determined. In particular, integration software sums the amplitude of each peak over adjacent pixels of the detector, distinguishing them from the background noise. Software like XDS also take into account crystal absorption and incident beam divergence effects.

Scaling

Even if theory states that symmetry-related spots must have the same amplitude (Friedel's law), this is not always the case after integration. Changes in beam intensity during data collection, crystal anisotropy and diffraction quality loss due to global radiation damage can cause this kind of problem. For this reason, a scaling step is needed. It consists in compensating the differences in amplitude between these reflections. The scaled amplitudes can then be used to compute the Fourier transform and reconstruct the protein structure.

3.1.6 The phase problem

As mentioned before, to obtain the electron densities we need calculate the inverse Fourier transform of the structure factor F(hkl). This requires the knowledge of diffracted amplitudes and phases. Now, we know that a detector measures intensities, rather than amplitudes. Amplitudes F are related to measured intensities I by the simple relation 3.6

$$I = |F|^2 \tag{3.6}$$

So we can see that the phase is not directly measurable, being $|e^{i\phi}|^2 = 1$. A number of methods have been developed to solve this problem, which is central in crystallography. Since phase information is lost by measuring intensities instead of amplitudes, a function, named Patterson function was defined. It is defined as the autoconvolution of the observed structure factors, and its peaks are related to the interatomic distances.

Isomorphous replacement methods (Green et al., 1954) consist in binding to the protein ligands with a high diffraction power, typically heavy atoms. Considering the number of different heavy atom complexes, we can distinguish between MIR (Multiple Isomorphous Replacement) and SIR (Single Isomorphous Replacement). The phasing is based on calculating the differences between the structure factors for the native and the derivative crystals. The heavy atom sites will then be easy to find on a Patterson map. The information on their position will then be used to phase the derivative crystal and, then, knowing their contribution to the structure factors, the native crystal.

Another family of methods is based on the use of anomalous scatterer (Hendrickson, 1991), *i.e.* atoms that absorb and fluoresce at a certain wavelength. These methods, based on the number of wavewlengths used in the experiments, are called MAD (Multiple Anomalous Dispersion) and SAD (Single Anomalous Dispersion). The role of the anomalous scatterers is to introduce a difference between centrosymmetric reflections, that will not obey any more the Friedel's law. This information is then used to retrieve the phases for the dataset. Anomalous diffusion is dependent on the X-ray wavelength, and for this reason it requires to collect multiple datasets at different energies.

The method of molecular replacement is based on obtaining initial phases from known analogues of the protein of interest, or at least of a part of it. In this work, all the crystallographic structures were solved by the molecular replacement method, using the program Phaser (McCoy et al., 2007). Molecular replacement is based on confronting the measured Patterson function with one other calculated from a protein on known structure, which should be similar to the one we are looking for. The method consists then in calculating the rotation and translation of the map that maximizes the superposition between the two Patterson functions (Rossmann and Blow, 1962). Phases are then retrieved from this initial model. The initial guess is usually good enough if there is at least a sequence identity of 25% between the

replacement model and the protein of interest.

3.1.7 Structure refinement and validation

Once the initial phase guess is obtained, we can compute an electron density map, in which we can fit a first model of our protein. Since the first electron density is based on the MR homologous, the model needs then to be refined in order to calculate new phases and improve the final electron density map.

This is usually an iterative process, where the protein atoms are fitted into the electron density and then the improved model is used to calculate the new phases to improve the density. Various refinement techniques exists, and they usually differ by how the model is modified. To asses if a model improves the structure or not, we usually consider the variation of the R-factor defined in 3.7, where F_o is the measured structure factor and F_c is the structure factor calculated from the model for the same reflection.

$$R = \frac{\sum_{reflections} |F_o - F_c|}{\sum_{reflections} |F_o|} \tag{3.7}$$

Another widely used criterion is the R-free (Brünger, 1992), defined as the R-factor, but considering only a subset of randomly chosen reflections, usually between 5% and 10% of the total number of reflections, which are then not included in the refinement. This value reduces the model bias implicitly included in the R factor.

3.2 Molecular dynamics simulations

Proteins are not static objects. They are animated by internal dynamics, which is fundamental for their functioning. The picture we get from crystallographic structures represents simply the most stable conformations of the protein, at a temperature, 100 K, which is far below the domain of biology. The approach of molecular dynamics simulation consists in solving the equation of motion (EOM) for each atom of the protein and of the solvent at room temperature (300 K), taking the crystallographic structure as a starting point. In this paragraph I will introduce the equation of motion of bound atoms in the most general way, and then show how we can use force fields to solve it in the case of protein simulations.

3.2.1 The equation of motion (EOM) of atoms in a protein

An atom, submitted to a force, moves following Newton's equation

$$F = m\frac{\partial^2 x}{\partial t^2} \tag{3.8}$$

This equation stands for only one atom. Depending on the kind of dependency the force has on the other variables, especially the spatial coordinates, it has an analytic solution.

A multibody problem, especially in the case where forces are dependent on the relative position of the atoms, is more complex and has in general no analytical solution. The most common strategy is to calculate a numerical, approximated solution. Most numeric integration algorithms work by calculating the value of the force at each time t, and use it to calculate the coordinates and velocities at the time $t + \Delta t$ If we make the assumption that the force stays constant between the time steps, which is equivalent to say that we are integrating time over steps Δt that are small enough, the equation of motion of each atom can be integrated using for example the Verlet algorithm.

$$\begin{aligned} x(t + \Delta t) &= x(t) + v(t)\Delta t + \frac{F(t)}{2m}\Delta t^2 \\ v(t + \Delta t) &= v(t) + \frac{F(t) + F(t + \Delta t)}{2m}\Delta t^2 \end{aligned} \tag{3.9}$$

This particular algorithm is interesting because it is the time reversible form of the most classical Euler integrator

$$\begin{aligned} x(t + \Delta t) &= x(t) + v(t)\Delta t + \frac{F(t)}{2m}\Delta t^2 \\ v(t + \Delta t) &= v(t) + \frac{F(t)}{m}\Delta t^2 \end{aligned} \tag{3.10}$$

It is easy to observe that, to use this kind of approach, we simply have to calculate the parameter $F(t)$ at each time step and for each atom. But we need to consider that $F(t)$ in reality depends on the relative positions and velocities, being therefore $F(x(t), v(t))$. This implies that the exact form of the force term is too heavy, depending, in a system of N atoms, on 6N parameters at least. For this

reason, approximate forms of the force, called force field, have been developed.

3.2.2 Force field parameters to integrate the EOM

A force field is a set of parameters that defines the forces that an atom of a particular kind, is subjected to. Considering that the force is the derivative of the potential along the spatial directions,

$$F = -\frac{\partial V}{\partial x_i} \tag{3.11}$$

the force field defines the potential rather than the force itself. The general form of a force field, equation 3.12, includes multiple parameters for the different kinds of forces acting on an atom in a protein.

$$\begin{aligned} V = & \sum_{ij}^{bonds} \frac{1}{2} k_{ij}(r_{ij} - r_{ij}^{eq})^2 \\ & + \sum_{ijk}^{angles} \frac{1}{2} k_{ij}(\theta_{ijk} - \theta_{ijk}^{eq})^2 \\ & + \sum_{ijkl}^{dihedrals} k_{\phi,ijkl}(1 + cos(n\phi_{ijkl} - \phi_0)) \\ & + \sum_{i>j} \left[\frac{q_i q_j}{\epsilon r_{ij}} + \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} \right] \end{aligned} \tag{3.12}$$

We assume here that the potential is defined as the sum of multiple, independent terms. The first three terms in the forcefield account for bonded interactions. Their name is due to the fact that these terms contain bond lengths, angles and dihedrals for bound atoms. This set of contributions describes movements such as vibrations of chemical bonds around their equilibrium position, angles and dihedral angles modifications and improper parameters, accounting for groups that can have movements other than the ones listed above. The last line of the force field includes the long range interaction, or non-bonded parameters. The main contribution to this term comes from electrostatic potential and van der Waals interaction. They are usually taken into account as a Lennard-Jones potential.

The general definition described above is valid for all of the force fields commonly used, but each of them uses different functional forms. In particular, equation 3.12

is based on the Amber force field functional form (Cornell et al., 1995). Practically, the force field contains, for each atom type, equilibrium values of bond, angle and dihedral angle and the constants for its functional form. Other common force fields like CHARMM (Brooks et al., 2009) or GROMOS, use a different functional form, as well as a different definition of parameters.

3.2.3 Solvent models

Proteins in a living system are in contact with water. To take this into account during MD simulations we include the protein into a solvation box. Ideally, we should use an infinite environment to calculate properties of the system without any bias, but this is computationally unfeasible. The use of a box with periodic boundary conditions overcomes this problem.

Multiple models of water have been developed during the years, to try to reproduce solvent properties in the most accurate way. They are divided in two main categories: implicit and explicit water models.

Implicit models consider the solvent as a continuous medium. They use a functional approach to calculate as a numerical value the influence of the environment on each atom, in particular the contribution of water to the free energy of the system. The two principal models are in this case the Poisson-Boltzmann and the Generalized Born (GB) model, which is a linear approximation of the first one. Both of these models are implemented in molecular dynamics software, like Gromacs, and the GB model is by far the most used one.

In the explicit solvent approach, we characterize each single water molecule within the system. This increases enormously the amount of atoms for which simulation is carried out, but also allows to describe more precisely the interactions between the protein and solvent. Different explicit water models have been developed over the years. The most simple are the three site models SPC, TIP3P (Jorgensen et al., 1983) and SPC-E (Berendsen et al., 1987), all of which describe water by its three atomic sites H-O-H. Bond parameter and partial charges on water atoms are slightly different in these three models. Four site models, such as the common TIP4P (Jorgensen et al., 1983) include a fourth dummy atom, whose role is to adjust the electrostatic potential around the molecule.

The chosen water model can influence the final output of a simulation, both

concerning bulk solvent properties and protein dynamics (van der Spoel et al., 1998; Tarek and Tobias, 2000). Using more accurate explicit models should be preferred, but their computational cost is heavy. In any case, the access to massively parallel computational resources should make the explicit solvent possible for any length of simulation.

3.3 Molecular docking

One aim of structural biology is to participate in the design of new drugs. To achieve this goal, techniques of molecular docking, i.e. a large spectrum of methodologies to predict the interactions between molecules are of great help. (Sousa et al., 2013). In the framework of this thesis work, we are interested in the interaction between proteins and ligands.

3.3.1 Principles

In a docking experiment, we look for the best conformation, or binding mode, of a ligand inside a target protein, which is called the receptor. The docking program will look for candidate binding modes using different algorithms, which explore the conformational space of the ligand-receptor system and extracts various poses, usually hundreds or thousands. These poses are then scored using a function, which is typical for each program and is probably the most important feature of a docking pipeline. Finally, all the poses are ranked by their scoring, to extract the most probable one, which should also describe the optimal binding mode for the ligand within the receptor.

3.3.2 Docking algorithms

The algorithms used by docking programs follow different methods (Lengauer and Rarey, 1996), depending on the required exhaustiveness and on the use of computational resources.

The method of shape complementarity consists in a calculation of the affinity between the ligand and the protein based on the mutual interaction between their surfaces. In particular, this approach looks at the interaction between the calculated

electrostatic fields of the two partners. This docking method is the simplest one, but has lower accuracy and is also less suited to give flexibility to the ligand, and even less to use a flexible receptor.

The second, most complex method is the simulation. It consists in simulating the binding process by calculating the mutual interaction between the atoms of the receptor and the ligand. For its functioning, it is almost costless to take flexibility into account with this method. The main drawback is the computational cost of exploring a larger energy landscape. But modern optimized algorithms and the always increasing computing power have made simulation docking a routine.

Finally, two of the most commonly used programs, GOLD (Jones et al., 1997) and AutoDock (Goodsell et al., 1996), use a genetic algorithm. The method in this case is to evolve a random initial conformation, usually for both the ligand and the receptor, trying to optimize some variable. This method leads to the most exhaustive exploration of the energy landscape of the complex.

3.3.3 Scoring function

Scoring functions (Jain, 2006) have the role to rank the number of docking poses generated by the search algorithm. To perform this task, a scoring function needs to handle the 3-dimensional structure of a protein, to analyze the receptor, the shape of the ligand, and the binding mode between the two. The methods used to score docking poses fall in three main families.

Force fields can be used, as in molecular dynamics simulations, to estimate the affinity between protein and ligand. The use of force fields provides a direct estimation of the binding energy, because of their ability to calculate forces between atom pairs of ligand and receptors. Using force fields, we can also calculate the energy contribution due to internal strains both within ligand and receptor, that could give a hint on the relative probability to find them in a particular conformation. Finally they can also be used to calculate desolvation energy using the same models of implicit solvent of molecular dynamics simulations.

Empirical scoring functions are simpler, and they work by counting the number of interactions between molecules (Böhm, 1998). Some interactions add points to the final score, like the number of possible hydrogen bonds in every pose, or the hydrophobic/hydrophilic interactions between ligand and receptor. When the ligand

is considered as being flexible, a penalty score will be given also to each mobile rotamer that is fixed in a non-optimal conformation.

Knowledge-based functions, or statistical potentials, are the last type of scoring functions (Muegge and Martin, 1999; Muegge, 2006). They start from the observation of most occurring interactions in large databases, such as the PDB, to derive a potential of mean force, which describes each interaction of interest.

3.3.4 Flexible docking

Docking programs usually try to dock a series of conformations of a ligand, for which the rotamers are modified, into the binding site of a rigid receptor. This assumption is however too stringent, and a few methods exist to take intrinsic receptor flexibility into account. The large number of degrees of freedom of receptors is rapidly too computationally heavy, so the flexibility of the receptor is reduced to that of a few amino acids of the predicted binding site. Even in this case, to speed up calculations, approximate and faster methods are needed. The most common method consists in using rotamer libraries, that the program will explore to generate a large amount of conformations for the receptor.

Flexible receptors can also be derived from experimental techniques. NMR ensembles are an example of protein structure ensembles that have successfully been used as flexible receptors in docking (Totrov and Abagyan, 2008). More recently, hidden conformations in crystallographic structures have also been used to generate multiple docking receptors with different conformations (Fischer et al., 2014).

Other methods rely on molecular dynamics simulations to generate receptors, by extracting snapshots from a simulation (Wong et al., 2005). But this requires the production of lengthy and computationally consuming simulations. Recent development of faster graphic processor-based simulation algorithms, along with the use of implicit solvent models has partially solved this issue, but it is still uncommon to use MD snapshots as docking receptors.

Chapter 4

Optimization of KM297, a neutral AChE reactivator

The main goal of this thesis work was to participate in the design of bifuncitonal reactivators of OP-inhibited AChE. To this end, we investigated the structural dynamics of AChE during ligand binding and solved structures of native and OP inhibited AChE in complex with the newly designed reactivators. A bifunctional reactivator (figure 4.1), is a molecule featuring an oxime and a PAS ligand, connected by a flexible linker, usually formed by carbon atoms.

Starting from a previously known PAS binding molecule, we designed a series of bifunctional reactivators, that were then screened *in vitro* to find a lead molecule (section 4.1), with the best reactivation capability. This lead was characterized by flexible docking, to analyze its possible binding modes, and molecular dynamics, to define an optimal linker length (sections 4.2 and 4.3). By solving the crystal structure of *Tc*AChE, both in its native form and inhibited by an OP, in complex

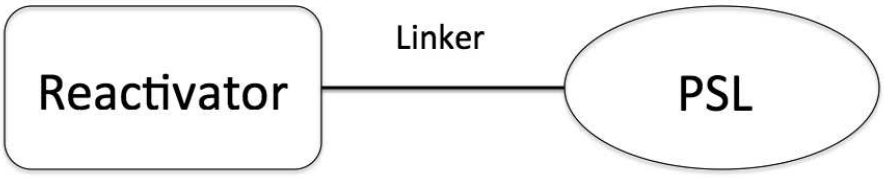


Figure 4.1: Schematic representation of a bifunctional reactivator. An oxime function is connected to a PAS ligand (PSL) through a flexible linker.

with our lead molecule we found that the chosen PAS ligand could also bind in the active site (section 4.4). Therefore we modified the lead, in order to force it to bind exclusively at the PAS. The family of modified derivatives designed to this end was also characterized by flexible docking and by solution of the crystal structures of their complexes with AChE (section 4.5).

All the crystallographic structures were solved using *Tc*AChE, but since our reactivators are aiming at reactivating AChE in humans, human AChE would be a more appropriate choice. We managed to obtain crystals of recombinant hAChE in the absence of fasciculin (section 4.6), but their diffraction quality was lower than that of *Tc*AChE crystals.

4.1 Design, synthesis and *in vitro* evaluation of the reactivator KM297

In the development of an efficient bifunctional reactivator, knowledge of the structure of the enzyme and of the binding modes of its ligands is fundamental. Not only can ligands bind to the main conformation of the enzyme, determined by X-ray crystallography, but also to alternative conformations of binding sites, as introduced in chapters 1 and 2. Rearrangements of the PAS region upon ligand binding to AChE were observed in crystallographic structures (Bourne et al., 2004; Colletier et al., 2006b; Rydberg et al., 2006).

Tetrahydroacridine (tacrine), known to bind at the active site of AChE (Harel et al., 1993), has also been incorporated into bifunctional inhibitors (Harel et al., 2000; Bourne et al., 2010a). The crystal structures of AChE in complex with ligands bearing a bis-tacrine function (Rydberg et al., 2006; Colletier et al., 2006b) revealed that tacrine can also bind to an alternative conformation of the PAS Trp279 in *Tc*AChE (figure 4.2). The ability to bind at the PAS is a key feature for a bifunctional reactivator, and we decided to exploit it.

A series of molecules, all featuring a tacrine moiety as a PAS ligand, was synthesized by the joint effort of our collaborators at the Université de Strasbourg and Université de Rouen. These molecules were tested *in vitro* and we found that KM297 (figure 4.3), featuring a 3-hydroxy-2-pyridine aldoxime as reactivating function and a linker length of 4 carbons, was the most efficient in the *in vitro* reactivation tests

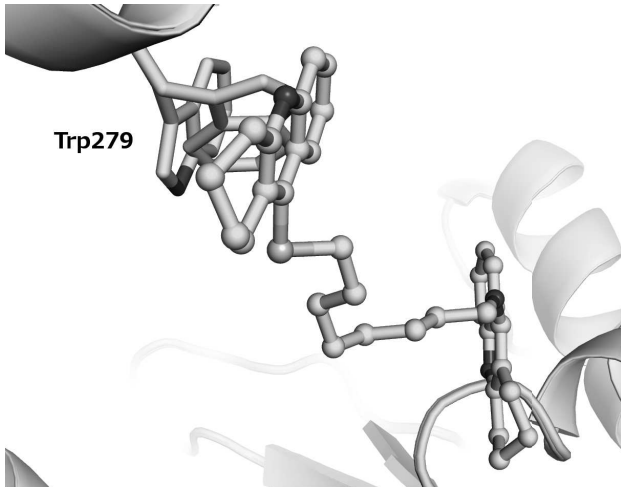


Figure 4.2: Crystal structure of *Tc*AChE in complex with the bis-tacrine inhibitor NF595 shown in cyan (PDB acces code 2cek) superposed to the structure of native *Tc*AChE (PDB acces code 1ea5), shown in green. We can observe the rotation of the PAS residue Trp279 around the χ_1 angle, to allow the binding of the inhibitor.

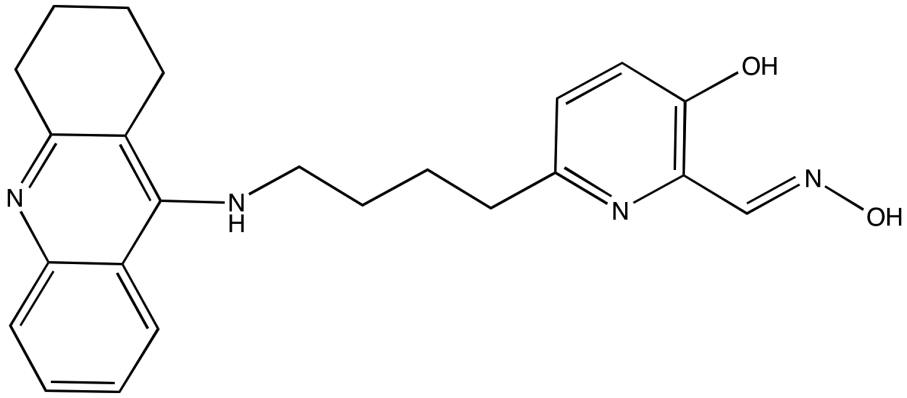


Figure 4.3: Lead molecule KM297.

(Kliachyna et al., 2014).

In vitro results showed that KM297 can reactivate VX-inhibited hAChE more efficiently than any available reactivator, with also a good efficiency against tabun and paraoxon. Its bimolecular reactivation constant (table 4.1) against VX is twice that of obidoxime and about 3 fold that of HI-6, the best reactivator available so far. Against tabun and paraoxon, for which HI-6 has no measurable reactivation capacity, KM297 is respectively 20 and 1.5 fold more efficient than obidoxime. Reactivation

Reactivator	VX	Tabun	Paraoxon
2-PAM	0.28	nd	nd
HI-6	9	nd	nd
Obidoxime	11	0.16	20
KM297	22	3	31

Table 4.1: Comparison of the bimolecular reactivation constants of different reactivators for hAChE inhibited by VX, tabun and paraoxon. All values are expressed in $mM^{-1}min^{-1}$.

efficiency tests, not reported here, have also shown that the homologue with a 5 carbon linker has a lower *in vitro* reactivation power than KM297, with its 4 carbon linker.

Due to its high efficiency in reactivating several OP inhibited hAChE, KM297 was chosen as the lead molecule of this family.

4.2 Flexible docking of KM297 into hAChE-VX

We performed flexible docking experiments of the reactivator KM297 to the non-aged form of VX-inhibited hAChE. These experiments were aimed at verifying that KM297 could bind to OP-inhibited hAChE, and the choice of flexible docking was due to the fact that tacrine can bind to alternative conformations of the PAS, as observed by X-ray crystallography for NF595 (Colletier et al., 2006b). Finally, docked conformations can help in elucidating the mechanism of reactivation, and were used as a starting point to perform molecular dynamics simulations of the complex hAChE-VX-KM297. The choice of VX was made due to the necessity to characterize in detail how KM297 can reactivate so well hAChE inhibited by this particular OP, and also because VX is considered the most dangerous of the OPs, due to its high toxicity and persistence.

4.2.1 Material and methods

Calculations were performed using AutoDock Vina (Trott and Olson, 2010). The receptor was prepared by superposition of native hAChE crystal structure (Cheung et al. (2012) PDB acces code 4ey4), and VX-inhibited non aged mouse AChE (PDB

acces code 2y2u). The active site of the inhibited mouse enzyme was used as a reference to model the receptor active site in the human enzyme. The choice of mAChE as a template is justified by the identity of the residues within the gorge of mAChE and hAChE.

We used as flexible residues the closest to the supposed position of the ligand into the enzyme, Tyr72[1], Asp74, Trp86, Tyr124, Ser125, Trp286, Tyr337, Phe338, Tyr341, along with the ethyl group of the non-aged VX. These residues were provided as a list to the MGL AutoDockTools (Morris et al., 2009) script that generates flexible receptors from a PDB file.

Ligands were designed and optimized by processing their SMILES string (Weininger, 1988) with phenix.eLBOW (Moriarty et al., 2009). This program optimized the ligand's partial charges and protonation state, writing the output to a PDB file, which was then pre processed with the appropriate script from the AutoDockTools set. The procedure has been automated with bash scripts to screen ligands with different characteristics simultaneously. Docking experiments were performed following the Vina docking protocol.

A parallelepipedal docking box of about 17200 $Å^3$, smaller than the recommended upper limit of 27000 $Å^3$, was centered at the middle of the gorge, at the level of the bottleneck residues, so as to consider the whole gorge volume as the docking region and include all the side chains of the flexible residues listed above. We chose to generate nine output conformers, that were then observed qualitatively using molecular graphics programs like PyMOL (Schrödinger, 2010).

4.2.2 Ligand scoring and conformational analysis

Vina scores obtained for the best 9 conformers of ligand KM297 are shown in table 4.2. As stated by the Vina documentation, we can conclude that KM297, having an estimated affinity in the sub-nanomolar range, is a very good ligand for the enzyme. Unfortunately, this calculated affinity values did not allow to properly discriminate between docked conformations, due to the fact that they are all of the same order of magnitude. A more detailed observation of the output structures is required in this case.

The interest in performing flexible docking experiments lies in the possibility

[1] hAChE numbering.

Conformation	Affinity(kcal/mol)	Affinity(Mol)
1	-11.7	3.36E-10
2	-11.4	5.88E-10
3	-11.3	7.09E-10
4	-11.3	7.09E-10
5	-11.2	8.53E-10
6	-11.2	8.53E-10
7	-11.1	1.03E-09
8	-11.1	1.03E-09
9	-10.9	1.49E-09

Table 4.2: Affinities for the nine best conformers of KM297 docked into non-aged hAChE-VX complex.

of finding affinities for alternative conformations of the enzyme. In particular, we wanted to verify that our working hypothesis of binding to a non-native conformation of the tryptophan at the PAS was supported by the calculations. Among the nine output conformations, six presented the enzyme in its native conformation, while three featured a conformational rearrangement of the PAS. Number 4 was particularly interesting, because it features an aromatic stacking between Trp286[2] and Tyr72, similar to the crystal structure of the *Tc*AChE-NF595 complex (PDB acces code 2cek). As we can observe in the insert of figure 4.4, Trp286 also rotates by about 90 degrees around the χ_1 angle from its native conformation, but in the opposite direction that in the 2cek structure, thus laying in the same plane. This generates a very similar binding mode for the tacrine moiety in the two conformations, also due to the fact that the very stable residue Tyr72 (Tyr70 in *Tc*AChE) is not observed to move, neither in the crystal structure nor in the docking results, being therefore parallel to Trp286. The only difference arising from the opposite rotation is that the tacrine moiety is flipped so that its double ring faces each aromatic residues and forms the appropriate aromatic stacking. The flexible docking tests confirmed our hypothesis of tacrine moiety of KM297 binds at the PAS of VX-inhibited AChE, and in particular to a non-native conformation of Trp286.

[2]hAChE numbering.

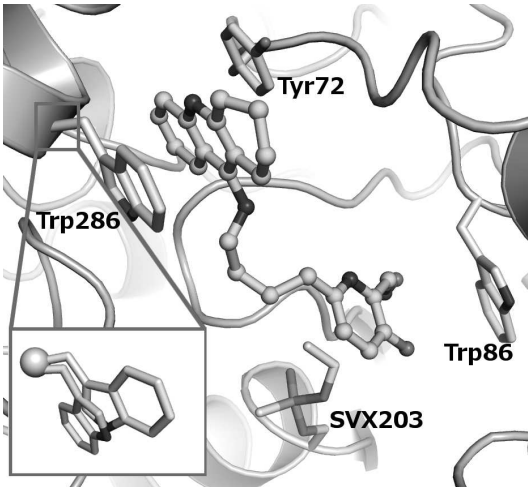


Figure 4.4: Docked conformation number four of KM297 within the gorge of VX-inhibited hAChE. We can observe the tacrine moiety at the PAS stacked in between Tyr72 and Trp286 in an alternative conformation. In the insert we show a superposition of the conformation of Trp286 in our output (grey) and the one observed for Trp279 in the crystal structure of the *Tc*AChE-NF595 complex (cyan). We observe that they rotate in two opposite directions, but lay in the same plane for the aromatic stacking of the tacrine moiety.

4.3 Simulation of hAChE-VX-KM297

We simulated a hAChE-VX-KM297 complex observed from docking experiments to get useful information for ligand optimization. In particular, we determined that the observed binding mode was effective for stabilizing the ligand and we characterized the influence of the carbon linker length on the distance between the oxime oxygen and the OP phosphorus. This process required several steps, from the calculation of force field parameters for the modified catalytic serine to the analysis of reactivation conditions in MD simulations. In this section we will discuss the details and results of this set of simulations.

4.3.1 Calculation of force field parameters for OP-inhibited AChE

As explained in Chapter 3, we need force field parameters describing each residue if we want to perform molecular dynamics simulations. Force fields are capable of

handling all principal native 20 amino-acids, and there are simple methods to include small ligand molecules in the simulation system. However, OPs bind covalently to the catalytic serine of AChE, creating a new phosphylated residue, which needs to be defined in the force field. For this reason, we calculated force field parameters for phosphylserines resulting from the reaction with the most common organophosphate nerve agents. Inhibited-serine models were extracted from the PDB, since there are a few structures of non-aged AChE displaying the conformations of OPs bound to the catalytic serine. The phosphylated serine of interest was capped by methylamine at the C-terminus and acetyl group at the N-terminus and minimized using Chimera (Pettersen et al., 2004). Two main chain conformations of the capped residue, the α-helix and β-sheet, were taken into account for the RESP (Restrained ElectroStatic Potential) charges (Cornell et al., 1995) calculation for the system using the REDS-2 server (Vanquelef et al., 2011). Following the instruction of the server maintainer, calculations were performed using Gaussian09 with fixed side chain angles and using all the options compatible with the parm99 force field set. After obtaining the partial charges of the residues, they were included into the Amber99sb force field, with the proper form of the bonded and non-bonded parameters for each atom. The missing parameters were actually defined into the General Amber Force Field (GAFF) (Wang et al., 2004), wich was created for the purpose of describing small organic molecules compatible with force fields of the Amber family. GAFF contained all the bond lengths, angles and dihedral constant for the non-standard serine-OP conjugates, such as the dihedral angle C-N-P-O in the serine-tabun adduct.

In order to verify the validity of our parameter set, it was tested by running molecular dynamics simulations of inhibited AChE. We first modified the Amber99sb force field description files to include our newly defined residues. Then, the catalytic serine of native human AChE was modified accordingly, as well as the conformation of His447, in order to match the conformational changes observed in the structures of mAChE or *Tc*AChE, when available. All calculations were performed using Gromacs4.5 (Hess et al., 2008) and the experimental protocol described in section (4.3.2). The dynamics of the system was then simulated for 5 ns to verify the stability of the constructed residue. Having a force field description for the VX-inhibited serine means that we can simulate how KM297 binds to AChE inhibited by VX.

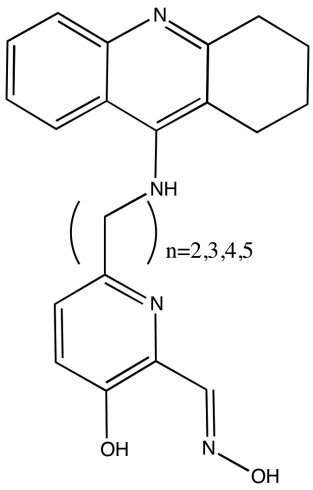


Figure 4.5: General structures of molecules of the KM family. The number of carbon atoms in the linker (n) was varied between two and five.

4.3.2 Optimization of the ligand linker length

We involved molecular dynamics simulations in the optimization process of KM297 to improve our knowledge of the reactivation mechanism. The goal was to better understand the mechanism of interaction between the oxime and the inhibited enzyme. We decided to simulate several KM297 derivatives (figure 4.5), differing in the carbon linker length, to observe how this parameter influences the distance between the oxime oxygen and the phosphor of VX. The evolution of this distance along the simulation was considered as being directly related to the reactivation properties of the molecule.

Methods

The model structure of the non-aged hAChE-VX complex in the conformation resulting from flexible docking was simulated using the amber99sb force field, with the VX parameters calculated as described above. The constructed protein model was included in a cubic box of $10 \times 10 \times 10$ nm^3, with periodic boundary conditions. The KM-ligand topology was prepared using acpype (Sousa da Silva and Vranken, 2012; Wang et al., 2006), to be included in the simulation system. The ligand in docked conformation number 4 was added to the system, which was then solvated in simple point charge water. After charge equilibration with the appropriate number

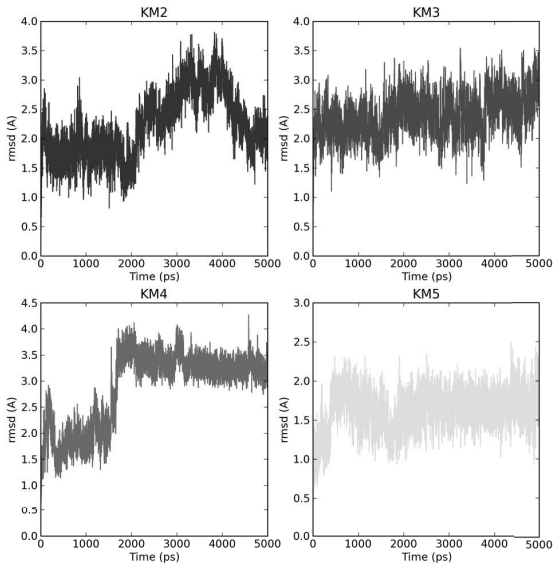


Figure 4.6: RMSD of bifunctional KM reactivators within the gorge of VX-inhibited hAChE calculated from a 5 ns MD simulation. The four molecules have linker lengths of 2, 3, 4 and 5 carbons respectively.

of counter ions, the energy was minimized. The temperature was kept at 300 K by coupling to a Berendsen thermostat at 300 K with a coupling time of $\tau T = 0.1$ ps (Berendsen et al., 1984). Pressure was maintained at 1 bar by coupling to a Berendsen barostat with compressibility of 4.5×10^{-5} bar and $\tau P = 1.0$ ps. Bond lengths were restrained with the LINCS method (Hess et al., 1997) and the electrostatic interactions were calculated using the particle-mesh Ewald algorithm (Essmann et al., 1995; Darden et al., 1993). The dynamics of the system was then simulated for 5 ns and output coordinates were stored every picosecond. Stability of the ligand conformation was checked by calculating its RMSD with the program g_rmsd, fitting at each timestep the positions of protein $C\alpha$. Distances were monitored with the program g_distance, defining the OP phosphor atom and the oxime oxygen as two separated groups. Data were then processed and plotted using python modules matplotlib and Numpy.

Results and discussion

The RMSD of the ligand position within the enzyme was calculated, after having aligned at each MD simulation step the protein $C\alpha$ positions. The alignment of the protein backbone $C\alpha$ atoms means that the RMSD value we are calculating is directly proportional to the movement of the ligand within the gorge of the enzyme. Results showed that all of the ligands remain near their initial position during the simulation (figure 4.6), because the average RMSD for each of the molecules is of the order of 2 Å. This means that tacrine, the PAS ligand moiety of the bifunctional reactivator, is adapted to stabilize KM297 and its analogues in the docked conformation. The molecule with a 4-carbon linker produced a steep variation of RMSD after 2 ns, but visual observation of the trajectory revealed that it was due to a movement of the oxime moiety, and that the PAS ligand remained in its initial position all along the simulation. The stability of the PAS conformation also confirms that the distance between the VX phosphor atom and the oxime oxygen is a good parameter for investigating the movements of the linker and the oxime within the gorge of the enzyme.

After measuring the distances at each 1 ps step of the simulations, we plotted their distributions in a histogram with bins of 1 Å (figure 4.7). Average distances reported for each molecule in table 4.3 are smaller for the 4 and 5 carbon linker molecules. It is straightforward to conclude that the two longer linkers better fit the spatial dimension of the gorge, and are, therefore, more efficient than the shorter ones in positioning the oxime near the OP.

The reactivation event is a chemical process that can occur only if the two atoms get close enough. The minimal distance between phosphor and oxygen observed during the simulation trajectory is, therefore, a direct indicator of a possible reactivation event in our set of simulations and it was used to discriminate which of the linkers was the most appropriate. The 4-carbon linker molecule is the one that brings the oxime group closest to the phosphor atom, with a minimum distance of 3.5 Å, which is also close to the limit achievable in an MD simulation (the sum of the van der Waals radii of oxygen and phosphorus is 1.8 Å+1.5 Å=3.3 Å).

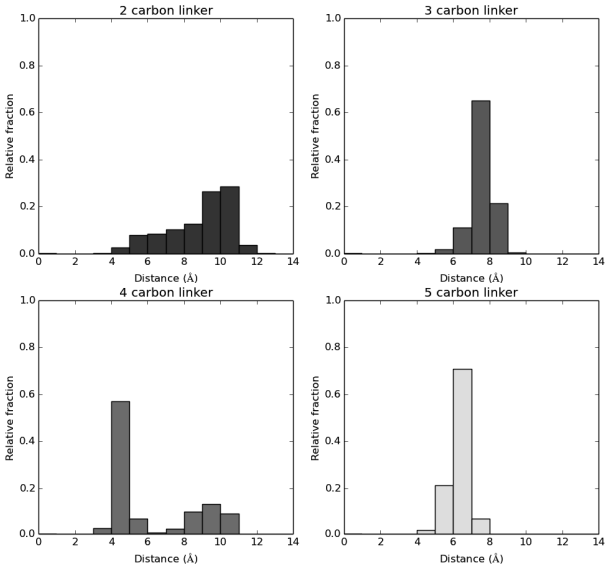


Figure 4.7: Distribution of the distances between the oxime oxygen and phosphor atom of VX calculated from a 5 ns molecular dynamics simulation for KM297 analogs with several linker lengths. We can observe how the distribution for the 4-carbon linker molecule is the only one showing a sharp peak at a shortest distance. It is also the data set where the closest distance, between 3 and 4 Å, is reached more often than in the other.

Carbon linker length	Average distance(Å)
2	8.82
3	7.58
4	6.15
5	6.33

Table 4.3: Average distances between the oxime oxygen and the phosphor atom of VX for different carbon linker lengths. Four and 5 carbon linkers have the lowest value, which is presumably more favorable to the reactivation reaction.

4.4 X-ray crystallography

A series of crystallographic experiments was performed with the aim of structurally characterizing the interactions between AChE and the reactivators of the KM family. Using hAChE for structural studies would have been ideal. Unfortunately, the form

of hAChE available in our laboratory (untruncated recombinant form) crystallizes only in the presence of FAS-2, an inhibitor that occupies the PAS and blocks the access to the gorge. The presence of FAS-2 is therefore an obstacle for ligand soaking experiments. Since good quality crystals of truncated recombinant hAChE (Cheung et al., 2012) were not yet available to us, we decided to proceed using *Tc*AChE.

*Tc*AChE yields well diffracting crystals. Also, the crystal packing leaves the PAS and gorge freely accessible, making it possible for ligands to enter the active site during soaking experiments. For these reasons, AChE from this species is a common choice for performing crystallographic experiments. The main goal of this structural study was to obtain a crystallographic structure of ternary *Tc*AChE-OP-reactivator complexes, in order to determine, from a structural point of view, how our reactivators interact with the non-aged form of the AChE-OP complex. We chose tabun instead of VX as a nerve agent for these crystallographic experiments. The reason for this choice is that a tabun-inhibited hAChE is reactivated about 10 times slower by KM297 than VX- or paraoxon-inhibited hAChE, as observed *in vitro* (Kliachyna et al., 2014). Thus, the probability that reactivation may occur inside the crystalline enzyme during the soaking time is reduced, and the ternary complex should accumulate and be easier to observe. Slow aging of tabun-AChE was also a favorable consideration. Since we do not know how the crystal environment would affect reactivation constants of KM297, a slow aging OP allows us to soak ligands for several hours, without significant aging. We first solved the structure of the non aged *Tc*AChE-tabun complex, and used it later as the starting structure for the refinement of the ternary complexe.

4.4.1 Material and methods

*Tc*AChE was provided by Prof. I. Silman's laboratory at the Weizmann Institute of Science (Israel). It is purified from the electric organs of *Torpedo californica* fish according to the protocol presented in Sussman et al. (1988). The crystallization condition used in the experiment work at 4°C, so all the manipulations have been performed in a cold room. Crystallization experiments were done using the hanging drop method with drops of 1 μl of precipitant and 1 μl of protein, at a concentration between 10 and 15 mg/ml. *Tc*AChE is known to crystallize in 36% PEG200, 150 mM MES pH 5.6. In any case, the percentage of PEG200 was refined by screen-

ing for each new purification batch. We generally obtain crystals for concentrations spanning from 30 to 40% PEG200. They usually grew in 2-10 days and can be found in three different crystal forms. The two most commonly obtained are trigonal[3] and orthorhombic-1[4] ones. The first appear as cubes of a few hundred micrometers in dimension, and diffract up to 1.8Å while the latter grow in the form of multiple, superposed plates and they diffract to about 2 Å. The latter needed to be separated before flash-cooling, in order to reduce the chances of hitting multiple crystals when collecting a data set. A third crystal form is also orthorhombic[5], but its diffraction quality is significantly lower (2.4 Å resolution). The active site gorge is not accessible, being closed by a symmetry-related molecule.

Crystal soaking

Crystal soakings were performed by fishing a crystal and putting it into a drop of the mother liquor, containing the ligand we want to study in a concentration depending on its solubility and its affinity.

To study OPs and their surrogates, *Tc*AChE crystals were soaked in a 2 μl drop of the mother liquor containing 0.5 mM of OP. The concentration and volume both allow manipulating only small amounts of the toxic OP, while also minimizing the quantity of isopropanol, in which the OP is stored, in the soaking conditions. We performed experiments mainly with sarin, VX and tabun, with soaking times spanning from 5 minutes to 10 days, depending on the need to trap either the non-aged or the aged form. Half time of aging in solution is 3 hours for sarin, 27 hours for tabun and 36 hours for VX. For tabun, complete aging in solution requires soaking times larger than 48 hours. The pH of soaking (5.6) lower than at physiological conditions and spontaneous hydrolysis of the organophosphate in the crystallization conditions suggested reducing the soaking time to a few minutes to trap the structure of the non-aged complex. In any case, due to the very high bimolecular reaction rates, inhibition *in crystallo* is mainly rate-limited by the diffusion of the ligand into the solvent channels.

The soaking of neutral reactivators presents more difficulties. Due to their lack of permanent charge, these molecules are weakly soluble in water (a few hundred

[3]Space group 152 (P3121) , 112 112 138 Å, 90° 90° 120°
[4]Space Group 19 (P212121), 91 107 151 Å, 90° 90° 90°
[5]Space Group 19 (P212121), 80 112 164 Å, 90° 90° 90°

μM). We managed to solubilize KM297 and its homologs at a concentration of 20 mM in 100% dimethylsulfoxide (DMSO) at room temperature. DMSO deteriorates crystal quality, so we needed to screen for optimal soaking conditions in the presence of this solvent. Our optimized conditions were found by testing dilutions between 1 and 20% of DMSO and collecting diffraction data sets on soaked crystals. We obtained the best results by mixing 5μl of reactivator-DMSO solution with 95μl of mother liquor, to obtain a final concentration of 1 mM reactivator with 5% DMSO. Crystals were soaked in these conditions for a time spanning between 2 minutes and 2 days.

For non-aged ternary complexes, crystals were at first soaked for a few hours with tabun, according to the protocol described above. Then they were fished again and soaked for 2-20 minutes into a second solution, containing mother liquor and 1 mM of reactivator and 0.5 mM tabun, with 5% DMSO and 5% isopropanol. The aim of adding tabun is to re-inhibit the enzyme that was eventually reactivated inside the crystals.

Ligand co-crystallization

An alternative to soaking for solving structures of proteins in complex with ligands is the co-crystallization method. It consists in growing crystals in a drop of mother liquor containing the ligand. In this way, the ligand ideally binds to the enzyme before crystallization actually occurs.

We performed co-crystallization experiments of *Tc*AChE in complex with the reactivators KM297 and its homolog molecule JDS207. Following the crystallization protocol described above, hanging drops were prepared with $1 - 1.5$ μl of protein at a concentration of 10 to 15 mg/ml and the same volume of precipitant solution containing 1 mM of reactivator and 5% DMSO.

Data collection and processing

Owing to the cryo-protectantxs power of PEG200 at concentrations higher than 20%, *Tc*AChE crystals can be cryo-cooled without any additional cryoprotectant. Diffraction data were collected at the ESRF, at the MX beam lines ID14-4, ID23 and ID29. Each diffraction image is collected by rotating the crystal during exposure. In our case, we collected either with an oscillation range of 1° per image, or using

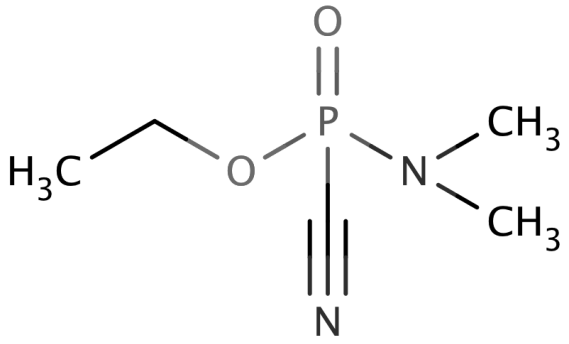


Figure 4.8: Nerve agent tabun (GA)

the fine slicing collection method, in which images are collected with an oscillation range of 0.1° instead of 1°. For the orthorhombic space group, images should be collected over at least 90° of rotation, in order to obtain a complete data set. The higher symmetry of the trigonal space group requires, in theory, only 60°. Usually more images than the minimum required by the theory are collected, to maximize the chances of obtaining a complete data set. These days, synchrotron beam lines offer an automatic data processing routine that was used to quickly inspect data sets and check the presence of ligands in the structures. Along with using auto-processing results, most data sets were also reprocessed by us using XDS (Kabsch, 2010). Knowing the crystal form, we processed directly in the proper space group. Once the reflection data set was scaled and truncated at a proper resolution using the AIMLESS-POINTLESS pipeline in CCP4, the native *Tc*AChE monomer (PDB acces code 1EA5) was used as the initial model to perform molecular replacement for trigonal crystals. The native *Tc*AChE dimer (PDB acces code 1w75) was used for the orthorhombic crystals. Molecular replacement was done using Phaser (McCoy et al., 2007). Finally, refinement was performed with either PHENIX (Adams et al., 2010) or REFMAC (Murshudov et al., 2011). Coot (Emsley et al., 2010) was employed to observe and manipulate the models.

Data	
Resolution range (Å)	45.74 - 2.2 (2.28 - 2.2)
Space group	P 31 2 1
Unit cell (a,b,c (Å); α, β, γ, (°))	111.8 111.8 137.2 90 90 120
Total reflections	274775 (27951)
Unique reflections	50213 (5024)
Multiplicity	5.5 (5.6)
Completeness (%)	98.84 (98.55)
Mean I/sigma(I)	13.7 (2.20)
CC 1/2	0.999 (0.918)
Refinement	
R-work (%)	17.5 (32.7)
R-free (%)	21.6 (33.7)
RMS bond lengths (Å)	0.008
RMS bond angles (°)	1.08
Atoms in asymmetric unit	4784
Ramachandran favored (%)	96
Ramachandran outliers (%)	0.19
Average B-factor (Å^2)	42.00

Table 4.4: Refinement and data statistics for *Tc*AChE-tabun conjugate. Values in parenthesis correspond to the highest resolution shell.

4.4.2 Crystallographic structure determination of the non-aged *Tc*AChE-tabun conjugate.

Of all the OP nerve agent-AChE conjugates, AChE-tabun is the most resistant to reactivation, while having a relatively long half-life in the non-aged form (27 h, Worek et al. (2004)). It was thus the OP of choice as a starting point in the crystallographic study of reactivators because of its low reactivability by oximes, which means that we don't expect reactivation of tabun-inhibited enzyme to happen within the crystal. Along with this, aging time longer than the soaking experiment, ensures that no significant aging occurs over the experiment time scale. The structure of the non-aged AChE-tabun was solved with mouse AChE(Carletti et al., 2008), but the *Tc*AChE-tabun structure remained unsolved.

Data collection and processing

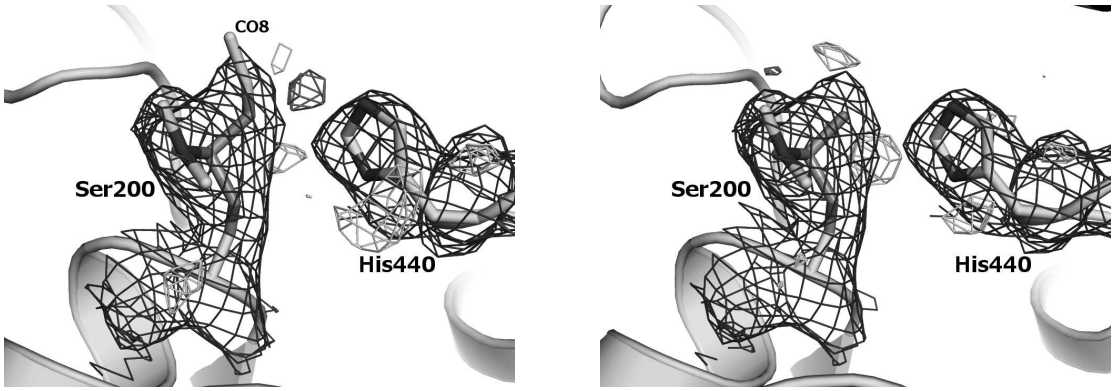
Trigonal *Tc*AChE crystals were soaked for 5 minutes in a solution containing 36% PEG200, 100 mM MES pH 5.6, 0.5 mM tabun and 1% isopropanol. Data were collected on the ID14-4 beam line at the ESRF. We collected 90 images with oscillation range of 1° per image at a temperature of 100 K. Data were processed with XDS (Kabsch, 2010) using the trigonal space group unit cell parameters and were truncated using XSCALE at a resolution of 2.2 Å. Complete statistics for data processing and refinement are shown in table 4.4.

Structure refinement

Refinement was performed using phenix.refine (Adams et al., 2010). Atom positions were refined, along with individual atomic B-factors. Restraints for the tabun substituent were calculated with phenix.eLBOW (Moriarty et al., 2009), and the length and angles for the binding to the catalytic serine were obtained from the output of quantum mechanical calculations previously used for the force field parameter calculation (section 4.3.1), performed on the REDS server (Vanquelef et al., 2011). Water molecules were added, using the automatic water search in Coot (Emsley et al., 2010) and a comparison with the ones observed in the high resolution native *Tc*AChE structure 1ea5. Water molecules were then validated with MolProbity (Chen et al., 2010). After several cycles of restrained refinement and manual reconstruction in Coot, refinement statistics, converged (table 4.4).

Structure analysis

After molecular replacement, a strong signal at about 11σ was present in the Fo-Fc electron density map in front of the catalytic serine. This signal is the key signature of the presence of the organophospate within the structure, as it is in agreement with the presence of a phosphorus atom at covalent bonding distance from the Ser200-OG atom. The tabun moiety was subsequently modeled in the active site as shown in figure 4.9. The electron density map around the position of the ethyl group of tabun is not clearly defined in the non-aged *Tc*AChE-tabun complex (figure 4.9 a). Considering that aging of tabun involves the loss of this group, the structure has also been refined in the aged form. However, a peak appeared in the Fo-Fc electron



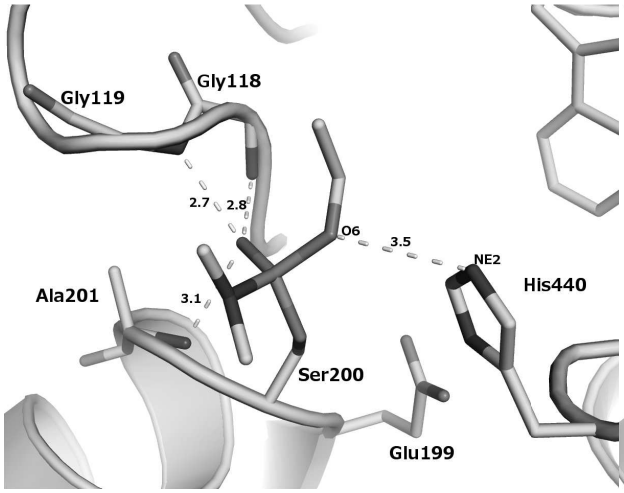
(a) non-aged complex (b) aged complex

Figure 4.9: Electron density map of the catalytic serine inhibited by tabun either in (a) non-aged and (b) aged form. Blue mesh represents the 2Fo-Fc map at 1σ level and green and red mesh the Fo-Fc map at +3 and -3 σ respectively. The poor definition of the 2Fo-Fc map around the ethyl group in the non-aged model could suggest either partial aging of the complex. Refinement of the aged model (b) presents a peak in the Fo-Fc map, at the position of the ethyl group.

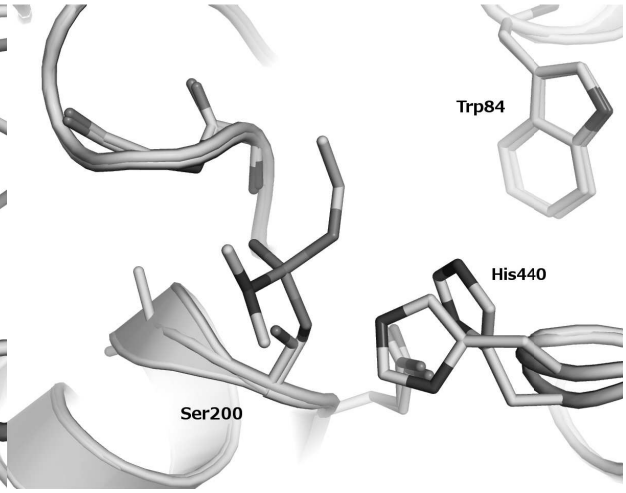
density map at the position of the C08 carbon atom of the ethyl leaving group (figure 4.9 b) suggesting that the non-aged form is the most occupied in the crystal. It was observed, for mAChE and hAChE (Carletti et al., 2008) that aging, along with the dealkylation of the ethyl group, involves a conformational change in the position of the catalytic histidine, that returns to its native conformation. The catalytic histidine could not be modeled in its native conformation, even when refining the aged form of the *Tc*AChE-tabun complex, thus supporting the hypothesis of the complex being non-aged, if aging procedes similarly in mAChE and *Tc*AChE. Comparison with the structure of the non-aged mAChE-tabun complex (figure 4.10 c) shows that catalytic histidine assumes the same conformation in the presence of tabun.

Tabun interacts with the environment of the active site (figure 4.10 a) by forming four hydrogen bonds. The first three are between the O2 atom of tabun and the oxyanion hole residues Gly118-Gly119 and Ala201, similarly to what was observed for substrate analog OTMA (Chapter 1, Colletier et al. (2006a)). The other oxygen atom in tabun (O6) also forms a hydrogen bond with the catalytic triad residue His440, which is, as already anticipated, observed in a non-native conformation. We

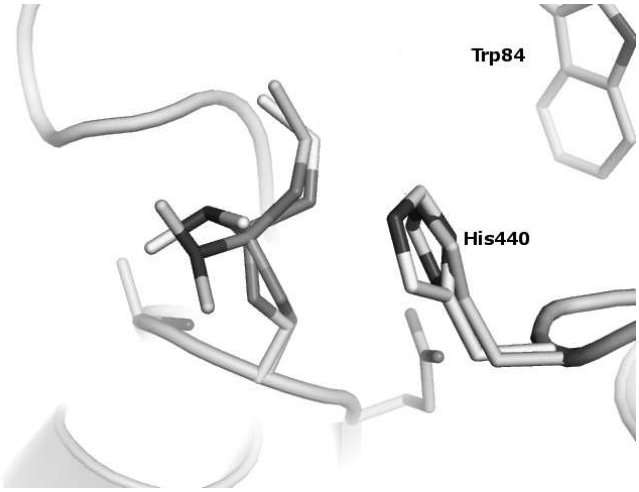
can appreciate the amplitude of this conformational change in the superposition of the non-aged *Tc*AChE-tabun conjugate and the native *Tc*AChE structure (figure 4.10 b). The side chain of His440 rotates 90° around the χ_1 angle, while the backbone atoms are shifted away from the catalytic Ser200 (1 Å for $C\alpha$). The orientation of the dimethyl-amine substituent seems to be different from the one observed in non-aged mAChE-tabun conjugate, (PDB acces code 3dl4; Carletti et al. (2008); figure 4.10 c). But in our case, the resolution permitted to fit the two methyls in a unique conformation, while the lower resolution of mAChE crystals (2.5 Å) made it more difficult to unequivocally model this moiety.



(a) Interactions tabun-AChE



(b) Native and tabun inhibited *Tc*AChE



(c) Inhibited *Tc*AChE and mAChE

Figure 4.10: a: Polar contacts between tabun and the residues in the *Tc*AChE gorge. The main interaction is through hydrogen bond between Tabun O2 atom and the oxy-anion hole residues Gly118-119 and Ala201. These bonds stabilize the substrate in native enzyme. There is also another hydrogen bond between tabun O6 and the His440 NE2 atom. This interaction pushes His440 in a non-native conformation. Distances are expressed in Å. b: Superposition of native (PDB acces code 1ea5, cyan) and inhibited by tabun (green) structures of *Tc*AChE, featuring an alternative conformation of His440. c: Superposition of non aged tabun-*Tc*AChE (green) and tabun-mAChE (yellow,(Carletti et al., 2008)). We can observe how the dimethyl-amine is oriented in two different ways in the enzyme from the two species, while the position of the leaving group is the same.

4.4.3 Crystallographic structure determination of the *Tc*AChE-KM297 complex

With the aim of structurally characterizing the interactions between KM297 and AChE, we solved the structure of *Tc*AChE-KM297 complex. *In vitro* inhibition tests have already shown that it has a good affinity for the enzyme (IC50=1.4 μM), which is positive for soaking tests. However, due to the low solubility of KM297 in aqueous solution, a big effort was devoted to finding the proper soaking conditions for formation of the complex. Data were collected at the ESRF for about 50 crystals during different experiments[6] to finally obtain a first structure.

Data collection and processing

Soakings trials revealed that the DMSO concentration in the soaking solutions must not be higher than 15%. Above this threshold, crystals either dissolved during the soaking procedure, or lost their diffraction quality. Due to the difficulties to grow trigonal crystals, all the soaking tests were performed with orthorhombic-1 plates. At lower DMSO percentages, between 5% and 10%, a few reflections were observed up to the known resolution limit for this kind of crystal (around 2.0 Å). *Tc*AChE orthorhombic-1 crystals were soaked for 24 hours in a solution containing 36% PEG, 100 mM MES pH 5.6, 1 mM KM297, 5% DMSO. A complete data set of 240 images, with an oscillation range of 0.5° per image was collected at a temperature of 100 K on the beam line ID 14-4 at the ESRF.

Automatically processed data were truncated at a resolution of 2.6 Å (complete statistics for data processing and refinement are reported in table 4.5). After performing molecular replacement with a *Tc*AChE dimer model (PDB access code 1w75), the electron density map could be observed to check for the ligand.

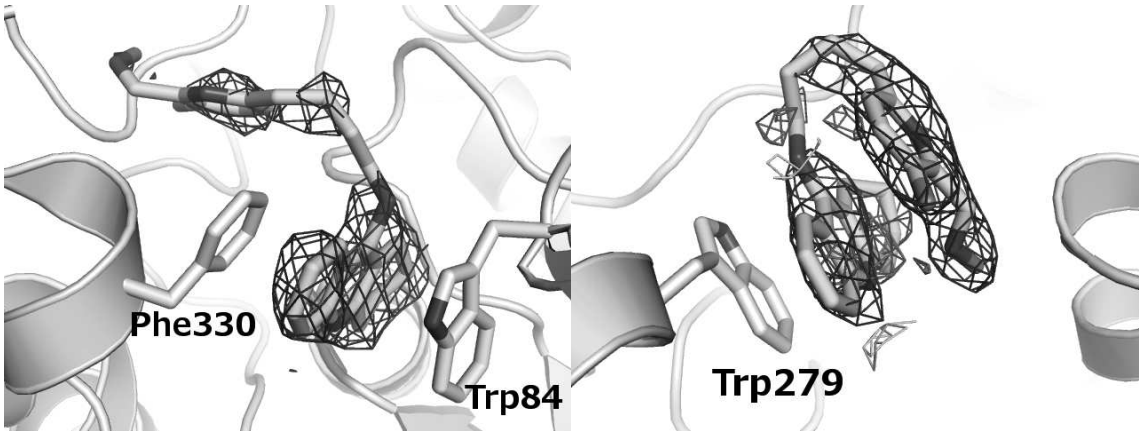
Structure refinement

Refinement was performed using phenix.refine. Atom positions were refined, along with atom individual B-factors, in several cycles of refinement and manual reconstruction using Coot. Restraints for the ligand KM297 were calculated with phenix.eLBOW. Water molecules were added, using the automatic water search

[6]Thanks to the IBS BAG experiment

Data	
Resolution range (Å)	45.91 - 2.63 (2.724 - 2.63)
Space group	P 21 21 21
Unit cell (a,b,c (Å); α, β, γ, (°))	91.3 106.2 150.5 90 90 90
Total reflections	214755 (21464)
Unique reflections	43964 (4317)
Multiplicity	4.9 (5.0)
Completeness (%)	99.49 (99.40)
Mean I/sigma(I)	9.22 (1.88)
$CC^{1/2}$	0.992 (0.739)
Refinement	
R-work (%)	19.0 (28.2)
R-free (%)	26.5 (37.2)
RMS bond lengths (Å)	0.009
RMS bond angles (°)	1.20
Atoms in asymmetric unit	9139
Ramachandran favored (%)	94
Ramachandran outliers (%)	0.28
Average B-factor ($Å^2$)	35.00

Table 4.5: Data and refinement statistics for *Tc*AChE-KM297 complex. Values in parenthesis refer to the highest resolution shell.



(a) CAS (b) PAS

Figure 4.11: 2Fo-Fc electron density maps at 1 σ level (blue mesh) and Fo-Fc at +3 and -3 σ level (green and red red mesh respectively) for KM297 in he gorge of *Tc*AChE. Panels show the ligand at CAS (a), with the position of the catalytic serine shown in red, and PAS (b) of *Tc*AChE.

in Coot and were compared with the ones observed in the high resolution native *Tc*AChE structure 1w75. MolProbity tools were used to validate the structure.

Structure analysis

A first inspection of the structure suggested the possible presence of a KM297 molecule bound at the active site. This observation was also corroborated by a conformational modification of the Phe330 side chain. A more detailed analysis of the electron density map was performed using the program phenix.ligand, to look for other possible KM297 molecules bound to the enzyme. We were able to find, in each of the two protein monomers in the asymmetric unit, two molecules of KM297, located in the gorge of the enzyme. Both of them have a well defined tacrine moiety, while the electron density map of the linker is less clearly defined (figure 4.11). The molecule in the CAS has also a poor electron density around the oxime moiety, especially the oxime C=N-O function. This can be either due to a higher mobility of this part of the molecule, or to a partial hydrolysis of the oxime, which could be less stable at pH 5.8 of the crystallographic conditions than at pH 7.

We can observe in figure 4.12 how KM297 is bound to the peripheral and active sites of the enzyme. Binding to the PAS is fundamental for a bifunctional reactivator.

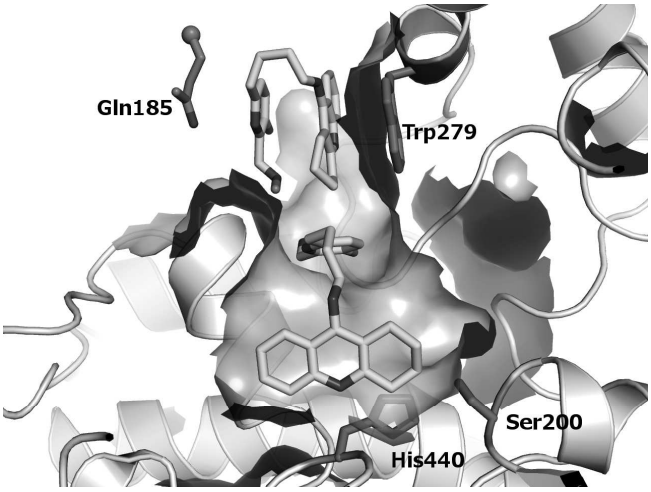


Figure 4.12: Binding of two KM297 molecules in the gorge of *Tc*AChE. We can observe that the presence of a KM297 molecule in the active site prevents the oxime function of the molecule at the PAS to enter into the gorge as originally designed for the reactivation reaction.

This observation supports, therefore, the choice of tacrine as a PAS binding moiety for KM297. However, no conformational rearrangement of the PAS residues was necessary to accommodate the ligand, as the orientation of Trp279 is the same as observed in the native enzyme.

One should note that the orientation of the pyridine aldoxime of the KM297 molecule bound at the PAS is not suitable for reactivation. Instead of going down into the gorge the oxime is stacked to the tacrine moiety, with the carbon linker forming a U-turn. In this conformation the oxime moiety forms an hydrogen bond with the main chain carbonyl of Tyr334. A decisive contribution to the stabilization of the oxime in this conformation at the PAS is the presence of residue Gln185 from a symmetry related enzyme in the crystal, whose side chain lies parallel to the pyridine aldoxime at 3.6 Å (figure 4.13). This distance is proper to the packing of orthorhombic-1 crystals. Trigonal crystals are arranged differently, thus soaking in this crystal form should not be affected by the vicinity of the symmetry-related enzyme. However, this experiment was not performed, due to the difficulty to grow trigonal *Tc*AChE crystals. No KM297 molecule was found at the PAS of crystals grown in presence of KM297.

As suggested from a first observation of the electron density, the second molecule bound to the enzyme is oriented upside-down in the CAS, with the tacrine moiety

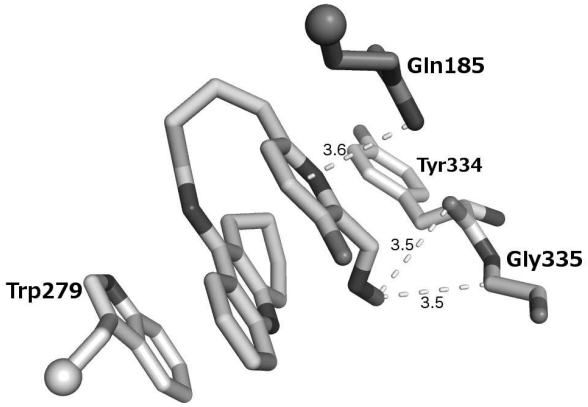


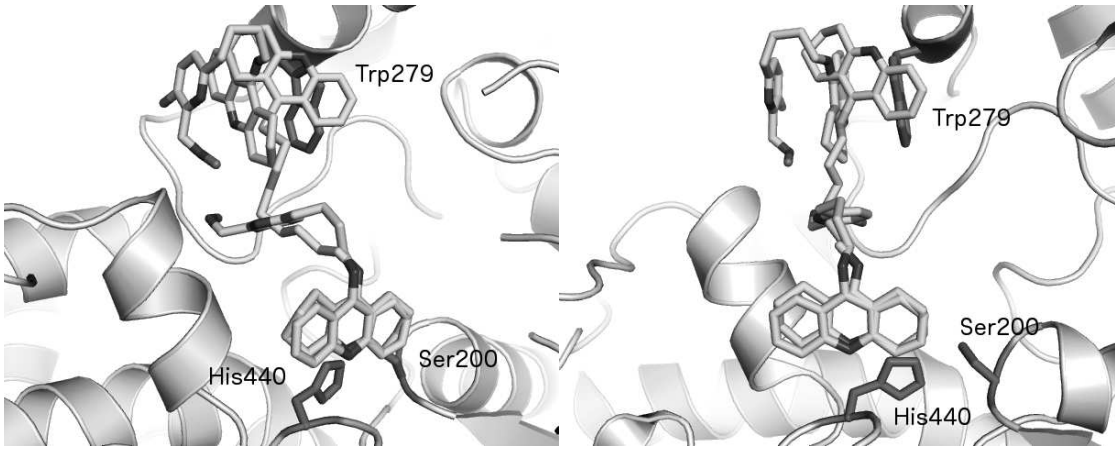
Figure 4.13: Interactions between KM297 and the PAS region of *Tc*AChE. We can observe the aromatic stacking between the tacrine moiety and Trp279, and the hydrogen bonds with Tyr334 and Gly335. In magenta we show the residue Gln185 from a symmetry related enzyme in the crystal, whose amide group could interact with the pyridine $\pi-$system. Distances are expressed in Å.

sandwiched between Trp84 and Phe330 and the carbon linker pointing towards the PAS. The position of the tacrine moiety in the active site is analogous to what is observed both in the structure of the *Tc*AChE-tacrine complex (PDB acces code 1acj) and also in similar molecules featuring a tacrine moiety, such as A8 and NF595 (PDB acces codes 1odc and 2cek, respectively).

It should be noted that the oxime function does not lie in the plane of the pyridine but appears to be slightly twisted. This observation indicates that the oxime function is not necessarily conjugated to the pyridine π-system when bound to the enzyme. It also justifiesthe fact that we did not impose any constraint of coplanarity for the oxime and pyridine during docking and MD simulations.

In figure 4.14, we show a superimposition of the *Tc*AChE-KM297 structure and *Tc*AChE-NF595 structure (PDB acces code 2cek). The tacrine moiety of the KM297 bound in front of Trp84 superimposes remarkably well to that of the NF595 molecule. The situation is different at the PAS, where both tacrine moieties do not superimpose. This can be explained by the fact that the PAS residue Trp279 stays in its native conformation and does not rotate as in the NF595 structure. For this reason, the planes of both tacrine moieties lay perpendicular one to each other.

The structure of this complex provide us with good conditions to perform soaking



(a) (b)

Figure 4.14: Superimposition of the KM297 and NF595 (PDB acces code 2cek) *Tc*AChE complexes. The tacrine moiety of KM297 observed in the active site adopts the same conformation as that of NF595, while the tacrine moieties at the PAS are not superimposable. In particular, we don't observe the conformational change of the side chain of Trp279. Panels a and b are rotated by 45 degrees around the vertical axis.

KM297 so as to obtain the ternary complex conjugate.

4.4.4 Crystallographic structure determination of non-aged *Tc*AChE-NEDPA-KM297 complex

With the aim of obtaining the structure of KM297 in the putative conformation assumed during reactivation of inhibited AChE, we solved the structure of the ternary complex *Tc*AChE-NEDPA-KM297. Due to the difficulty to access OP nerve agents, and more generally to avoid manipulating these dangerous molecules, the following experiments were performed using OP surrogates that are orders of magnitude less toxic than the actual chemical warfare agents, yet produce the same conjugate of the inhibited serine. In particular, we used the tabun surrogate 4-nitrophenyl ethyl dimethylphosphoramidate (NEDPA; Meek et al. (2012)), synthesized by the group of Prof. P.Y. Renard from the University of Rouen (figure 4.15).

The reduced toxicity is obtained by replacing the cyanide group of the leaving group of tabun, by a p-nitrophenol group. This substitution has two effects: it reduces the accessibility of the molecule to the active site and makes it non-volatile,

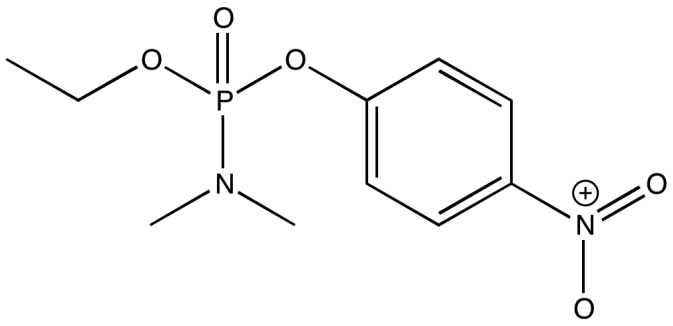


Figure 4.15: Tabun surrogate NEDPA. Cyanide, the leaving group of tabun, is replaced by paranitrophenol.

allowing a safer manipulation. Also, such of surrogates are not submitted to the chemical warfare convention and can be manipulated in any laboratory. *In vitro* tests performed in the laboratory[7] have shown that this surrogate inhibits AChE 740-fold slower than tabun ($k_i = 1.0 \pm 0.1 10^4\ M^{-1}min^{-1}$vs $7.4 \pm 0.2 10^7\ M^{-1}min^{-1}$). As we will show below, it produces the same adduct as tabun on the catalytic serine, and is therefore strictly equivalent to tabun for crystallographic and reactivation studies.

Data collection and processing

Orthorhombic-1 *Tc*AChE crystals were in a first step soaked in a solution containing 36% PEG200, 100 mM MES pH 5.8, 0.1 mM NEDPA and 1% isopropanol for 4 hours. Then the were soaked again for 10 minutes in 36% PEG200, 100 mM MES pH 5.8, 0.1 mM NEDPA and 1% isopropanol, 5% DMSO and 1 mM KM297. A complete data set was collected on beam line ID23-1 at the ESRF. We collected with the fine slicing method, 1270 images with a rotation angle of 0.1° per image at 100 K. Data were processed with XDS, using the orthorhombic-1 unit cell parameters. Using AIMLESS (Evans and Murshudov, 2013), we truncated them at a resolution of 2.9 Å. Molecular replacement was performed using a *Tc*AChE dimer (PDB entry 1w75).

[7]M.T. Froment, personal communication

Data	
Resolution range (Å)	58.14 - 2.9 (3.004 - 2.9)
Space group	P 21 21 21
Unit cell (a,b,c (Å); α, β, γ, (°))	91.5 106.4 150.5 90 90 90
Total reflections	150232 (15096)
Unique reflections	32837 (3215)
Multiplicity	4.6 (4.7)
Completeness (%)	98.71 (99.14)
Mean I/sigma(I)	9.20 (2.3)
$CC^{1/2}$	0.99 (0.85)
Refinement	
R-work (%)	18.5 (25.4)
R-free (%)	26.1 (33.2)
RMSD bond lengths (Å)	0.009
RMSD bond angles (°)	1.20
Atoms in asymmetric unit	8777
Ramachandran favored (%)	92
Ramachandran outliers (%)	0.47
Average B-factor ($Å^2$)	39.50

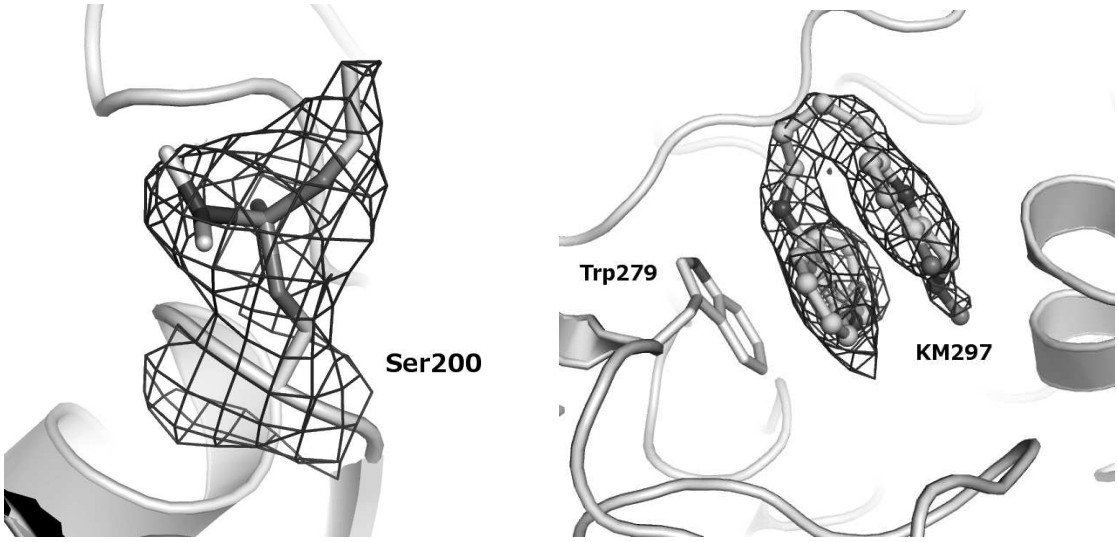
Table 4.6: Data collection and refinement statistics for *Tc*AChE-NEDPA-KM297 complex. Values in parenthesis are relative to the highest resolution shell.

Structure refinement

Refinement was performed using phenix.refine. Atom positions were refined, along with atom individual B-factors. Restraints for the tabun substituent and for KM297 were calculated with phenix.eLBOW. Bond length and angles for the covalent bond between tabun and the catalytic serine were obtained from the model used in the calculations of force field parameters presented above (section 4.3.1). Water molecules were added, using the automatic water search in Coot and a comparison with the ones observed in the high resolution native *Tc*AChE structure 1ea5. Statistics for the refined structure can be found in table 4.6.

Structure analysis

As for the *Tc*AChE-tabun complex, the presence of the OP in the active site was indicated by a high positive peak in the Fo-Fc map after molecular replacement. We managed to model the OP correctly in both monomers, in particular in monomer



(a) Inhibited serine and tabun surrogate (b) PAS

Figure 4.16: a: 2Fo-Fc electron density map at 1 σ level (blue mesh) of Ser200 modified by the tabun surrogate NEDPA from monomer B of *Tc*AChE. Differently from the non-aged *Tc*AChE-tabun complex, the ethyl group of NEDPA is well defined in the maps; b: 2Fo-Fc electron density map at 1 σ level (blue mesh) and Fo-Fc map at +3 σ and -3 σ (green and red mesh respectively) of KM297 at the PAS of NEDPA-inhibited *Tc*AChE. As for the complex with native *Tc*AChE, the KM297 molecule at the PAS is folded and the oxime does not enter the gorge.

B, where the electron density of the ethoxy substituent was well defined compared to that in monomer A (figure 4.16 a). Aging of the adduct can be dismissed, given that His440 was found in a single conformation, corresponding to that observed in the non-aged form of the mAChE-tabun complex. We found electron density corresponding to a reactivator molecule in monomer A, close to Trp279 (figure 4.16-b). The tacrine moiety could be modeled in this density in a conformation identical to that binding at the PAS in the structure of *Tc*AChE-KM297 discussed above. The pyridine aldoxime part of the molecule is similarly folded to stack on the aromatic part of the tacrine.

The important result is that the reactivator molecule did not bind in the active site as in native *Tc*AChE (figure 4.17). At a more detailed level of observation, we can conclude that the ethyl group of non-aged tabun would overlap with the tacrine moiety of the molecule located in the active site of the *Tc*AChE-KM297 complex. For this reason the reactivator can not bind deep into the active site gorge

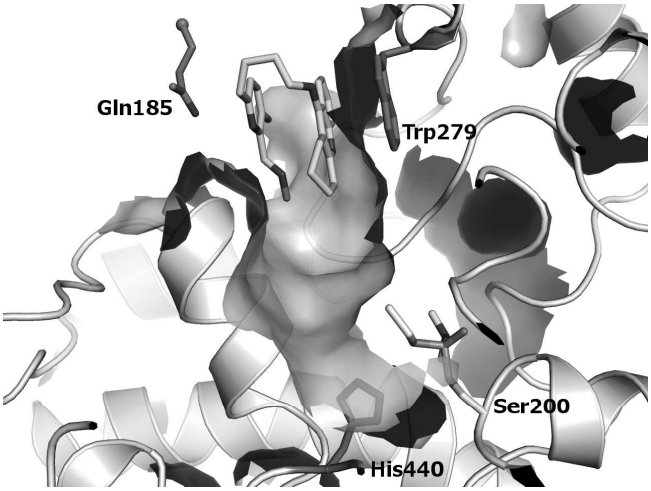


Figure 4.17: Position of KM297 in the gorge of NEDPA-inhibited *Tc*AChE.

of the inhibited enzyme. At the PAS, KM297 is stabilized by a hydrogen bond with Tyr334 and by the vicinity of symmetry related Gln185, whose influence on the binding mode was discussed above. If we compare the surface of the gorge with the one observed in the complex with native enzyme (figure 4.12), we can conclude that the presence of the tabun surrogate, together with the non-native conformation of the catalytic histidine between the native and tabun-inhibited structure, reduces the accessible volume of the active site. This confirms our hypothesis that the ligand binds only at the PAS in the non-aged form of the enzyme inhibited by an OP. The oxime function does not point down into the gorge Thus, this structure does not represent the functional state of the ternary complex. Nevertheless we have a direct confirmation that our strategy of targeting the PAS ligand of the OP-inhibited enzyme with a reactivator was correct.

Some electron density is found in the gorge, spanning from Trp84 to Tyr334. Such density has already been observed in many *Tc*AChE structures, but no consensus was found on its origin. It was possible to model it as a small PEG molecule, triethylene glycol, which is present in the mother liquor.

We demonstrated, through docking and crystallographic studies that binding of the tacrine moiety at the PAS can occur. In this mode, simulation showed that the 4-carbon linker of KM297 is the most suited to span the gorge and bring the oxime close to the OP-inhibited serine. Yet, we could not confirm the action mode of the reactivator by crystallographic study, because of an artifact induced by crystal

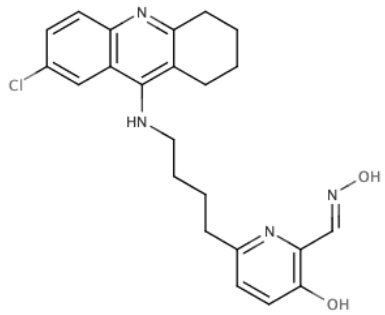


Figure 4.18: A derivative of KM297, featuring a chlorine substituent at position 7 on the tacrine moiety of the molecule, generating a steric hindrance, which should prevent binding within the active site of the enzyme.

packing. This could be assessed by performing soakings in another crystal form, *i.e.* the trigonal one, featuring a different crystal packing that the orthorhombic-1 form. However, we did not manage to grow enough crystals in this form to perform the experiments. Still we can suggest some chemical modifications to our lead molecule, in order to maintain its reactivation properties, while limiting its binding to the active site of the uninhibited enzyme.

4.5 Lead optimization: the JDS family

The high affinity of KM297 for the native enzyme (IC 50=0.25 ± 0.05 μM against hAChE) can be potentially a drawback in the design of a highly efficient reactivator. An ideal reactivator should have a high affinity for the inhibited enzyme, and a low affinity for the native enzyme, to avoid inhibition of AChE and subsequent secondary effects within the nervous systems . Based on the structure of the *Tc*AChE-KM297 complex, we designed a new family of molecules, carrying a chlorine atom at the position 7 of the tacrine moiety (figure 4.18). The role of this substituent is to decrease the affinity for the active site by creating a steric clash between the ligand and the environment of the gorge, while maintaining the affinity for the PAS.

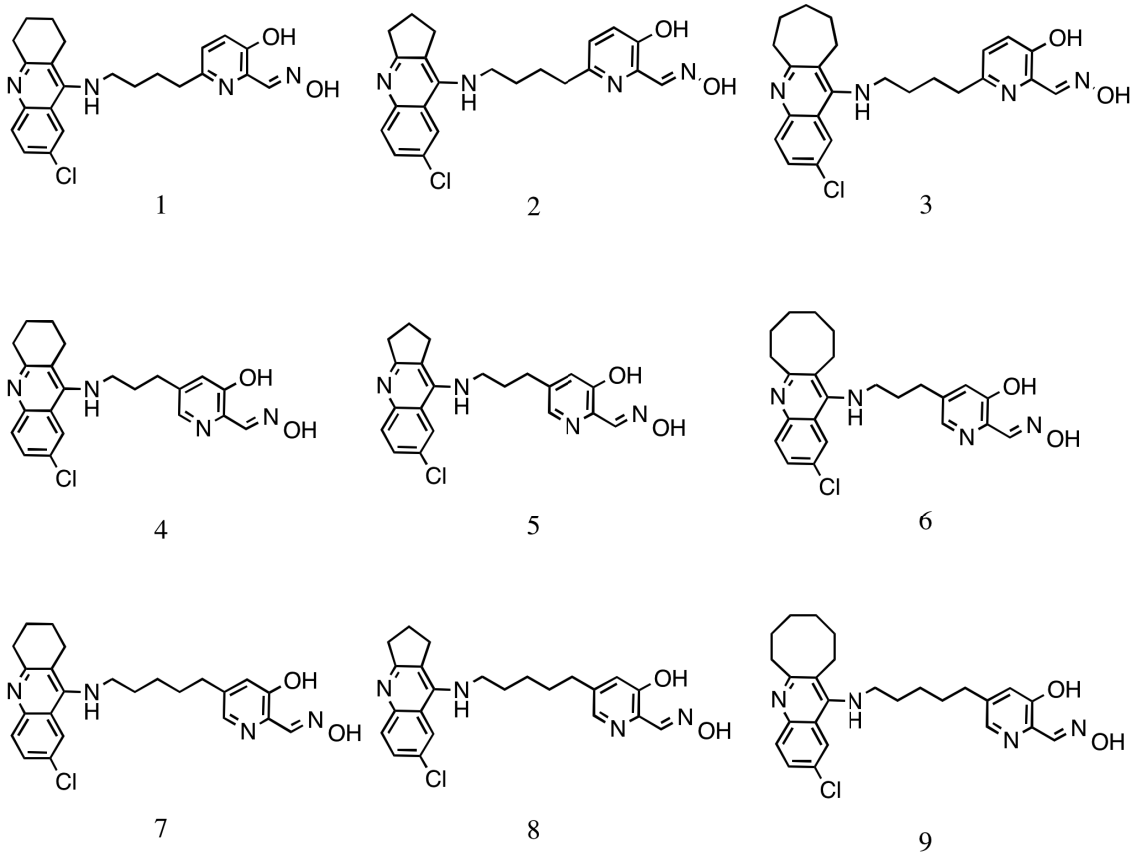


Figure 4.19: A family of nine chlorinated derivatives of KM297. They all bear a chlorine atom at position 7 of the tacrine group. Carbon linker lengths span from 3 to 5 atoms and saturated cycle sizes from 5 to 7 carbons.

4.5.1 Docking of chlorinated derivatives of KM297

We used our docking protocol to screen nine possible modifications of KM297 (figure 4.19) in order to asses the influence of the presence of chlorine, of the size of the saturated ring and of the linker length (3, 4 or 5 carbon atoms), on the binding to native and inhibited AChE. As predicted, calculated affinities (table 4.7) are good for all of the compounds, the tacrine moiety of the reactivator interacting with the PAS tryptophan. The docking score is particularly good for molecule number 1, featuring a 4 carbon linker and a 6-carbon saturated cycle similarly to KM297. Julien de Sousa from the University of Strasbourg then synthesized this molecule, named JDS207 hereafter. It was studied in more detail by flexible docking, crystallography and *in vitro* reactivation tests.

Molecule	Binding energy (kcal/mol)
1	-13.1
2	-12.6
3	-11.6
4	-12.5
5	-11.7
6	-10.3
7	-11.6
8	-11.9
9	-13.0

Table 4.7: Docking scores of the KM derivative series in non-aged hAChE-VX. Molecule number 1, featuring a 4 carbon linker and a 6-carbon saturated cycle is the one with the best score in docking experiments.

4.5.2 Docking of JDS207

JDS207 (figure 4.19-1), the chlorinated analog that has the best affinity for the non-aged form of hAChE-VX complex in docking experiments. We decided to examine its docked conformations in more detail, also because, from the set of nine suggested modifications, it is structurally closest to the lead molecule KM297. Docking results reported in table 4.8 show that it has a calculated affinity lower than KM297 for the non-aged hAChE-VX complex (-13.1 kcal M^{-1} vs -11.7 kcal M^{-1} for KM297). However, docking to the native form of hAChE lead to the conclusion that KM297 is more affine than JDS207 for the native enzyme (-12.4 kcal M^{-1} for JDS207 vs -13.3 kcal M^{-1} for KM297). This observation is also corroborated by *in vitro* affinity measurements that show the IC50 of KM297 is 10 times higher than JDS207 for native hAChE.

Structural observation of the binding modes to both the native and VX-inhibited form of hAChE shows that none of the output conformations has a tacrine in the active site, but they all feature binding at the PAS. None of the outputs features the same rotation of Trp286 (Trp279 in *Tc*AChE) as for KM297, but the lowest energy solution has an analogous stacking of the tacrine between Trp286 and Tyr124. This difference in stacking modes seems to be due to the presence of the chlorine substituent. This atom would lead to multiple clashes with the residue of the PAS region if we try to model it in the docked conformation of KM297 previously discussed.

Conformation	Binding energy hAChE-VX (kcal/mol)	Binding energy hAChE(kcal/mol)
1	-13.1	-12.4
2	-13.1	-11.7
3	-12.8	-11.5
4	-12.7	-11.4
5	-12.6	-11.2
6	-12.0	-10.9
7	-11.6	-10.8
8	-11.1	-10.5
9	-11.1	-10.4

Table 4.8: Docking scores for the 9 output conformers of JDS207 in non-aged hAChE-VX and native hAChE.

4.5.3 Crystallographic structure determination of TcAChE-JDS207 complex

Data collection and processing

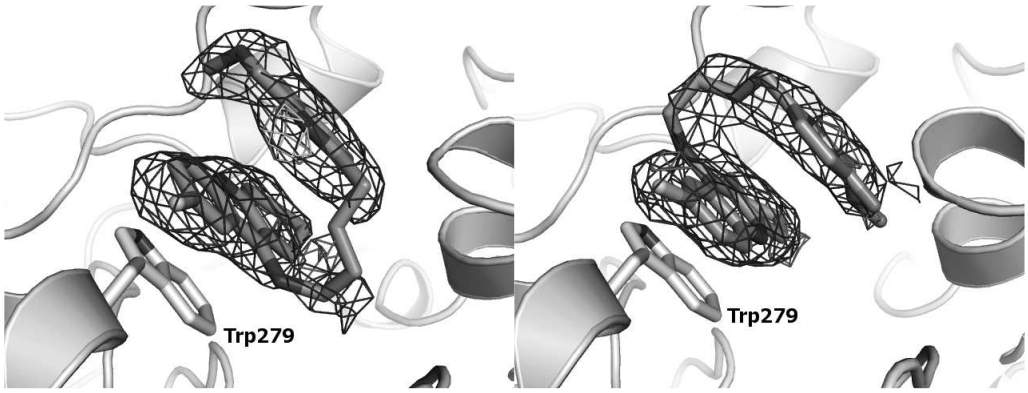
*Tc*AChE orthorhombic-1 crystals were soaked at 4°C for 15 minutes in mother liquor (36% PEG200, 100 mM MES pH 5.8) plus1 mM JDS207 and 5% DMSO. 180 images with oscillation range of 1° per image were collected on the ID29 beam line at the ESRF at a temperature of 100 K. Data, processed with XDS, were truncated with AIMLESS at a resolution of 2.9 Å. Molecular replacement was performed using the *Tc*AChE dimer, PDB entry 1w75, as a template. Complete statistics are reported in table 4.9.

Structure refinement

Refinement was performed using phenix.refine. Atom positions were refined, along with atom individual B-factors. Restraints for the ligand JDS207 were calculated using phenix.eLBOW. Water molecules were added, using the automatic water search in Coot and a comparison with the ones observed in the high resolution native *Tc*AChE structure 1ea5. After several refinement cycles and manual reconstruction in Coot, refinement statistics converged (table 4.9).

Data	
Resolution range (Å)	46.3 - 2.9 (3.0 - 2.9)
Space group	P 21 21 21
Unit cell (a,b,c (Å); α, β, γ, (°))	91.8 107.2 150.4 90 90 90
Total reflections	212283 (21557)
Unique reflections	33562 (3300)
Multiplicity	6.3 (6.5)
Completeness (%)	99.9 (99.9)
Mean I/sigma(I)	10.1 (2.9)
$CC^{1/2}$	0.99 (0.90)
Refinement	
R-work(%)	19.4 (27.2)
R-free (%)	26.6 (31.5)
RMSD bond lengths (Å)	0.097
RMSD bond angles (°)	2.28
Atoms in asymmetric unit	9051
Ramachandran favored (%)	92
Ramachandran outliers (%)	0.19
Average B ($Å^2$)	32.5

Table 4.9: Data and refinement statistics for *Tc*AChE-JDS207 complex. Values between parenthesis refer to the highest resolution shell.



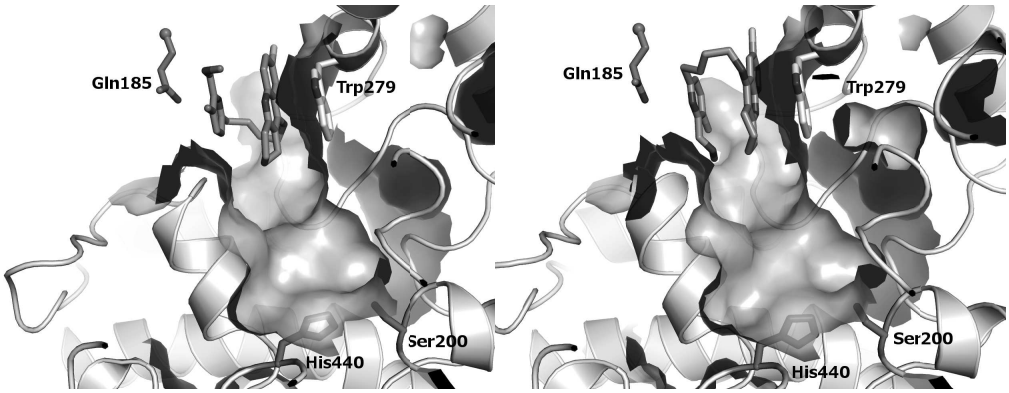
(a) (b)

Figure 4.20: 2Fo-Fc electron density map of JDS207 bound at the peripheral site of *Tc*AChE, represented at 1 σ level by the blue mesh. Green and red meshes show the Fo-Fc electron density map at +3 σ and -3 σ level respectively. JDS207 is folded on itself with the tacrine moiety in between the aromatic ring of Trp279 and that of the pyridine aldoxime. Monomers A and B present the molecule in two different orientations, as shown in panels a and b, respectively.

Structure analysis

The electron density map showed the presence of a tacrine moiety at the PAS, stacking in front of Trp279, which is again found in its native conformation (figure 4.20). Orienting JDS207 is easier than KM297, due to the presence of the chlorine atom, which is clearly defined, being the strongest peak in the electron density map. As for the structure of *Tc*AChE-KM297 described above, the pyridine aldoxime part is stacked against the tacrine part, forming a triple aromatic stacking at the PAS. Also in this case, the presence of the residue Gln185 from a symmetry related enzyme of the crystal seems to help in the stabilization of this conformation (figure 4.21), along with a hydrogen bond with Tyr334. To overcome this bias, soaking with a trigonal crystal was performed, but the ligand was not observable in the electron density map. The electron density for the linker is better defined in monomer B than in the case of KM297, which helps in fitting the ligand in the proper orientation. There is no electron density for the linker in monomer A. This allowed us to model the reactivator in an alternative orientation, with the NH group of the tacrine pointing towards the outside of the gorge.

In contrast to what was observed in the *Tc*AChE-KM297 structure (figure 4.12),



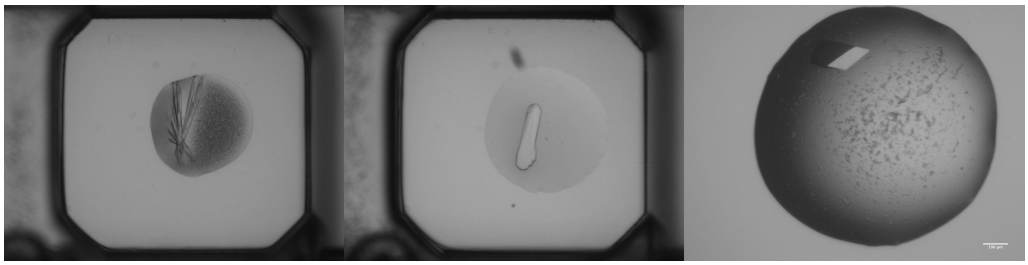
(a) (b)

Figure 4.21: Position of the JDS207 molecules in the gorge of *Tc*AChE. Monomers A and B feature two different orientation, with the one in a seeming more prone to reactivation.

there is no sign of the presence of the reactivator in the active site of the enzyme (figure 4.21). A small electron density peak in front of Trp84 is not different from what is also generally observed in crystals of native *Tc*AChE, and could be related to the presence of a low-weight PEG molecule from the mother liquor solution.

4.6 Structure determination of native hAChE

The vast majority of the crystallographic structures of AChE were obtained, over the last two decades, either from *Tc*AChE or mouse AChE. Only a small subset is from hAChE. Also until 2010, the only known way to obtain good-quality crystals of the human enzyme was to crystallize the complex with the snake toxin fasciculine-2. This toxin has the inconvenience of occupying the PAS and restraining access to the active site gorge. Efforts were made to produce crystals of native hAChE using truncated forms of the enzyme (Cheung et al., 2012; Dvir et al., 2010). Cheung's hAChE, which was truncated at the N-terminus position 1 and C-terminus position 544 yielded the best crystals to date, allowing to solve structures at 2.1 Åresolution. In the meantime, and as described in the following, we also managed to obtain crystals of hAChE truncated at residue 544, thus cutting the long C-terminal helix, and solved a structure of this form of hAChE.



(a) Condition 1 (b) Condition 2 (c) Condition 3

Figure 4.22: Hits from crystallization screening. Condition 1: 100 mM Bicine pH 9, 20% PEG 6000; Condition 2: 50 mM Hepes pH 7, 1.6 M Li_2SO_4, 50 mM $MgSO_4$; Condition 3: 0.1 M trisodiumcitrate pH 5.6, 0.1 M Li_2SO_4, 12% PEG6000

4.6.1 Enzyme production

Frozen CHO-K1 cells previously transfected with a pGS-hAChE-L544Stop vector carrying the glutamine synthetase gene marker, and stably expressing variant L544Stop hAChE, were thawed and cultured in UltracultureTM medium (Lonza, Belgium) containing $50\mu M$ methionine sulphoximine (MSX) for selection pressure. The cells were expanded in flasks before being transferred to roller-bottles for enzyme production. The culture medium, 1/3 Ultraculture (Ozyme) and 2/3 Ham's F12 (Dutscher), containing the secreted enzyme, was changed every 3-4 days. Our form of enzyme does not contain His tag and was purified to homogeneity as described in Carletti et al. (2008). Shortly, protein was precipitated from the media by ammonium sulfate. After dialysis, the enzyme contained in the pellets was purified on a sepharose-4B/procainamide affinity column then a monoQ anion-exchange chromatography column. Fractions containing high hAChE activity were pooled, buffer was exchanged to match that used in Cheung et al. (2012) (10 mM HEPES pH 7.0 and 10 mM NaCl), and the enzyme was finally concentrated to 12-16 mg/mL.

4.6.2 Crystallogenesis

Crystallization conditions were at first screened with the Hampton screens 1-6 at the EMBL high throughput crystallization laboratory (HTXLab) in Grenoble. After two months of monitoring the plates, we found crystals in three conditions (figure 4.22). We tried to reproduce them in hanging drop 24 wells plates, but only condition number 2 yielded crystals. Screening the concentrations of Li_2SO_4 and $MgSO_4$, we

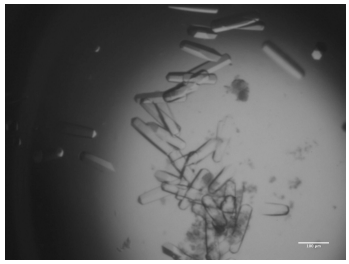


Figure 4.23: Crystals of hAChE-L544Stop from optimized conditions (100 mM Hepes pH7, 1.7M Li_2SO_4, 60 mM $MgSO_4$) Scale bar correspond to 100 μm.

managed to grow bigger crystals of about 100 μm in their longest dimension in large numbers (figure 4.23). Optimized conditions are 100 mM Hepes pH 7, 1.7 M Li_2SO_4, 60 mM $MgSO_4$. These crystals were cryo-protected with mother liquor containing 18% glycerol and flash-cooled in liquid nitrogen to be taken to the synchrotron beam lines.

4.6.3 Data collection and processing

We collected data from several crystals obtained in the optimized condition number 2 on the beam line ID23-1 at the ESRF, at a temperature of 100 K. Due to the high concentration of Li_2SO_4, about half the collected data sets presented large noise from powder diffraction of salt micro-crystals. It was possible for iMOSFLM (Battye et al., 2011) to automatically index them in P6 group with unit cell parameters of 211, 211, 115 Å, 90° 90° 120° similar to that reported by Dvir et al. (2010) for the L544Stop variant expressed in Drosophila cells. A total of 120 images was collected, with an oscillation range of 1° per image. Data were successfully processed with XDS with the P6 space group unit cell parameters and integrating 80 images yielded a complete data set. Using AIMLESS-POINTLESS in CCP4, we found that the actual space group was P61, and that the data set was truncated at a resolution of 3.4 Å. Matthews coefficient calculation suggests that the most probable asymmetric unit composition is of 4 monomers, but after molecular replacement, only one dimer of AChE was found. This yields a solvent content in the crystals of 75%. The low protein content of this crystal form may be the cause of the low diffraction power (Kantardjieff and Rupp, 2003). For this reason, our efforts are still devoted to the reproduction of the screening hits number 1 and 3.

Data	
Resolution range (Å)	49.0 - 3.4 (3.5 - 3.4)
Space group	P 61
Unit cell (a,b,c (Å); α, β, γ, (°))	210.9 210.9 116.3 90 90 120
Total reflections	288467 (29202)
Unique reflections	40653 (4058)
Multiplicity	7.1 (7.2)
Completeness (%)	100.00 (99.98)
Mean I/sigma(I)	10.9 (2.0)
$CC^{1/2}$	0.995 (0.664)
Refinement	
R-work(%)	17.3 (30.4)
R-free (%)	21.0 (34.5)
RMSD bond lengths (Å)	0.010
RMSD bond angles (°)	1.35
Ramachandran favored (%)	83
Ramachandran outliers (%)	4.1
Clashscore	12.4
Average B-factor	94.6

Table 4.10: Data collection and refinement statistics for hAChE-L544 stop. Values between parenthesis refer to the highest resolution shell.

4.6.4 Structure refinement

Refinement was performed using phenix.refine. Atom positions were refined, along with atom individual B-factors. Water molecules were added, using the automatic water search in Coot and a comparison withe the ones observed in the high resolution native hAChE structure 4ey4. The flexible loop 492-498, not present in the replacement model, has been reconstructed with the fit_loop utility in Coot and adjusted into the electron density map by real space refinement. After several refinement cycles and manual reconstruction in Coot, refinement statistics (table 4.10) converged and provided us with interpretable structures.

4.6.5 Structural analysis

As observed in orthorhombic crystals of *Tc*AChE, and also for hAChE (Dvir et al., 2010), an asymmetric unit contains a dimer of the enzyme. The two monomers interact through the 4-helix bundle. The truncation of the C-terminal helix before

Figure 4.24: A dimer of hAChE

residue 544 rigidifies the helix and favorizes packing of the enzymes in the crystals.

Crystal packing is analogous to that observed in the structure of Dvir et al. (2010). In this arrangement of molecules, the gorge of monomer A is blocked by the presence of the symmetry related enzyme in the crystal. In particular, the flexible loop 492-498, which is for instance not observed for monomer B by Cheung et al. (2012), is visible in our case and participates in crystal contact, by being trapped at the entrance of the gorge of monomer A of the symmetry related molecule, with Arg493 interacting with Trp286 and Tyr72 (figure 4.25).

Clear electron density corresponding to sugars are observed for the two monomers in front of Asn350 and for monomer B close to Asn265, two known sites of glycosylation and was modeled with NAG molecules. The electron density map corresponding to these sites for monomer A was too weak to be modeled with a particular sugar.

During the refinement process, a strong positive peak at 6 σ level in the Fo-Fc electron density map suggested that some ligand could be situated in front of the catalytic serine in both of the two monomers. We modeled an acetate group in this density, covalently bound to the catalytic serine. Its presence, already presented in a structure of native mAChE (PDB acces code 2xud, (Bourne et al., 2010b)), is able to explain the electron density and to slightly improve the statistics of the refinement.

Another unknown density is present along the gorge, spanning from the PAS to the level of Trp86. This is reminiscent of the density usually modeled as a PEG

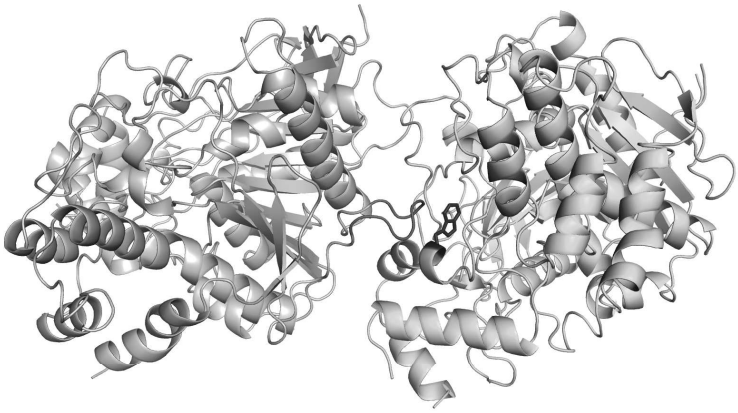


Figure 4.25: Crystal contact between two hAChE molecules. We can observe the flexible loop 490-495 of chain B, in red, interacting with the PAS of chain A, in green.

molecule in *Tc*AChE structures, except than no PEG was present, neither in the crystallization conditions, nor in the cryo-protecting solution. We decided to model it with a series of dummy atoms.

4.7 Conclusions

In this chapter we have shown the process leading from the rational design of a neutral reactivator, KM297, to the optimization of its derivatives, in particular JDS207. This process has involved docking studies, molecular dynamics simulations and x-ray crystallography. The molecule KM297 and all its potential derivatives were patented by our collaboration (Kliachyna et al., 2013).

We managed to observe through flexible docking that KM297 can bind at the PAS of VX-inhibited AChE, both to its native state and to alternative conformations of Trp286, analogous to the one observed for AChE inhibitors NF595 (PDB acces code 2cek), A8B (PDB acces code 1odc) and TZ5 (PDB acces code 1Q83).

Molecular dynamics simulations confirmed that tacrine is a PAS ligand adapted for the design of bifunctional reactivators, due to the stability of KM297 within the enzyme. Measuring the distance between the oxygen atom of pyridine-aldoxime and the VX phosphor atom along the simulation, we were able to conclude that a 4

carbon linker has the optimal length to promote reactivation.

We managed to solve multiple crystal structures, in order to characterize the interactions between our reactivators and *Tc*AChE, both native and inhibited by an OP. KM297 has two binding sites in native *Tc*AChE, one at the peripheral site and the other into the active site. Only binding at the PAS, although biased by a crystal artifact, can lead to reactivation, but this requires that the gorge is not occupied by another ligand. In our structures, the presence of another KM297 molecule bound at the active site makes impossible to observe the oxime moiety of the KM297 at the PAS reaching the catalytic triad.

Solving the structure of non-aged *Tc*AChE-tabun has allowed us to show that the presence of tabun induces a reorientation of the catalytic histidine in *Tc*AChE, also reducing the accessible volume in the active site.

We deduce that, as already shown by docking studies, KM297 cannot bind within the active site of the inhibited enzyme, but only at the PAS. This was confirmed by solving the structure of the ternary complex *Tc*AChE-NEDPA-KM297. NEDPA is a tabun surrogate that leads to the same adduct on the catalytic serine, but is less dangerous to manipulate. In the ternary complex, KM297 was observed to bind only at the PAS, but again the oxime could not enter the gorge, again due to a packing artifact.

Without a structure of the ternary complex with the reactivator in the proper position, we decided that the first optimization to make on the molecule was to force binding only at the PAS. The family of JDS molecules was designed from KM297 by adding a chlorine substituent on the tacrine moiety. Docking studies have shown that this modification is effective, and this was also confirmed by the crystallographic structure of the *Tc*AChE-JDS207 complex, where the ligand is only found at the PAS of the native enzyme.

All the crystallographic structures of *Tc*AChE in complex with the reactivators KM297 and JDS207 presented in this chapter were obtained by soaking orthorhombic-1 crystals. As mentioned above, the reactivator conformation is affected by the presence of a symmetry related enzyme, creating a crystallographic artifact in *Tc*AChE-reactivator complex structures. In order to overcome this bias, we could either perform co-crystallization or soakings in trigonal crystal of *Tc*AChE, this space group being characterized by a different contact between symmetry related enzymes in the PAS region. However, the presence of the ligand was not observed at the

PAS in *Tc*AChE crystals grown in the presence of KM297, and we didn't manage to obtain crystals in the presence of JDS207. Trigonal crystals, which grow in the same conditions as the orthorhombic-1 ones, were not obtained in our crystallization experiments, and we did not succeed in performing micro-seeding from old crystals of this form.

Both KM297 and JDS207 were tested *in vitro* to know their reactivation constants against OPs and IC50 for native hAChE. They both present a good reactivation capability for VX-inhibited hAChE, with a reactivation constant $k_{r2} = 22$ $\text{mM}^{-1}min^{-1}$ for KM297 and $k_{r2} = 13$ $\text{mM}^{-1}min^{-1}$. As a comparison, HI-6, the best molecule on the market today, has a $k_{r2} = 9$ $\text{mM}^{-1}min^{-1}$ against VX. JDS207 has an IC50 for hAChE ($IC_{50} = 2.33 \pm 0.18$ μM) 10 times higher then that of KM297 ($IC_{50} = 0.25 \pm 0.02$ μM). As discussed, a lower affinity for the native enzyme is a positive point in the design of a reactivator, so JDS207 appears to be a promising optimization of our lead KM297.

The efforts in crystallizing human AChE were focused on the mutant hAChE544stop, that was found to crystallize in three different conditions. Crystals grow in Li_2SO_4 similar to those obtained by Dvir et al. (2010). These crystals couldn't be used for soaking of our reactivators, due to the inaccessibility of the active site of one of the two monomers and to the difficulties of solubilizing hydrophobic molecules in high salt concentrations. The high solvent content of these crystals (75% with respect to an average of 43% in the PDB) is a limiting factor. For this reason, it will be important in the future to reproduce the other screening hits and verify if they can lead to better diffracting crystals like the ones obtained by Cheung et al. (2012).

Chapter 5

Building a flexible docking receptor library

The utility of structural dynamics in drug design is introduced in chapter 2. While in chapter 4 we presented a practical case of rational drug design targeting alternative conformation of a specific residue, an in-depth characterization of the conformations accessible to side chain is needed for more general cases. In this chapter we will present how information from molecular dynamics simulations can be used to design a library of conformers. Simulations were ran both on the human AChE and on TcAChE, for which many X-ray structures already exist. In this chapter we will discuss about simulation of native hAChE and *Tc*AChE. Comparing different simulation strategies, one long simulation versus a combination of multiple shorter simulations, we will argue that combining multiple short simulations is preferable for sample the populations of side chains rotamers. With the multiple simulations approach, we compared the dynamics of key residues in the gorge of *Tc*AChE and hAChE, revealing differences in their dynamical behavior. Flexible docking methods are usually not specific to the dynamics of the receptor, but are rather based on statistical properties of a particular amino acid. We developed a method that takes advantage of MD simulations to generate a receptor library that reflects the motions proper to specific residues of AChE.

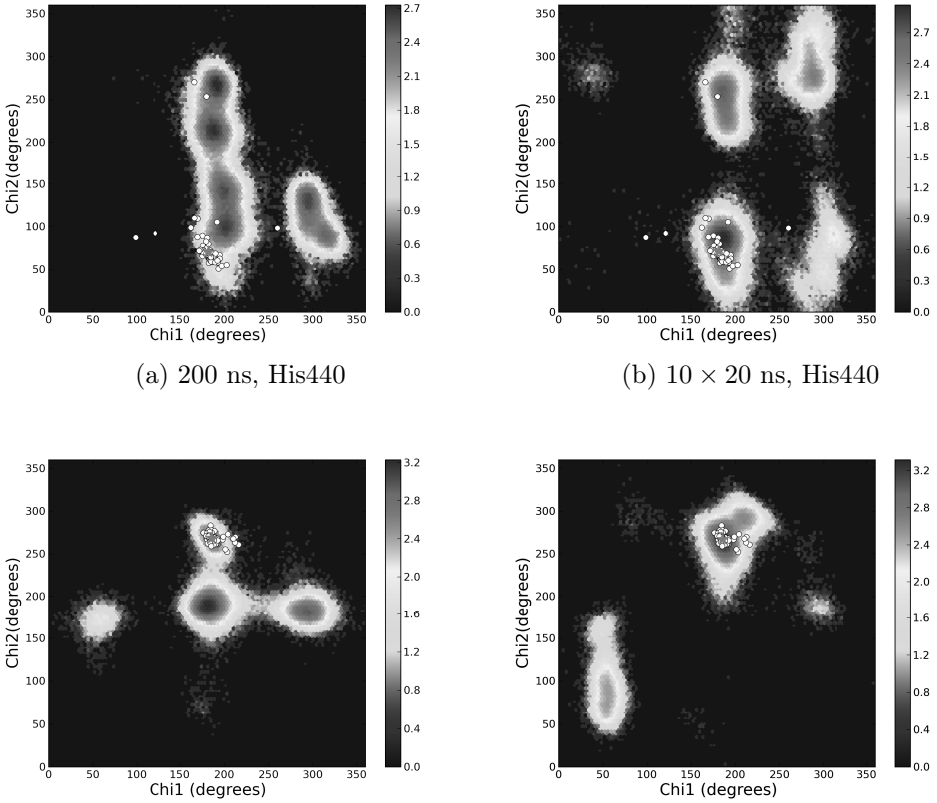
5.1 Simulations of native AChE

5.1.1 Material and methods

The simulations presented in this chapter were performed using as a starting model either the native *Tc*AChE monomer (PDB acces code 1ea5) or the native hAChE monomer (PDB acces code 4ey4) as a starting model. The simulation methodology is exactly the same for the two enzymes for the sake of comparison. Protein models was inserted in a cubic box of $10 \times 10 \times 10$ nm^3, with periodic boundary conditions and solvated in simple point charge waters (SPC-E). After charge equilibration with the appropriate number of counter ions, the energy was minimized. The temperature was kept at 300 K by coupling to a Berendsen thermostat at using with a coupling time of $\tau T = 0.1$ ps (Berendsen et al., 1984). Pressure was maintained at 1 bar by coupling to a Berendsen barostat with compressibility of 4.5×10^{-5} bar and $\tau P = 1.0$ ps. Bond lengths were restrained with the LINCS method (Hess et al., 1997) and the electrostatic interactions were calculated using the particle-mesh Ewald algorithm (Essmann et al., 1995; Darden et al., 1993). Output coordinates were written every 10 ps. Using this methodology, a set of 20 trajectories of 20ns were produced for both TcAChE and hAChE. Initial velocities differed in all simulations, owing to the use of a different random seed. For both TcAChE and hAChE, we also ran a 200 ns trajectory using the same protocol as described above. Simulations were performed using GROMACS and the GROMOS56 force field.

Dihedral angles for the side chains of all residues in the protein were calculated with the same time resolution using the program g_chi and the -rama option. Conformations of residues like tyrosine and phenylalanine are identical when rotating by 180° around χ_2. This has to taken into account when comparing dihedral sampling. For comparison of the sampling between the long and the multiple short trajectories, every crystallographic structure of *Tc*AChE available in the PDB was downloaded and assembled into an multi-state PDB file. This file format can be processed with the same tools used for the calculation of dihedral angles from MD simulation trajectories.

Figures were produced using Python and matplotlib (Hunter, 2007).



(a) 200 ns, His440

(b) 10×20 ns, His440

(c) 200 ns, Glu199

(d) 10×20 ns, Glu199

Figure 5.1: Dihedral angles for side chains of *Tc*AChE residues His440 and Glu199, calculated from a 200 ns MD simulation (a, c) and 10×20 ns simulations (b, d). The heat map represents the density of points on the plane, and white dots correspond to the conformations of the crystallographic structures of *Tc*AChE , both native and in presence of ligands, in the PDB. Native structure is indicated by the green star.

5.1.2 Sampling of side chain conformations in one long versus multiple short simulations

We compared the sampling of side chain dihedral angles in one long simulation (200 ns) versus a combination of multiple shorter ones (10×20 ns). Evaluation of sampling efficiency was based on the inspection of (χ_1, χ_2) plots calculated for different amino acids in the gorge of *Tc*AChE. The choice of *Tc*AChE over hAChE was led by the fact that there are 79 *Tc*AChE structures in the PDB as of April

2014, and 9 hAChE structures devoid of FAS-2. Comparison with crystal structures would therefore be more meaningful for *Tc*AChE. We chose a set of amino acids that occupy key regions in the gorge of *Tc*AChE, and which were observed previously to have different degrees of flexibility. We chose the active site residues His440, Trp84 and Glu199, the peripheral site Trp279, the bottleneck residue Phe330, and Tyr442, a residue involved in the dynamics of the backdoor region (Sanson et al., 2011). Comparison was made between two data sets of 20000 (χ_1, χ_2) points each. One is calculated from a 200 ns simulation ('long' data set) and the second from a set of 10 different 20 ns simulations ('short' data set), randomly chosen from the 20 trajectories we produced. To complete the analysis, sampling has been compared to all the 79 crystallographic structures of *Tc*AChE available in April 2014 in the PDB. Results are shown in figures 5.1, 5.2, 5.3, where the heatmap represents the simulation results and the white dots correspond to the conformations found in the different crystal structures.

We can observe that almost all the white dots are found over occupied regions of the heatmap both for long and short simulations. The most striking exceptions are the two conformers of His440 (χ_1, χ_2)=(100,90) and (χ_1, χ_2)=(120,95), which are only rarely found in the 10×20ns simulations set and never in the 200 ns simulation (figure 5.1). This particular conformational state is found in PDB structures 1vxr and 1gqr. In the first, it is induced by the presence of nerve agent VX at the active site that, as discussed above for tabun, features a conformational change of the catalytic histidine, induced by the presence of the OP. In the second, the conformational change is induced by the presence of rivastigmine in the acitve site (Bar-On et al., 2002).

Given the equivalency of the two simulation methods for reproducing conformational side chain eterogeneity in known crystal structures, we can now investigate how they perform in terms of sampling of the whole conformation space for each residue. Two characteristics are compared: the number of observable conformations and the separation between them. The general trend we observe is that using many short simulations we can find more and better separated occupied regions in the (χ_1, χ_2) plane.

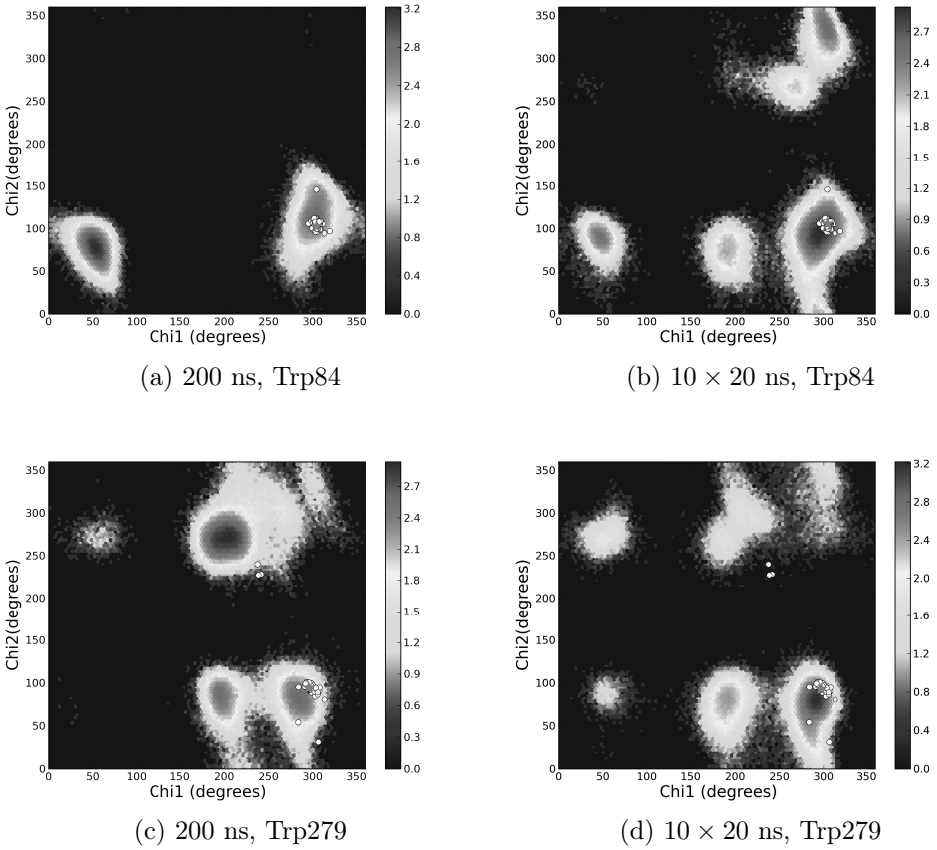
Histidine 440 and glutamate 199 His440 was found to display more conformations in the multiple short simulations dataset, as compared to the long one (figure

5.1 a-b). Based on occupancies, the first of these two conformers (300,300) appears the most probable. As mentioned already, the coverage by simulations of conformations observed in crystallographic structures is roughly equivalent in the short and long datasets. The short data set is however less biased by initial velocities, allowing to single out false positives in the long dataset. For example, the transition from peak (200,50) to peak (200,250) is not a proper local minimum of energy, but only due to a slow rotation rate in the 200 ns simulation. Concerning Glu199 (figure 5.1 c-d) is less straightforward. Its conformations are better separated in the long simulation data set, and the recovering of the different areas is different between the two simulations sets, contrarly to what is observed for all the other residues.

Tryptophan 84 and 279 An interesting feature is observed for the active site residue Trp84 in figure 5.2 a-b. In the long simulation, only two rotamers are observed, while in the many short simulation data set at least five major conformations appear. Movements of this residues were observed in MD simulations to be related to the flexibility of the Ω loop, requiring timescales longer than the 20 ns of the simulation. The high flexibility of this residue in MD simulations was already observed by Xu et al. (2008), and has been ascribed to the influence of crystal packing on the mobility of the Ω loop region. In the many short simulations approach the different velocities seeds accelerate this evolution by sampling different regions of the conformational space one at the time.

The case for the peripheral site Trp279, in figure 5.2 c-d, known to be one of the most flexible residues in the enzyme is different. In this case, the long and short simulations cover approximatively the same conformational space, with the exception of the (50,100) region for 10×20. The many short simulation approach in this case allows more precise separation the various conformations, that appear less broadend in the (χ_1, χ_2) plots.

Phenylalanine 330 and tyrosine 442 Phenylalanine and tyrosine demand a more careful observation. As mentioned above, we have to consider that their conformations are identical if rotated by 180 degrees around their χ_2 angle. We show an example in figure 5.3, where we compare dihedral angles for Phe330, the bottleneck residue and Tyr442, a residue involved in the opening of the backdoor. In both cases, no major differences were observed, between the 200 ns and the 10×20 ns



(a) 200 ns, Trp84

(b) 10×20 ns, Trp84

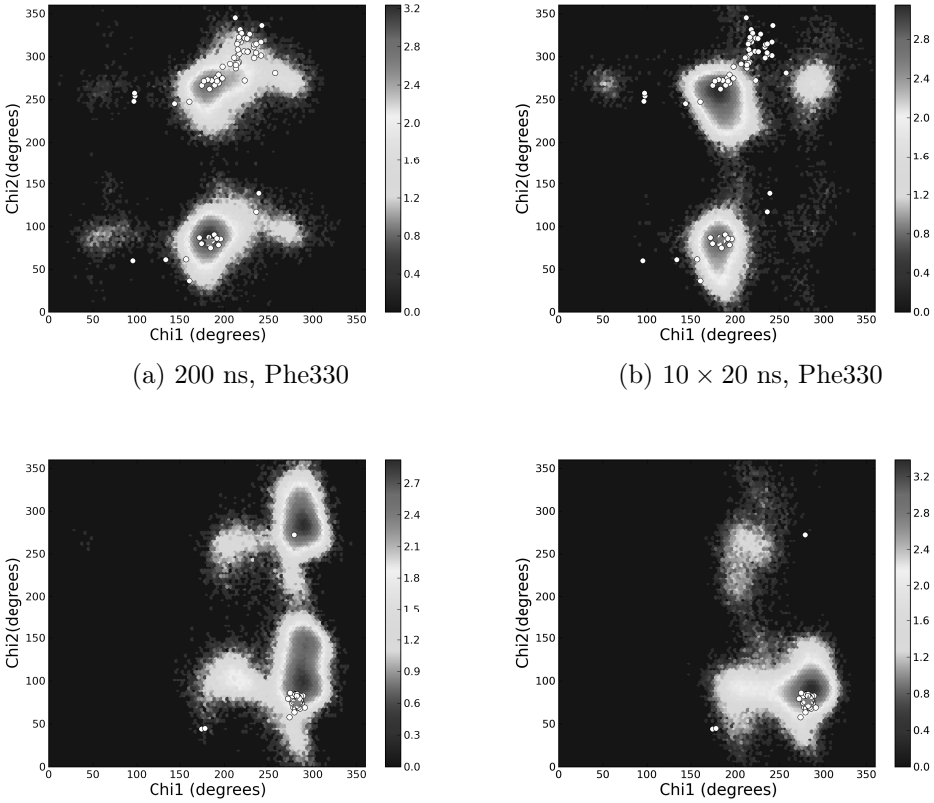
(c) 200 ns, Trp279

(d) 10×20 ns, Trp279

Figure 5.2: Dihedral angles for side chains of *Tc*AChE residues Trp84 and Trp279, calculated from a 200 ns MD simulation (a, c) and 10×20 ns simulations (b, d). The heat map represents the density of points on the plane, and white dots correspond to the conformations of the crystallographic structures of *Tc*AChE , both native and in presence of ligands, in the PDB. Native structure is indicated by the green star.

simulations, the region at peak (300,300) found in the 200 ns simulation for Tyr442 being identical to the one at (300,100) in the 10x20 ns simulation. The two methods of sampling account equivalently for the crystallographic conformations and the same regions are covered by the heatmap.

This qualitative analysis confirmed that the method of combining multiple short simulations is more adapted than a single long simulation to reproduce the conformational heterogeneity of amino acid side chains.



(a) 200 ns, Phe330

(b) 10 × 20 ns, Phe330

(c) 200 ns, Tyr442

(d) 10 × 20 ns, Tyr442

Figure 5.3: Dihedral angles for side chains of *Tc*AChE residues Phe330 andTyr442, calculated from a 200 ns MD simulation (a, c) and 10×20 ns simulations (b, d). The heat map represents the density of points on the plane, and white dots correspond to the conformations of the crystallographic structures of *Tc*AChE , both native and in presence of ligands, in the PDB. Native structure is indicated by the green star.

5.1.3 Convergence of the multiple short simulations method

After observing qualitatively that the approach of combining multiple different short simulations is more appropriate to sample side chain conformers, we investigated if the sum of short trajectories converges, or if adding more trajectories systematically and significantly improves the coverage of the conformational space. We developed an iterative method to evaluate the evolution of the surface of (χ_1, χ_2) plots as a function of the number N of trajectories used to generate the observed dihedral angle distribution. We consider the quantity I (equation 5.1), defined as the relative

increase in the occupied surface S_N of the (χ_1, χ_2) when adding a new trajectory to the data set.

$$I = \frac{S_{N+1} - S_N}{S_N} \tag{5.1}$$

The Python functions reported below (listing 5.1) were used for this analysis. The first function, named countOnes, is used to calculate how many pixels are occupied, *i.e* different from 0, in a (χ_1, χ_2) plot, thus determining the surfaces S_N. The second one, ControlSurface, iteratively combines the 20 trajectories in random order, measuring the surfaces S_N and S_{N+1} calculated from N and N+1 trajectories. If the increment I calculated with the two surface values is greater than a chosen threshold, the new trajectory is accepted as useful for additional sampling. For the next iteration, the program will consider $N = N+1$, and continue the analysis. After 20 iterations, the final value of N_{tot}, representing the number of useful trajectories, is stored. The same test is performed multiple times choosing the trajectories in random order, to ensure that the result is not dependent on the order in which trajectories were summed. Finally, the average N_{tot} obtained from the repetitions of the test is printed for each residue, along with the average value of the surface $S_{N_{tot}}$.

```
def countOnes(H):
    b = 0
    for x in H:
        for y in x:
            if y == 1:
                b+=1
    return b

def ControlSurface(filename):
    for start in head:
        print start
        modX,modY = np.loadtxt(filename+start, usecols=(0,1), unpack=
            True)
        B, xmod, ymod = np.histogram2d(modX,modY,bins=(360,360), normed
            = False)
        model = ["a","b","c","d","e","f","g","h","i","j","k","l","m","n
            ","o","p","q","r","s","t"]
        model.remove(start)
        counter = 1
        for s in range(0,19):
            ind = random.choice(model)
            model.remove(ind)
            x,y = np.loadtxt(filename+ind, usecols=(0,1), unpack=True)
```

```
            H, xed, yed = np.histogram2d(x,y,bins=(360,360), normed =
                False)
            BinH = H > 0
            BinB = B > 0
            CompareH = BinH*I
            CompareB = BinB*I
            Result = CompareH - CompareB
            threshold = countOnes(B)*float(cutoff)/100
            B +=H
            if countOnes(Result) > threshold:
            counter += 1
        counters.append(counter)
        surface.append(countOnes(BinB))
    return (np.average(counters),np.average(surface))
```

Listing 5.1: Python code for the surface convergence test

The output of the test is represented in figure 5.4. Each point represents, for a single residue[1], and the average surface S occupied on its (χ_1, χ_2) plot as a function of the average number of trajectories which passed the convergence test. 15 trajectories would be accepted with a threshold of 1% (panel A) and 12 if we chose a threshold of 2% (panel B). The sum of 10 trajectories satisfies a threshold condition between 2% and 5%, *i.e.* a situation between panels B and C. A rough estimate indicates that a single conformational island accounts for 5% of the total surface. The choice of 10x20 ns simulations[2] thus fully samples the conformational space of each residue[3]. Interestingly, the number of trajectories required to converge does not depend on the residue considered, but only on the chosen threshold.

[1]We analyzed the convergence of the sampling for the aromatic residues in the gorge: Tyr70, Trp84, Phe120, Tyr121, Trp233, Trp279, Phe288, Phe290, Phe330, Phe331, Trp432, Tyr442

[2]Used to compare with a 200 ns simulation

[3]This conclusion stands for residues with dihedral orders up to 2 up to χ_2, but has to be demonstrated for higher dihedral orders

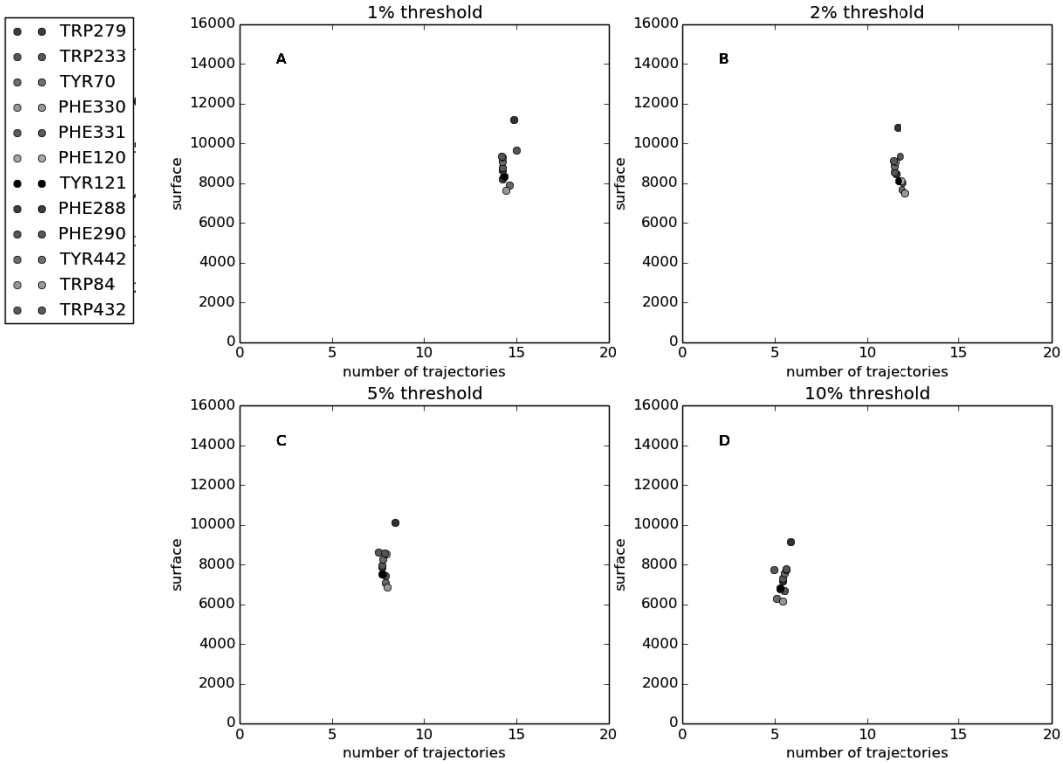
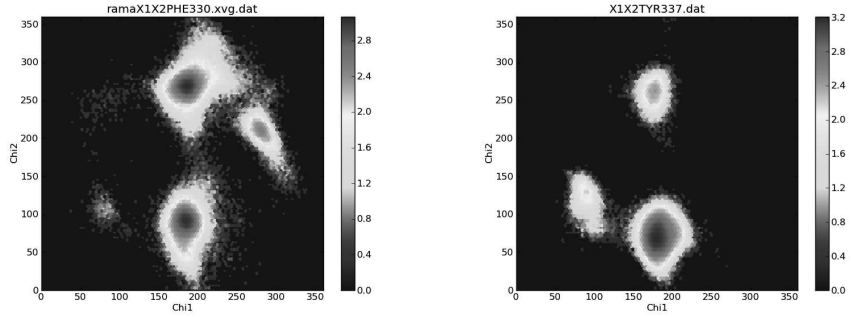


Figure 5.4: Results the convergence test for the many short simulations method. Each panel represents the results for a different threshold value. Each dot represents, for a single residue, the average surface covered for a residue as a function of the average number of trajectories required to obtain it. We can see that the cutoff choice provides similar results, independently from the kind of residues used.



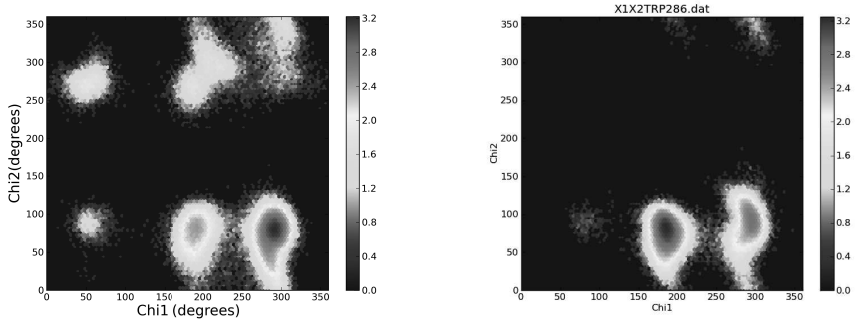
(a) *Tc*AChE-Phe330 (b) hAChE-Tyr337

Figure 5.5: Dihedral angles for Phe330 (*Tc*AChE) and Tyr337 (hAChE) as assessed in 10x20 ns MD simulations. The heat map represents the occurrence of each dihedral pair in logarithmic scale.

5.1.4 Comparison of human and *Torpedo californica* AChE dynamics

The comparison between human and *Torpedo californica* AChE is fundamental to determine if, using *Tc*AChE as a model for drug design, we eventually miss some fundamental conformational or dynamical differences that could influence the design of molecules targeting hAChE. It is important to remember that the gorge environment in the two species is very similar, with the only difference being residue Phe330 in *Tc*AChE that replaces Tyr337 in the human enzyme. This difference could account for subtle alteration of ligand binding modes, observed for example by Cheung et al. (2012) for donepezyl. They observe that the nitrogen of the piperidine of donepezyl in hAChE interacts with Tyr337 through an hydrogen bond with a water molecule. In *Tc*AChE, this interaction occurs with Tyr121, on the opposite side of the bottleneck, leading to a different orientation of the piperidine in the two species. We decided to observe whether or not this different binding mode could be due to a difference in side chains dynamics. The analysis is based on the comparison of 10×20ns simulations for both hAChE and *Tc*AChE.

At first, a clear difference is observed (figure 5.5), as expected, for the flexibility of Phe330 in *Tc*AChE with respect to Try337 in hAChE. Conformations of Phe330 accessed *in silico* are more numerous and broadened than those of Tyr337. Also, Tyr337 does not assume a $\chi_1 > 200$ in contrast to Phe330. The size of each



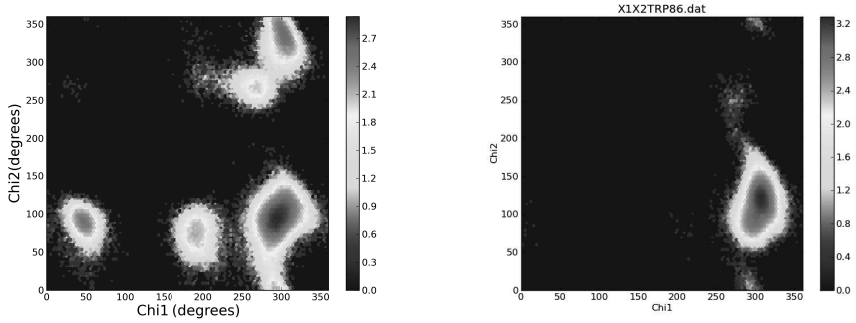
(a) *Tc*AChE-Trp279 (b) hAChE-Trp286

Figure 5.6: Dihedral angles for Trp279 (*Tc*AChE) and hAChE Trp286 (hAChE) as assessed in 10x20 ns MD simulations. The heat map represents the occurrence of each dihedral pair in logarithmic scale.

conformational island provides an idea of the stability of this particular dihedral conformation: the smaller the area, the more rigid the conformation. We can explain this difference by observing that the phenoxyl group of Tyr337 forms a hydrogen bond with the phenoxyl of its neighbor Tyr124. An analogous tyrosine is found also in *Tc*AChE (Tyr121) but in the absence of a hydroxyl group for Phe330, no hydrogen bond can be formed, leaving thus more freedom to Phe330, which can explore a larger conformational space.

As discussed above, one of the most flexible residues in AChE is the peripheral site tryptophan. Observing the results for Trp279 in *Tc*AChE and Trp286 in hAChE simulations (figure 5.6) we can see that they assume the same 5 main conformations, but their relative distribution is different. A sixth conformation is observed only for Trp279 at (50, 270). Trp279 tends to be found more frequently in its native crystallographic position (χ_1, χ_2)=(300,90), represented on the heatmap by a star, while the most probable position for Trp286 is found at (χ_1, χ_2)=(180, 90). Also, the two conformational islands at the top of the plots ($\chi_2 > 200$) appear more populated for *Tc*AChE than for hAChE. As discussed in the previous chapter, the reactivators we developed (KM297 and JDS207) target an alternative conformation of the peripheral site tryptophan, which was observed in crystallography only for *Tc*AChE. This difference in the flexibility of the residue in the two species, as observed in MD simulations, has to be taken into account carefully in the design process.

Another fundamental residue for the functioning of the enzyme is Trp84 (Trp86



(a) *Tc*AChE-Trp84 (b) hAChE-Trp86

Figure 5.7: Dihedral angles for Trp84 (*Tc*AChE) and Trp86 (hAChE) as assessed in 10x20 ns MD simulations. The heat map represents the occurrence of each dihedral pair in logarithmic scale.

in hAChE), which is involved in the stabilization of the substrate and in the traffic of reaction products. We can observe in figure 5.7 that Trp86 in hAChE remains in its native conformation. Even if its χ_2 angle can span a bit more than 100 degrees, its χ_1 angle is fixed in its initial position. On the other hand, Trp84 in *Tc*AChE can move around its χ_1 angle, assuming five main conformations, the most interesting being the one at (χ_1, χ_2)=(50,70) and corresponding to a configuration in which the backdoor channel is completely open, as observed in MD simulations (Xu et al., 2010; Sanson et al., 2011). The different dynamics of the active site tryptophan between hAChE and *Tc*AChE in particular is intriguing. No major structural differences are present in this region of the enzyme between the two species, but the hydrogen bond between Trp86 and Tyr449 appears to be more stable in hAChE than in *Tc*AChE. There maybe some backbone movements of the Ω-loop that could differ between the two species, leading to different dynamics on the short 20 ns timescale.

We can conclude that hAChE and TcAChE present key dynamical differences in residues which are fundamental for ligand binding and trafic in the gorge, like Trp84, Trp279 and Phe330[4]. Their different behavior has to be taken into account when designing ligands targeting hAChE.

[4]Trp86, Trp286 and Tyr337 respectively in hAChE

5.2 Generation of receptors library

Once we verified that the ensemble of 10 short trajectories is suitable for modeling side chain dynamics, we searched for a method to produce a receptor library for flexible docking, based on the dihedral angles distributions calculated from the simulations. At first, we looked for methods to find the most probable rotamers for a set of residues of interest. These rotamers have then been used to generate the docking receptors. Attention has been paid to possible clashes that could arise between neighboring residues from this generation method and to the role of water molecules in the process.

We have chosen to work with *Tc*AChE instead of hAChE for a first validation of the method. Even if we have shown that the two species have important dynamical differences, the choice of *Tc*AChE is justified by the fact that more structures are available in the PDB, so that the validation of our results with the available structures can be more conclusive.

5.2.1 Extraction of the most probable side chain conformations from MD simulation trajectories

Data sets from 10×20 ns trajectories are stored by Gromacs to a file containing (χ_1, χ_2) pairs on each line. They are easy to be processed with a Python code, in particular by creating Numpy N-dimensional arrays, with N being the maximum dihedral order for each residue. We tested two methods for extracting the most probable conformations for the residues chosen to be flexible in the subsequent docking phase.

The first method is restricted to N=2, because the usage of image processing consumes too much computational resources for arrays of more than 2 dimensions. At first, we populate a 2-dimensional histogram with the different (χ_1, χ_2) pairs for each residue. We chose a bin size of 3°. We consider that this choice of binning is fine enough. For example, for a tyrosine rotating by an angle of 3° around its χ_1 angle, the oxygen of the phenol group moves by 0.3 Å only. Once the histogram is filled, we apply a Gaussian filter to smothen the peaks and finally search for the position of local maxima with a photographic filter from the SCikit-NDImage module (van der Walt et al., 2014). An example of the output of this function is

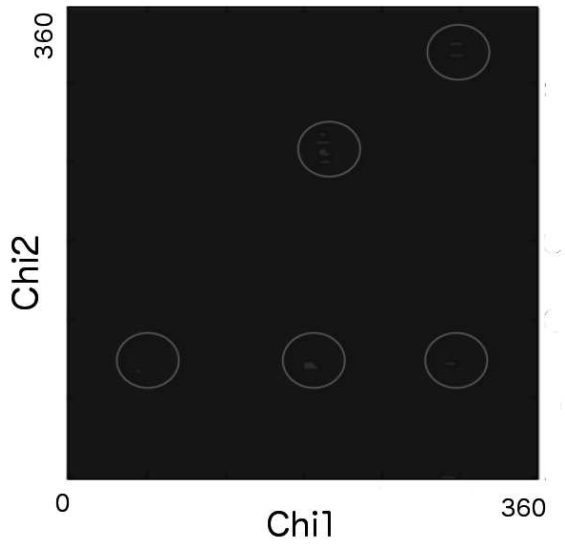


Figure 5.8: Peak coordinates for *Tc*AChE residue Trp279, as found from our code. Red dots represent the coordinates of all the maxima found by the program. For each circle, only the coordinates of the highest peak are considered.

shown in figure 5.8 for the peripheral site residue Trp279. The program then simply writes the coordinates of each of the maxima to a file that can be passed as an input to the receptor generation program.

The second method is easier and therefore less heavy. It is also not limited to residues with only two dihedral. At first we generate a normalized N-dimensional array containing the distribution of observed χ_N angles. A normalized array can be considered as a probability distribution $P(\chi_1, \chi_2)$ for the occurrence of a (χ_1, χ_2) pair during the MD simulations. We then generate random pairs following this distribution by using a rejection sampling algorithm. Figure 5.9 shows an example of 100 random points generated following the simulated dihedral angles distribution of Trp279.

```
def randomHist(pdf):
    normPdf=pdf/pdf.sum() #normalize
    check = False
    while check==False:
        chi1= np.random.randint(0, 120)
        chi2= np.random.randint(0, 120)
        z = np.random.uniform(0,1)
        if z<normPdf[chi1,chi2]:
            check = True
```

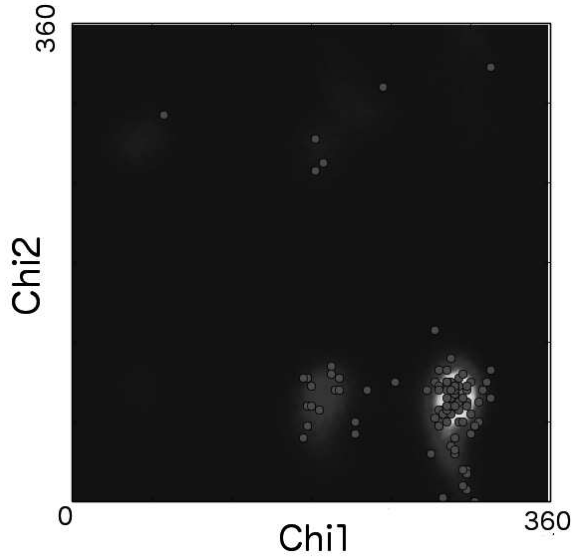


Figure 5.9: Results for the rejection sampling method, applied to *Tc*AChE residue Trp279. 100 green dots are generated following the heat map as a distribution of probability.

```
    return 3*chi1,3*chi2
```

Listing 5.2: Python code for the rejection sampling method.

The function randomHist (listing 5.2) for the two-dimensional case, generates a random (χ_1, χ_2) pair using the array as a distribution of probability.

5.2.2 Modification of *Tc*AChE structure to generate receptors

Two codes, described in detail in appendix A, were developed to modify the side chains of *Tc*AChE residues according to dihedral pairs obtained with either the peak finder or the rejection sampling method. Both of them are able to produce receptor structures, but while the peak finder implies a combinatory approach to the problem, which means the number of receptors grows exponentially with the number of residues included[5], the rejection sampling allows the user to chose the number of output structures. The rejection sampling method is faster and more easily adaptable, especially to the case of side chains with dihedral order larger than

[5] $N_{receptors} = \prod_{i=1}^{n} X_i$ where X_i is the number of possible peaks per residue.

2. We produced, using the rejection sampling method, 300 structures and performed a first validation of the methodology. Results for some ligands are reported below (section 5.4).

5.3 Validation of the method: docking routines

Once we verified that the method for the production of the library is reliable, a docking routine has been created to use the library conformers as a set of receptors and validate the approach. The validation consisted in evaluating the capability, for our method,to reproduce the crystallographic structures of the *Tc*AChE-ligand complexes present on the PDB, in particular the complexes in which ligand binding occurs to an alternative conformation of the enzyme. For this first validation, we chose to generate 300 receptors with the rejection sampling method, considering as flexible residues whose side chain conformation are heterogeneous in crystal structures: Trp86, Trp279, Phe288, Phe330, and His440.

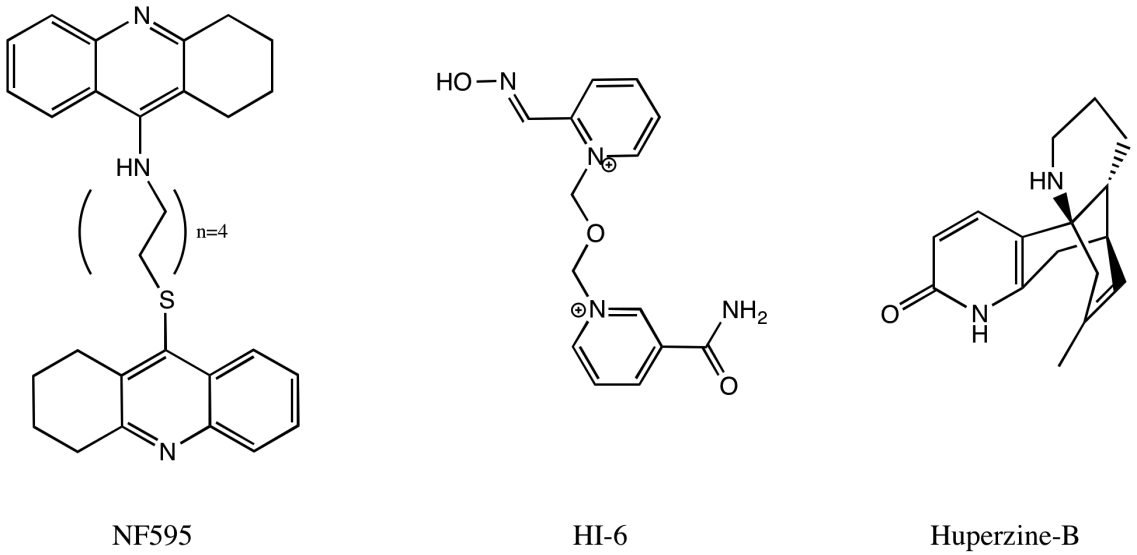
Our method of receptor generation is completely independent from the choice of the docking program, and the adoption of AutoDock Vina lays in its large popularity between the users and in the simplicity of its usage. It has also been observed that the scoring function of Vina is the most appropriate to describe docking in the gorge of AChE[6].

First, we used receptors generated by our method to screen for ligands of which the structure in complex with ACHE was known. After downloading the component definition files of each of these ligands in mmCIF format, we processed them using phenix.elbow to generate a file in the proper PDB. The ligand pdb files were then processed with the script prepare_ligand4.py from the AutoDockTools set, to create a pdbqt ligand that can be provided to AutoDock Vina as input.

A configuration file for docking experiment was created through the AutoDock plugin for PyMOL. The docking box was centered at the middle of the active site gorge and its dimension was chosen so that the box includes every conformer of the flexible residues.

The program performs a single docking experiment of each of the ligands against each one of the receptors. A single flexible docking experiment is in this case trans-

[6]W. Song and Y. Xu, unpublished



NF595 HI-6 Huperzine-B

Figure 5.10: Molecules used fot validation

formed in a series of a few hundreds rigid dockings, one for each receptor in the library. Outputs are saved to a single folder, bearing the ligand name, and that contains all of the output log files and structures of the docking experiments.

5.4 Validation of the method: analysis of the results

The raw output of the docking routine consists in a series of separated ligand structure files, each of them having a code name referring to the specific receptor they were docked to. The first operation performed in the analysis is to sort them following their docking binding energy. A text file containing the affinities and the name of each structure is created and serves as a reference for the analysis process.

In order to simplify the analysis of the outputs, each of the solutions, ranked from the highest binding energy to the lowest, are combined in a multistate pdb file containing the ligand within the corresponding receptor, which can later be observed with molecular graphics programs. The analysis routine also calculates the binding energy distributions of the solutions and their RMSD with respect to the lowest energy solution, which are then useful for interpretation of the results. We report here the analysis of three molecules (figure 5.10) that led to representative results: the bifuncitonal inhibitor NF595 (Savini et al., 2003), the reactivator HI-6 (Oldiges

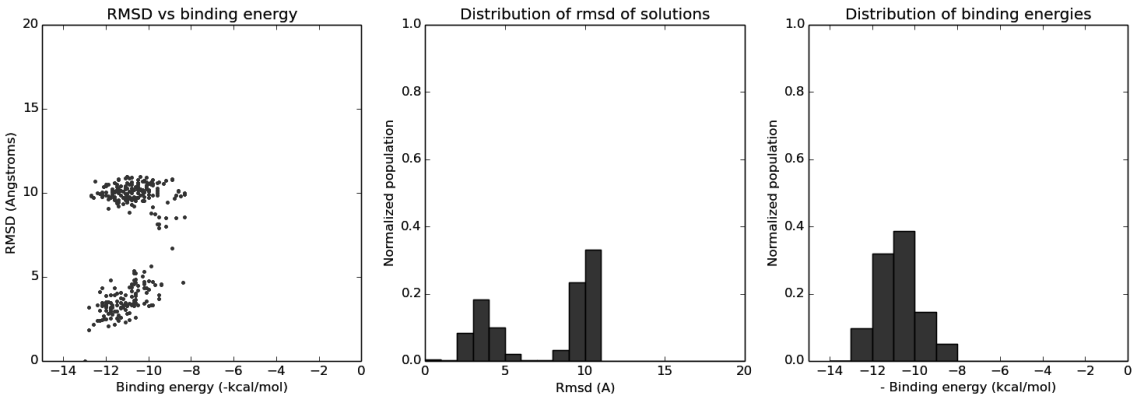


Figure 5.11: Docking results for the bifunctional inhibitor NF595 within 300 *Tc*AChE structures generated with the rejection sampling method. We observe how the resulting conformation presents two main spatial arrangements, with the same docking binding energy.

and Schoene, 1970) and the natural inhibitor Huperzine-B (Liu et al., 1986).

5.4.1 Quantitative analysis of docking solutions

For the analysis process, the solution presenting the lowest binding energy is chosen as a reference. This reference solution is used to calculate ligand RMSD, a value that is probably the most simple description of conformational differences between the docking solutions. The ranked solutions for each ligand are represented on a cartesian plane, as a function of their binding energy and their RMSD from the best solution. We also calculated the distributions of binding energies and ligand RMSD in order to have a quantitative evaluation of the homogeneity of docked solutions. These data were used to generate a scatter plot, in which they can be clustered in families, and two histograms, containing the normalized number of solutions per value of RMSD and binding energy.

The first concerns the bifunctional inhibitor NF595 (figure 5.11). We can observe that the ligand takes two different major spatial conformations, which share the same range of binding energy value. The fact that the two conformations account for half of the solutions is easily related to the similarity of the two extremities of the molecule that can be docked upside-down without significant change in the binding energy.

A smaller, non flexible ligand is the AChE inhibitor Huperzine-B (HUB, figure

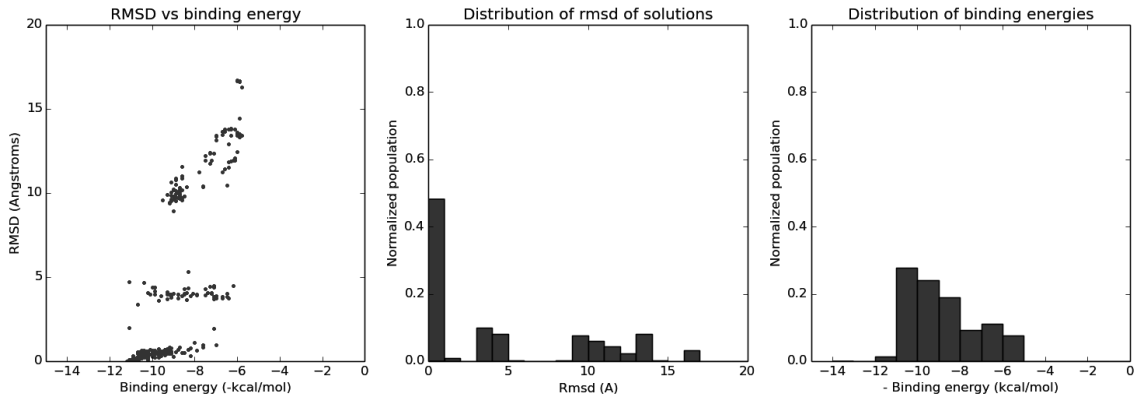


Figure 5.12: Docking results for Huperzine-B within 300 *Tc*AChE structures generated with the rejection sampling method. 50% of the solutions have an RMSD from the lowest energy solution of less than 1 Å, meaning that the binding site and the orientation of this molecule are not modified by the conformational changes of the side chains of most of the chosen flexible residues.

5.12). In this case, the binding conformation of the ligand is better defined than for NF595, as can be deduced from the fact that around 50% of the solutions have an RMSD lower than 1 Å with respect to the lowest energy one. This means that, even if the conformation of the receptor changes, the molecule binds with a large probability at the same place and mostly in the same orientation.

A third case is the reactivator HI-6, which also features two similar moieties and has a more flexible structure than Huperzine-B. These two characteristics can help in interpreting the results of figure 5.13. In fact, the RMSD values are widely spread without any particularly favorable orientation, while the binding energy values for this molecule fall within a narrow window. We can say that the flexible structure of the ligand, along with the aromatic nature of its pyridiniums rings, that can interact with most of the residues of the gorge wall, helps in always finding a proper orientation inside the gorge with a reasonably low binding energy.

5.4.2 Structural analysis of docking solutions

The goal of generating multiple receptor structures was to improve flexible docking procedures for ligands for which a conformational change in the gorge was revealed in the crystal structure of the complex. This structural validation is only preliminery, since it has beeen carried on only for a few molecules. In particular, we would need

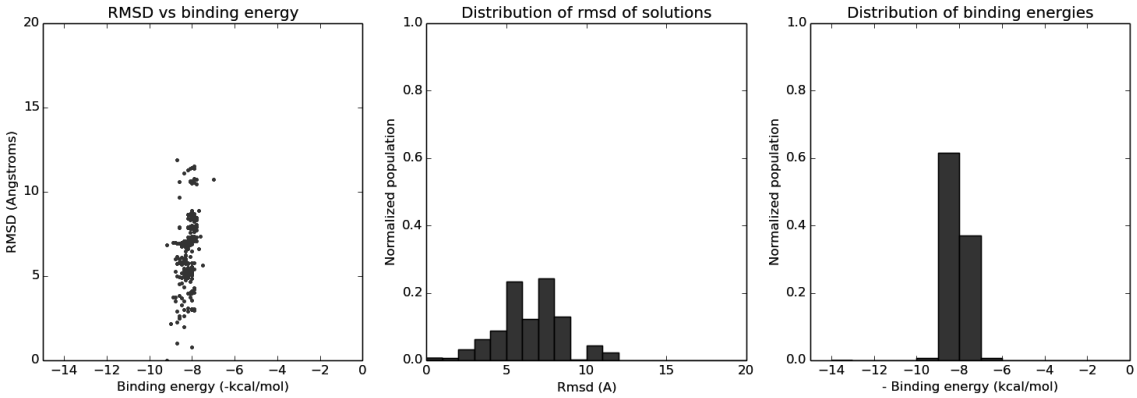
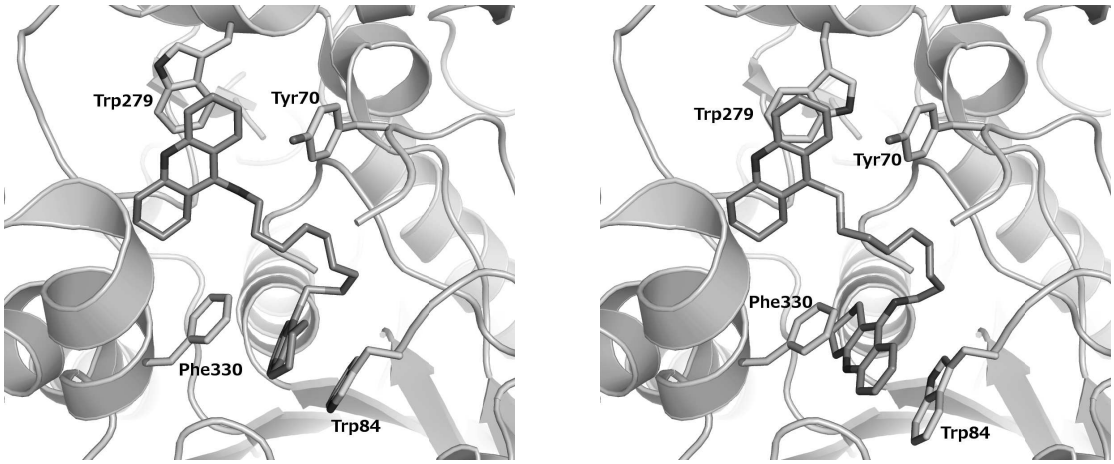


Figure 5.13: Docking results for the reactivator HI-6 within 300 *Tc*AChE structures generated with the rejection sampling method. Solutions present a uniform spatial dispersion, while the calculated binding energy is the same for all of them.



(a) Solution cluster A (b) Solution cluster B

Figure 5.14: Representative conformations for the two solution clusters for NF595. They are very similar, but the docked molecule is found in two opposite orientations, with the sulfur atom either at the PAS or at the bottom of the gorge. Images are in the same orientation.

to compare our output also with AutoDock Vina flexible docking results, and not only with the available crystal structures.

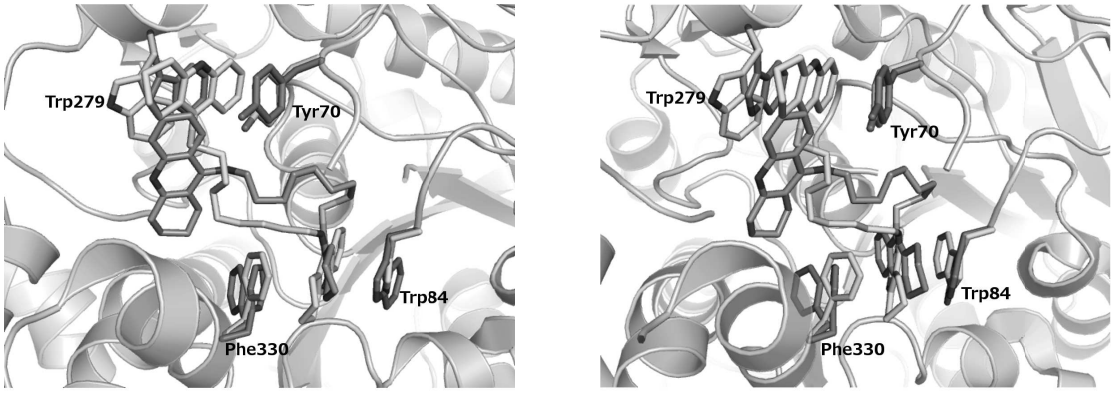
For NF595, we obtained two main conformational clusters of the docked solutions (figure 5.11). In the two panels if figure 5.14, containing a representative solution for each of the clusters, we can see that the RMSD separation between the two families is due to the upside-down orientation of the ligand, with either the tetrahydroacridine

bearing the sulfur atom at the PAS of the enzyme or in the CAS. If we compare the two structures, we can observe that the tetrahydroacridine at the active site is always in aromatic stacking in between Trp84 and Phe330. The one at the peripheral site is also found in aromatic stacking with Trp279. The relative similarity in the conformation of the solutions conforts the reliability of the method in finding binding modes for this type of ligand.

We can compare the solution from one of the clusters with the crystallographic structure of the complex *Tc*AChE-NF595 (pdb code 2cek). The superimposition of the docking output and the crystallographic structure is shown in figure 5.15. The main features of the binding mode at the active site are well reproduced by the docking solution, in particular the orientation of Phe330, found in an alternative conformation, and Trp84, which correctly maintains its native orientation. The tacrine appears to be correctly sandwiched between these two residues, leading to a correct orientation of the unsaturated cycle. At the peripheral site, however, the docking results cannot be superimposed to the crystallographic structure. The conformation of the ligand is different, because the different displacement of Trp279 in the receptor with respect to the alternative conformation observed in the crystallographic structure prevents the formation of a double aromatic stacking involving Trp279 and Tyr70. The conformation of Trp279 observed in the crystallographic structure is also present in the library, but to reproduce the aromatic sandwich, we would need an exact parallelism between the peripheral site tryptophan and Tyr70. However, a divergence of 3 degrees from the crystallographic position, due to the binning used for the dihedral generation, would be enough to avoid a perfect parallelism between Trp279 and Tyr70. This problem could probably be avoided by an accurate minimization of the docking solutions, which has not been included in our experimental procedure for the sake of reducing computational time.

Analysis of binding modes of HI-6 (not shown) has shown that this reactivator is docked in the active site, rather than assuming the position observed in crystal structure, half way through the gorge. The AChE inhibitor Donepezyl correctly binds with its active site moiety, but again the peripheral site part of the moelcule is not in the same orientation than in the crystallographic structure.

None of the analysed structures was comparable with the crystal structure. A comparison including Vina flexible docking results would be necessary, in order to asses the utility of our receptors generation method.



(a) (b)

Figure 5.15: Comparison between docking solution from cluster A and crystal structure 2cek (Colletier et al., 2006b). The docked ligand is in orange, flexible receptor residues in green and the crystallographic structure in cyan (ligand) and magenta (residues). While the two structures present the same binding mode in the active site, between Phe330 and Trp84, they differ in the conformation of both the ligand and Trp279 at the peripheral site. Panels (a) and (b) are rotated by 20 degrees around the vertical axis.

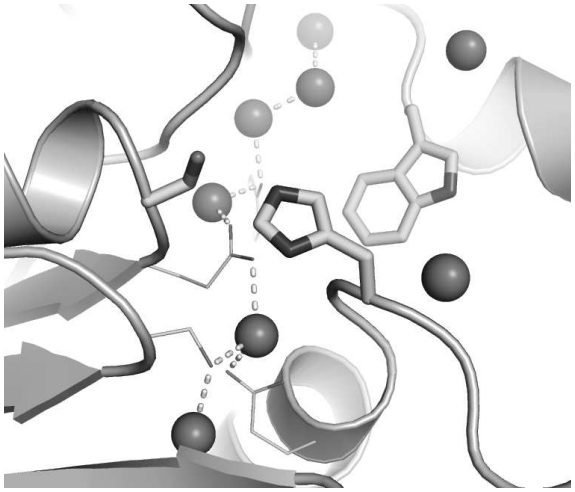


Figure 5.16: Conserved waters in the gorge of *Tc*AChE (Koellner et al., 2000) included in the docking receptors library.

5.4.3 The role of water molecules

The results described in the previous section were obtained without any water molecules in the gorge of AChE. We know, however, that some of these molecules are highly conserved (Koellner et al., 2000) and actively participate in ligand binding, so we include them in the initial model for receptor generation. The docking procedure was the same as reported before and data were processed in the same way.

As observed above, there was no clear definition of a favored conformer for HI-6 in the set of receptors (figure 5.13). The presence of water molecules inside the active site, while not improving the binding energy score, restrict the orientation of the ligand, leading to a single conformer accounting for more than 80% of the solutions (figure 5.17). While having an impact on the statistical distribution of solutions, the presence of water molecules did not improve the prediction of crystal structures, since again we could not reproduce them for any of the ligands.

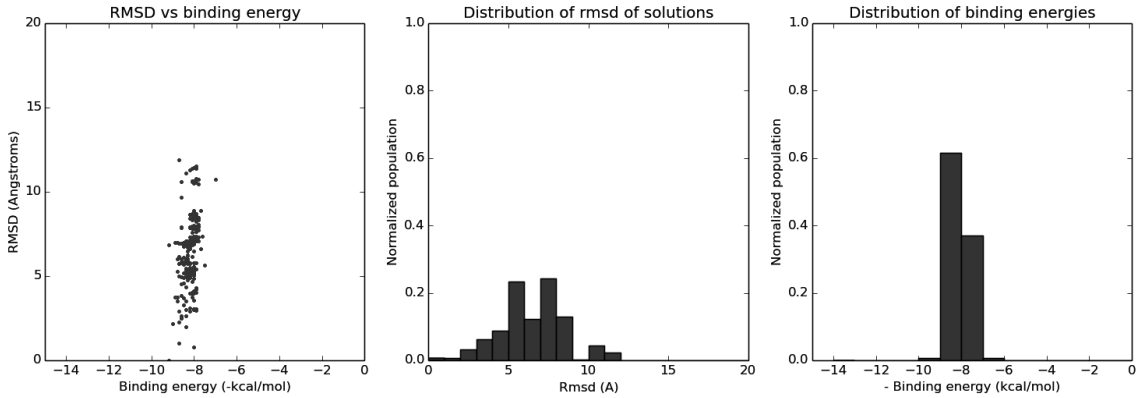
5.5 Conclusions

The comparison between the sampling of side chain conformers in one long or many short simulations has shown that the strategy of producing 10 trajectories of 20 ns each is appropriate to completely explore the space of (χ_1, χ_2) orientation for each residue.

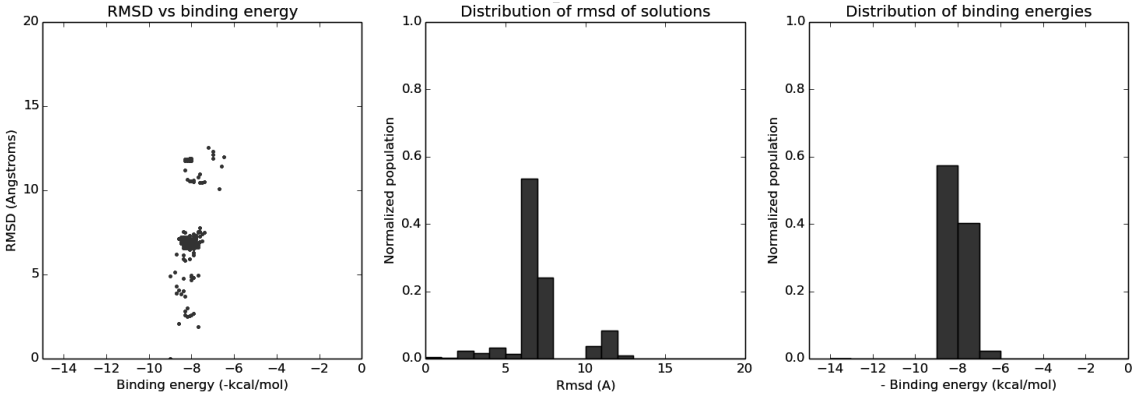
This strategy of simulating the dynamics of the enzyme was used to compare the conformational heterogeneity of some key residues in the gorge of hAChE and *Tc*AChE, and to produce a library of receptors for flexible docking.

The comparison of side chain conformational heterogeneity obtained from MD simulation for hAChE and *Tc*AChE has shown that the human enzyme has a reduced conformational heterogeneity with respect to the torpedo one both at the peripheral site and in the active site. The striking difference of the dynamics of Trp84 (*Tc*AChE) and Trp86 (hAChE) suggests some different behavior concerning the traffic of reaction products. The peripheral site tryptophan of hAChE, on the other hand, presents the same set of conformations, only with different relative occurrence during the MD simulations.

The production of receptor libraries was tested with two methods, a combination



(a) No waters



(b) Waters

Figure 5.17: Docking results compared for HI-6 within 300 *Tc*AChE structures generated with the rejection sampling method, with and without conserved water molecules in the receptor. The presence of water molecules leads to the exclusion of some conformational regions for the ligand, as observable from the sharper distribution of RMSD values.

of peaks conformers and a rejection sampling method. The combination of peaks has some major drawbacks. It is computationally intensive, not adapted to be ported to dihedral orders higher than 2 and does not consider the fact that a conformation is not sharply defined, but can be broadened over a large region in the (χ_1, χ_2) plot. For this lack of intrinsic flexibility, when checking for clashes a lot of conformations are excluded and lost in the final sampling. The rejection sampling method works differently. It can handle a large number of residues, with no limitation in dihedral order. Using all the information in the simulation data sets, rather than only the peaks in the dihedral distributions, helps in accounting for the spreading and relative

population of different conformers.

A docking method was implemented to test the efficiency of this new approach of sampling side chains conformers. Screening all the receptors in our library with a standard rigid docking procedure, it requires the same computational time as a standard flexible docking experiment with AutoDock Vina.

The analysis of docking results shows that ligand docking scores (binding energy) and conformational variability logically depends on the size and the degree of flexibility of each ligand. Small rigid molecules are more univocally docked than large, highly flexible ones. Calculated affinities can help in discriminating between molecules whose binding depends or not on a specific conformation of the side chains in the gorge. The comparison of the docking solution with crystallographic structure remains of limited success. We were not able to fully reproduce the binding of molecules like NF595, as discussed above in the results section. Other molecules, not presented in the manuscript like A8 (Rydberg et al., 2006) or ortho-7 (Hörnberg et al., 2010), binding *in cristallo* to the same alternative conformation of Trp279 as NF595, could also not be docked properly at the peripheral site. Even in cases where the receptor conformation was present in our library, docking procedure failed in properly positioning the ligand inside it. This discrepancy does not appear to be due to the lack of water molecules in the original test. Including conserved waters into the calculation didn't allow to better reproduce the crystallographic structures, but helped in restricting the number of solutions for some flexible ligands, like HI-6 as discussed above. As discussed for Trp84 in *Tc*AChE, the difference in binding mode observed in docking and crystallography could also arise from the fact that the observed crystallographic structure, affected by the crystal packing, might not represent the predominant conformation of the protein in solution, that might be better sampled by MD simulations.

We cannot conclude on the real efficiency of our docking method. Yet we believe that it is a clear improvement over methods using a single conformation of the enzyme or limited conformations found in crystal structures. The problem with our method is that the final receptor is rigid and that only the ligand can adapt to the receptor during the docking, which could lead to imperfect minimization of the energy of the system. A possible evolution of the methodology would be to use the approach of sampling side chains through MD simulation as a method to generate the initial poses directly inside the docking code. A genetic algorithm could then be

able to evolve both the ligand orientation and the angles of the side chains within the conformational space identified in the MD simulation.

Chapter 6

General conclusions and perspectives

We presented two main research projects in this manuscript. The first one, concerning the design and evaluation of bifunctional neutral reactivators, has allowed producing a first lead molecule, KM297, and an optimized derivative, JDS207. This process involved docking studies, molecular dynamics simulations and X-ray crystallography. The second one, concerning computational sampling of side chains conformations, aimed at producing a receptor library for flexible docking on *Tc*AChE. Results concerning the efficiency of the docking method were not conclusive, but the process allowed us to provide insight into the structural dynamics of hAChE and *Tc*AChE, and on the method for taking account this flexibility in the modeling.

6.1 Design and optimization of KM297

The design of KM297, chosing tacrine as the peripheral site binder has been followed by its successful synthesis(Kliachyna et al., 2014). The reactivation capability of KM297 has been tested *in vitro* and was proven to be more efficient than 2-PAM (the standard reference in literature) and HI-6, the last generation molecule to reach the market. Flexible docking studies of KM297 on human AChE have comforted the design hypothesis of binding to an alternative conformation of the peripheral site, targeting in particular aromatic stacking of the tacrine moiety between Trp286 and Tyr72. The complex hAChE-KM297 showed a conformation analogous to that observed with the bis-tacrine inhibitor NF595, which was the initial model used for the design of KM297.

Molecular dynamics simulations demonstrated the binding stability of tacrine at the peripheral site in the docked conformation. Analysis of the distance during simulation between the oxime and the phosphorous atom of an OP, VX in this case, has been used to determine an optimal linker length between tacrine and the reactivion function. A 4-carbon linker appears to be most appropriate to correctly position the oxime for reactivation, supported by *in vitro* tests, which showed that reactivation by KM297 is better than KM175, the analogue with a 5-carbon linker length.

We solved structures of KM297 in complex with both native and tabun-inhibited *Tc*AChE. The complex *Tc*AChE-KM297 has allowed to observetwo possible binding sites for the reactivator in the native enzyme, one at the peripheral site and another in the active site, with the oxime pointing towards the peripheral site. By contrast, when the catalytic serine was phosphylated by NEDPA, a surrogate forming the same conjugate as tabun, KM297 could not enter the active site. In this structure, only one molecule of KM297 has been observed in the gorge, bound at the peripheral site.

A series of new derivatives has been subsequently designed, to prevent binding within the active site. In particular, introducing a chlorine substituent at position 7 of the tacrine moiety has emerged as a chemically simple and structurally efficient solution. According to molecular docking, the presence of the chlorine in this particular position should create a steric clash preventing binding of the tacrine moiety in front of Trp84. The chlorinated derivative of KM297, named JDS207, has been synthesized and studied in detail. X-ray crystallography confirmed that JDS207 binds only at the peripheral site.

KM297 and JDS207 can be compared for their *in vitro* reactivation and inhibition properties. They both present good reactivation capability fo VX-inhibited hAChE, with a reactivation constant of $k_{r2} = 22mM^{-1}min^{-1}$ for KM297 and of $k_{r2} = 13mM^{-1}min^{-1}$ for JDS207. As a comparison HI-6, the last molecule to reach the market, has a $k_{r2} = 9mM^{-1}min^{-1}$. The IC50 of JDS207 for hAChE, (2.33 ± 0.18 μM) is 10 times larger than that of KM297 (0.25 ± 0.02 μM). The chlorination of the tacrine moiety thus decreased the global affinity of the molecule for non-inhibited AChE by limiting binding to peripheral site. Yet, it is not clear at this point if inhibition of the non-inhibited enzyme by a reactivator is a bad property for a reactivator. This should be determined during *in vivo* experimentation.

Computational studies are all based on models of hAChE, while the crystallographic structures presented were obtained with *Tc*AChE. Yet, there could be a serious bias between *Tc*AChE and hAChE structures due to non conserved residues across species (Cheung et al., 2012). For the sake of uniformity, we tried to obtain hAChE crystals that would allow solving the structure of hAChE complexes. We obtained crystals in conditions similar to those obtained by (Dvir et al., 2010). But these crystals could not be used for soaking of our reactivators, due to the high salt concentrations found in the crystallization conditions, in which KM297 and JDS207 precipitate.

6.2 Side chains dynamics and flexible docking

We assessed two simulation strategies with the aim of improving the characterization of side chains mobility in the gorge of AChE and build an optimal conformational library. Comparing one long (200 ns) with many combined short (10x20 ns) simulations, we concludes that both approaches are adequate to reproduce the conformational diversity observed in crystal structures. Yet, the multiple short simulations approach is more suitable to obtain more and better defined conformational islands as represented on plots of the (χ_1, χ_2) angles that characterize each conformation.

Therefore, we chose this latter approach (10x20 ns simulations) to compare the dynamic behavior of some key residues in the gorge of *Tc*AChE and hAChE. We were able to show that the conformational variability is reduced in the human enzyme with respect to the torpedo enzyme. For the aim of designing reactivators, the similarity of the PAS sampling between the two species was encouraging. In particular, we observed that the PAS tryptophan assumes the same conformations in both enzymes, though with different occurrence rates.

We developed two methods to generate a library of docking receptors from the dihedral data sets generated from simulations. Combining the various peak coordinates has not proven to be efficient, even if is conceptually easier. At first, it is complicated to extend this approach to dihedral angles of order larger than 2. Also, all the information included in the spread of (χ_1, χ_2) regions is lost by considering only one point per conformational island. Eventually, exclusion of clashes related to sub-optimal relative position of dihedrals can lead to the loss of an entire conformer.

Including less strict cutoff criteria and minimizing the output structures does not appear as a valid option.

Therefore, we tested a rejection sampling method, in which only one specific dihedral pair, rather than the whole conformational island, is excluded in case of a clash. This approach has the advantage of taking into account the relative occurrence of the different conformations, so that the most probable ones recur more often in the receptor library.

Once the receptor is prepared, we implemented an automated procedure to dock a set of ligands into them and verify the efficiency of the side chain sampling. Docking results show that ligand docking scores (binding energy) and conformational variability logically depend on the size and the degree of flexibility of each ligand. Small rigid molecules are more univocally docked than large, highly flexible ones.

For a given molecule the docked conformation is most often different from the crystallographic structure of the complex suggesting that some key features are missing in the receptor library. For example we could not reproduce the double aromatic stacking of the tacrine moiety of NF595 at the peripheral site of AChE despite the proper receptor conformation, with the rotation of Trp279, is present in the receptors library. This implies that the limiting factor lays in the docking procedure rather than the receptor library itself.

Adding conserved crystallographic water molecules helps in disambiguating docking results for highly affine and flexible molecules, such as HI-6, but this does not help in reproducing the crystallographic structures.

6.3 Perspectives

We have shown how the structure-based development of reactivators allowed us to obtain efficient molecules. In this work we presented molecules based on a tacrine PAS ligand, but many additional PAS ligands have been tested and will be. In particular, we will target the synthesis of molecules that enhance Blood-Brain-Barrier crossing like sugars and/or allow to better solubilization in aqueous medium. A better solubility will be of great advantage for *in vivo* experiments. Until now, preliminary surviving tests in a mouse model have not been conclusive, since the amount of KM297 we could inject to a mouse is limited by the solubility of the

compound [1]. Galenic formulation must be explored to improve the bioavailability of the molecules, for example by making chlorhydrate salts of them.

Crystallographic results have all been obtained with *Tc*AChE, rather than hAChE. As we have shown, the binding modes of ligands within the two enzymes can differ, due to the different dynamics and composition of the active site gorge. Ideally we will work exclusively with hAChE, and our effort to obtain well diffracting crystal, such as the ones by (Cheung et al., 2012), are still ongoing. We managed to obtain some crystals in the same space group, but they diffracted only up to 3.2 Å. Optimized crystallization conditions along with a wider screen of cryo protecting solutions, should allow us to push this limit farther, towards the current resolutions of *Tc*AChE crystals, i.e close to the 2 Å-limit. The current production rate of the hAChE-544stop mutant in CHO cells will yield sufficient amount of enzyme to perform exhaustive crystallization screening tests.

The docking pipeline has proven not to be adapted to take advantage of the improved specificity for side chain conformer generation. Probably the best way to benefit from this approach should be to include it directly into the docking program. Each conformation generated this way will then be used to produce an initial pose for the ligand that current docking algorithms are able to optimize in order to find a properly docked conformation for both the ligand and the enzyme. This big effort, however, goes beyond the scope of this thesis.

[1] G. Calas, unpublished results

Chapter 7

Resumé de la thèse en français

7.1 Introduction

7.1.1 L'acetylcholinesterase

L'acétylcholinestérase (AChE) est une des enzymes des plus rapides de la nature. Elle intervient dans la transmission du signal nerveux, en hydrolysant le neurotransmetteur acétylcholine (ACh) dans les synapses et jonctions neuromusculaires. Son efficacité catalytique est proche de la limite théorique imposée par la vitesse de diffusion des substrats et produits. Sa structure tridimensionnelle a initialement surpris par le fait que le site actif se trouve enfoui au fond d'une gorge étroite et profonde. Mais cette particularité génère un environnement optimal pour favoriser l'hydrolyse du substrat par la triade catalytique.

En raison de ce rôle, l'AChE est la cible de différentes familles de toxiques, notamment les neurotoxiques organophosphorés (OP). Cette large famille de molécules comprend des pesticides, comme le paraoxon, et des toxiques de guerre, comme le sarin, qui est redevenu d'actualité lors de la guerre civile syrienne. Les OPs inhibent l'AChE en formant une liaison covalente avec la sérine catalytique, et l'empêchent donc de jouer son rôle de nucléophile dans l'activité d'hydrolyse de l'ACh. Les conséquences physiologiques de cette inhibition sont mortelles pour l'homme. L'activité catalytique de l'enzyme peut être restaurée en utilisant les oximes, une classe de puissant nucléophiles qui sont capables de libérer la serine catalytique en déplaçant l'OP. Cependant, après une durée variable suivant la nature de l'OP, les complexes

AChE-OP subissent une réaction secondaire dit de vieillissement, après laquelle la réactivation par une oxime n'est plus possible.

Les études structurales des complexes AChE-OP-oxime ont montré que le positionnement l'oxime dans la structure n'est pas optimal pour la réactivation de l'enzyme inhibée par les toxiques organophosphorés. Nos travaux ont eu pour objectif de concevoir des oximes plus efficaces en adoptant la stratégie des réactivateurs bifonctionnels: la partie réactivatrice est connectée à un ligand connu pour se fixer au site périphérique de l'AChE.

L'étude des réactivateurs bifonctionnels a été le coeur de ce travail de thèse, et les résultats sont détaillés dans ce manuscrit.

7.1.2 Conception de médicaments et dynamique des proteines

La conception raisonnée des médicaments est un processus basé sur la connaissance de la structure des protéines cibles, qui permet de créer des molécules qui auront une action sur ces mêmes cibles. Plusieurs techniques existent, la plupart desquelles utilisent de larges ressources de calcul. Ces techniques présentent le défaut de reposer sur un nombre limité de structures cibles et donc de souvent négliger la nature dynamique des protéines.

Les protéines sont des macromolécules, qui existent dans plusieurs états conformationnels à un instant donné. Cette description, dite du paysage conformationnel, nous permet de comprendre que chaque protéine peut être considérée comme un objet dynamique, qui à l'équilibre peut passer d'un état conformationnel à un autre, avec des échelles temporelles allant de la picoseconde jusqu'à la seconde.

Pour cette raison, un large nombre de techniques existe pour caractériser la dynamique structurale des protéines à une résolution atomique. Toutes ces informations, obtenues notamment par cristallographie au rayons X, résonance magnétique nucléaire ou simulations par dynamique moléculaire, sont utilisées dans les travaux de conception de médicaments.

7.2 Resultats

7.2.1 Conception et optimisation d'un oxime bifonctionel

Dans le cadre de nos projets, plusieurs molécules ont été synthétises. Parmi elles, la molécule KM297, composée d'une tacrine comme ligand du site périphérique de l'AChE et d'une 2-hydroxypyridine aldoxime comme fonction réactivatrice, a donné les meilleur résultats dans les tests *in vitro*. Pour cette raison, elle a été l'objet des études structurales plus détaillées, pour identifier quel type de modification pourrait l'améliorer. La caractérisation par des méthodes computationelles a permis de montrer que la longueur du linker de carbones de la molécule KM297 était la meilleure pour bien positionner l'oxime dans le site actif de l'AChE inhibée par le VX. La structure cristallographique du complexe *Tc*AChE-KM297 illustre que la molécule présente deux sites de fixation dans la gorge de l'enzyme non inhibée. Le premier site de fixation est à l'entrée de la gorge au niveau du site périphérique. Mais la présence du partenaire symétrique dans le cristal impose une conformation qui empêche la pénétration de la partie réactivatrice dans la gorge. La deuxième molécule se trouve avec la partie tacrine positionnée dans le site actif, en interaction avec le Trp84.

L'observation de la structure de ce complexe nous a permis de concevoir et synthétiser un dérivé de KM297 ne se liant qu'au site périphérique de l'AChE, le JDS207. La molécule JDS207 présente en position 7 de la tacrine un atome de chlore, qui devrait créer un encombrement stérique lorsque la molécule se lie au fond du site actif par interaction avec Trp84. Cette hypothèse a été d'abord confirmée par docking moléculaire avant d'être validée expérimentalement par cristallographie. Les deux méthodes ont permis de mettre en évidence que le seul mode de fixation pour le JDS207 est au niveau du site périphérique, en interaction directe avec le résidu Trp279.

Toutes les structures cristallographiques présentées dans ce manuscrit ont été obtenues avec l'enzyme de Torpedo californica. Pour avoir des résultats plus applicables à l'homme, nous avons concentré nos efforts sur l'obtention de cristaux de l'AChE humaine. Les cristaux obtenus jusqu'à présent diffractent à trop faible résolution (3.2 Å) pour révéler le niveau de détail nécessaire à la conception de nouvelles molécules. De plus, le site actif est bloqué pour une molécule d'enzyme sur

deux en raison du type d'empilement cristallin. Ils se sont donc montrés inadaptés pour les études structurales de complexes AChE-réactivateurs. Mais les efforts se poursuivent pour obtenir des cristaux de qualité suffisante.

7.2.2 Developement d'un outil pour le docking flexible

L'étude en dynamique moléculaire des chaînes latérales des résidus aromatiques localisés dans la gorge de la *Tc*AChE nous a permis de montrer que la méthode la plus appropriée pour couvrir au mieux l'espace conformationnel est de combiner plusieurs trajectoires courtes, chacune avec des vitesses initiales différentes. Une étude plus détaillée nous a permis de montrer que 10 trajectoires était suffisantes pour efficacement échantillonner l'espace conformationnel de ces chaînes latérales.

Cette technique nous a d'abord permis de montrer que les chaînes latérales des résidus de la gorge ont une dynamique différente entre l'hAChE et la *Tc*AChE. Cette remarque est importante, car la plupart des informations structurales obtenues à ce jour sont sur la *Tc*AChE. Le mode de liaison des molécules dans les deux enzymes peut donc significativement varier en raison de cette hétérogénéité dynamique et conformationnelle.

Une bibliothèque de récepteurs couvrant l'ensemble de l'espace conformationnel des résidus du site actif a ensuite été générée à partir des données produites par les calculs de dynamique moléculaire. La bibliothèque a été testée avec tous les ligands de AChE pour lesquels une structure cristallographique du complexe *Tc*AChE/ligand est connue. Les résultats ont montré une claire dépendance de l'énergie de liaison estimée à la conformation de chaque récepteur. Toutefois, nous n'avons pas réussi à reproduire la structure cristallographique observée, du moins à l'aide du logiciel utilisé pour exploiter la bibliothèque. La présence ou non, lors du docking, des molécules d'eau très conservées dans les différentes structures cristallographiques de l'AChE n'a pas permis non plus de retrouver les complexes observés dans les structures cristallographiques. Mais leur présence permet de réduire le nombre de conformations pour un ligand donné et aide donc à mieux isoler la meilleure conformation prédite par docking.

7.3 Conclusions generales

En conclusion de ce travail, nous avons conçu des réactivateurs bifonctionnels avec un noyau tacrine comme ligand périphérique de l'AChE qui sont supérieures en terme d'activité *in vitro* à l'activité des molécules actuellement en commerce. Leur solubilité limitée en milieu acqueux sous leur formulation actuelle est toutefois un problème, et elle sera adressée dans la continuité de ces travaux. Deux pistes sont envisagés: 1) la conception de nouvelles molécules avec des ligands du PAS monochargés à base quinolinium, sachant que la présence d'une seule charge pourrait être suffisante pour permettre le passage de la barrière hémato-encéphalique 2) la formulation des ligands sous forme de sel de chlorhydrate ou mésylate, pour optimiser leur biodisponibilité. L'utilisation de l'enzyme *Tc*AChE est sans doute un facteur biaisant dans les études structurales. Pour cette raison, nous avons travaillé à l'obtention de cristaux d'hAChE similaires à ceux obtenus par (Cheung et al., 2012) à partir d'une forme d'hAChE tronquée au résidu 544. À ce jour, nous avons obtenu des cristaux identiques mais ne diffractant pas au delà de 3.2 Å, encore loin des 2 Å espérés. Les efforts poursuivent à l'aide d'une forme moléculaire strictement identique à celle utilisée par Cheung et al.

Malgré l'utilisation orginale de notre bibliothèque conformationnelle, notre méthode de docking n'a pas permis de reproduire de manière fiable les complexes observés dans les structures cristallographiques lorsqu'ils sont accompagnés de changements conformationnels. On observe, par contre, une très nette dépendance de l'énergie de liaison estimée vis-à-vis des différentes conformations du récepteur, ce qui prouve l'intérêt de la bibliothèque pour ne passer à côté d'une conformation pour laquelle le ligand a des interactions préférentielles. Une méthode pour bénéficier au mieux de notre bibliothèque conformationelle serait de directement inclure notre méthode de génération des récepteurs flexibles dans le logiciel de docking. Un algorithme moderne pourrait ainsi faire varier la conformation initiale du récepteur à partir des informations contenues dans la bibliothèque, pour obtenir des résultats plus représentatifs de la diversité conformationelle intrinsèque du récepteur.

Bibliography

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography*, 66(Pt 2):213–21.

Bajorath, J. (2002). Integration of virtual and high-throughput screening. *Nature reviews. Drug discovery*, 1(11):882–94.

Bar-Even, A., Noor, E., Savir, Y., Liebermeister, W., Davidi, D., Tawfik, D. S., and Milo, R. (2011). The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry*, 50(21):4402–10.

Bar-On, P., Millard, C. B., Harel, M., Dvir, H., Enz, A., Sussman, J. L., and Silman, I. (2002). Kinetic and structural studies on the interaction of cholinesterases with the anti-Alzheimer drug rivastigmine. *Biochemistry*, 41(11):3555–64.

Bartolucci, C., Perola, E., Cellai, L., Brufani, M., and Lamba, D. (1999). "Back door" opening implied by the crystal structure of a carbamoylated acetylcholinesterase. *Biochemistry*, 38(18):5714–9.

Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R., and Leslie, A. G. W. (2011). iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta crystallographica. Section D, Biological crystallography*, 67(Pt 4):271–81.

Berendsen, H. J. C., Grigera, J. R., and Straatsma, T. P. (1987). The missing term in effective pair potentials. *The Journal of Physical Chemistry*, 91(24):6269–6271.

Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81(8):3684.

Böhm, H.-J. (1998). Prediction of binding constants of protein ligands: A fast method for the prioritization of hits obtained from de novo design or 3D database search programs. *Journal of Computer-Aided Molecular Design*, 12(4):309.

Bourgeois, D. and Royant, A. (2005). Advances in kinetic protein crystallography. *Current opinion in structural biology*, 15(5):538–47.

Bourgeois, D., Vallone, B., Schotte, F., Arcovito, A., Miele, A. E., Sciara, G., Wulff, M., Anfinrud, P., and Brunori, M. (2003). Complex landscape of protein structural dynamics unveiled by nanosecond Laue crystallography. *Proceedings of the National Academy of Sciences of the United States of America*, 100(15):8704–9.

Bourgeois, D. and Weik, M. (2009). Kinetic protein crystallography: a tool to watch proteins in action. *Crystallography Reviews*, 15(2):87–118.

Bourne, Y., Kolb, H. C., Radić, Z., Sharpless, K. B., Taylor, P., and Marchot, P. (2004). Freeze-frame inhibitor captures acetylcholinesterase in a unique conformation. *Proceedings of the National Academy of Sciences of the United States of America*, 101(6):1449–54.

Bourne, Y., Radic, Z., Sulzenbacher, G., Kim, E., Taylor, P., and Marchot, P. (2006). Substrate and product trafficking through the active center gorge of acetylcholinesterase analyzed by crystallography and equilibrium binding. *The Journal of biological chemistry*, 281(39):29256–67.

Bourne, Y., Radić, Z., Taylor, P., and Marchot, P. (2010a). Conformational remodeling of femtomolar inhibitor-acetylcholinesterase complexes in the crystalline state. *Journal of the American Chemical Society*, 132(51):18292–300.

Bourne, Y., Radić, Z., Taylor, P., and Marchot, P. (2010b). Conformational remodeling of femtomolar inhibitor-acetylcholinesterase complexes in the crystalline state. *Journal of the American Chemical Society*, 132(51):18292–300.

Bourne, Y., Renault, L., and Marchot, P. (2014). Crystal Structure of Snake Venom Acetylcholinesterase in Complex with Inhibitory Antibody Fragment Fab410 bound at the Peripheral Site: Evidence for Open and Closed States of a Backdoor Channel. *The Journal of biological chemistry*.

Bourne, Y., Taylor, P., and Marchot, P. (1995). Acetylcholinesterase inhibition by fasciculin: Crystal structure of the complex. *Cell*, 83(3):503–512.

Brazzolotto, X., Wandhammer, M., Ronco, C., Trovaslet, M., Jean, L., Lockridge, O., Renard, P.-Y., and Nachon, F. (2012). Human butyrylcholinesterase produced in insect cells: huprine-based affinity purification and crystal structure. *The FEBS journal*, 279(16):2905–16.

Brooks, B. R., Brooks, C. L., Mackerell, A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M., and Karplus, M. (2009). CHARMM: the biomolecular simulation program. *Journal of computational chemistry*, 30(10):1545–614.

Brown, D. and Superti-Furga, G. (2003). Rediscovering the sweet spot in drug discovery. *Drug Discovery Today*, 8(23):1067–1077.

Brünger, A. T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, 355(6359):472–5.

Bui, J. M. and McCammon, J. A. (2006). Protein complex formation by acetylcholinesterase and the neurotoxin fasciculin-2 appears to involve an induced-fit mechanism. *Proceedings of the National Academy of Sciences of the United States of America*, 103(42):15451–6.

Bui, J. M., Radic, Z., Taylor, P., and McCammon, J. A. (2006). Conformational transitions in protein-protein association: binding of fasciculin-2 to acetylcholinesterase. *Biophysical journal*, 90(9):3280–7.

Burmeister, W. P. (2000). Structural changes in a cryo-cooled protein crystal owing to radiation damage. *Acta Crystallographica Section D Biological Crystallography*, 56(3):328–341.

Carletti, E., Colletier, J.-P., Schopfer, L. M., Santoni, G., Masson, P., Lockridge, O., Nachon, F., and Weik, M. (2013). Inhibition pathways of the potent organophosphate CBDP with cholinesterases revealed by X-ray crystallographic snapshots and mass spectrometry. *Chemical research in toxicology*, 26(2):280–9.

Carletti, E., Li, H., Li, B., Ekström, F., Nicolet, Y., Loiodice, M., Gillon, E., Froment, M. T., Lockridge, O., Schopfer, L. M., Masson, P., and Nachon, F. (2008). Aging of cholinesterases phosphylated by tabun proceeds through O-dealkylation. *Journal of the American Chemical Society*, 130(47):16011–20.

Carlson, H. A. and McCammon, J. A. (2000). Accommodating Protein Flexibility in Computational Drug Design. *Mol. Pharmacol.*, 57(2):213–218.

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta crystallographica. Section D, Biological crystallography*, 66(Pt 1):12–21.

Cheung, J., Rudolph, M. J., Burshteyn, F., Cassidy, M. S., Gary, E. N., Love, J., Franklin, M. C., and Height, J. J. (2012). Structures of human acetylcholinesterase in complex with pharmacologically important ligands Structures of human acetylcholinesterase in complex with pharmaco- logically important ligands.

Colletier, J.-P., Bourgeois, D., Sanson, B., Fournier, D., Sussman, J. L., Silman, I., and Weik, M. (2008). Shoot-and-Trap: use of specific x-ray damage to study structural protein dynamics by temperature-controlled cryo-crystallography. *Proceedings of the National Academy of Sciences of the United States of America*, 105(33):11742–7.

Colletier, J.-P., Fournier, D., Greenblatt, H. M., Stojan, J., Sussman, J. L., Zaccai, G., Silman, I., and Weik, M. (2006a). Structural insights into substrate traffic and inhibition in acetylcholinesterase. *The EMBO journal*, 25(12):2746–56.

Colletier, J.-P., Royant, A., Specht, A., Sanson, B., Nachon, F., Masson, P., Zaccai, G., Sussman, J. L., Goeldner, M., Silman, I., Bourgeois, D., and Weik, M. (2007). Use of a 'caged' analogue to study the traffic of choline within acetylcholinesterase by kinetic crystallography. *Acta crystallographica. Section D, Biological crystallography*, 63(Pt 11):1115–28.

Colletier, J. P., Sanson, B., Nachon, F., Gabellieri, E., Fattorusso, C., Campiani, G., and Weik, M. (2006b). Conformational flexibility in the peripheral site of Torpedo californica acetylcholinesterase revealed by the complex structure with a bifunctional inhibitor. *Journal of the American Chemical Society*, 128(14):4526–7.

Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995). A Second Generation Force Field for the Simulation of Proteins , Nucleic Acids , and Organic Molecules. *Society*, (6):5179–5197.

Cousin, X., Créminon, C., Grassi, J., Méflah, K., Cornu, G., Saliou, B., Bon, S., Massoulié, J., and Bon, C. (1996). Acetylcholinesterase from Bungarus venom: a monomeric species. *FEBS Letters*, 387(2-3):196–200.

Dale, H. (1914). *The Action of Certain Esters and Ethers of Choline, and their Relation to Muscarine*. Wellcome p edition.

Darden, T., York, D., and Pedersen, L. (1993). Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *The Journal of Chemical Physics*, 98(12):10089.

Das, K., Clark, A. D., Lewi, P. J., Heeres, J., De Jonge, M. R., Koymans, L. M. H., Vinkers, H. M., Daeyaert, F., Ludovici, D. W., Kukla, M. J., De Corte, B., Kavash, R. W., Ho, C. Y., Ye, H., Lichtenstein, M. A., Andries, K., Pauwels, R., De Béthune, M.-P., Boyer, P. L., Clark, P., Hughes, S. H., Janssen, P. A. J., and Arnold, E. (2004). Roles of conformational and positional adaptability in structure-based design of TMC125-R165335 (etravirine) and related

non-nucleoside reverse transcriptase inhibitors that are highly potent and effective against wild-type and drug-resistant HIV-1 varian. *Journal of medicinal chemistry*, 47(10):2550–60.

Das, K., Lewi, P. J., Hughes, S. H., and Arnold, E. (2005). Crystallography and the design of anti-AIDS drugs: conformational flexibility and positional adaptability are important in the design of non-nucleoside HIV-1 reverse transcriptase inhibitors. *Progress in biophysics and molecular biology*, 88(2):209–31.

Dunbrack, R. L. and Cohen, F. E. (1997). Bayesian statistical analysis of protein side-chain rotamer preferences. *Protein science : a publication of the Protein Society*, 6(8):1661–81.

Dvir, H., Silman, I., Harel, M., Rosenberry, T. L., and Sussman, J. L. (2010). Acetylcholinesterase: from 3D structure to function. *Chemico-biological interactions*, 187(1-3):10–22.

Ekström, F., Akfur, C., Tunemalm, A.-K., and Lundberg, S. (2006a). Structural changes of phenylalanine 338 and histidine 447 revealed by the crystal structures of tabun-inhibited murine acetylcholinesterase. *Biochemistry*, 45(1):74–81.

Ekström, F., Hörnberg, A., Artursson, E., Hammarström, L.-G., Schneider, G., and Pang, Y.-P. (2009). Structure of HI-6*sarin-acetylcholinesterase determined by X-ray crystallography and molecular dynamics simulation: reactivator mechanism and design. *PloS one*, 4(6):e5957.

Ekström, F., Pang, Y.-P., Boman, M., Artursson, E., Akfur, C., and Börjegren, S. (2006b). Crystal structures of acetylcholinesterase in complex with HI-6, Ortho-7 and obidoxime: structural basis for differences in the ability to reactivate tabun conjugates. *Biochemical pharmacology*, 72(5):597–607.

Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010). Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography*, 66(Pt 4):486–501.

Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995). A smooth particle mesh Ewald method. *The Journal of Chemical Physics*, 103(19):8577.

Evans, P. R. and Murshudov, G. N. (2013). How good are my data and what is the resolution? *Acta crystallographica. Section D, Biological crystallography*, 69(Pt 7):1204–14.

Faerman, C., Ripoll, D., Bon, S., Le Feuvre, Y., Morel, N., Massoulié, J., Sussman, J. L., and Silman, I. (1996). Site-directed mutants designed to test back-door hypotheses of acetylcholinesterase function. *FEBS Letters*, 386(1):65–71.

Fischer, M., Coleman, R. G., Fraser, J. S., and Shoichet, B. K. (2014). Incorporation of protein flexibility and conformational energy penalties in docking screens to improve ligand discovery. *Nature chemistry*, 6(7):575–83.

Fleisher, J. H. and Harris, L. W. (1965). Dealkylation as a mechanism for aging of cholinesterase after poisoning with pinacolyl methylphosphonofluoridate. *Biochemical pharmacology*, 14(5):641–50.

Frauenfelder, H., Sligar, S. G., and Wolynes, P. G. (1991). The energy landscapes and motions of proteins. *Science*, 254(5038):1598–1603.

Garman, E. F. (2010). Radiation damage in macromolecular crystallography: what is it and why should we care? *Acta crystallographica. Section D, Biological crystallography*, 66(Pt 4):339–51.

Gilson, M. K., Straatsma, T. P., McCammon, J. a., Ripoll, D. R., Faerman, C. H., Axelsen, P. H., Silman, I., and Sussman, J. L. (1994). Open "back door" in a molecular dynamics simulation of acetylcholinesterase. *Science (New York, N.Y.)*, 263(5151):1276–8.

Goodsell, D. S., Morris, G. M., and Olson, a. J. (1996). Automated docking of flexible ligands: applications of AutoDock. *Journal of molecular recognition : JMR*, 9(1):1–5.

Green, D. W., Ingram, V. M., and Perutz, M. F. (1954). The Structure of Haemoglobin. IV. Sign Determination by the Isomorphous Replacement Method. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 225(1162):287–307.

Greenblatt, H. M., Guillou, C., Guénard, D., Argaman, A., Botti, S., Badet, B., Thal, C., Silman, I., and Sussman, J. L. (2004). The complex of a bivalent derivative of galanthamine with torpedo acetylcholinesterase displays drastic deformation of the active-site gorge: implications for structure-based drug design. *Journal of the American Chemical Society*, 126(47):15405–11.

Greer, J., Erickson, J. W., Baldwin, J. J., and Varney, M. D. (1994). Application of the Three-Dimensional Structures of Protein Target Molecules in Structure-Based Drug Design. *Journal of Medicinal Chemistry*, 37(8):1035–1054.

Griffiths, A. D. and Tawfik, D. S. (2003). Directed evolution of an extremely fast phosphotriesterase by in vitro compartmentalization. *The EMBO journal*, 22(1):24–35.

Hann, M. M. and Oprea, T. I. (2004). Pursuing the leadlikeness concept in pharmaceutical research. *Current opinion in chemical biology*, 8(3):255–63.

Harel, M., Kleywegt, G. J., Ravelli, R. B., Silman, I., and Sussman, J. L. (1995). Crystal structure of an acetylcholinesterase-fasciculin complex: interaction of a three-fingered toxin from snake venom with its target. *Structure (London, England : 1993)*, 3(12):1355–66.

Harel, M., Kryger, G., Rosenberry, T. L., Mallender, W. D., Lewis, T., Fletcher, R. J., Guss, J. M., Silman, I., and Sussman, J. L. (2000). Three-dimensional structures of Drosophila melanogaster acetylcholinesterase and of its complexes with two potent inhibitors. *Protein science : a publication of the Protein Society*, 9(6):1063–72.

Harel, M., Quinn, D. M., Nair, H. K., Silman, I., and Sussman, J. L. (1996). The X-ray Structure of a Transition State Analog Complex Reveals the Molecular Origins of the Catalytic Power and Substrate Specificity of Acetylcholinesterase. *Journal of the American Chemical Society*, 118(10):2340–2346.

Harel, M., Schalk, I., Ehret-Sabatier, L., Bouet, F., Goeldner, M., Hirth, C., Axelsen, P. H., Silman, I., and Sussman, J. L. (1993). Quaternary ligand binding to aromatic residues in the active-site gorge of acetylcholinesterase. *Proceedings of the National Academy of Sciences*, 90(19):9031–9035.

Harel, M., Sussman, J. L., Krejci, E., Bon, S., Chanal, P., Massoulié, J., and Silman, I. (1992). Conversion of acetylcholinesterase to butyrylcholinesterase: modeling and mutagenesis. *Proceedings of the National Academy of Sciences of the United States of America*, 89(22):10827–31.

Hartenfeller, M. and Schneider, G. (2011). De novo drug design. *Methods in molecular biology*, 672:299–323.

Hendrickson, W. (1991). Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. *Science*, 254(5028):51–58.

Henrich, B., Bergamaschi, A., Broennimann, C., Dinapoli, R., Eikenberry, E., Johnson, I., Kobas, M., Kraft, P., Mozzanica, A., and Schmitt, B. (2009). PILATUS: A single photon counting pixel detector for X-ray applications. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 607(1):247–249.

Henzler-Wildman, K. and Kern, D. (2007). Dynamic personalities of proteins. *Nature*, 450(7172):964–72.

Henzler-Wildman, K. a., Lei, M., Thai, V., Kerns, S. J., Karplus, M., and Kern, D. (2007). A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature*, 450(7171):913–6.

Hertzberg, R. P. and Pope, A. J. (2000). High-throughput screening: new technology for the 21st century. *Current Opinion in Chemical Biology*, 4(4):445–451.

Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. (1997). LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, 18(12):1463–1472.

Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008). GROMACS 4 Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *Journal of Chemical Theory and Computation*, 4(3):435–447.

Hörnberg, A., Artursson, E., Wärme, R., Pang, Y.-P., and Ekström, F. (2010). Crystal structures of oxime-bound fenamiphos-acetylcholinesterases: reactivation

involving flipping of the His447 ring to form a reactive Glu334-His447-oxime triad. *Biochemical pharmacology*, 79(3):507–15.

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing In Science & Engineering*, 9(3):90–95.

Irwin, J. J., Sterling, T., Mysinger, M. M., Bolstad, E. S., and Coleman, R. G. (2012). ZINC: a free tool to discover chemistry for biology. *Journal of chemical information and modeling*, 52(7):1757–68.

Jackson, C. J., Liu, J.-W., Carr, P. D., Younus, F., Coppin, C., Meirelles, T., Lethier, M., Pandey, G., Ollis, D. L., Russell, R. J., Weik, M., and Oakeshott, J. G. (2013). Structure and function of an insect α-carboxylesterase (αEsterase7) associated with insecticide resistance. *Proceedings of the National Academy of Sciences of the United States of America*, 110(25):10177–82.

Jacobsson, M., Lidén, P., Stjernschantz, E., Boström, H., and Norinder, U. (2003). Improving structure-based virtual screening by multivariate analysis of scoring data. *Journal of medicinal chemistry*, 46(26):5781–9.

Jain, A. N. (2004). Virtual screening in lead discovery and optimization. *Current opinion in drug discovery & development*, 7(4):396–403.

Jain, A. N. (2006). Scoring functions for protein-ligand docking. *Current protein & peptide science*, 7(5):407–20.

Jones, G., Willett, P., Glen, R. C., Leach, A. R., and Taylor, R. (1997). Development and validation of a genetic algorithm for flexible docking. *Journal of molecular biology*, 267(3):727–48.

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2):926.

Juers, D. H. and Matthews, B. W. (2004). Cryo-cooling in macromolecular crystallography: advantages, disadvantages and optimization. *Quarterly Reviews of Biophysics*, 37(2):105–119.

Kabsch, W. (2010). Xds. *Acta crystallographica. Section D, Biological crystallography*, 66(Pt 2):125–32.

Kantardjieff, K. A. and Rupp, B. (2003). Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein science : a publication of the Protein Society*, 12(9):1865–71.

Kliachyna, M., Nussbaum, V., Renou, J., Gianluca, S., Jacques-Philippe, C., Mélanie, A., Mélanie, T.-L., Sanson, B., Weik, M., Jean, L., Renard, P.-Y., Nachon, F., and Baati, R. (2013). Functional hybrid molecules for the reactivation of poisoned acetylcholinesterase.

Kliachyna, M., Santoni, G., Nussbaum, V., Renou, J., Sanson, B., Colletier, J.-P., Arboléas, M., Loiodice, M., Weik, M., Jean, L., Renard, P.-Y., Nachon, F., and Baati, R. (2014). Design, synthesis and biological evaluation of novel tetrahydroacridine pyridine- aldoxime and -amidoxime hybrids as efficient uncharged reactivators of nerve agent-inhibited human acetylcholinesterase. *European Journal of Medicinal Chemistry*, 78:455–467.

Koellner, G., Kryger, G., Millard, C. B., Silman, I., Sussman, J. L., and Steiner, T. (2000). Active-site gorge and buried water molecules in crystal structures of acetylcholinesterase from Torpedo californica. *Journal of molecular biology*, 296(2):713–35.

Kronman, C., Ordentlich, A., Barak, D., Velan, B., and Shafferman, A. (1994). The "back door" hypothesis for product clearance in acetylcholinesterase challenged by site-directed mutagenesis. *J. Biol. Chem.*, 269(45):27819–27822.

Kryger, G., Harel, M., Giles, K., Toker, L., Velan, B., Lazar, A., Kronman, C., Barak, D., Ariel, N., Shafferman, A., Silman, I., and Sussman, J. L. (2000). Structures of recombinant native and E202Q mutant human acetylcholinesterase complexed with the snake-venom toxin fasciculin-II. *Acta Crystallographica Section D Biological Crystallography*, 56(11):1385–1394.

Lengauer, T. and Rarey, M. (1996). Computational methods for biomolecular docking. *Current opinion in structural biology*, 6(3):402–6.

Leslie, A. G. W. (2006). The integration of macromolecular diffraction data. *Acta crystallographica. Section D, Biological crystallography*, 62(Pt 1):48–57.

Lin, J.-H., Perryman, A. L., Schames, J. R., and McCammon, J. A. (2002). Computational drug design accommodating receptor flexibility: the relaxed complex scheme. *Journal of the American Chemical Society*, 124(20):5632–3.

Liu, J.-S., Zhu, Y.-L., Yu, C.-M., Zhou, Y.-Z., Han, Y.-Y., Wu, F.-W., and Qi, B.-F. (1986). The structures of huperzine A and B, two new alkaloids exhibiting marked anticholinesterase activity. *Canadian Journal of Chemistry*, 64(4):837–839.

Mandal, S., Moudgil, M., and Mandal, S. K. (2009). Rational drug design. *European journal of pharmacology*, 625(1-3):90–100.

Marion, D. (2013). An introduction to biological NMR spectroscopy. *Molecular & cellular proteomics : MCP*, 12(11):3006–25.

Masson, P., Carletti, E., and Nachon, F. (2009). Structure, activities and biomedical applications of human butyrylcholinesterase. *Protein and peptide letters*, 16(10):1215–24.

Masson, P., Nachon, F., and Lockridge, O. (2010). Structural approach to the aging of phosphylated cholinesterases. *Chemico-biological interactions*, 187(1-3):157–62.

Massoulié, J., Bon, S., Perrier, N., and Falasca, C. (2005). The C-terminal peptides of acetylcholinesterase: cellular trafficking, oligomerization and functional anchoring. *Chemico-biological interactions*, 157-158:3–14.

Maxwell, D. M. and Brecht, K. M. (2001). Carboxylesterase: specificity and spontaneous reactivation of an endogenous scavenger for organophosphorus compounds. *Journal of applied toxicology : JAT*, 21 Suppl 1:S103–7.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007). Phaser crystallographic software. *Journal of applied crystallography*, 40(Pt 4):658–674.

McInnes, C. (2007). Virtual screening strategies in drug discovery. *Current opinion in chemical biology*, 11(5):494–502.

Meek, E. C., Chambers, H. W., Coban, A., Funck, K. E., Pringle, R. B., Ross, M. K., and Chambers, J. E. (2012). Synthesis and in vitro and in vivo inhibition potencies of highly relevant nerve agent surrogates. *Toxicological sciences : an official journal of the Society of Toxicology*, 126(2):525–33.

Mercey, G., Verdelet, T., Renou, J., Kliachyna, M., Baati, R., Nachon, F., and Jean, L. (2012). Reactivators of Acetylcholinesterase Inhibited by Organophosphorus Nerve Agents. *Accounts of Chemical Research*.

Mercey, G., Verdelet, T., Saint-André, G., Gillon, E., Wagner, A., Baati, R., Jean, L., Nachon, F., and Renard, P.-Y. (2011). First efficient uncharged reactivators for the dephosphylation of poisoned human acetylcholinesterase. *Chemical communications (Cambridge, England)*, 47(18):5295–7.

Michel, J. (2014). Current and emerging opportunities for molecular simulations in structure-based drug design. *Physical chemistry chemical physics : PCCP*, 16(10):4465–77.

Millard, C. B., Koellner, G., Ordentlich, A., Shafferman, A., Silman, I., and Sussman, J. L. (1999a). Reaction Products of Acetylcholinesterase and VX Reveal a Mobile Histidine in the Catalytic Triad. *Journal of the American Chemical Society*, 121(42):9883–9884.

Millard, C. B., Kryger, G., Ordentlich, A., Greenblatt, H. M., Harel, M., Raves, M. L., Segall, Y., Barak, D., Shafferman, A., Silman, I., and Sussman, J. L. (1999b). Crystal structures of aged phosphonylated acetylcholinesterase: nerve agent reaction products at the atomic level. *Biochemistry*, 38(22):7032–9.

Moriarty, N. W., Grosse-Kunstleve, R. W., and Adams, P. D. (2009). electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta crystallographica. Section D, Biological crystallography*, 65(Pt 10):1074–80.

Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., and Olson, A. J. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of computational chemistry*, 30(16):2785–91.

Muegge, I. (2006). PMF scoring revisited. *Journal of medicinal chemistry*, 49(20):5895–902.

Muegge, I. and Martin, Y. C. (1999). A general and fast scoring function for protein-ligand interactions: a simplified potential approach. *Journal of medicinal chemistry*, 42(5):791–804.

Murray, J. W., Rudiño Piñera, E., Owen, R. L., Grininger, M., Ravelli, R. B. G., and Garman, E. F. (2005). Parameters affecting the X-ray dose absorbed by macromolecular crystals. *Journal of synchrotron radiation*, 12(Pt 3):268–75.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta crystallographica. Section D, Biological crystallography*, 67(Pt 4):355–67.

Nachon, F., Brazzolotto, X., Trovaslet, M., and Masson, P. (2013). Progress in the development of enzyme-based nerve agent bioscavengers. *Chemico-biological interactions*, 206(3):536–44.

Nachon, F., Stojan, J., and Fournier, D. (2008). Insights into substrate and product traffic in the Drosophila melanogaster acetylcholinesterase active site gorge by enlarging a back channel. *The FEBS journal*, 275(10):2659–64.

Nar, H., Messerschmidt, A., Huber, R., van de Kamp, M., and Canters, G. W. (1991). Crystal structure analysis of oxidized Pseudomonas aeruginosa azurin at pH 5·5 and pH 9·0. *Journal of Molecular Biology*, 221(3):765–772.

Oldiges, H. and Schoene, K. (1970). [Pyridinium and imidazolium salts as antidotes for soman and paraoxon poisoning in mice]. *Archiv für Toxikologie*, 26(4):293–305.

Ollis, D. L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S. M., Harel, M., Remington, S. J., Silman, I., Schrag, J., Sussman, J. L., Verschueren, K. H., and Goldman, A. (1992). The α / β hydrolase fold. *"Protein Engineering, Design and Selection"*, 5(3):197–211.

Pang, Y.-P. and Kozikowski, A. P. (1994). Prediction of the binding sites of huperzine A in acetylcholinesterase by docking studies. *Journal of Computer-Aided Molecular Design*, 8(6):669–681.

Perrier, A. L., Massoulié, J., and Krejci, E. (2002). PRiMA: the membrane anchor of acetylcholinesterase in the brain.". *Neuron*, 33(2):275–285.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004). UCSF Chimera–a visualization system for exploratory research and analysis. *Journal of computational chemistry*, 25(13):1605–12.

Quinn, D. M. (1987). Acetylcholinesterase: enzyme structure, reaction dynamics, and virtual transition states. *Chemical Reviews*, 87(5):955–979.

Ravelli, R. B. and McSweeney, S. M. (2000). The 'fingerprint' that X-rays can leave on structures. *Structure (London, England : 1993)*, 8(3):315–28.

Rochu, D., Chabrière, E., and Masson, P. (2007). Human paraoxonase: a promising approach for pre-treatment and therapy of organophosphorus poisoning. *Toxicology*, 233(1-3):47–59.

Rogers-Evans, M., Alanine, A., Bleicher, K., Kube, D., and Schneider, G. (2004). Identification of novel cannabinoid receptor ligandsvia evolutionary de novo design and rapid parallel synthesis. *QSAR & Combinatorial Science*, 23(6):426–430.

Rossmann, M. G. and Blow, D. M. (1962). The detection of sub-units within the crystallographic asymmetric unit. *Acta Crystallographica*, 15(1):24–31.

Rydberg, E. H., Brumshtein, B., Greenblatt, H. M., Wong, D. M., Shaya, D., Williams, L. D., Carlier, P. R., Pang, Y.-P., Silman, I., and Sussman, J. L. (2006). Complexes of alkylene-linked tacrine dimers with Torpedo californica acetylcholinesterase: Binding of Bis5-tacrine produces a dramatic rearrangement in the active-site gorge. *Journal of medicinal chemistry*, 49(18):5491–500.

Sanson, B., Colletier, J.-P., Xu, Y., Lang, P. T., Jiang, H., Silman, I., Sussman, J. L., and Weik, M. (2011). Backdoor opening mechanism in acetylcholinesterase based on X-ray crystallography and molecular dynamics simulations. *Protein science : a publication of the Protein Society*, 20(7):1114–8.

Sanson, B., Nachon, F., Colletier, J.-P., Froment, M.-T., Toker, L., Greenblatt, H. M., Sussman, J. L., Ashani, Y., Masson, P., Silman, I., and Weik, M. (2009).

Crystallographic snapshots of nonaged and aged conjugates of soman with acetylcholinesterase, and of a ternary complex of the aged conjugate with pralidoxime. *Journal of medicinal chemistry*, 52(23):7593–603.

Savini, L., Gaeta, A., Fattorusso, C., Catalanotti, B., Campiani, G., Chiasserini, L., Pellerano, C., Novellino, E., McKissic, D., and Saxena, A. (2003). Specific targeting of acetylcholinesterase and butyrylcholinesterase recognition sites. Rational design of novel, selective, and highly potent cholinesterase inhibitors. *Journal of medicinal chemistry*, 46(1):1–4.

Schames, J. R., Henchman, R. H., Siegel, J. S., Sotriffer, C. A., Ni, H., and McCammon, J. A. (2004). Discovery of a novel binding trench in HIV integrase. *Journal of medicinal chemistry*, 47(8):1879–81.

Schneider, G. and Fechner, U. (2005). Computer-based de novo design of drug-like molecules. *Nature reviews. Drug discovery*, 4(8):649–63.

Schotte, F., Lim, M., Jackson, T. A., Smirnov, A. V., Soman, J., Olson, J. S., Phillips, G. N., Wulff, M., and Anfinrud, P. A. (2003). Watching a protein as it functions with 150-ps time-resolved x-ray crystallography. *Science*, 300(5627):1944–7.

Schrödinger, L. (2010). The {PyMOL} Molecular Graphics System, Version˜1.6.

Shaw, D. E., Bowers, K. J., Chow, E., Eastwood, M. P., Ierardi, D. J., Klepeis, J. L., Kuskin, J. S., Larson, R. H., Lindorff-Larsen, K., Maragakis, P., Moraes, M. A., Dror, R. O., Piana, S., Shan, Y., Towles, B., Salmon, J. K., Grossman, J. P., Mackenzie, K. M., Bank, J. A., Young, C., Deneroff, M. M., and Batson, B. (2009). Millisecond-scale molecular dynamics simulations on Anton. In *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis - SC '09*, page 1, New York, New York, USA. ACM Press.

Silman, I. and Sussman, J. L. (2005). Acetylcholinesterase: 'classical' and 'non-classical' functions and pharmacology. *Current opinion in pharmacology*, 5(3):293–302.

Sousa, S. F., Ribeiro, a. J. M., Coimbra, J. T. S., Neves, R. P. P., Martins, S. a., Moorthy, N. S. H. N., Fernandes, P. a., and Ramos, M. J. (2013). Protein-ligand

docking in the new millennium–a retrospective of 10 years in the field. *Current medicinal chemistry*, 20(18):2296–314.

Sousa da Silva, A. W. and Vranken, W. F. (2012). ACPYPE - AnteChamber PYthon Parser interfacE. *BMC research notes*, 5(1):367.

Sussman, J., Harel, M., Frolow, F., and Varon, L. (1988). Purification and Crystallization of a Dimeric Form of Acetylcholinesterase from Torpedo californica Subsequent to Solubilization with phosphatidylinositol-specific phospholipase C. *j mol Biol*, 203:821–823.

Sussman, J. L., Harel, M., Frolow, F., Oefner, C., Goldman, A., Toker, L., and Silman, I. (1991). Atomic structure of acetylcholinesterase from Torpedo californica: a prototypic acetylcholine-binding protein. *Science (New York, N.Y.)*, 253(5022):872–9.

Tai, K., Shen, T., Börjesson, U., Philippopoulos, M., and McCammon, J. a. (2001). Analysis of a 10-ns molecular dynamics simulation of mouse acetylcholinesterase. *Biophysical journal*, 81(2):715–24.

Tarek, M. and Tobias, D. J. (2000). The dynamics of protein hydration water: a quantitative comparison of molecular dynamics simulations and neutron-scattering experiments. *Biophysical journal*, 79(6):3244–57.

Teng, T. Y. (1990). Mounting of crystals for macromolecular crystallography in a free-standing thin film. *Journal of Applied Crystallography*, 23(5):387–391.

Totrov, M. and Abagyan, R. (2008). Flexible ligand docking to multiple receptor conformations : a practical alternative. *Current Opinion in Structural Biology*, 18(2):178–184.

Trott, O. and Olson, A. (2010). AutoDock Vina : Improving the Speed and Accuracy of Docking with a New Scoring Function , Efficient Optimization , and Multithreading. *Journal of computational chemistry*.

van der Spoel, D., van Maaren, P. J., and Berendsen, H. J. C. (1998). A systematic study of water models for molecular simulation: Derivation of water models optimized for use with a reaction field. *The Journal of Chemical Physics*, 108(24):10220.

van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., and Yu, T. (2014). scikit-image: image processing in Python. *PeerJ*, 2:e453.

Vanquelef, E., Simon, S., Marquant, G., Garcia, E., Klimerak, G., Delepine, J. C., Cieplak, P., and Dupradeau, F.-Y. (2011). R.E.D. Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments. *Nucleic acids research*, 39(Web Server issue):W511–7.

Wahlgren, W. Y., Pál, G., Kardos, J., Porrogi, P., Szenthe, B., Patthy, A., Gráf, L., and Katona, G. (2011). The catalytic aspartate is protonated in the Michaelis complex formed between trypsin and an in vitro evolved substrate-like inhibitor: a refined mechanism of serine protease action. *The Journal of biological chemistry*, 286(5):3587–96.

Wang, J., Wang, W., Kollman, P. A., and Case, D. A. (2006). Automatic atom type and bond type perception in molecular mechanical calculations. *Journal of molecular graphics & modelling*, 25(2):247–60.

Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. a., and Case, D. a. (2004). Development and testing of a general amber force field. *Journal of computational chemistry*, 25(9):1157–74.

Weik, M. and Colletier, J. P. (2010). Temperature-dependent macromolecular X-ray crystallography. *Acta crystallographica. Section D, Biological crystallography*, 66(Pt 4):437–46.

Weik, M., Ravelli, R. B., Kryger, G., McSweeney, S., Raves, M. L., Harel, M., Gros, P., Silman, I., Kroon, J., and Sussman, J. L. (2000). Specific chemical and structural damage to proteins produced by synchrotron radiation. *Proceedings of the National Academy of Sciences of the United States of America*, 97(2):623–8.

Weininger, D. (1988). SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *Journal of Chemical Information and Modeling*, 28(1):31–36.

Wilson, I. B. and Ginsburg, S. (1955). A powerful reactivator of alkylphosphate-inhibited acetylcholinesterase. *Biochimica et Biophysica Acta*, 18:168–170.

Wong, C. F., Kua, J., Zhang, Y., Straatsma, T. P., and McCammon, J. A. (2005). Molecular docking of balanol to dynamics snapshots of protein kinase A. *Proteins*, 61(4):850–8.

Worek, F., Seeger, T., Goldsmith, M., Ashani, Y., Leader, H., Sussman, J. S., Tawfik, D., Thiermann, H., and Wille, T. (2014a). Efficacy of the rePON1 mutant IIG1 to prevent cyclosarin toxicity in vivo and to detoxify structurally different nerve agents in vitro. *Archives of toxicology*, 88(6):1257–66.

Worek, F., Seeger, T., Reiter, G., Goldsmith, M., Ashani, Y., Leader, H., Sussman, J. L., Aggarwal, N., Thiermann, H., and Tawfik, D. S. (2014b). Post-exposure treatment of VX poisoned guinea pigs with the engineered phosphotriesterase mutant C23: A proof-of-concept study. *Toxicology letters*, 231(1):45–54.

Worek, F., Thiermann, H., Szinicz, L., and Eyer, P. (2004). Kinetic analysis of interactions between human acetylcholinesterase, structurally different organophosphorus compounds and oximes. *Biochemical pharmacology*, 68(11):2237–48.

Xu, Y., Colletier, J.-P., Jiang, H., Silman, I., Sussman, J. L., and Weik, M. (2008). Induced-fit or preexisting equilibrium dynamics ? Lessons from protein crystallography and MD simulations on acetylcholinesterase and implications for structure-based drug design. pages 601–605.

Xu, Y., Colletier, J.-P., Weik, M., Qin, G., Jiang, H., Silman, I., and Sussman, J. L. (2010). Long route or shortcut? A molecular dynamics study of traffic of thiocholine within the active-site gorge of acetylcholinesterase. *Biophysical journal*, 99(12):4003–11.

Xu, Y., Shen, J., Luo, X., Silman, I., Sussman, J. L., Chen, K., and Jiang, H. (2003). How does huperzine A enter and leave the binding gorge of acetylcholinesterase? Steered molecular dynamics simulations. *Journal of the American Chemical Society*, 125(37):11340–9.

Zimmermann, J. (1996). PYRIMIDINE DERIVATIVES AND PROCESSES FOR THEP REPARATION THEREOF.

Appendix A

Docking code

Generation of receptors from the peak finder algorithm

Modification of the TcAChE stracture to generate the receptors is done using python modules in PyMOL.

In particular, we took advantage of the set_chi function and of the 'rotamers.py' module written by Dan Kulp (Dunbrack and Cohen, 1997). Starting from the structure of native *Tc*AChE (1ea5) we modified side chains of a set of chosen residues following the (χ_1, χ_2) values extracted from molecular dynamics simulations as explained above. The code (reported in appendix 1) works in slightly different ways for the two methods of extractions of dihedrals.

For the method of histogram peak finding, we need to calculate all of the possible combinations for the (χ_1, χ_2) pairs for each residue of interest. This means that the total number of output structures is determined by the equation

$$N_{receptors} = \prod_{i=1}^{n} X_i \tag{A.1}$$

where X_i is the number of possible peaks per residue. It is evident that the number of output structures becomes exponentially large with each residue added . If we consider an average of five conformation peaks per residue, the modification of only three residues dihedrals would lead to 125 output structures, 625 output structures for four residues and so forth.

Each output structure is generated by modifying in sequence all the chosen residues and storing them in a separate pdb file. It is important to check at this point if the newly created structure contains some clashes between the rotated side chains and the rest of the protein. This is due to the fact that each conformational region in the (χ_1, χ_2), even if in reality it is spread over some tens of degrees, is represented by a single value for each dihedral angle, loosing all the information on the possible

adjustments the residue can make. Idealization of the structures in the docking library is performed in two steps. At first, a clash checking function, written using the function pymol.cmd.distance, excludes all the structures with residues getting closer then 1.5Å. If water molecules are included in the modeling, a second function deletes all the ones that are too close to a residue, normally at a distance lower than 2Å. These distances are smaller than what could be accepted in a structure, but they allow making a first selection of very problematic conformations. Output structures passing the first clash filter are then processed with a crystallography tool, phenix.geometry_minimization (Adams et al., 2010), to idealize the regions where contacts are likely to occur.

```
reinitialize

python

import numpy as np
import pymol
import rotamers


def loadData(filename):
    chi= []
    chi1,chi2 = np.loadtxt(filename, delimiter=(","), usecols=(0,1),
        unpack = True)
    for i in range(0,len(chi1)):
        tmpList=[]
        tmpList.append(chi1[i])
        tmpList.append(chi2[i])
        chi.append(tmpList)
    return chi


pymol.cmd.load("start_model.pdb")
pymol.cmd.select("views", selection = "res 331 or res 279 or res 330 ␣␣
    or res 84")

chi84 = loadData("trp84.dat")
chi279 = loadData("trp279.dat")
chi330 = loadData("phe330.dat")
chi331 = loadData("phe331.dat")

priority = 1
for i in chi84 :
    rotamers.set_rotamer("res 84", i[0], i[1])
    for j in chi279:
        rotamers.set_rotamer("res 279", j[0],j[1])
        for k in chi331:
            rotamers.set_rotamer("res 331", k[0],k[1])
            for l in chi330:
```

```
                rotamers.set_rotamer("res 330", l[0],l[1])
                pymol.cmd.select("toSave", selection = "chain A")
                pymol.cmd.select("reference", selection = "res 1-600
                    and not view")
                pymol.cmd.select("NewRef", selection = "reference and
                    not views and not neighbor views")
                ClashControl = pymol.cmd.distance("Clash", "views", "
                    NewRef", "1.5", "0")
                #
                #
                #Check if there are no clashes and controls all water
                    molecules
                #to delete the ones that are too close to the protein
                if ClashControl == 0:
                    for ind in range(2001,2740):
                        waterMol = "resn HOH and resi " + str(ind)
                        pymol.cmd.select("Water_model", waterMol)
                        distance = pymol.cmd.dist("Contact", "
                            Water_model", "toSave", "2.5", "0")
                        if distance != 0:
                            pymol.cmd.remove("Water_model")
                        pymol.cmd.delete("Contact")
                        pymol.cmd.delete("Water_model")
                    pymol.cmd.save("output"+str(priority)+".pdb","
                        toSave",1)
                    priority += 1
                pymol.cmd.delete("toSave")


python end

quit
```

Generation of receptors with the rejection sampling method

The method of rejection sampling is more flexible, both from the conceptual point of view and concerning the program structure. Its key feature is the ability to define a level of completeness for the user, that allows to choose the desired number of output structures. The program will generate as many dihedral sets as needed to satisfy the request. Differently from the peak finder method, we do not need to modify the code to include additional residues. We just need to supply a list of valid residue names and IDs. Obviously, the more residues we want to include in the library generation, the longer the calculation will take. The structure of the program allows to simplify clash checking. In this case, after each residue is modified, if the conformation clashes we simply reject it and start again the process from the beginning, generating a new set of rotamers. As soon as a valid conformation is found, the program tries to apply an analogous modification to the following residue in the input list. Each time all of the residues have been properly modified, a structure is stored and the program starts back with the next editing cycle.

Also, this second routine can handle water molecules, which allowed us to verify the role played by solvent molecules that are highly conserved in AChE structures on the reliability of docking experiments.

```
# -*- coding: utf-8 -*-
"""
Created on Wed Jan 23 10:12:19 2013

@author: gianluca Santoni
A pymol script to generate docking receptors from Molecular dynamics
    simulation
ensemble.
Uses pymol to be simple in MacOS. Links are a bit weird there :)
It supposes you already extracted the Chi1-Chi2 ensembles

Requires:
Numpy
Rotamers.py
"""
import __main__
import numpy as np
import pymol
import rotamers

import re
from math import *
'''
Here I define how many dihedrals
any residue has.

will be used later to search for the good options
```

```
'''

'''
The dihedrals dictionary is used as a reference
to know how many side-chain angles each
residue has.
Obviously, alanine is not considered.
'''
dihedrals = {}

dihedrals['ARG']=[['chi1'],
                  ['chi2'],
                  ['chi3'],
                  ['chi4']
                 ]

dihedrals['ASN']=[['chi1'],
                  ['chi2']
                 ]

dihedrals['ASP']=[['chi1'],
                  ['chi2']
                 ]

dihedrals['CYS']=[['chi1']
                  ]

dihedrals['GLN']=[['chi1'],
                  ['chi2'],
                  ['chi3']
                 ]

dihedrals['GLU']=[['chi1'],
                  ['chi2'],
                  ['chi3']
                 ]

dihedrals['HIS']=[['chi1'],
                  ['chi2']
                 ]

dihedrals['ILE']=[['chi1'],
                  ['chi2']
                 ]

dihedrals['LEU']=[['chi1'],
                  ['chi2']
                 ]
```

```
dihedrals['LYS']=[['chi1'],
                  ['chi2'],
                  ['chi3'],
                  ['chi4']
                 ]


dihedrals['MET']=[['chi1'],
                  ['chi2'],
                  ['chi3']
                 ]

dihedrals['PHE']=[['chi1'],
                  ['chi2']
                 ]

dihedrals['PRO']=[['chi1'],
                  ['chi2']
                 ]

dihedrals['SER']=[['chi1']
                 ]

dihedrals['THR']=[['chi1']
                 ]

dihedrals['TRP']=[['chi1'],
                  ['chi2']
                 ]

dihedrals['TYR']=[['chi1'],
                  ['chi2']
                 ]

dihedrals['VAL']=[['chi1']
                 ]


def populateHistogram(residue, dataPath):
    '''
    get residue name and populates an histogram
    of dihedrals distribution
    '''
    resn = residue[:3]
    if len(dihedrals[resn])==1:
        density = np.zeros((120), dtype=int)
        fileChi1 = open(dataPath+'/chi1'+residue+'.xvg')
        lines1 = fileChi1.readlines()
        fileChi1.close()
```

```
    elif len(dihedrals[resn])==2:
        density = np.zeros((120,120))
        fileChi1 = open(dataPath+'/chi1'+residue+'.xvg')
        fileChi2 = open(dataPath+'/chi2'+residue+'.xvg')
        lines1 = fileChi1.readlines()
        lines2 = fileChi2.readlines()
        fileChi1.close()
        fileChi2.close()
    elif len(dihedrals[resn])==3:
        density = np.zeros((120,120,120), dtype=int)
        fileChi1 = open(dataPath+'/chi1'+residue+'.xvg')
        fileChi2 = open(dataPath+'/chi2'+residue+'.xvg')
        fileChi3 = open(dataPath+'/chi3'+residue+'.xvg')
        lines1 = fileChi1.readlines()
        lines2 = fileChi2.readlines()
        lines3 = fileChi3.readlines()
        fileChi1.close()
        fileChi2.close()
        fileChi3.close()
    elif len(dihedrals[resn])==4:
        density = np.zeros((120,120,120,120), dtype=int)
        fileChi1 = open(dataPath+'/chi1'+residue+'.xvg')
        fileChi2 = open(dataPath+'/chi1'+residue+'.xvg')
        fileChi3 = open(dataPath+'/chi1'+residue+'.xvg')
        fileChi4 = open(dataPath+'/chi1'+residue+'.xvg')
        lines1 = fileChi1.readlines()
        lines2 = fileChi2.readlines()
        lines3 = fileChi3.readlines()
        lines4 = fileChi4.readlines()
        fileChi1.close()
        fileChi3.close()
        fileChi3.close()
        fileChi4.close()

    for x in lines1:
        reference = x.split()
        order = int(reference[0])
        chi1 = float(reference[1])
        if chi1<0:
            chi1+=360
        try:
            Chi2 = lines2[order].split()
            chi2 = float(Chi2[1])
            if chi2<0:
                chi2+=360
            try:
                Chi3 = lines3[order].split()
                chi3 = float(Chi3[1])
                if chi3<0:
                    chi3+=360
```

```
                try:
                    Chi4 = lines4[order].split()
                    chi4 = float(Chi4[1])
                    if chi4<0:
                        chi4+=360
                    density[floor(chi1/3),floor(chi2/3),floor(chi3/3),
                        floor(chi4/3)]+=1
                except:
                    density[floor(chi1/3),floor(chi2/3),floor(chi3/3)
                        ]+=1

            except:
                density[floor(chi1/3),floor(chi2/3)]+=1

        except:
            density[floor(chi1/3)]+=1
    return density

def TEST(resn, datafolder):
    file1=open(datafolder+"/chi1"+resn+".xvg")
    file2=open(datafolder+"/chi2"+resn+".xvg")
    density = np.zeros((120,120))
    lines1 = file1.readlines()
    lines2 = file2.readlines()
    limit = len(lines1)
    file1.close()
    file2.close()
    for x in range(0,limit):
        reference = lines1[x].split()
        order = int(reference[0])
        chi1 = float(reference[1])
        if chi1<0:
            chi1+=360
        Chi2 = lines2[x].split()
        chi2 = float(Chi2[1])
        if chi2<0:
            chi2+=360
        density[floor(chi1/3),floor(chi2/3)]+=1
    return density

def randomHist(pdf):
    '''
    picks a random conformation
    '''
    normPdf=pdf/pdf.sum() #normalize
    check = False
    while check==False:
        chi1= np.random.randint(0, 120)
        chi2= np.random.randint(0, 120)
        z = np.random.uniform(0,1)
```

```
        if z<normPdf[chi1,chi2]:
            check = True
    return 3*chi1,3*chi2

cyclesNumber = 300


dataFile = 'config.txt'
dataFolder = 'TcAChE_angles'


config=open(dataFile)
residues = config.readlines()
config.close()

Datasets = {}
Angles = {}

#creating a dictionary of input residues
#will be populated from input file
for x in residues:
    r = x.rstrip('\n')
    Datasets[r]=TEST(r, dataFolder)


# Starting the pymol part
#loop over all the lines of the histogram to generate a new enzyme
    conformer
#also checks clash at every modification

x=287
__main__.pymol_argv = [ 'pymol', '-qc']
pymol.finish_launching()

pymol.cmd.set("suspend_update", "1" )
pymol.cmd.load("waters_model.pdb")
while x < cyclesNumber:
    print x
    Scoring = True
    for res in residues:
        r = res.rstrip('\n')
        resn = int(r[3:])+3
        chi_A=randomHist(Datasets[r])
        rotamers.set_rotamer("res "+str(resn), chi_A[0], chi_A[1])
        pymol.cmd.select("Clash", selection ="res "+str(resn))
        pymol.cmd.select("Ref", selection="waters_model and not (Clash
            expand 1.5)" )
        distance = pymol.cmd.distance("Check", "Ref", "Clash", "2")
        if distance != 0:
            Scoring = False
```

```
                break
        if Scoring == True:
            pymol.cmd.save("Output"+str(x)+".pdb","waters_model", 1)
            x+=1
        pymol.cmd.delete("Check")

pymol.cmd.quit()
```

Appendix B

Published manuscripts

Organic & Biomolecular Chemistry

RSC Publishing

PAPER

Reaction site-driven regioselective synthesis of AChE inhibitors†

Cite this: DOI: 10.1039/c3ob42109k

Emilia Oueis,[a] Gianluca Santoni,[b,c,d] Cyril Ronco,[a] Olga Syzgantseva,[a] Vincent Tognetti,[a] Laurent Joubert,[a] Anthony Romieu,[a] Martin Weik,[b,c,d] Ludovic Jean,*[a] Cyrille Sabot,*[a] Florian Nachon[e] and Pierre-Yves Renard*[a]

Received 19th August 2013,
Accepted 24th October 2013

DOI: 10.1039/c3ob42109k

www.rsc.org/obc

The enzyme-directed synthesis is an emerging fragment-based lead discovery approach in which the biological target is able to assemble its own multidentate ligands from a pool of building blocks. Here, we report for the first time the use of the human acetylcholinesterase (AChE) as an enzyme for the design and synthesis of new potent heterodimeric huprine-based inhibitors. Both the specific click chemistry site within the protein and the regioselectivity of the Huisgen cycloaddition observed suggest promising alternatives in the design of efficient mono- and dimeric ligands of AChE. Finally, a detailed computational modelling of the click reaction was conducted to further understand the origin of this TGS selectivity.

Introduction

The target-guided synthesis (TGS) is a relatively unexplored but promising approach for lead discovery. More precisely, an enzyme acts as a template to autoassemble its own multidentate inhibitors from a pool of building blocks bearing complementary reacting fragments.[1] This *in situ* ligation of building blocks is often achieved through the reliable alkyne–azide Huisgen cycloaddition (called "click chemistry")[2] to yield the corresponding 1,2,3-triazole-containing inhibitors. In addition, acetylcholinesterase that plays a key role in the treatment of Alzheimer's disease (AD)[3] was the first enzyme used to validate the *in situ* click chemistry approach. It is well-known that the active site of AChE is composed of two distinct binding sites, named the peripheral (1) and the acylation sites (2), connected by a narrow hydrophobic channel, the AChE gorge (3) (Fig. 1A). Taking advantage of the enzyme structure, Sharpless and Kolb[4] reported the capacity of *mouse* AChE (*m*-AChE) to catalyse exclusively the formation of highly potent *syn*-triazole heterodimeric inhibitors such as (±)-**TZ2PIQA5** and (±)-**TZ2PIQA6** (Fig. 1B). More precisely, the *syn*-triazole formation occurred from a mixture of tacrine azide (catalytic site binder) and phenyl tetrahydroisoquinoline alkyne (peripheral site binder) building blocks, in a particularly favourable constricted region composed of side chains of Tyr124, Phe297, Tyr337, and Phe338 at the bottom of the gorge (*vide infra*).[5] The observed *syn*-regioselectivity was attributed to the parallel orientation of both azide and alkyne reactive groups that pointed toward the catalytic pocket.

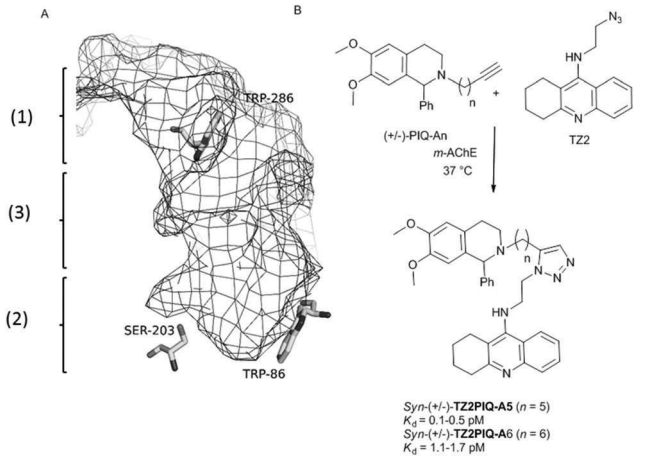


Fig. 1 (A) Description of the active site of *m*-AChE: (1) peripheral anionic site; (2) acylation site; (3) hydrophobic channel. (B) Hits that emerged from *in situ* click chemistry experiments described by Sharpless and Kolb.[4]

Following these interesting results, the *in situ* click chemistry was performed in other constricted active sites of enzymes (HIV protease,[6] carbonic anhydrase,[7] or the transcriptional repressor EthR[8]) or at a presumably more flexible

[a]*Normandie Univ, COBRA, UMR 6014 & FR 3038, Univ Rouen, INSA Rouen, CNRS 1, rue Tesnière 76821 Mont-Saint-Aignan, Cedex, France.*
E-mail: pierre-yves.renard@univ-rouen.fr; Fax: (+) 33 2 35 52 29 71
[b]*Univ Grenoble Alpes, IBS, F-38027 Grenoble, France*
[c]*CNRS, IBS, F-38027 Grenoble, France*
[d]*CEA, IBS, F-38027 Grenoble, France*
[e]*Institut de Recherche Biomédicale des Armées BP73, F-91993 Brétigny-sur-Orge, France*

†Electronic supplementary information (ESI) available. See DOI: 10.1039/c3ob42109k

environment such as the subunit interfaces of the nicotinic acetylcholine receptor nAChR.[9] To the best of our knowledge, only one click chemistry site was reported for each enzyme investigated to date, leading to one triazole regioisomer (either *syn* or *anti*). Indeed, not only the overall spacing between the two monomeric ligands but also the location of the triazole ring construction within the enzyme proved to be of crucial importance for the success of the *in situ* click chemistry approach.[2b] Consequently, in order to extend the scope of the TGS, we questioned whether the identification of a new reaction site within a protein would be possible. This should offer new opportunities in the rational design of new enzyme inhibitors by using various building block structures.

Thus, in this study we explored a novel AChE reaction site suitable for TGS, involving the use of bulky active site ligands such as huprine building blocks. Preliminary work from our laboratory suggests that the huprine moiety would be a suitable catalytic site binder for *in situ* click chemistry.[10] Huprine derivatives, developed by Camps *et al.*, are the best AChE active-site ligands reported to date, especially huprines X and Y that exhibit high affinity towards AChE's acylation site (IC_{50} = 0.75 nM and 0.78 nM respectively, 270-fold more potent than tacrine),[11] and can be readily functionalized at position 9 or 12.[12] Moreover, we selected the phenyl tetrahydroisoquinoline scaffold that already proved to be a convenient peripheral site binder in TGS.[4] Not to mention the currently growing interest in developing multi-target-directed drugs and particularly new potent dual-binding site AChE inhibitors able to exert a dual action,[13] *i.e.* interacting simultaneously with the AChE active site (inhibition of cholinesterasic activity) and the peripheral site (inhibition of AChE-mediated Aβ deposition).[14]

From this investigation, the following important outcomes may be expected: (1) the design and synthesis of new potent dual-binding site inhibitors directed for the first time by the flexible recombinant human *rh*-AChE;[15] (2) a better understanding of the relationship between the location of the click chemistry within the enzyme and the observed regioselectivity; and (3) a deeper insight into the fundamentals that govern the TGS strategy.

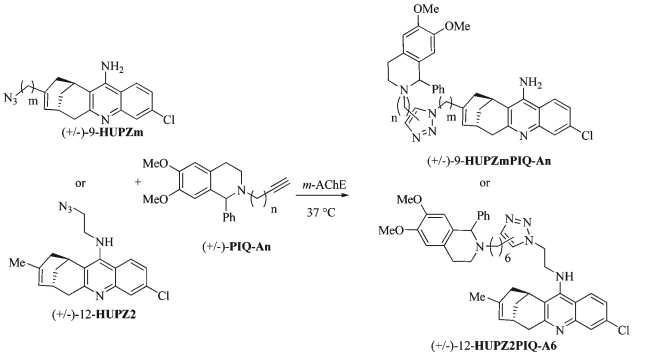
Results and discussion

(±)-12-**HUPZ2**,[16] functionalized by an amino ethyl azide group at position 12 on the quinoline ring, and (±)-**PIQ-A6** (n = 6),[17] bearing an alkyne group, were incubated in the presence of the model enzyme *m*-AChE (Table 1, entry 1). Interestingly, no trace of (±)-12-**HUPZ2PIQ-A6** was observed using HPLC coupled with ESI/MS detection, whereas the incubation under similar conditions of (±)-**PIQ-A6** and **TZ2**, the latter sharing important structural similarities with (±)-12-**HUPZ2**, afforded the corresponding dimeric inhibitor *syn*-(±)-**TZ2PIQ-A6** (Fig. 1B) as reported by Sharpless and Kolb.[4] Huprine and tacrine share similar binding zones within the active site of the enzyme, but the additional carbobicyclic ring system in the huprine scaffold hinders the access to the esteratic site, which would prevent the formation of the triazole of (±)-12-**HUPZ2-PIQ-A6** in this sequestered enzyme site. Moreover, the moderate IC_{50} value of (±)-12-**HUPZ2PIQ-A6** (*i.e.* 7.0 nM) is in agreement with our previous studies showing that functionalization of huprines at the quinoline moiety (position 12) resulted in a significant decrease of their inhibitory potency. Indeed, the free amino group on the quinoline ring is involved in key hydrogen bonds with the nearby structural water molecule[12d] (Fig. 2).

This preliminary result emphasizes the importance of finding alternative reaction sites within an enzyme, especially to widen the scope of bivalent inhibitors designed *via in situ* click chemistry. Furthermore, we recently reported that suitable modifications at position 9 of the huprine derivative leave the water network unchanged, and could increase their affinity toward AChE. Interestingly, our previous docking studies suggested that the ligation to a peripheral site binder could be carried out at position 9.[10] Consequently, the enzymatic triazole synthesis could be envisioned in the *a priori* less favorable mid-gorge of *m*-AChE, *i.e.* closer to the peripheral anionic site. Yet, the higher degree of enzyme flexibility expected at this location may disadvantage the click chemistry process.

Thus, huprines bearing alkyl azide (m = 2 and 4), and phenyl tetrahydroisoquinolines decorated with alkyl acetylene

Table 1 *m*-AChE-directed synthesis of huprine-based heterodimers



Entry	Ligand	m	n	Click *in situ*	IC_{50} [a] (nM)
1	(±)-12-**HUPZ2PIQ-A6**	n.a.[b]	6	—	7.0 ± 1.0
2	(±)-9-**HUPZ4PIQ-A1**	4	1	—	1.5 ± 0.3
3	(±)-9-**HUPZ4PIQ-A2**	4	2	*anti*[c]	0.4 ± 0.1
4	(±)-9-**HUPZ4PIQ-A3**	4	3	*syn*/*anti*[d]	1.2 ± 0.3
5	(±)-9-**HUPZ4PIQ-A4**	4	4	—	3.1 ± 0.6
6	(±)-9-**HUPZ2PIQ-A4**	2	4	—	13.0 ± 2.5

[a] Corresponding to the *anti* isomer. [b] Not applicable. [c] Detected after a two-day incubation period. [d] Detected after a seven-day incubation period.

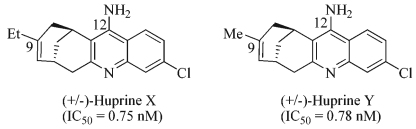


Fig. 2 (±)-Huprine X, Y.

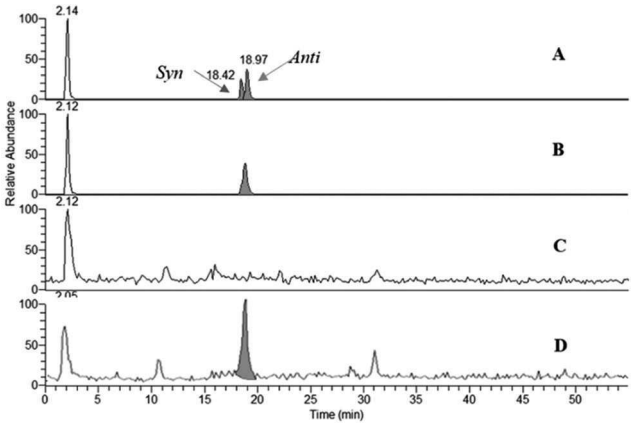


Fig. 3 LCMS/SIM traces of (±)-9-**HUPZ4PIQ-A2**: (A) a mixture of *anti* and *syn*-(±)-9-HUPZ4PIQ-A2 obtained under the thermal Huisgen process; (B) *anti*-(±)-9-**HUPZ4PIQ-A2** formed by catalytic CuAAC; (C) trace of the blank experiment without *m*-AChE; and (D) trace of the click experiment with *m*-AChE.

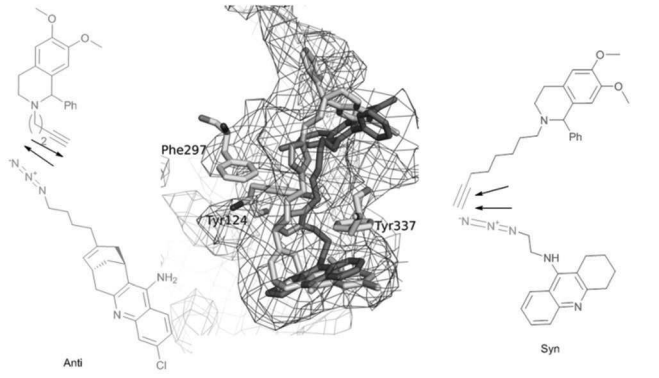


Fig. 4 Superimposition of *in situ* auto-assembled ligand *syn*-(±)-**TZ2PIQ-A6** (in red) and *anti*-(±)-9-**HUPZ4PIQ-A2** (in blue) docked in the active site of *m*-AChE (pdb code 1Q83). The triazole unit of each ligand is colored in orange.

of varying chain lengths (n = 1–4), were incubated in the presence of *m*-AChE. Control experiments were also performed without an enzyme (or with BSA protein). The reaction mixtures were next analysed by RP-HPLC through electrospray ionisation mass spectrometry in the positive selected ion monitoring mode (SIM) (Fig. 3), and compared to authentic samples obtained by thermal or copper-catalysed 1,3-dipolar cycloaddition.[18]

From ten heterodimers that could theoretically be formed *in situ* (Table 1, entries 2–6), only the *anti*-9-**HUPZ4PIQ-A2** and the regioisomeric mixture of *syn*- and *anti*-9-**HUPZ4PIQ-A3** (≈ 1 : 2) were observed (Table 1, entries 3 and 4). These results show that the acetylene moiety of the peripheral site ligand (±)-**PIQ-A2**, with a two carbon-atom chain length, would be located at an optimal distance from the azide unit of (±)-9-**HUPZ4** to adopt a specific conformation leading to a unique regioisomer (*i.e. anti*). However, the presence of an additional methylene unit in (±)-**PIQ-A3** would enable the alkyne chain to fold, and lead to the subsequent *syn*-triazole isomer as well. Interestingly, no trace of (±)-9-**HUPZ2PIQ-A4** was formed (entry 6), even though it has the same overall spacing between the peripheral and catalytic site binders as (±)-9-**HUPZ4PIQ-A2**. In this case, the triazole unit would be unfavourably located at the bottleneck of the active site gorge.

It is noteworthy that this study, together with Sharpless and Kolb's results, shows that AChE is able to template either the *anti* isomer of (±)-9-**HUPZ4PIQ-A2** (entry 3) or the *syn* isomer of (±)-**TZ2PIQ-A6**, depending on the location of the click chemistry within the active site of the enzyme. Indeed, as depicted in Fig. 4, these dual-binding site inhibitors are templated at two clearly distinct sites of *m*-AChE, located on either side of the enzyme bottleneck. *syn*-(±)-**TZ2PIQ-A6** is assembled in the constricted region located in proximity to the acylation site of the enzyme, resulting from reacting functionalities extended in a parallel orientation. *anti*-(±)-9-**HUPZ4PIQ-A2** is ligated closer to the rim of the gorge, an unprecedented reaction site of AChE, from complementary functional groups lying in an antiparallel orientation.[19]

We next examined whether these promising results could be extrapolated to the more flexible *rh*-AChE. In fact, the symptomatic approach to treat neurological disorders such as AD consists of controlling the enzymatic activity of human AChE. It is noteworthy that switching from *m*- to *rh*-AChE template is not a trivial task as both enzymes have marked differences in dynamic properties despite fully conserved active site residues.[15b] The same series of binary mixtures was incubated in the presence of *rh*-AChE to direct the synthesis of heterodimeric huprine-based inhibitors (Table 2). Gratifyingly, *in situ* click chemistry was observed and similar results were obtained with the clinically relevant enzyme. Note that longer reaction times were necessary before detecting the triazole-adducts (±)-9-**HUPZ4PIQ-A2** and (±)-9-**HUPZ4PIQ-A3**, probably due to the greater flexibility of the *rh*-AChE active site compared with the *m*-AChE one.

Consequently, these *in situ* click chemistry results highlighted that the functionalization of the huprine scaffold at position 9 was an appropriate strategy to access the peripheral anionic site of AChE (Table 2, entry 1 *vs.* 3 and 4). As expected, the overall chain length between the catalytic and the peripheral site ligand (entries 2 and 5 *vs.* 3 and 4), as well as the location of the triazole ring into the gorge, should be considered (entry 3 *vs.* 6). The latter information would be particularly valuable to build efficient triazole-containing dimeric and

Table 2 *rh*-AChE-directed synthesis of huprine-based heterodimers

Entry	Ligand	*m*	*n*	Click *in situ*	IC_{50} [a] (nM)
1	(±)-12-**HUPZ2PIQ-A6**	n.a.[b]	6	—	6.1 ± 0.8
2	(±)-9-**HUPZ4PIQ-A1**	4	1	—	3.3 ± 0.2
3	(±)-9-**HUPZ4PIQ-A2**	4	2	*anti*[c]	0.6 ± 0.1
4	(±)-9-**HUPZ4PIQ-A3**	4	3	*syn*/*anti*[d]	0.6 ± 0.1
5	(±)-9-**HUPZ4PIQ-A4**	4	4	—	0.8 ± 0.1
6	(±)-9-**HUPZ2PIQ-A4**	2	4	—	5.3 ± 0.3

[a] Corresponding to the *anti* isomer. [b] Not applicable. [c] Detected after a nine-day incubation period. [d] Detected after a thirteen-day incubation period.

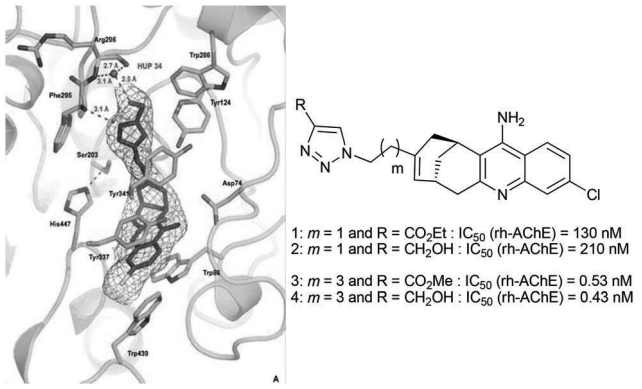


Fig. 5 X-ray crystal structure of the 4-*m*-AChE complex (pdb code 4A16) and comparative IC_{50} values of huprines bearing a triazole unit separated by a two- and four-carbon chain length.

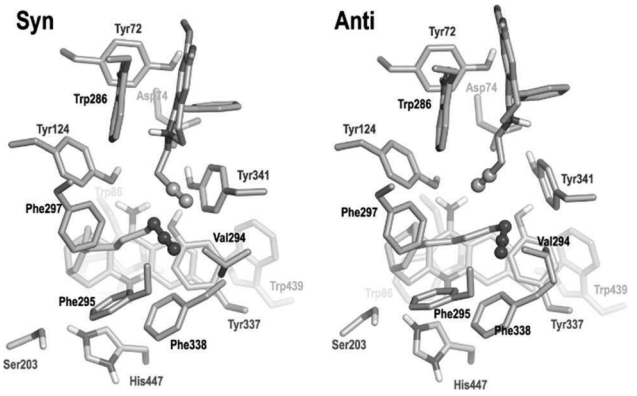


Fig. 6 Pre-reactive complexes of *syn*-(±)-9-**HUPZ4PIQ-A2** and *anti*-(±)-9-**HUPZ4PIQ-A2** in the active site of *rh*-AChE. The main chain of active site residues is represented as grey sticks, with oxygen in red and nitrogen in blue. Atoms of the reactive azido and ethyne groups are represented as spheres.

monomeric inhibitors. Indeed, the present TGS approach confirmed our previous observations[10] showing that ligands bearing a triazole unit at four carbon atoms away from the huprine scaffold, such as compounds **3** and **4**, were much more potent than their two carbon atom chain length analogue **1** or **2**, due to better accommodation midway in the active site gorge, and notably favourable interactions of the triazole with Tyr341 and the main chain of Phe295 (Fig. 5).[20] Compound **4** constitutes one of the best *rh*-AChE active-site ligands reported to date (IC_{50} = 0.43 nM).

In order to investigate the potential reasons for the regioselective formation of the *anti*-(±)-9-**HUPZ4PIQ-A2** (*i.e.* the role of the residues in close proximity to the reaction site, the reaction barriers and the total energy of the protein–substrate complex), additional theoretical studies were undertaken. To this aim, the structures of protein–ligand complexes for *syn*- and *anti*-molecules were obtained by docking and MD equilibration. The structure of the complexes was used as the basis to model the click reaction between substrates, using the combined quantum mechanics/molecular mechanics approach (QM/MM). A detailed computational procedure is given in ESI.†

First, pre-reactive complexes of *syn*-(±)-9-**HUPZ4PIQ-A2** and *anti*-(±)-9-**HUPZ4PIQ-A2** in the active site of *rh*-AChE were analysed (Fig. 6). The huprine moiety is firmly maintained by aromatic stacking in between Trp86 and Tyr337, while the azido group is confined in an aromatic pocket formed by Phe295, Phe297, Tyr337 and Phe338.

The tetrahydroisoquinoline substrate is bound to the peripheral site, stacking parallel to Trp286 with additional T-stacking involving its phenyl ring and Tyr341. In the pre-*syn* complex, the alkynyl group lies away from Trp286 (4.6 Å) parallel to the azido group and makes favorable π–π interaction with Tyr341 (3.4 Å). By contrast, the alkynyl group lies orthogonally to the azido group in the pre-*anti* complex, at closer distance from Trp286 than Tyr341 (3.4 Å *vs.* 4.0 Å). However, no clear conclusion could be drawn at this stage from this preliminary analysis.

Consequently, reaction barriers for *syn* and *anti* substrates in the protein environment were next investigated from a quantum chemistry viewpoint by fully optimizing the corresponding transition states, and were found to be 16.3 kcal mol^{-1} and 20.5 kcal mol^{-1}, respectively. These values (both accessible at ambient temperature) suggest a slight but not definitive kinetic preference for the *syn* product. It can be noticed that Senapati and co-workers[21] reported a similar difference (4 kcal mol^{-1}) between the activation barriers in favour of the experimentally non-observed product. The strong catalytic role of the enzyme can also be enlightened by remarking that the activation energies calculated using the same computational protocol for the formation of the heterodimeric inhibitors without AChE are about twice (35 kcal mol^{-1}) the ones calculated in the presence of the enzyme.

From a thermodynamic point of view, a direct comparison of the total energies of protein–substrate complexes shows that the *anti*-complex is more stable than the *syn*-complex by 173 kcal mol^{-1} at the MM level and 144 kcal mol^{-1} at the QM/MM level. This means that once the system is trapped in the *anti* conformation, it cannot be converted to the *syn* isomer. Thus, the selective formation of *anti*-9-**HUPZ4PIQ-A2** is mainly the result of accumulation of protein–substrate complexes and interactions in the *anti* configuration. It should be mentioned that similar findings were reported earlier by Senapati and co-workers[21] on the selective formation of the observed product in *m*-AChE. It thus seems reasonable to conclude that the physicochemical factors driving the selectivity of the process in *rh*-AChE are the same as those in *m*-AChE.

Finally, the total energy difference (*i.e.* 144 kcal mol^{-1}) between *syn* and *anti* products can be decomposed into three contributions: substrate and protein internal energy differences and binding interaction (protein–substrate) differences between both conformations. It was found that 9.6 kcal mol^{-1} is brought about by the stabilization of the *anti*-substrate with respect to the *syn*-substrate, 17.3 kcal mol^{-1} is due to a higher binding interaction energy of the *anti*-isomer with the protein and the larger contribution of 117.5 kcal mol^{-1} comes from

the protein internal energy difference between both conformations. Interestingly, these three contributions have cumulative effects, favouring the formation of the *anti*-complex. The main contribution comes clearly from the protein internal energy differences, and thus highlights the induced fit contribution to the triazole formation. These results, which are in agreement with experiments, also highlight the relevance of such a QM/MM approach.

Conclusions

Herein, we have reported the first use of human AChE as the template for *in situ* synthesis of subnanomolar dual binding site inhibitors. To achieve this outcome, the identification of a novel click chemistry site within the AChE gorge that enables the use of bulky huprine building blocks as an acylation site binder was required. Whilst the *syn*-triazole isomer was previously reported by Sharpless and Kolb to be preferentially formed in the sequestered region within the AChE gorge, interestingly, we observed a reversal of regioselectivity moving the reaction toward the upper part of the gorge, in the mid-gorge site of AChE. This selectivity is thermodynamically driven and is principally related to the protein internal energy difference between both conformations, and shows for the first time that an enzyme could govern the absolute regioselectivity of the Huisgen reaction (either *syn* or *anti* isomer), depending on the reaction site selected and the overall spacing between the two monomeric ligands.

Importantly, this study shows that the TGS, sustained by docking experiments, is a flexible tool capable of highlighting different click chemistry sites within an enzyme, depending on the combination of building blocks selected. Consequently, the identification of a broader diversity of inhibitors may be expected.

Acknowledgements

This work was supported by Agence Nationale pour la Recherche (through ANR_06_BLAN_0163 DETOXNEURO and ANR_09_BLAN_0192 ReAChE programs), and the Région Haute Normandie (the CRUNCh program, and a PhD fellowship for EO). We thank C. Thiot-Sabot for helpful discussions.

Notes and references

1 (*a*) D. Rideout, T. Calogeropoulou, J. Jaworski and M. Mccarthy, *Biopolymers*, 1990, **29**, 247–262; (*b*) G. E. Boldt, T. J. Dicherson and K. D. Janda, *Drug Discovery Today*, 2004, **9**, 143–148; (*c*) A. V. Eliseev, *Drug Discovery Today*, 2004, **9**, 348–349; (*d*) K. B. Sharpless and R. Manetsh, *Expert Opin. Drug Discovery*, 2006, **1**, 525–538.

2 (*a*) W. G. Lewis, L. G. Green, F. Grynszpan, Z. Radič, P. R. Carlier, P. Taylor, M. G. Finn and K. B. Sharpless, *Angew. Chem., Int. Ed.*, 2002, **41**, 1053–1057; (*b*) R. Manetsch, A. Krasinski, Z. Radič, J. Raushel, P. Taylor, K. B. Sharpless and H. C. Kolb, *J. Am. Chem. Soc.*, 2004, **126**, 12809–12818; (*c*) T. Suzuki, Y. Ota, Y. Kasuya, M. Mutsuga, Y. Kawamura, H. Tsumoto, H. Nakagawa, M. G. Finn and N. Miyata, *Angew. Chem., Int. Ed.*, 2010, **49**, 6817–6820; (*d*) T. Suzuki, Y. Ota, Y. Kasuya, M. Mutsuga, Y. Kawamura, H. Tsumoto, H. Nakagawa, M. G. Finn and N. Miyata, *Angew. Chem., Int. Ed.*, 2010, **122**, 6969–6972; (*e*) S. K. Mamidyala and M. G. Finn, *Chem. Soc. Rev.*, 2010, **39**, 1252–1261; (*f*) X. Hu and R. Manetsch, *Chem. Soc. Rev.*, 2010, **39**, 1316–1324; (*g*) S. W. Millward, R. K. Henning, G. A. Kwong, S. M. Pitram, H. D. Agnew, K. M. Deyle, A. Nag, J. E. Hein, S. S. Lee, J. Lim, J. A. Pfeilsticker, K. B. Sharpless and J. R. Heath, *J. Am. Chem. Soc.*, 2011, **133**, 18280–18288; (*h*) C. Peruzzotti, S. Borrelli, M. Ventura, R. Pantano, G. Fumagalli, M. S. Christodoulou, D. Monticelli, M. Luzzani, A. L. Fallacara, C. Tintori, M. Botta and D. Passarella, *ACS Med. Chem. Lett.*, 2013, **4**, 274–277; (*i*) W. Tieu, T. P. Soares da Costa, M. Y. Yap, K. L. Keeling, M. C. J. Wilce, J. C. Wallace, G. W. Booker, S. W. Polyak and A. D. Abell, *Chem. Sci.*, 2013, **4**, 3533–3537.

3 P. Anand and B. Singh, *Arch. Pharmacal. Res.*, 2013, **36**, 375–399.

4 A. Krasinski, Z. Radič, R. Manetsch, J. Raushel, P. Taylor, K. B. Sharpless and H. C. Kolb, *J. Am. Chem. Soc.*, 2005, **127**, 6686–6692.

5 For the crystal structure of the *syn*-(±)-**TZ2PIQA6**-mouse AChE complex, see: Y. Bourne, H. C. Kolbe, Z. Radic, K. B. Sharpless, P. Taylor and P. Marchot, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 1449–1454.

6 M. Whiting, J. Muldoon, Y.-C. Lin, S. M. Silverman, W. Lindstrom, A. J. Olson, H. C. Kolb, M. G. Finn, K. B. Sharpless, J. H. Elder and V. V. Fokin, *Angew. Chem., Int. Ed.*, 2006, **45**, 1435–1439.

7 V. P. Mocharla, B. Colasson, L. V. Lee, S. Röper, K. B. Sharpless, C.-H. Wong and H. C. Kolb, *Angew. Chem., Int. Ed.*, 2005, **44**, 116–120.

8 N. Willand, M. Desroses, P. Toto, B. Dirié, Z. Lens, V. Villeret, P. Rucktooa, C. Locht, A. Baulard and B. Deprez, *ACS Chem. Biol.*, 2010, **5**, 1007–1013.

9 N. P. Grimster, B. Stump, J. R. Fotsing, T. Weide, T. T. Talley, J. G. Yamauchi, Á. Nemecz, C. Kim, K.-Y. Ho, K. B. Sharpless, P. Taylor and V. V. Fokin, *J. Am. Chem. Soc.*, 2012, **134**, 6732–6740.

10 C. Ronco, E. Carletti, J.-P. Colletier, M. Weik, F. Nachon, L. Jean and P.-Y. Renard, *ChemMedChem*, 2012, **7**, 400–405.

11 (*a*) P. Camps, R. El Achab, D. M. Görbig, J. Morral, D. Muñoz-Torrero, A. Badia, J. E. Baños, N. M. Vivas, X. Barril, M. Orozco and F. J. Luque, *J. Med. Chem.*, 1999, **42**, 3227–3242; (*b*) P. Camps, R. El Achab, J. Morral, D. Muñoz-Torrero, A. Badia, J. E. Baños, N. M. Vivas, X. Barril, M. Orozco and F. J. Luque, *J. Med. Chem.*, 2000, **43**, 4657–4666.

12 (*a*) C. Ronco, G. Sorin, F. Nachon, R. Foucault, L. Jean, A. Romieu and P.-Y. Renard, *Bioorg. Med. Chem.*, 2009, **17**,

4523–4536; (*b*) C. Ronco, L. Jean and P.-Y. Renard, *Tetrahedron*, 2010, **66**, 7399–7404; (*c*) C. Ronco, L. Jean, H. Outaabout and P.-Y. Renard, *Eur. J. Org. Chem.*, 2011, 302–310; (*d*) C. Ronco, R. Foucault, E. Gillon, P. Bohn, F. Nachon, L. Jean and P.-Y. Renard, *ChemMedChem*, 2011, **6**, 876–888; (*e*) F. Nachon, E. Carletti, C. Ronco, M. Trovaslet, Y. Nicolet, L. Jean and P.-Y. Renard, *Biochem. J.*, 2013, **453**, 393–399.

13 A. Cavalli, M. L. Bolognesi, A. Minarini, M. Rosini, V. Tumiatti, M. Recanatini and C. Melchiorre, *J. Med. Chem.*, 2008, **51**, 347–371.

14 (*a*) P. Camps, X. Formosa, D. Muñoz-Torrero, J. Petrignet, A. Badia and M. V. Clos, *J. Med. Chem.*, 2005, **48**, 1701–1704; (*b*) P. Camps, X. Formosa, C. Galdeano, T. Gómez, D. Muñoz-Torrero, M. Scarpellini, E. Viayna, A. Badia, M. V. Clos, A. Camins, M. Pallàs, M. Bartolini, F. Mancini, V. Andrisano, J. Estelrich, M. Lizondo, A. Bidon-Chanal and F. J. Luque, *J. Med. Chem.*, 2008, **51**, 3588–3598; (*c*) E. Viayna, T. Gómez, C. Galdeano, L. Ramírez, M. Ratia, A. Badia, M. V. Clos, E. Verdaguer, F. Junyent, A. Camins, M. Pallàs, M. Bartolini, F. Mancini, V. Andrisano, M. P. Arce, M. I. Rodríguez-Franco, A. Bidon-Chanal, F. J. Luque, P. Camps and D. Muñoz-Torrero, *ChemMedChem*, 2010, **5**, 1855–1870.

15 (*a*) J. Peters, M. Trovaslet, M. Trapp, F. Nachon, F. Hill, E. Royer, F. Gabel, L. van Eijck, P. Masson and M. Tehei, *Phys. Chem. Chem. Phys.*, 2012, **19**, 6764–6770; (*b*) M. Trovaslet, M. Trapp, M. Weik, F. Nachon, P. Masson, M. Tehei and J. Peters, *Chem.–Biol. Interact.*, 2013, **1**, 14–18.

16 (−)-Huprine derivatives are about 2-fold more active than their racemic mixtures (*cf.* ref. 11*b*).

17 The configuration of the chiral center of PIQ has little effect on the binding affinity of the corresponding heterodimers (*cf.* ref. 4).

18 We failed to prepare regioselectively the *syn*-triazole heterodimers under ruthenium-catalyzed 1,3-dipolar cycloaddition.

19 For information, the most potent ligands of *Drosophila melanogaster* were *anti* substituted triazoles: Z. Radič, R. Manetsch, D. Fournier, K. B. Sharpless and P. Taylor, *Chem. Biol. Interact.*, 2008, **175**, 161–165.

20 Crystallization of the *m* or *rh*-AChE and *anti*-(±)-9-**HUPZ4-PIQ-A2** proved unsuccessful.

21 S. Senapati, Y. Cheng and J. A. McCammon, *J. Med. Chem.*, 2006, **49**, 6222–6230.

European Journal of Medicinal Chemistry 78 (2014) 455–467

Contents lists available at ScienceDirect

European Journal of Medicinal Chemistry

journal homepage: http://www.elsevier.com/locate/ejmech

Original article

Design, synthesis and biological evaluation of novel tetrahydroacridine pyridine- aldoxime and -amidoxime hybrids as efficient uncharged reactivators of nerve agent-inhibited human acetylcholinesterase

Maria Kliachyna [a], Gianluca Santoni [b], Valentin Nussbaum [a], Julien Renou [c], Benoit Sanson [b], Jacques-Philippe Colletier [b], Mélanie Arboléas [d], Mélanie Loiodice [d], Martin Weik [b,*], Ludovic Jean [c], Pierre-Yves Renard [c,*], Florian Nachon [b,d,*], Rachid Baati [a,*]

[a] Université de Strasbourg, Faculté de Pharmacie, CNRS/ UMR 7199 BP 24, 74 route du rhin 67401 Illkirch, France
[b] Commissariat à l'Energie Atomique, Institut de Biologie Structurale, F-38054 Grenoble; CNRS, UMR5075, F-38027 Grenoble; Université Joseph Fourier, F-38000, Grenoble, France
[c] Normandie University, COBRA, UMR 6014 and FR 3038; University of Rouen; INSA of Rouen; CNRS, 1 rue Tesniere 76821 Mont-Saint-Aignan, Cedex, France
[d] Département de Toxicologie, Institut de Recherche Biomédicale des Armées BP7391993 Brétigny/s/Orge, France

ARTICLE INFO

Article history:
Received 27 November 2013
Received in revised form
3 March 2014
Accepted 14 March 2014
Available online 15 March 2014

Keywords:
Organophosphorus nerve agents
Reactivation of acetylcholinesterase
Pyridinaldoximes
Oximes
Hybrids
Tacrine

ABSTRACT

A series of new uncharged functional acetylcholinesterase (AChE) reactivators including heterodimers of tetrahydroacridine with 3-hydroxy-2-pyridine aldoximes and amidoximes has been synthesized. These novel molecules display *in vitro* reactivation potencies towards VX-, tabun- and paraoxon-inhibited human AChE that are superior to those of the mono- and bis-pyridinium aldoximes currently used against nerve agent and pesticide poisoning. Furthermore, these uncharged compounds exhibit a broader reactivity spectrum compared to currently approved remediation drugs.

1. Introduction

Organophosphorus compounds (OP) include the extremely toxic chemical warfare agents (CWA) (sarin, soman, cyclosarin, tabun, methylphosphonothioate VX) and pesticides (paraoxon, parathion, tetraethyl pyrophosphate (TEPP)) (Fig. 1). Their acute toxicity results from the irreversible inhibition of acetylcholinesterase (AChE) through phosphylation of its catalytic serine [1]. Accumulation of neurotransmitter acetylcholine (ACh) at cholinergic synapses ensues, leading to nervous and respiratory failures. Depending on the class of OP and on the administrated dose, death can occur within minutes [2].

Due to the similarity between the chemical precursors of CWA and pesticides, and to the relatively simple chemistry involved in their synthesis, efforts to control the proliferation of these agents have proved of limited success [3]. Illustrative examples include the terrorist attack in the Tokyo subway in 1995, the bombing of Kurd civilians during the Iraq–Iran war in 1988, and that of civilians in Syria, as reported in August 2013. Additionally, despite the international efforts aimed at regulating and lessening the use of these environmentally toxic compounds, ca. 100 different OP are still used intensively as pest control agents, with only anecdotal monitoring. This results in about 3,000,000 acute intoxications per year, 200,000 of which lead to death [4,5]. Therefore, the development of effective measures to counteract OP poisoning remains a challenging issue to protect and treat both civilian and military populations [6].

* Corresponding authors.
E-mail addresses: martin.weik@ibs.fr (M. Weik), pierre-yves.renard@univ-rouen.fr (P.-Y. Renard), florian@nachon.net (F. Nachon), rachid.baati@unistra.fr (R. Baati).

http://dx.doi.org/10.1016/j.ejmech.2014.03.044
0223-5234/

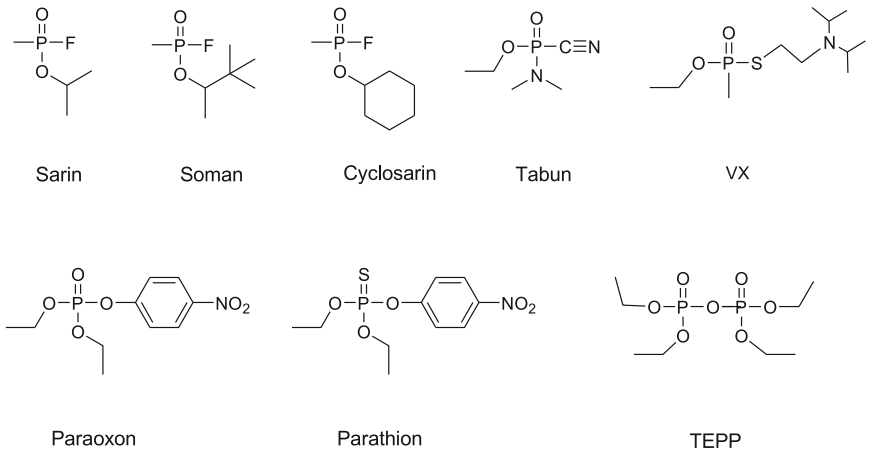


Fig. 1. Structures of organophosphorus CWA and pesticides.

The current treatment against OP poisoning consists in the administration of a combination of atropine (antimuscarinic agent) and diazepam (anticonvulsant drug), to limit convulsions, and of a standard pyridinium oxime (pralidoxime, trimedoxime, HI-6, obidoxime, and HLö-7, Fig. 2) to reactivate AChE. Oximes exert their action on OP-inhibited AChE by attacking the phosphorus atom of the phosphylated serine, yielding to the removal of the phosphonate and recovery of the enzyme's catalytic activity. To this end, pyridinium oximes must display high nucleophilicity, which is generally attained by the formation of an oximate anion at physiological pH. As of today, however, not a single oxime has proven equally effective against all types of OP-inhibited AChE, and most are ineffective for the reactivation of tabun-inhibited hAChE [6].

Another weakness of currently approved pyridinium aldoximes is their difficulty in crossing the blood–brain barrier (BBB) owing to the permanent charge carried by the oximate [7]. For example, it was estimated using *in vivo* rat brain microdialysis coupled with HPLC/UV, that the BBB penetration of the most commonly used oxime, 2-PAM, is only 10% [8]. Therefore, oximes reactivating AChE in the peripheral nervous system are not effective in the brain and, consequently, do not protect against the neurological effects of OP-exposure.

To overcome this obstacle, we recently reported the synthesis of a series of uncharged oxime-based compounds, both able to cross the BBB and to reactivate OP-inhibited AChE in the CNS [6]. Monoisonitrosoacetone (MINA) and diacetylmonooxime (DAM) respectively bearing the α-ketoaldoxime and ketoxime moieties (Fig. 3) were also reported to cross the BBB, yet their *in vitro* reactivation potency towards OP-inhibited AChE was lower than that exhibited by 2-PAM [9]. Connection of an α-ketoaldoxime function to a piperidine-derived peripheral site ligand (PSL) allowed increasing the affinity for AChE, resulting in higher reactivation rates of sarin- and VX- inhibited AChE. Yet, these compounds were still less efficient than currently used pyridinium oximes towards the latter, and totally inefficient towards tabun-inhibited AChE [10]. Hydroxyimino acetamides **1a**–**b** showed *in vitro* reactivation efficacy superior to MINA and DAM, but remained less effective than 2-PAM, except for cyclosarin [11]. Likewise, amidine-oxime reactivators **2a,b** were less potent than 2-PAM in reactivating AChE *in vitro* [12]. Low reactivation potency of these new uncharged reactivators is likely due to their higher pKa, which results in inefficient deprotonation at physiological pH and thus, in a reduced nucleophilicity (for instance, for MINA, pKa = 8.3) [13].

Recently we found that 3-hydroxy-2-pyridine aldoximes **3** and amidoxime **4** (Fig. 4) exhibited high reactivation first order rate constant towards VX-inhibited hAChE ($k_r = 0.5 \pm 0.1$ min^{-1} and 0.08 ± 0.1 min^{-1}, respectively) [14,15], *i.e.* values either similar or one order of magnitude larger than that displayed by 2-PAM ($k_r = 0.06 \pm 0.01$ min^{-1}). Owing to the lower affinity of **3** and **4** for VX-hAChE (second order reactivation rate constants $k_{r2} = k_r/K_D$ are 0.015 and 0.0026 mM^{-1} min^{-1} for **3** and **4**, respectively, to compare to 0.28 mM^{-1} min^{-1} for 2-PAM), much higher concentrations are however required to reach the desirable reactivation of the enzyme.

We have shown that connection of 3-hydroxy-2-pyridinaldoxime reactivators with a PSL ligand, such as phenyltetrahydroisoquinoline (PIQ) [16], results in a dramatically improved affinity for the enzyme and, consequently, in increased

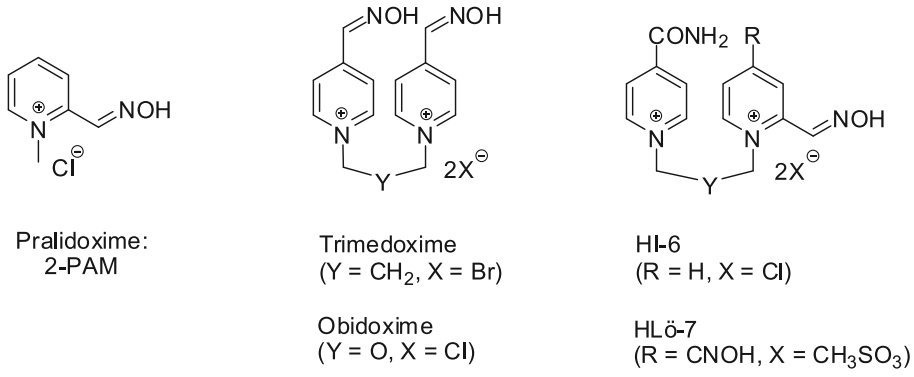


Fig. 2. Structures of developed pyridinium aldoxime reactivators.

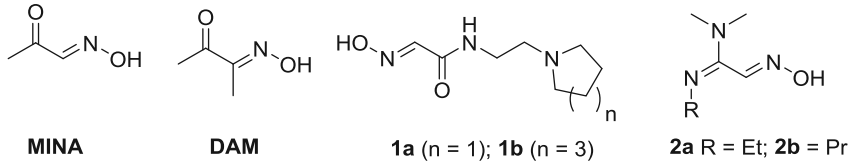


Fig. 3. Neutral reactivators: α-ketoaldoximes, ketoxime, hydroxyimino acetamides, and amidine-oximes.

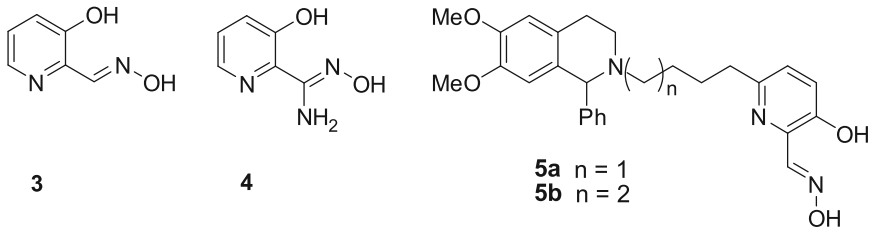
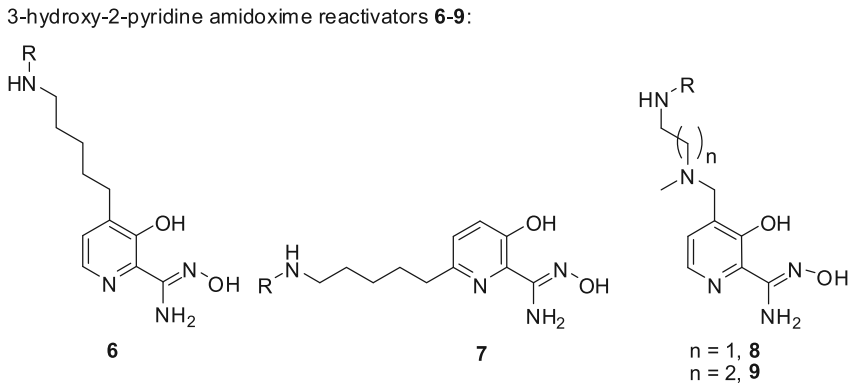


Fig. 4. 3-hydroxy-2-pyridine aldoxime, 3-hydroxy-2-pyridine amidoxime, and phenyl-tetrahydroisoquinoline-3-hydroxy-2-pyridine aldoxime AChE reactivators.

reactivation potency towards OP-inhibited hAChE [17]. Recently, similar beneficial effects have been reported for tryptoline-3-hydroxypyridine-2-aldoxime hybrids for the reactivation of OP-inhibited h-BChE [17c]. PIQ-functionalized 3-hydroxy-2-pyridine aldoximes indeed displayed reactivation potency either equalling or exceeding those of HI-6 and obidoxime. For example, **5a** was 2.3 and 2.8 times more efficient than obidoxime and HI-6, respectively, and 5 times more potent than trimedoxime, the best known reactivator of tabun-inhibited hAChE ($k_{r2} = 3.4$ mM^{-1} min^{-1} for **5a**, as compared to 0.7 mM^{-1} min^{-1} for trimedoxime).

One of the important drawbacks when using pyridinium aldoximes (especially for pyridinium-4-aldoximes) is that the



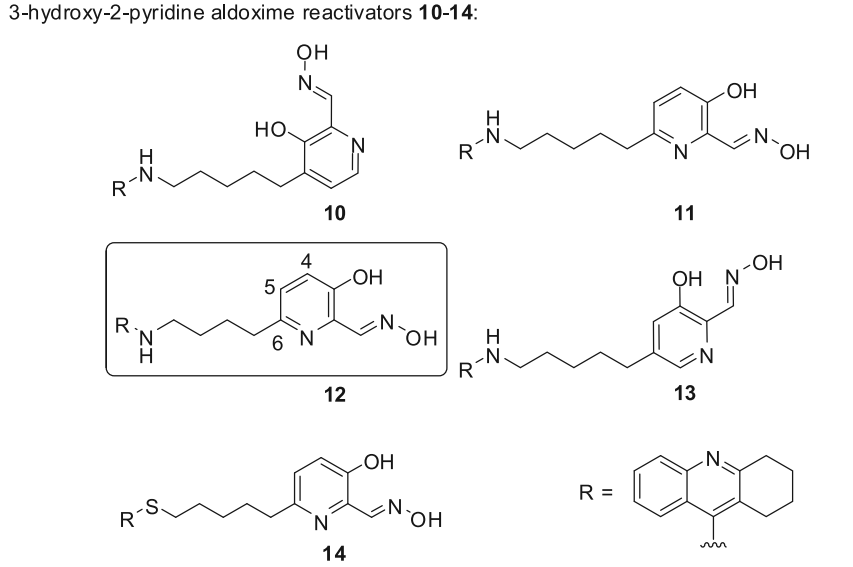


Fig. 5. Structures of novel uncharged AChE hybrid reactivators investigated.

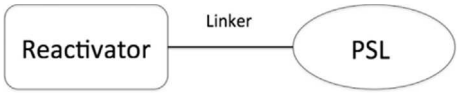


Fig. 6. General structure of targeted hetero-bifunctional reactivators.

reactivation products, phosphyloximes, can themselves display high affinity for the enzyme, and thus prolong AChE inhibition (recapture phenomenon) [18]. In the case of oximes **3**, **5a**, **5b**, as well as in that of amidoxime **4**, the presence of the 3-OH group in α-position of the oxime/amidoxime moiety allows subsequent intramolecular cyclisation, resulting in isoxasole formation and thus in the bypass of this possible complication [14,15].

In the following, we describe the rational design, synthesis and *in vitro* evaluation of a new class of uncharged reactivator molecules. Their design is based on a 1,2,3,4-tetrahydroacridine (tacrine) substructure, able to bind to the peripheral site of AChE [19,20], coupled with various linkers to either 3-hydroxy-2-pyridine aldoxime or 3-hydroxy-2-pyridine amidoxime (Fig. 5).

2. Results and discussion

2.1. Rational design of bifunctional reactivator molecules

Acetylcholinesterase (AChE) features an active site buried at the bottom of a deep and narrow gorge [21]. At the entrance of the latter, a peripheral binding site serves as a transient binding location for substrates and products *en route* to/from the active site [22], but also as an anchor point for ligands [23]. Bifunctional inhibitors have been developed that display higher affinity, and which are based on two moieties binding at either site that are connected by a flexible linker [19,20,24]. By analogy, we designed bifunctional reactivators that comprise a peripheral site ligand (PSL) connected to a reactivator function by a covalent linker (Fig. 6). In addition to increasing the affinity of the reactivator for the enzyme, the use of a PSL further allows preventing unproductive binding to the active site as it was, for example, observed with 2-PAM [25].

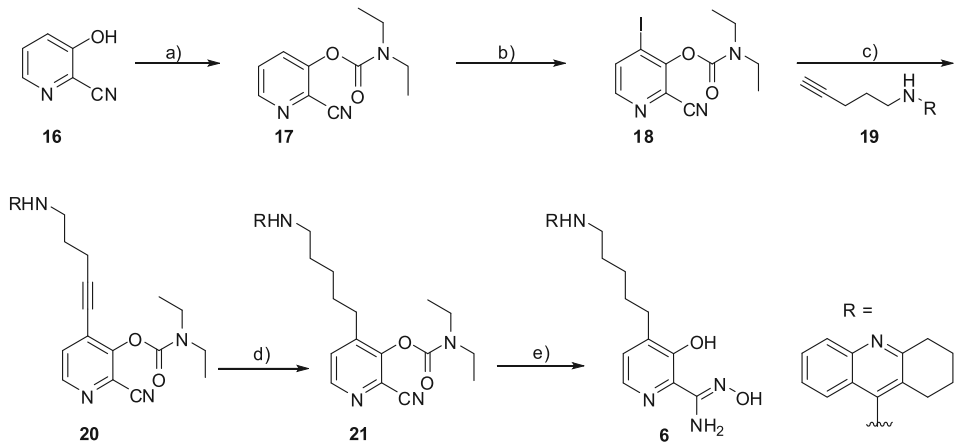
In practice, we used as a starting model the X-ray crystal structure of *Torpedo californica* (*Tc*)AChE in complex with a bis-tacrine inhibitor [19], which allowed predicting how the tacrine moiety would bind to the peripheral site. Obviously, the length of the linker had to be optimized so that the oxime moiety could approach the enzyme's phosphylated serine without unbinding from the peripheral site. *In silico* optimization of the linker-length was performed to maximize the reactivation potency of the bifunctional reactivators, and is described in Section 3.

3. Synthesis of the hybrid molecules

3.1. Synthesis of 1,2,3,4-tetrahydro-9-aminoacridine α-hydroxypyridine amidoximes

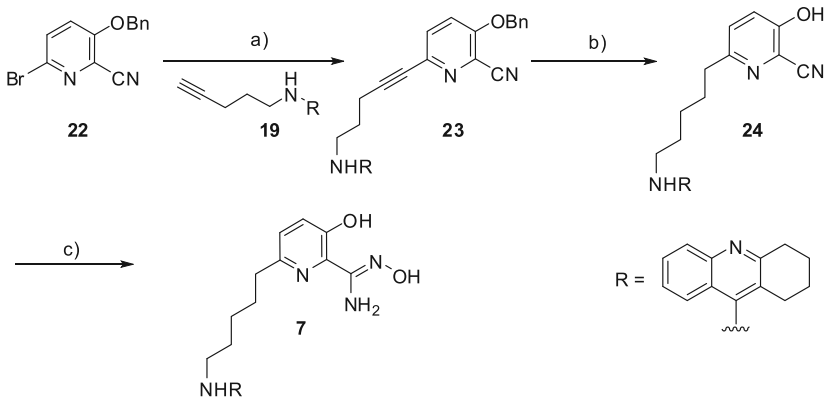
The 4-substituted α-hydroxypyridine amidoxime **6** was synthesized starting from commercially available 3-hydroxy-2-cyanopyridine **16** (Scheme 1). The introduction of the *ortho*-directing *N*-(diethylamino)carbamate group (both as ortho directing group for the metallation, and as further phenol protecting group) was achieved in standard conditions allowing to obtain compound **17** with a good yield (95%). Subsequent o*rtho*-lithiation and substitution with iodine gave the key building block **18**. The Sonogashira coupling with alkyne **19** [26] proceeded cleanly and allowed producing the hybrid compound **20** with 87% yield. Reduction of the alkyne **20** in the presence of Pd/C yielded the corresponding saturated compound **21**. Attempts to selectively deprotect the *N*-(diethylamino)carbamate group of **21** in either acidic or basic conditions proved unsuccessful, due to the competing hydrolysis or degradation of the cyano group. After a set of screening conditions, this critical reaction was successfully achieved upon the treatment of **21** with a large excess (30 equiv.) of hydroxylamine. Under these conditions, the *N*-(diethylamino)carbamate was removed in a *one-pot* process, and the cyano group was concomitantly converted to the desired amidoxime **6**. Thus, the final hybrid molecule **6** was synthesized in a convergent 5-steps protocol with 13% overall yield.

The synthesis of α-hydroxyamidoxime tetrahydroacridine conjugate **7** is shown in Scheme 2. It differs from that of conjugate **6** in that the substitution on the pyridine ring was performed using the optimized modular approach described previously for the Sonogashira coupling of bromopyridine **22** [27] with alkyne **19** (Scheme 2). Indeed, this coupling worked nicely and gave **23** in excellent yield (86%). The *one-pot* alkyne reduction and *O*-debenzylation proceeded cleanly to give nitrile **24**, which was directly reacted



Reagents and conditions: a) *N*-diethylamino carbamate, pyridine, 0°C, **95%**; b) THF, *n*-BuLi, -78°C, then I_2, **48%**; c) **19**, $Pd(PPh_3)_4$, CuI, Et_3N, THF, 20 °C, **87%**; d) H_2, Pd/C, EtOAc, **99%**; e) $NH_2OH.HCl$, pyridine, EtOH reflux, **32 %**

Scheme 1. Synthesis of amidoxime **6**.



Reagents and conditions: a) **19**, $Pd(PPh_3)_4$, CuI, Et_3N, THF, 20°C, **86%**; b) H_2, $Pd(OH)_2/C$, MeOH, **80%**; c) $NH_2OH.HCl$, Na_2CO_3, EtOH, H_2O, **65%**.

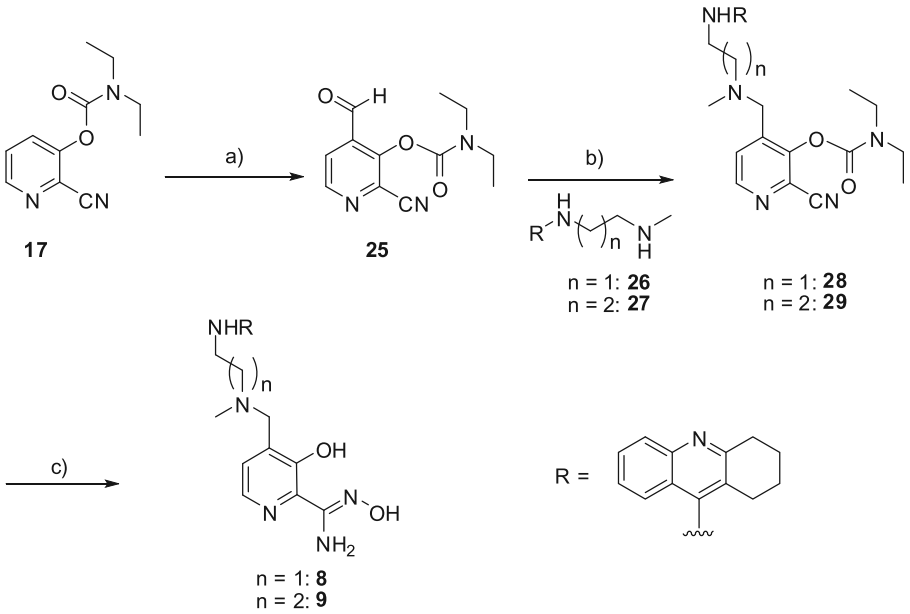
Scheme 2. Synthesis of amidoxime **7**.

with excess of hydroxylamine hydrochloride to give expected amidoxime **7** with 45% overall yield over the 3 steps.

Docking studies (unpublished results) suggested that the presence of a tertiary amine in the linker could increase the affinity of the molecule for AChE -and thus presumably enhance its reactivation efficacy -owing to the amine protonation at physiological pH that would allow cation-π interaction with the tyrosine residues in AChE gorge [28]. To verify this hypothesis, we therefore prepared hybrids **8** and **9**. The synthesis of these α-hydroxypyridine amidoxime hybrids connected with tetrahydroacridine by *N*–Me linker was based on the reductive amination of aldehyde **25** (obtained by *ortho*-lithiation and subsequent formylation of **17**) with amines **26** and **27** [29] (Scheme 3). The treatment of **28** and **29** with excess hydroxylamine gave α-hydroxypyridine amidoximes **8** and **9** in 13% and 19% overall yields, respectively, over the 3 steps syntheses.

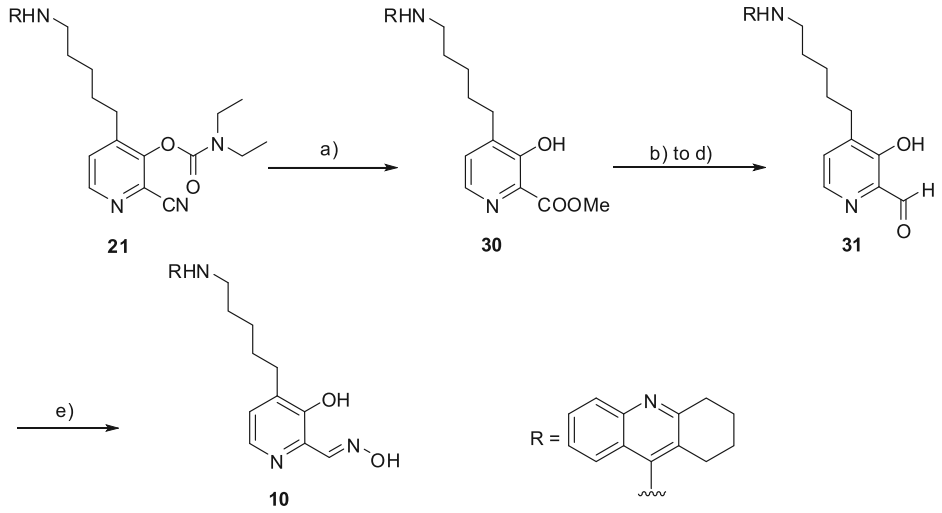
3.2. Synthesis of 1,2,3,4-tetrahydro-9-aminoacridine α-hydroxypyridine aldoximes

For the synthesis of α-hydroxypyridine oxime hybrids, we first started with the synthesis of 3-hydroxypyridine aldoxime with the linker in position 4, using the common intermediate **21**. For these molecules, we privileged the *one-pot* carbamate group removal and methanolysis of cyano- function to give methyl ester **30** with a reasonable yield of 42% (Scheme 4). Next, **30** was converted to **31** *via* a three-step [17] sequence that included: (a) TBDMS protection of phenol function; (b) reduction of methyl ester to aldehyde with DIBAL-H; (c) and subsequent deprotection of the silyl-oxy function using TBAF (33% over three steps). Reaction of the resulting aldehyde **31** with NH_2OH resulted in the formation of the desired pyridinaldoxime **10**, with 7% overall yield.



Reagents and conditions: a) *n*-BuLi, THF, -78°C, then DMF, **53%**; b) **26** or **27**, $NaBH(OAc)_3$, THF, 20 °C, **75-99%**; c) $NH_2OH.HCl$, pyridine, EtOH, reflux, **32-37%.**

Scheme 3. Synthesis of amidoximes **8** and **9**.



Reagents and conditions: a) MeOH, H_2SO_4, reflux, 24 h, **42%**; b) TBDMSCl, DMF, rt; c) DIBAL-H,CH_2Cl_2, -78°C; d) TBAF, THF, 0°C, **33%** over 3 steps; e) NH_2OH.HCl, CH_3COONa, EtOH, **51%**.

Scheme 4. Synthesis of oxime **10**.

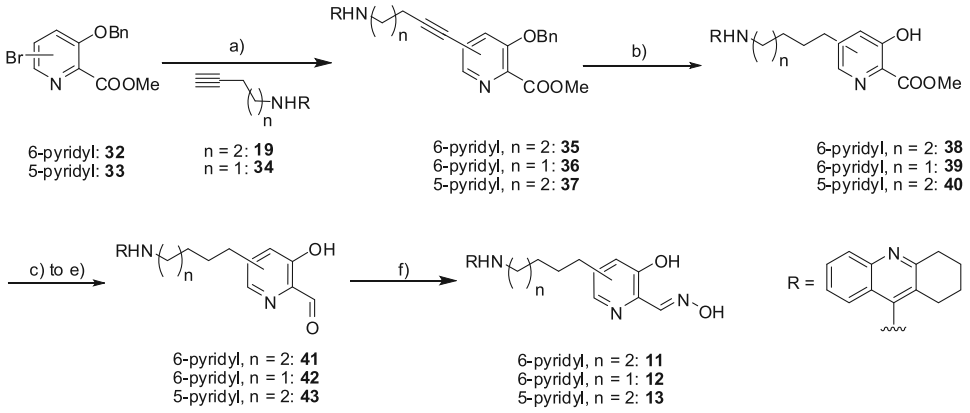
The 5- and 6-substituted α-hydroxypyridine oximes analogues **11**–**13** were obtained using the modular synthetic approach, consisting in the Sonogashira coupling reaction of 6-bromopyridine **32** [17] or 5-bromopyridine **33** [30] with alkynes **19** or **34** [31], respectively, followed by reduction of the alkyne and finally *O*-debenzylation. The subsequent three-step transformation of methyl ester to aldehyde, ensued by reaction with hydroxylamine (Scheme 5), yielded the desired α-hydroxypyridine oximes **11**–**13** in 15, 28 and 18% yields.

3.3. Synthesis of 1,2,3,4-tetrahydro-9-thioacridine α-hydroxypyridine aldoxime

In order to tune the affinity of hybrid reactivator for hAChE, the structure of the PSL was changed by the substitution of NH at position 9 of 1,2,3,4-tetrahydroacridine with a sulphur atom. In short, conjugate **14** bearing 1,2,3,4-tetrahydro-9-thioacridine as PSL was synthesized according to Scheme 6 by alkylating 1,2,3,4-tetrahydro-9-thioacridine **44** [32] with **45** to give **46**, which was then used for Sonogashira coupling reaction with **47** to provide compound **48**. Contrary to **35**–**37**, the attempts to reduce the triple bond of **48** using Pd/C or $Pd(OH)_2$/C failed, possibly due to the poisoning of catalyst by the reaction with thioether groups. The use of only 1.5 equiv of PtO_2 yet resulted in successful hydrogenation, allowing to obtain compound **49** with 64% yield. The further three-step sequence comprising protection of the phenol group, reduction of methyl ester with DIBAL-H into aldehyde, and the following deprotection using TBAF generated the desired aldehyde **50**, with 42% yield. Finally, reaction of **50** with NH_2OH resulted in the formation of pyridinaldoxime **14**, with 6% overall yield.

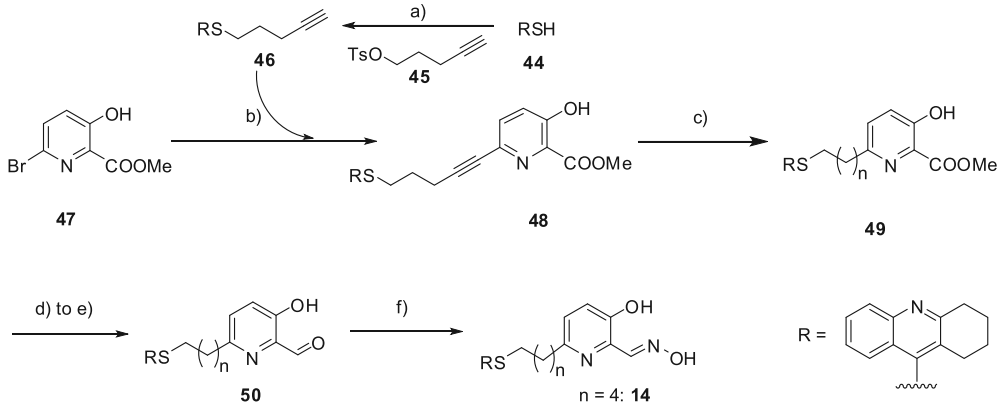
4. Docking and molecular dynamics simulations

Ternary complexes between VX-inhibited hAChE and 3-hydroxypyridine aldoxime with 2, 3, 4 (compound **12**) or 5 (compound **11**) carbons linker were investigated by molecular docking and molecular dynamics simulations (see Supporting Information). Flexible docking experiments were undertaken, in which the side chains of active site gorge residues key to the binding of the reactivators were allowed to wander from their native position (see



Reagents and conditions: a) **19** or **34**, $Pd(PPh_3)_4$, CuI, Et_3N, THF, 20°C, **78-98%** ; b) H_2, $Pd(OH)_2$/C, MeOH, **89-96 %**; c) TBDMSCl, DMF, rt; d) DIBAL-H, CH_2Cl_2, -78°C; e) TBAF, THF, 0°C, **43%-54%** over 3 steps; f) NH_2OH.HCl, CH_3COONa, EtOH, **55%-62%**.

Scheme 5. Synthesis of oximes **11**, **12**, and **13**.



Reagents and conditions: a) **45**, $(n\text{-Bu})_4NBr$, 50% aq. NaOH, THF/H_2O, **56%**; b) **46**, $Pd(PPh_3)_2Cl_2$, CuI, Et_3N, THF, 50°C, **90%**; c) H_2, PtO_2, 5 atm, AcOEt, rt, **64 %**; d) TBDMSCl, DMF, rt; d) DIBAL-H, CH_2Cl_2, -78°C; e) TBAF, THF, 0°C, **42%** over 3 steps; f) $NH_2OH.HCl$, CH_3COONa, EtOH, **57%**.

Scheme 6. Synthesis of oxime **14**.

Supporting Information). In all four docking experiments, a conformational change was observed in the side chain of peripheral-site Trp286 (Fig. 7), faithfully reproducing the crystallographic observation from the bis-tacrine/*Tc*AChE complex (PDB code 2CEK) [19]. This conformational change results in a π–π sandwiching of the tacrine moiety between the aromatic side chains of Tyr72 and Trp286. The scoring function of the docking algorithm yielded similar binding energies for the 4 molecules of about − 9 kcal/mol, indicating that the linker is sufficiently long to prevent unproductive binding of the reactivators in the gorge of the inhibited enzyme.

The structures resulting from the 4 independent dockings were then subjected to molecular dynamics (MD) simulations for 5 ns at a temperature of 300 K. For each, the distance between the aldoxime oxygen and the VX-AChE phosphorus atom was monitored along the simulation trajectory and the distributions of distances were analyzed (Fig. 8). The shorter the distance, the higher the probability to form a covalent bond between the reactivator oxygen and the phosphorus atom, that is, the more likely should reactivation occur at a fast rate. Molecule **12** brings the aldoxime oxygen closer to the phosphorus atom and more frequently (Fig. 8), suggesting that a 4-carbon linker between the PSL and the aldoxime is optimum for reactivation.

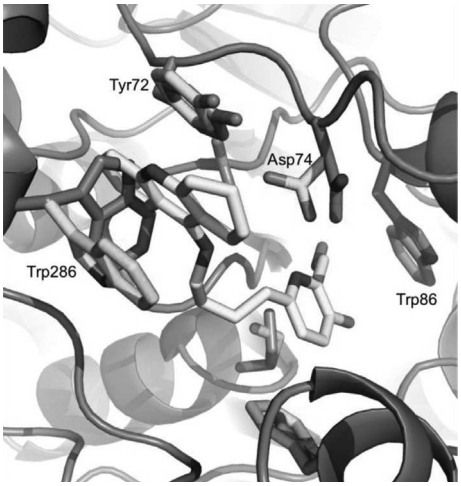


Fig. 7. Docked conformation of molecule **12** (yellow) at the peripheral site of VX-hAChE. The docked conformations of the flexible residues are shown in cyan and the native conformations (pdb code 1b41) in blue. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

5. *In vitro* reactivation of VX-, tabun and POX-inhibited hAChE

The abilities of amidoximes **6**–**9**, as well as of aldoximes **10**–**14** to reactivate *in vitro* VX-, tabun- and ethyl-paraoxon inhibited hAChE were investigated by spectrophotometry using the Ellman's reaction. Reactivation kinetics were compared to those of known reactivators, including 2-PAM, obidoxime, HLö-7, trimedoxime (TMB4) and HI-6. 6-Substituted α-hydroxypyridine amidoxime **7** and 4-substituted α-hydroxypyridine oxime **10** were both totally inefficient in reactivating either VX-inhibited hAChE or tabun-inhibited hAChE. (Table 1)

In contrast, 4-substituted tetrahydroacridine α-hydroxypyridine amidoximes **6** (C5 alkyl linker), **8** and **9** (*N*–Me linkers) displayed improved affinity and higher reactivation efficiencies (k_{r2}) towards VX-inhibited hAChE than both 2-PAM and trimedoxime. While they remained less efficient than obidoxime, HLö-7, and HI-6 in reactivating VX-inhibited hAChE, it is noteworthy that these compounds are the first amidoxime-based reactivators that are efficient towards VX-inhibited hAChE. Quite remarkably, aldoximes **11** and **12** substituted in position 6 of the pyridine ring by the C5 and C4 alkyl linker proved more efficient in reactivating VX-inhibited hAChE than 2-PAM, obidoxime, trimedoxime, and HI-6. Although oximes **11** and **12** do not surpass the reactivating efficiency of HLö-7, nor that of our previously described compound **5a**, they represent an interesting new starting point for further improvement. It must be stressed, in particular, that since they neither bear a permanent charge nor a tertiary amine, they should be able to cross the BBB quite efficiently.

Also noteworthy is the fact that oxime **12** is able to reactivate tabun-inhibited hAChE (Table 2), an OP complex known to be reluctant to reactivation due to its weak electrophilicity and to the steric hindrance imposed on the phosphorus atom in the tabun-hAChE adduct. In fact, compound **12** exhibited higher *in vitro* reactivation potency than all quaternary pyridinium aldoximes described to date. Despite a lower reactivation rate constant (k_r) compared to obidoxime and trimedoxime, compound **12** was 4-fold more efficient than trimedoxime, the best pyridinium aldoxime reactivator of tabun-inhibited hAChE. Again, this increase in

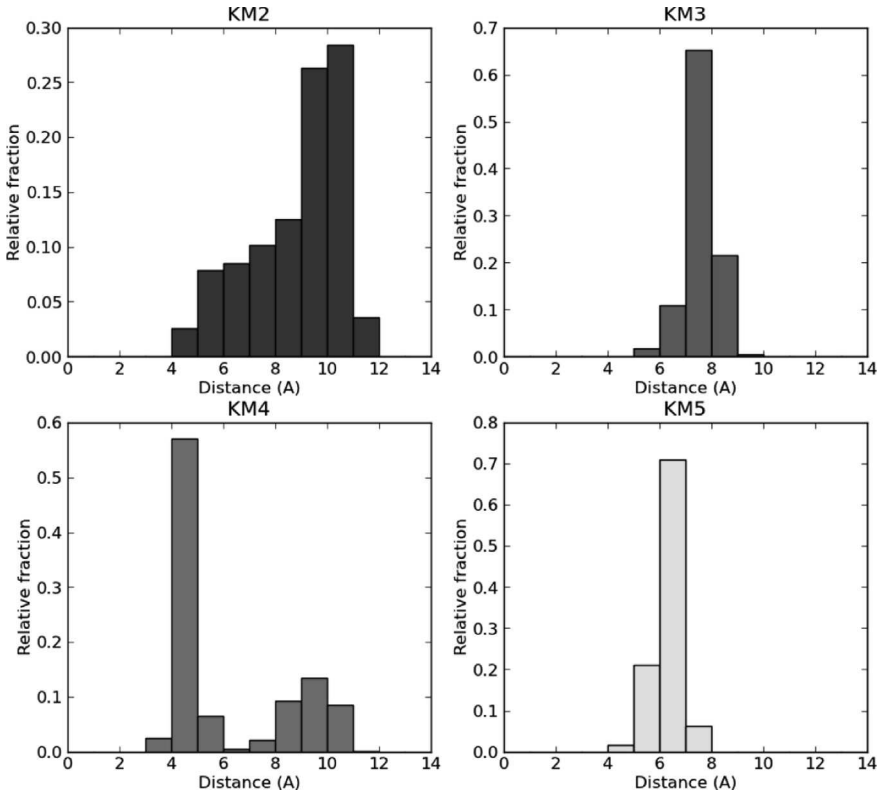


Fig. 8. Distribution of distances between the aldoxime oxygen and the VX phosphorus analyzed along the simulation trajectory for the four different linker lengths (KM 2: $-CH_2-CH_2-$; KM 3: $-CH_2-CH_2-CH_2-$, KM 4: $-CH_2-CH_2-CH_2-CH_2-$, **12**; KM 5: $-CH_2-CH_2-CH_2-CH_2-CH_2-$, **11**).

Table 1
In vitro reactivation of VX-hAChE.

Reactivator	k_r (min^{-1})	K_D (μM)	k_{r2} (mM^{-1} min^{-1})[b]
2-PAM	0.06 ± 0.01	215 ± 75	0.28
Obidoxime	0.60 ± 0.05	54 ± 12	11
HLö-7	0.49[a]	7.8[a]	63[a]
Trimedoxime	nd[c]	nd[c]	0.50 ± 0.02[c]
HI-6	0.44 ± 0.15	50 ± 26	9
6	nd[c]	nd[c]	0.65[c]
8	0.0094 ± 0.0004	15 ± 2	0.6
9	0.032 ± 0.001	30 ± 3	1.1
11	0.56 ± 0.12	41 ± 11	13
12	0.72 ± 0.07	31 ± 6	22
13	nd[d] (k_{obs} = 0.0038 ± 0.00006 min^{-1} at 1 μM)		
14	nd[d] (k_{obs} = 0.0076 ± 0.0005 min^{-1} at 200 μM)		

[a] From the ref. [33].
[b] $k_{r2} = k_r/K_D$.
[c] If [reactivator] $\ll K_D$, then there is a linear dependence between k_{obs} and [reactivator]: $k_{obs} = (k_r/K_D)$ [reactivator]. In this case, k_r and K_D cannot be determined, but $k_{r2} = k_r/K_D$ is the slope of the line.
[d] Not determined if k_{obs} is <0.01 min^{-1} at practical concentration.

Table 2
In vitro reactivation of tabun-hAChE.

Reactivator	k_r (min^{-1})	K_D (μM)	k_{r2} (mM^{-1} min^{-1})[a]
Obidoxime	0.04 ± 0.006	250 ± 110	0.16
HLö-7	0.020 ± 0.0007	106 ± 15	0.2
Trimedoxime	0.085 ± 0.005	145 ± 25	0.7
8	Less than 10% reactivation at 100 μM in 2 h		
9			
11	0.0075 ± 0.0005	30 ± 6	0.27
12	0.021 ± 0.001	7.1 ± 1.5	3

[a] $k_{r2} = k_r/K_D$.

efficiency is likely to be ascribed to its significantly greater affinity towards the tabun-hAChE conjugate (K_D). The better reactivity of compound **12** on both VX- and tabun-inhibited hAChE, as compared to compound **11**, suggest that a 4-carbon linker is optimum for both a good binding to the inhibited enzyme and an accurate orientation of the oxime function needed for displacing of the phosphyl or phosphoramidyl moiety.

As mentioned already, most organophosphate intoxication results from pesticide poisoning, which involves the formation of a diethylphosphoryl conjugate of hAChE. Testing molecules for their ability to reactivate ethylparaoxon-inhibited hAChE, therefore, allows evaluating them for their general efficacy against pesticide poisoning. Interestingly, compounds **11** and **12** proved more efficient than obidoxime and HLö-7 towards paraoxon-inhibited hAChE, (Table 3). Much like the trend observed in the case of VX, decreasing the linker length from 5 to 4 carbons led to a 1.5-fold increase in kr.

We note that compounds **7**, **10**, **11** and **12** are also able to inhibit native hAChE, with IC_{50} similar to tacrine and tacrine-pyridyl heterodimer **51** (Table 4). While these inhibitory properties could be envisaged as a drawback of our strategy aiming at increasing the efficacy of reactivators by increasing their affinities for the target enzyme, we argue conversely. Indeed, it remains to be proven that significant reversible inhibition of AChE at a concentration similar to that used to reactivate of OP-inhibited AChE is really problematic *in vivo*. Also, it was shown that reversible inhibitors have a protective effect on AChE, preventing its phosphylation by transiently occupying the active site [35].

Interestingly, 4-substituted pyridine aldoxime **10** showed no reactivation of VX-, tabun- and paraoxon-inhibited hAChE and a higher inhibition compared to tacrine. This is in strong contrast with 6-substituted compound **11** and suggests that 4-substitution of pyridine ring favours the binding of tacrine moiety to the

Table 3
In vitro reactivation of paraoxon-hAChE.

Reactivator	k_r (min^{-1})	K_D (μM)	k_{r2} (mM^{-1} min^{-1})[b]
Obidoxime	0.81[a]	32.2[a]	20[a]
HLö-7	0.63 ± 0.04	210 ± 31	3
Trimedoxime	0.34 ± 0.02[a]	47.8 ± 6.9[a]	17[a]
8	Less than 10% reactivation at 100 μM in 2 h		
9	nd[d] (k_{obs} = 0.0043 ± 0.0004 min^{-1} at 100 μM)		
11	0.0228 ± 0.0006	1.1 ± 0.1	20.9
12	0.111 ± 0.002	3.6 ± 0.2	31

[c]Not determined if k_{obs} is <0.01 min^{-1} at practical concentration.
[a] From the ref. [34].
[b] $k_{r2} = k_r/K_D$.

catalytic site of the enzyme, rather than at its peripheral site. The substitution of tacrine with thiotacrine as PSL resulted in three order of magnitude lower inhibition of native hAChE by model compound **52** and **14**, as compared to compounds **51** and **11**, respectively. However, oxime **14** exhibited very poor reactivation of VX-hAChE, demonstrating that the key to the improvement of the reactivation potency is finding the optimal balance between affinity of the bifunctional reactivator for the enzyme and positioning of the oxime function relative to the phosphorus atom of the OP-hAChE adduct.

6. Conclusion

We have described the synthesis and the *in vitro* biological properties of a series of novel, promising, non-permanently charged AChE reactivators, including differently substituted tacrine α-hydroxypyridine oximes and amidoximes. We demonstrated that the amidoxime function can be used as a reactivator function, and allows producing hybrid molecules that equal or surpass currently approved mono- and bis-pyridinium oximes, both in terms of reactivation efficiency and broadness of their activity spectrum. While it remains to be evaluated if the molecules are centrally active, or require further modifications to be so, this new family of oximes holds great promise for the medical treatment of OP-poisoning.

7. Experimental part

7.1. General materials and methods

All reactions were carried out in a flame dried glassware under an argon atmosphere with dry solvents, under anhydrous conditions unless otherwise indicated. The chemicals used in the synthesis and solvents were purchased from Aldrich, Acros, Alfa-Aesar or ABCR and used without further purification. Thin layer chromatographies (TLC) were carried out using Merck silica gel plates (silica gel 60 F_{254}). TLC plates were visualized using UV light (254 nm), followed by detection with cerium sulphate and ninhydrin. The column chromatographies were performed using Merck silica gel 60 (0.063–0.200 mm, 70–230 mesh, ASTM) according to a standard procedure. Analytical HPLC chromatography (purity >95%) were performed using SPD-20A Shimadzu apparatus equipped with a Zorbax 300SB-C18 column, using a mixture of acetonitrile and water as eluent (0.1% TFA). ^{1}H NMR and ^{13}C NMR spectra were recorded on Brucker DPX300, DPX400 and DPX500. Chemical shifts (δ) are quoted in parts per million and are calibrated relative to solvent residual peaks. Multiplicities are reported as follows: s = singlet, d = doublet, t = triplet, m = multiplet, q = quartet, qui = quintet. Coupling constant are reported in Hertz. Mass spectra were recorded on an MS/MS high resolution Micromass ZABSpecTOF spectrometer.

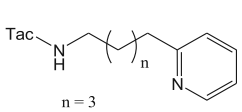
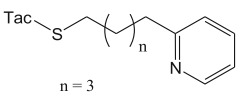
7.2. General procedures and analytical data of key compounds ***35***, ***36***, ***38***, ***39***, ***41***, ***42***, ***11***, and ***12***

General procedure 1 for the Sonogashira coupling reaction (synthesis of compounds **20**, **23**, **35**, **36**, and **37**). The flask containing solution of iodo- or bromopyridine (1 equiv) in THF/Et_3N was evacuated and filled with Ar three times before the addition of catalysts $Pd(PPh_3)_4$ (0.1 equiv) and CuI (0.2 equiv). After degazation with Ar the mixture was stirred at the room temperature for 5 min, then the degazed solution of alkyne (1.1 equiv) was added and the reaction mixture was stirred during 20 h at the room temperature. After concentration at reduced pressure the residue was purified by the column chromatography.

General procedure 2 for *N*-(diethylamino)carbamate group removal and nitrile conversion to amidoxime (synthesis of compounds **6**, **8**, and **9**). The solution of **21**, **28** or **29** (1 equiv), $NH_2OH.HCl$ (30 equiv), and pyridine (30 equiv) in 2 mL of ethanol was refluxed during 14 h. The solvent was removed in *vacuo*, the residue was subjected to preparative HPLC chromatography.

General procedure 3 for the one-pot alkyne reduction and O-debenzylation (synthesis of compounds **24**, **38**, **39**, and **40**). To the degazed solution of alkyne (1 equiv) in ethyl acetate or methanol Pearlman's catalyst $Pd(OH)_2/C$ (20%, moisture 50%, 0.3 equiv) was added. After the degazation the solution was bubbled with H_2. The reaction mixture was stirred at RT under H_2 (1 atm) during 24 h.

Table 4
Inhibition of hAChE.

Compound	IC_{50} (μM)	Compound	IC_{50} (μM)
Tacrine	0.20 ± 0.02 (Human erythrocyte AChE)[a]	**12**	1.40 ± 0.03 (hAChE)
 **51**	0.15 ± 0.01 (hAChE)	**13**	0.163 (hAChE)
7	0.26 ± 0.01 (hAChE)	 **52**	130 ± 15 (hAChE)
10	0.077 ± 0.003 (hAChE)	**14**	307 (hAChE)
11	0.26 ± 0.01 (hAChE)		

[a] From ref. [36].

The solution was filtered through celite, the solvent was evaporated, and the product was dried in *vacuo*.

General procedure 4 for the three-step transformation of the methyl ester into aldehyde (synthesis of compounds **31**, **41**, **42**, **43** and **50**). The solution of methyl ester **30**, **38**, **39**, **40** or **49** (1 equiv), imidazole (4.15 equiv), and TBDMSCl (2.4 equiv) in dry DMF was stirred at the room temperature under argon during 2 h. Ethyl acetate was added, and the organic phase was washed with water 3 times, dried over Na_2SO_4 and concentrated in *vacuum*. After drying in *vacuum* the compound obtained was subjected without purification to the following reduction of methyl ester with DIBAL-H. To the solution of the compound obtained (1 equiv) in dry CH_2Cl_2 at −78 °C DIBAL-H (1 M solution in CH_2Cl_2, 2 equiv) was added dropwise. The reaction mixture was stirred at −78 °C for 12 min, then it was quenched with MeOH, and the cooling bath was removed. When the mixture was warmed to the room temperature, the organic phase was washed with aqueous solution of NaOH (1 M), dried over Na_2SO_4 and concentrated under reduced pressure. After drying in *vacuum* the product obtained was subjected to the following reaction with TBAF. To the solution of the compound obtained (1 equiv) in dry THF at 0 °C TBAF (1 M solution in THF, 1.1 equiv) was added, and the reaction mixture was stirred at this temperature for 1 h. After removal of the solvent under reduced pressure the residue was purified by the column chromatography (silica gel).

General procedure 5 for synthesis of oximes **10**–**14.** The solution of aldehyde **31**, **41**, **42, 43** or **50** (1 equiv), $NH_2OH.HCl$ (1.5 equiv), and CH_3COONa (1.5 equiv) in dry ethanol was stirred during 8 h. After concentration under reduced pressure the crude product was purified by the column chromatography or preparative HPLC.

*7.2.1. Methyl 3-(benzyloxy)-6-{5-[(1,2,3,4-tetrahydroacridin-9-yl) amino]pent-1-yn-1-yl}pyridine-2-carboxylate (**35**)*

Compound **35** was obtained accordingly to the ***general procedure 1*** for Sonogashira coupling using **32** (0.147 g, 0.46 mmol, 1 equiv), **19** (0.132 g, 0.5 mmol, 1.1 equiv), $Pd(PPh_3)_4$ (0.041 g, 0.046 mmol, 0.1 equiv), CuI (0.017 g, 0.09 mmol, 0.2 equiv), Et_3N (3 mL), and THF (6 mL). The crude product was purified by the column chromatography (silica gel, gradient from ethyl acetate to ethyl acetate/MeOH = 8/2, $R_f = 0.24$). Yield 180 mg, 78%. ^{1}H NMR (300 MHz, $CDCl_3$, δ): 1.76–1.84 (4H, m), 1.87–1.97 (2H, qui, $J = 6.7$ Hz), 2.49 (2H, *t*, $J = 6.7$ Hz), 2.64 (2H, m), 3.03 (2H, m), 3.67 (2H, *t*, $J = 6.9$ Hz), 3.88 (3H, s, CH_3O), 5.14 (2H, s, $PhCH_2O$), 7.18–7.38 (8H, m, ArH), 7.48 (1H, *t*, $J = 7.6$ Hz), 7.90–7.93 (2H, m). ^{13}C NMR (100 MHz, $CDCl_3$, δ): 17.29; 22.79; 23.13; 25.05; 29.98; 33.79; 48.42; 52.88; 71.10; 80.64; 88.84; 120.24; 121.96; 122.93; 124.19; 127.12; 128.37; 128.47; 128.63; 128.75; 128.83; 128.95; 130.11; 132.24; 132.34; 135.23; 135.66; 146.91; 150.97; 153.20; 158.23; 164.99.

*7.2.2. Methyl 3-hydroxy-6-{5-[(1,2,3,4-tetrahydroacridin-9-yl) amino]pentyl}pyridine-2-carboxylate (**38**)*

Compound **38** was obtained accordingly to the ***general procedure 3*** for the reduction of the triple bond and O-debenzylation using **35** (0.180 g, 0.35 mmol), $Pd(OH)_2/C$ (20%, 0.075 g, 0.11 mmol, 0.3 equiv), and 2.5 mL of MeOH. Yield 0.140 g, 94%. $R_f = 0.34$ (CH_2Cl_2/MeOH = 10/1). ^{1}H NMR (300 MHz, $CDCl_3$, δ): 1.40–1.50 (2H, m), 1.67–1.78 (4H, m), 1.87–1.92 (4H, m), 2.63 (2H, m), 2.78 (2H, *t*, $J = 7.9$ Hz), 3.11 (2H, m), 3.56 (2H, *t*, $J = 7.2$ Hz), 4.00 (3H, s, CH_3O), 7.23 (1H, d, $J = 8.6$ Hz), 7.28 (1H, d, $J = 8.6$ Hz), 7.31–7.36 (1H, m), 7.54–7.59 (1H, m), 7.95 (1H, d, $J = 8$ Hz), 8.05 (1H, d, $J = 8.6$ Hz). ^{13}C NMR (100 MHz, $CDCl_3$, δ): 20.87; 22.11; 23.78; 26.38; 28.60; 29.62; 31.09; 37.02; 48.61; 53.55; 110.88; 115.99; 121.66; 124.19; 125.49; 127.88; 129.65; 132.63; 139.27; 152.03; 153.45; 155.55. HRMS (ESI, *m/z*): calcd. for $C_{25}H_{30}N_3O_3$ $[M + H]^+$, 420.22871; found, 420.22836.

*7.2.3. 3-Hydroxy-6-{5-[(1,2,3,4-tetrahydroacridin-9-yl)amino] pentyl}pyridine-2-carbaldehyde (**41**)*

Compound **41** was obtained by the ***general procedure 4*** from **38** with yield 37% (over three steps). $R_f = 0.23$ (ethyl acetate/MeOH = 10/3). ^{1}H NMR (400 MHz, $CDCl_3$, δ): 1.40–1.48 (2H, m), 1.67–1.79 (4H, m), 1.85–1.90 (4H, m), 2.64 (2H, m), 2.76 (2H, *t*, $J = 7.7$ Hz), 3.05 (2H, m), 3.52 (2H, *t*, $J = 7.3$ Hz), 7.23 (1H, d, $J = 8.7$ Hz), 7.25 (1H, d, $J = 8.7$ Hz), 7.29–7.33 (1H, m, ArH), 7.51–7.55 (1H, m, ArH), 7.92–7.95 (2H, m), 9.99 (1H, s, CHO). ^{13}C NMR (100 MHz, $CDCl_3$, δ): 22.73; 23.11; 24.88; 26.70; 29.48; 31.76; 33.48; 37.28; 49.45; 115.48; 123.10; 124.02; 126.64; 127.99; 129.01; 129.88; 131.06; 135.95; 151.45; 154.66; 157.23; 157.88; 198.83. HRMS (ESI, *m/z*): calcd. for $[M + H]^+$, $C_{24}H_{28}N_3O_2$, 390.21815; found, 390.21892.

*7.2.4. 2-[(1E)-(hydroxyimino)methyl]-6-{5-[(1,2,3,4-tetrahydroacridin-9-yl)amino]pentyl}pyridin-3-ol (**11**)*

Compound **11** was obtained accordingly to the ***general procedure 5*** using **41** (0.050 g, 0.13 mmol, 1 equiv), $NH_2OH.HCl$ (0.013 g, 0.19 mmol, 1.5 equiv), CH_3COONa (0.016 g, 0.19 mmol, 1.5 equiv), and 1.2 mL of ethanol. After evaporation of the solvent and purification by column chromatography (ethyl acetate/MeOH = 8/2) product was obtained with yield 55%. $R_f = 0.44$ (ethyl acetate/MeOH = 2/1). ^{1}H NMR (400 MHz, CD_3OD, δ): 1.37–1.44 (2H, m, CH_2), 1.68–1.76 (2H, qui, $J = 7.4$ Hz, CH_2), 1.77–1.85 (2H, qui, $J = 7.4$ Hz, CH_2), 1.90–1.94 (4H, qui, $J = 3.3$ Hz, CH_2), 2.63 (2H, m, CH_2), 2.70 (2H, *t*, $J = 7.4$ Hz, CH_2), 3.00 (2H, m, CH_2), 3.87 (2H, *t*, $J = 6.9$ Hz, CH_2), 7.08 (1H, d, $J = 8.4$ Hz, ArH), 7.20 (1H, d, $J = 8.4$ Hz, ArH), 7.50–7.54 (1H, m, ArH), 7.76–7.81 (2H, m, ArH), 8.17 (1H, s, N=CH), 8.30 (1H, d, $J = 8.7$ Hz). ^{13}C NMR (100 MHz, CD_3OD, δ): 22.31; 23.30; 25.25; 27.05; 30.30; 30.53; 30.91; 31.42; 37.56; 117.95; 121.70; 125.51; 126.08; 126.18; 133.42; 136.35; 152.90; 153.85; 154.49. MS *m/z* (ESI) calcd. for $C_{24}H_{29}N_4O_2$ $[M + H]^+$, 405.22; found, 405.2.

*7.2.5. Methyl 3-(benzyloxy)-6-{4-[(1,2,3,4-tetrahydroacridin-9-yl) amino]but-1-yn-1-yl}pyridine-2-carboxylate (**36**)*

Compound **36** was obtained accordingly to the ***general procedure 1*** for Sonogashira coupling reaction using **32** (0.292 g, 0.91 mmol, 1 equiv), **34** (0.250 g, 1 mmol, 1.1 equiv), $Pd(PPh_3)_4$ (0.081 g, 0.091 mmol, 0.1 equiv), CuI (0.035 g, 0.18 mmol, 0.2 equiv), Et_3N (6 mL), and THF (12 mL). The crude product was purified by the column chromatography (silica gel, gradient from ethyl acetate/MeOH = 10/1 to ethyl acetate/MeOH = 8/2), $R_f = 0.43$ (ethyl acetate/MeOH = 4/1). Yield 0.440 g, 98%. ^{1}H NMR (400 MHz, $CDCl_3$, δ): 1.82–1.89 (4H, m), 2.67 (2H, *t*, $J = 6.3$ Hz), 2.79 (2H, *t*, $J = 5.9$ Hz), 3.06 (2H, *t*, $J = 5.9$ Hz), 3.70–3.75 (2H, dt, $J = 6.4$ Hz), 3.95 (3H, s), 4.48 (1H, *t*, $J = 6.4$ Hz, NH), 5.19 (2H, s, OCH_2Ph), 7.25 (1H, d, $J = 8.8$ Hz), 7.28 (2H, d, $J = 8.8$ Hz), 7.31–7.42 (6H, m, ArH), 7.52–7.56 (1H, m), 7.97 (2H, *t*, ArH, $J = 8.3$ Hz). ^{13}C NMR (100 MHz, $CDCl_3$, δ): 22.10; 22.59; 22.99; 24.93; 29.89; 33.29; 47.34; 52.93; 71.10; 77.43; 81.98; 86.71; 116.94; 120.15; 121.92; 123.00; 124.56; 127.14; 127.66; 128.51; 128.97; 129.30; 130.09; 134.74; 135.59; 140.67; 146.04; 151.08; 153.37; 157.78; 164.98. HRMS (ESI, *m/z*): calcd. for $C_{31}H_{30}N_3O_3$ $[M + H]^+$, 492.22871; found, 492.22851.

*7.2.6. Methyl 3-hydroxy-6-{4-[(1,2,3,4-tetrahydroacridin-9-yl) amino]butyl}pyridine-2-carboxylate (**39**)*

Compound **39** was obtained accordingly to the ***general procedure 3*** using **36** (0.43 g, 0.87 mmol), $Pd(OH)_2/C$ (20%, 0.184 g, 0.26 mmol, 0.3 equiv), and 3 mL of MeOH. Yield 0.34 g, 96%. $R_f = 0.43$ (CH_2Cl_2/MeOH = 10/1). ^{1}H NMR (300 MHz, CD_3OD, δ): 1.67–1.74 (4H, m), 1.87–1.91 (4H, m), 2.68 (2H, *t*, $J = 5.3$ Hz), 2.72 (2H, *t*, $J = 7$ Hz), 2.96 (2H, *t*, $J = 5.5$ Hz), 3.62 (2H, *t*, $J = 6.7$ Hz), 3.96 (3H, s, CH_3O), 7.23 (1H, d, $J = 8.6$ Hz), 7.27 (2H, d, $J = 8.6$ Hz), 7.34–

7.40 (1H, m), 7.57–7.62 (1H, m), 7.74 (1H, d, $J = 8.4$ Hz), 8.10 (1H, d, $J = 8.4$ Hz). ^{13}C NMR (100 MHz, CD_3OD, δ): 23.25; 23.87; 25.96; 28.22; 31.31; 32.84; 37.27; 53.34; 115.73; 120.15; 125.07; 125.29; 125.74; 128.52; 130.46; 130.65; 131.09; 145.48; 154.49; 154.54; 156.94; 158.65; 170.62. HRMS (ESI, *m/z*): calcd. for $C_{24}H_{28}N_3O_3$ $[M + H]^+$, 406.21306; found, 406.21307.

7.2.7. 3-Hydroxy-6-{4-[(1,2,3,4-tetrahydroacridin-9-yl)amino] butyl}pyridine-2-carbaldehyde (**42**)

Compound **42** was obtained from **39** accordingly to the ***general procedure 4*** after purification by the column chromatography (silica gel, ethyl acetate/MeOH = 8/2) with yield 54% (over 3 steps). $R_f = 0.21$ (ethyl acetate/MeOH = 7/3). ^{1}H NMR (400 MHz, $CDCl_3$, δ): 1.66–1.75 (2H, m), 1.78–1.89 (6H, m), 2.67 (2H, m), 2.79 (2H, *t*, $J = 7$ Hz), 3.04 (2H, m), 3.52 (2H, *t*, $J = 7$ Hz), 7.20 (1H, d, $J = 8.6$ Hz), 7.24 (1H, d, $J = 8.6$ Hz), 7.31 (1H, td, $J = 6.8$ Hz, $J = 1$ Hz), 7.53 (1H, td, $J = 6.8$ Hz, $J = 1$ Hz), 7.89–7.93 (2H, m), 9.98 (1H, s, CH=O). ^{13}C NMR (100 MHz, $CDCl_3$, δ): 22.89; 23.20; 25.01; 26.96; 31.37; 33.92; 36.98; 49.32; 116.06; 120.29; 122.94; 123.95; 126.69; 128.68; 129.87; 136.00; 147.27; 150.99; 154.29; 157.29; 158.43; 198.74. HRMS (ESI, *m/z*): calcd. for $C_{23}H_{26}N_3O_2$ $[M + H]^+$, 376.20250; found, 376.20201.

7.2.8. 2-[1-(Hydroxyimino)methyl]-6-{4-[(1,2,3,4-tetrahydroacridin-9-yl)amino]butyl}pyridin-3-ol (**12**)

Compound **12** was obtained accordingly to the ***general procedure 5*** using **42** (0.090 g, 0. 24 mmol, 1 equiv), $NH_2OH.HCl$ (0.025 g, 0.36 mmol, 1.5 equiv), CH_3COONa (0.029 g, 0.36 mmol, 1.5 equiv), and 2.5 mL of ethanol. Compound **12** was purified by the column chromatography (silica gel, CH_2Cl_2/MeOH = 8/2) with yield 55%. $R_f = 0.6$ (ethyl acetate/MeOH = 2/1). ^{1}H NMR (400 MHz, CD_3OD, δ): 1.80–1.83 (4H, m, $CH_2CH_2CH_2CH_2$), 1.90–1.95 (4H, m, $CH_2CH_2CH_2CH_2$), 2.62 (2H, *t*, $J = 5.5$ Hz, CH_2), 2.71 (2H, *t*, $J = 6.4$ Hz, CH_2), 2.98 (2H, m, CH_2), 3.87 (2H, *t*, $J = 5.6$ Hz, CH_2), 7.04 (1H, d, $J = 8.4$ Hz, ArH), 7.09 (1H, d, $J = 8.4$ Hz); 7.47–7.51 (1H, m, ArH), 7.71–7.79 (2H, m, ArH), 8.08 (1H, s, CH=N), 8.24 (1H, d, $J = 8.8$ Hz). ^{13}C NMR (100 MHz, CD_3OD, δ): 22.04; 23.09; 25.03; 27.55; 29.84; 30.60; 37.11; 113.35; 117.54; 121.08; 125.43; 125.83; 126.12; 126.19; 133.51; 136.24; 140.59; 152.44; 152.75; 153.65; 154.12; 157.25. HRMS (ESI, *m/z*): calcd. for $C_{23}H_{27}N_4O_2$ $[M + H]^+$ 391.21340; found, 391.21337.

7.3. Biological essays

7.3.1. IC_{50} measurements

Recombinant hAChE was produced and purified as previously described (see reference: http://www.ncbi.nlm.nih.gov/pubmed/18975951). Oximes were dissolved in MeOH to make 5- or 10-mM stock solution. Recombinant hAChE activity was measured spectrophotometrically (absorbance at 412 nm) in the presence of various concentrations of oximes in 1 mL Ellman's buffer (phosphate 0.1 M, pH 7.4, 0.1% BSA, 5% MeOH, 0.5 mM DTNB, 25 °C). Measurements were performed at least in duplicate for each concentration tested. The concentration of oxime producing 50% of enzyme inhibition was determined by non-linear fitting using ProFit (Quantumsoft) using the standard IC_{50} equation: % Activity = $100^*IC_{50}/(IC_{50} + [Ox])$.

7.3.2. Inhibition of hAChE by OPs

VX, tabun were from DGA maîtrise NRBC (Vert le Petit, France). Paraoxon-ethyl was purchased from Sigma–Aldrich. Stock solution of VX and tabun were 5 mM in isopropanol. The inhibition of 120 μM hAChE is realized with a 5-fold excess of OPs and was performed in tris buffer (20 mM, pH 7.4, 0.1% BSA) at 25 °C. After a 20-min incubation, inhibited hAChE was desalted on PD-10 column (GE Healthcare).

7.3.3. Reactivation of hAChE inhibited by OPs

Hlö-7 was from DGA maîtrise NRBC (Vert le Petit, France), HI-6 from Pharmacie Centrale des Armées (Orléans, France) and obidoxime was from CRSSA (Lyon, France; synthesized by Bernard Desiré). OP-inhibited hAChE was incubated at 37 °C with at least 4 concentrations of oxime in phosphate buffer (0.1 M, pH 7.4, 0.1% BSA). The final concentration of MeOH in the incubation mix was below 2% and had no influence on the enzyme stability. At time intervals ranging from 1 to 10 min depending on the reactivation rate, 10-μL aliquots of each solutions containing the different concentrations of oxime were transferred to cuvettes containing 1 mM acetylthiocholine in 1 mL Ellman's buffer (phosphate 0.1 M, pH 7.4, 0.1% BSA, 0.5 mM DTNB, 25 °C) for measurement of hAChE activity. The enzyme activity in the control (uninhibited enzyme + oxime) remained constant during the experiment. Oximolysis was undetectable.

The percentage of reactivated enzyme ($\%E_{react}$) was calculated as the ratio of the recovered enzyme activity and activity in the control. The apparent reactivation rate k_{obs} for each oxime concentration, the dissociation constant K_D of inhibited enzyme–oxime complex (E–PO_x) and the maximal reactivation rate constant k_r, were calculated by non-linear fit with ProFit (Quantumsoft) using the standard oxime concentration-dependent reactivation equation derived from the following scheme:

$$E - P + \mathrm{Ox} \overset{K_D}{\rightleftarrows} E - \mathrm{POx} \xrightarrow{k_r} E + \mathrm{POx}$$

$$\%E_{\mathrm{react}} = 100 \cdot \left(1 - e^{k_{\mathrm{obs}} \cdot t}\right) \quad \text{and} \quad k_{\mathrm{obs}} = \frac{k_r[\mathrm{Ox}]}{K_D + [\mathrm{Ox}]}$$

Detailed reactivator concentrations used for each set of experiments are given in the Supplementary Information.

7.4. Molecular modelling

Flexible dockings have been performed using autodock vina [37] and by preparing the system in PyMOL (Schrödinger) using the plug-in developed by Daniel Seeliger (http://wwwuser.gwdg.de/~dseelig/adplugin.html). VX-hAChE was constructed from the apo form (pdb code 4EY4) by homology to the mAChE-VX structure (pdb code 2Y2U), keeping in the active site all the usually conserved water molecules. Residues in the gorge (Tyr72, Asp74, Trp86, Tyr124, Ser125, Trp286, Tyr337, Phe338, Tyr341) have been chosen as flexible, along with the ethyl group of VX. A docking box of 60 × 60 × 60 Å was chosen, centered at the bottom of the gorge between Tyr124 and Trp86. Ligands were built and optimized from SMILEs string using Phenix elbow [38]. The default parameter set of Autodock vina was used to generate 9 docking poses per molecule.

Molecular dynamics simulations were carried out using GROMACS 4.5.6 [39] and the Amber99sb forcefield [40]. The topological description of each KM molecule was built using acpype and the general amber forcefield [41]. The hAChE-VX complex in the conformation obtained from flexible dockings together with crystal water molecules and the different KM molecules, was immersed in a periodic water box of cubic shape with a minimal distance of 10 Å to any edge and periodic boundary conditions. The box was solvated using the TIP3P solvation model and chloride and sodium counter ions were added to neutralize the simulation system. After energy minimization using a 500-step steepest decent method, the system was subjected to equilibration at 1 bar for 50 ps under the conditions of position restraints for heavy atoms. The solvent, the counter ions, and the protein were coupled separately to a temperature bath at 300 K. The Lennard–Jones interactions were cutoff at 1.4 nm. The long-range electrostatic interactions

were handled using particle-mesh Ewald method for determining long-range electrostatics (9 Å cutoff). Temperature was set to 300 K and was kept constant using a Berendsen thermostat [42] (with a coupling time constant of 0.1 ps). Pressure with a reference value of 1 bar was controlled by a Berendsen barostat (with a coupling time constant of 1 ps and a compressibility of 4.5 10^5 bar). Full MD simulation was performed for 5 ns at 300 K, using 2 fs timesteps. All bond lengths were constrained using the LINCS algorithm allowing an integration step of 2 fs [43]. Coordinates were saved every 250 steps (every 0.5 ps).

Acknowledgements

The ANR (Detoxneuro ANR-06-BLAN-163, ReAChE ANR-09-BLAN-192), DGA (DGA/STTC PDH-2-NRBC-4-C-403, DGA-REI 2009-34-0023) and DTRA (HDTRA1-11-C-0047) are gratefully acknowledged for financial support. This work has been partially supported by INSA Rouen, Rouen University, CNRS, Labex SynOrg (ANR-11-LABX-0029) and région Haute-Normandie (CRUNCh network). We also thank Dr. Patrick Wehrung, Pascale Buisine and Cyril Antheaume (UdS, Faculty of Pharmacy, Service Commun d'Analyse) for mass and NMR analyses.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.ejmech.2014.03.044.

References

[1] (a) P. Taylor, Goodman, Gilman's, in: J.G. Hardman, L.E. Limbird, A.G. Goodman (Eds.), The Pharmacological Basis of Therapeutics, tenth ed., McGraw-Hill, New York, 2001, p. 175;
(b) R.E. Langford, Introduction to Weapons of Mass Destruction, Radiological, Chemical and Biological, Wiley, New York, 2004.

[2] (a) T.C. Marrs, Organophosphate poisoning, Phamacology Therapy 58 (1993) 51–66;
(b) F.R. Sidell, J. Borak, Chemical warfare agents: II. Nerve agents, Annals of Emergency Medicine 21 (1992) 865–871.

[3] Convention on the Prohibition of the Development, Production, Stockpiling and Use of Chemical Weapons and on Their Destruction; the Technical Secretariat of the Organization for the Prohibition of Chemical Weapons, Den Hague, Netherlands, 1993.

[4] M. Eddleston, N.A. Buckley, P. Eyer, A.H. Dawson, Management of acute organophosphorus pesticide poisoning, Lancet 371 (2008) 597–607.

[5] M. Eddleston, L. Karalliedde, N. Buckley, R. Fernando, G. Hutchinson, G. Isbister, F. Konradsen, D. Murray, J.C. Piola, N. Senanayake, R. Sheriff, S. Singh, S.B. Siwach, L. Smit, Pesticide poisoning in the developing world – a minimum pesticides list, Lancet 360 (2002) 1163–1167.

[6] G. Mercey, T. Verdelet, J. Renou, M. Kliachyna, R. Baati, F. Nachon, L. Jean, P.-Y. Renard, Reactivators of acetylcholinesterase inhibited by organophosphorus nerve agents, Accounts of Chemical Research 45 (2012) 756–766.

[7] D.E. Lorke, H. Kalasz, G.A. Petroianu, K. Tekes, Entry of oximes into the brain: a review, Current Medicinal Chemistry 15 (2008) 743–753.

[8] K. Sakurada, K. Matsubara, K. Shimizu, H. Shiono, Y. Seto, K. Tsuge, M. Yoshino, I. Sakai, H. Mukoyama, T. Takatori, Pralidoxime iodide (2-PAM) penetrates across the blood-brain barrier, Neurochemical Research 28 (2003) 1401–1407.

[9] T.M. Shih, J.W. Skovira, J.C. O'Donnell, J.H. McDonough, Central acetylcholinesterase reactivation by oximes improves survival and terminates seizures following nerve agent intoxication, Advanced Studies in Biology 1 (2009) 155–196.

[10] M.C. de Koning, M. van Grol, D. Noort, Peripheral site ligand conjugation to a non-quaternary oxime enhances reactivation of nerve agent-inhibited human acetylcholinesterase, Toxicology Letters 206 (2011) 54–59.

[11] (a) Z. Radić, R.K. Sit, Z. Kovarik, S. Berend, E. Garcia, L. Zhang, G. Amitai, C. Green, B. Radić, V.V. Fokin, K.B. Sharpless, P. Taylor, Refinement of structural leads for centrally acting oxime reactivators of phosphylated cholinesterases, Journal of Biological Chemistry 287 (2012) 11798–11809;
(b) R.K. Sit, Z. Radić, V. Gerardi, L. Zhang, E. Garcia, M. Katalinić, G. Amitai, Z. Kovarik, V.V. Fokin, K.B. Sharpless, P. Taylor, New structural scaffolds for centrally acting oxime reactivators of phosphylated cholinesterases, Journal of Biological Chemistry 286 (2011) 19422–19430.

[12] (a) J. Kalisiak, E.C. Ralph, J. Zhang, J.R. Cashman, Amidine-oximes: reactivators for organophosphate exposure, Journal of Medicinal Chemistry 54 (2011) 3319–3330;
(b) J. Kalisiak, E.C. Ralph, J.R. Cashman, Nonquaternary reactivators for organophosphate-inhibited cholinesterases, Journal of Medicinal Chemistry 55 (2012) 465–474.

[13] W.P. Jencks, J. Carriuolo, Reactivity of nucleophilic reagents toward esters, Journal of the American Chemical Society 82 (1960) 1778–1786.

[14] G. Saint-André, M. Kliachyna, S. Kodepelly, L. Louise-Leriche, E. Gillon, P.-Y. Renard, F. Nachon, R. Baati, A. Wagner, Design, synthesis, and evaluation of new alpha-nucleophiles for the hydrolysis of organo-phosphorus nerve agents: application to the reactivation of phosphorylated AChE, Tetrahedron 67 (2011) 6352–6361.

[15] L. Louise-Leriche, E. Păunescu, G. Saint-André, R. Baati, A. Romieu, A. Wagner, P.-Y. Renard, A HTS assay for the detection of organophosphorus nerve agent scavengers, Chemistry European Journal 16 (2010) 3510–3523.

[16] A. Krasiński, Z. Radić, R. Manetsch, J. Raushel, P. Taylor, K.B. Sharpless, H.C. Kolb, In situ selection of lead compounds by click chemistry: target-guided optimization of acetylcholinesterase inhibitors, Journal of the American Chemical Society 127 (2005) 6686–6692.

[17] (a) G. Mercey, T. Verdelet, G. Saint-André, E. Gillon, A. Wagner, R. Baati, L. Jean, F. Nachon, P.-Y. Renard, First efficient uncharged reactivators for the dephosphylation of poisoned human acetylcholinesterase, Chemical Communications 47 (2011) 5295–5297;
(b) G. Mercey, J. Renou, T. Verdelet, M. Kliachyna, R. Baati, E. Gillon, M. Arboleas, M. Loiodice, F. Nachon, L. Jean, P.-Y. Renard, Phenyltetrahydroisoquinoline-pyridinaldoxime conjugates as efficient uncharged reactivators for the dephosphylation of inhibited human acetylcholinesterase, Journal of Medicinal Chemistry 55 (2012) 10791–10795;
(c) J. Renou, M. Loiodice, M. Arboléas, R. Baati, L. Jean, F. Nachon, P.-Y. Renard, Tryptoline-3-hydroxypyridinaldoxime conjugates as efficient reactivators of phosphylated human acetyl and butyrylcholinesterases, Chemical Communications 50 (2014) 3947–3950.

[18] Y. Ashani, A.K. Bhattacharjee, H. Leader, A. Saxena, B.P. Doctor, Inhibition of cholinesterases with cationic phosphonyl oximes highlights distinctive properties of the charged pyridine groups of quaternary oxime reactivators, Biochemical Pharmacology 66 (2003) 191–202.

[19] J.P. Colletier, B. Sanson, F. Nachon, E. Gabellieri, C. Fattorusso, G. Campiani, M. Weik, Conformational flexibility in the peripheral site of *Torpedo californica* acetylcholinesterase revealed by the complex structure with a bifunctional inhibitor, Journal of the American Chemical Society 128 (2006) 4526–4527.

[20] E.H. Rydberg, B. Brumshtein, H.M. Greenblatt, D.M. Wong, D. Shaya, L.D. Williams, P.R. Carlier, Y.P. Pang, I. Silman, J.L. Sussman, Complexes of alkylene-linked tacrine dimers with *Torpedo californica* acetylcholinesterase: binding of bis5-tacrine produces a dramatic rearrangement in the active-site gorge, Journal of Medicinal Chemistry 49 (2006) 5491–5500.

[21] J.L. Sussman, M. Harel, F. Frolow, C. Oefner, A. Goldman, L. Toker, I. Silman, Atomic structure of acetylcholinesterase from *Torpedo californica*: a prototypic acetylcholine-binding protein, Science 253 (1991) 872–879.

[22] J.P. Colletier, D. Fournier, H.M. Greenblatt, J. Stojan, J.L. Sussman, G. Zaccai, I. Silman, M. Weik, Structural insights into substrate traffic and inhibition in acetylcholinesterase, Embo Journal 25 (2006) 2746–2756.

[23] Y. Bourne, P. Taylor, Z. Radic, P. Marchot, Structural insights into ligand interactions at the acetylcholinesterase peripheral anionic site, Embo Journal 22 (2003) 1–12.

[24] Y. Bourne, H.C. Kolb, Z. Radic, K.B. Sharpless, P. Taylor, P. Marchot, Freeze-frame inhibitor captures acetylcholinesterase in a unique conformation, Proceedings of the National Academy of Sciences of the United States of America 101 (2004) 1449–1454.

[25] B. Sanson, F. Nachon, J.P. Colletier, M.T. Froment, L. Toker, H.M. Greenblatt, J.L. Sussman, Y. Ashani, P. Masson, I. Silman, M. Weik, Crystallographic snapshots of nonaged and aged conjugates of soman with acetylcholinesterase, and of a ternary complex of the aged conjugate with pralidoxime, Journal of Medicinal Chemistry 52 (2009) 7593–7603.

[26] W.G. Lewis, L.G. Green, F. Grynszpan, Z. Radic, P.R. Carlier, P. Taylor, M.G. Finn, K.B. Sharpless, Click chemistry in situ: acetylcholinesterase as a reaction vessel for the selective assembly of a femtomolar inhibitor from an array of building blocks, Angewantde Chemie International Edition 41 (2002) 1053–1057.

[27] (a) T. Kawaguchi, H. Akatsuka, T. Iijima, T. Watanabe, J. Murakami, T. Mitsui, 2006, JP 2006 04 5220A.
(b) T. Kawaguchi, H. Akatsuka, T. Iijima, T. Watanabe, J. Murakami, T. Mitsui, 2004, WO 2004063202 A1.

[28] S. Butini, G. Campiani, M. Borriello, S. Gemma, A. Panico, M. Persico, B. Catalanotti, S. Ros, M. Brindisi, M. Agnusdei, I. Fiorini, V. Nacci, E. Novellino, T. Belinskaya, A. Saxena, C. Fattorusso, Exploiting protein fluctuations at the active-site gorge of human cholinesterases: further optimization of the design strategy to develop extremely potent inhibitors, Journal of Medicinal Chemistry 51 (2008) 3154–3170.

[29] C. Ronco, L. Jean, H. Outaabout, P.-Y. Renard, Palladium-catalyzed preparation of *N*-Alkylated tacrine and huprine compounds, European Journal of Organic Chemistry (2011) 302–310.

[30] T. Verdelet, G. Mercey, N. Correa, L. Jean, P.-Y. Renard, Straightforward and efficient synthesis of 3-benzyloxy-4-bromopicolinate ester and 3-benzyloxy-5-bromopicolinate ester, common building blocks for pharmaceuticals and agrochemicals, Tetrahedron 67 (2011) 8757–8762.

[31] R. Manetsch, A. Krasiński, Z. Radić, J. Raushel, P. Taylor, K.B. Sharpless, H.S. Kolb, In situ click chemistry: enzyme inhibitors made to their own specifications, Journal of the American Chemical Society 126 (2004) 12809–12818.
[32] G. Campiani, C. Fattorusso, S. Butini, A. Gaeta, M. Agnusdei, S. Gemma, M. Persico, B. Catalanotti, L. Savini, V. Nacci, E. Novellino, H.W. Holloway, N.H. Greig, T. Belinskaya, J.M. Fedorko, A. Saxena, Development of molecular probes for the identification of extra interaction sites in the mid-gorge and peripheral sites of butyrylcholinesterase (BuChE). Rational design of novel, selective, and highly potent BuChE inhibitors, Journal of Medicinal Chemistry 48 (2005) 1919–1929.
[33] F. Worek, H. Thiermann, L. Szinicz, P. Eyer, Kinetic analysis of interactions between human acetylcholinesterase, structurally different organophosphorus compounds and oximes, Biochemical Pharmacology 68 (2004) 2237–2248.
[34] F. Worek, T. Wille, M. Koller, H. Thiermann, Reactivation kinetics of a series of related bispyridinium oximes with organophosphate-inhibited human acetylcholinesterase-Structure-activity relationships, Biochemical Pharmacology 83 (2012) 1700–1706.
[35] Y. Wang, Y. Wei, S. Oguntayo, B.P. Doctor, M.P. Nambiar, A combination of [+] and [−]-Huperzine A improves protection against soman toxicity compared to [+]-Huperzine A in guinea pigs, Chemico-Biological Interactions 203 (2013) 120–124.
[36] (a) P.R. Carlier, E.S.-H. Chow, Y. Han, J. Liu, J.E. Yazal, Y.-P. Pang, Heterodimeric tacrine-based acetylcholinesterase inhibitors: investigating ligand-peripheral site interactions, Journal of Medicinal Chemistry 42 (1999) 4225–4231;
(b) P. Camps, R. El Achab, J. Morral, D. Muñoz-Torrero, A. Badia, J.E. Baños, N.M. Vivas, X. Barril, M. Orozco, F.J. Luque, New tacrine-huperzine A hybrids (huprines): highly potent tight-binding acetylcholinesterase inhibitors of interest for the treatment of Alzheimer's disease, Journal of Medicinal Chemistry 43 (2000) 4657–4666.
[37] E. Carletti, H. Li, B. Li, F. Ekström, Y. Nicolet, M. Loiodice, E. Gillon, M.-T. Froment, O. Lockridge, L.M. Schopfer, P. Masson, F. Nachon, Aging of cholinesterases phosphylated by tabun proceeds through O-dealkylation, Journal of the American Chemical Society 130 (2008) 16011–16020.
[38] P.D. Adams, P.V. Afonine, G. Bunkóczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.-W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, P.H. Zwart, PHENIX: a comprehensive Python-based system for macromolecular structure solution, Acta Crystallographica D66 (2010) 213–221.
[39] B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation, Journal of Chemical Theory and Computation 4 (2008) 435–447.
[40] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, Comparison of multiple amber force fields and development of improved protein Backbone parameters, Proteins 65 (2006) 712–725.
[41] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, Development and testing of a general amber force field, Journal of Computational Chemistry 25 (2004) 1157–1174.
[42] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, Journal of Chemical Physics 81 (1984) 3684–3690.
[43] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS: a linear constraint solver for molecular simulations, Journal of Computational Chemistry 18 (1997) 1463–1472.

Chemical Research in Toxicology

Article

pubs.acs.org/crt

Inhibition Pathways of the Potent Organophosphate CBDP with Cholinesterases Revealed by X-ray Crystallographic Snapshots and Mass Spectrometry

Eugénie Carletti,[†,‡,§,∥] Jacques-Philippe Colletier,[†,‡,§] Lawrence M. Schopfer,[⊥] Gianluca Santoni,[†,‡,§,∥] Patrick Masson,[†,‡,§,∥,⊥] Oksana Lockridge,[⊥] Florian Nachon,*[,∥] and Martin Weik*[,†,‡,§,#]

[†]Institut de Biologie Structurale J.P. Ebel, Commissariat à l'Energie Atomique, 41, rue Jules Horowitz, F-38027 Grenoble, France

[‡]CNRS, UMR5075, F-38027 Grenoble, France

[§]Université Joseph Fourier, F-38000 Grenoble, France

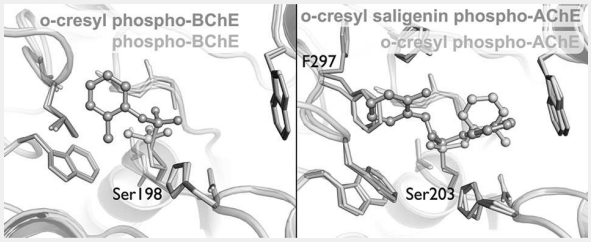
[∥]Département de Toxicologie, Institut de Recherche Biomédicale des Armées, 24 avenue des Marquis du Grésivaudan, 38702 La Tronche, France

[⊥]Eppley Institute, University of Nebraska Medical Center, Omaha, Nebraska 68198-5950, United States

[#]ESRF, 6 rue Jules Horowitz, BP 220, 38043 Grenoble Cedex, France

S *Supporting Information*

ABSTRACT: Tri-*o*-cresyl-phosphate (TOCP) is a common additive in jet engine lubricants and hydraulic fluids suspected to have a role in aerotoxic syndrome in humans. TOCP is metabolized to cresyl saligenin phosphate (CBDP), a potent irreversible inhibitor of butyrylcholinesterase (BChE), a natural bioscavenger present in the bloodstream, and acetylcholinesterase (AChE), the off-switch at cholinergic synapses. Mechanistic details of cholinesterase (ChE) inhibition have, however, remained elusive. Also, the inhibition of AChE by CBDP is unexpected, from a structural standpoint, i.e., considering the narrowness of AChE active site and the bulkiness of CBDP. In the following, we report on kinetic X-ray crystallography experiments that provided 2.7−3.3 Å snapshots of the reaction of CBDP with mouse AChE and human BChE. The series of crystallographic snapshots reveals that AChE and BChE react with the opposite enantiomers and that an induced-fit rearrangement of Phe297 enlarges the active site of AChE upon CBDP binding. Mass spectrometry analysis of aging in either $H_2{}^{16}O$ or $H_2{}^{18}O$ furthermore allowed us to identify the inhibition steps, in which water molecules are involved, thus providing insights into the mechanistic details of inhibition. X-ray crystallography and mass spectrometry show the formation of an aged end product formed in both AChE and BChE that cannot be reactivated by current oxime-based therapeutics. Our study thus shows that only prophylactic and symptomatic treatments are viable to counter the inhibition of AChE and BChE by CBDP.



■ INTRODUCTION

Tri-*o*-cresyl phosphate (TOCP), a toxic isomer of tricresyl phosphate (TCP), is suspected to play an important role in aerotoxic syndrome. TCP is added to commercial jet engine lubricants for its antiwear and flame-retardant properties.[1−3] When oil seals in jet-airplanes leak, the cabin bleed-air becomes contaminated with TOCP. TOCP is inhaled and metabolized by liver microsomal cytochrome P450 and serum albumin into 2-(*o*-cresyl)-4*H*-1,3,2-benzodioxaphosphoran-2-one (CBDP), also called cresyl saligenin phosphate.[4−6]

CBDP is a bicyclic organophosphorus compound (OP), which irreversibly inhibits both human acetylcholinesterase (hAChE; EC 3.1.1.7) and butyrylcholinesterase (hBChE; EC 3.1.1.8).[7] hAChE is the enzyme responsible for the termination of nerve-impulse transmission at cholinergic synapses, and as such, it is the primary target of chemically synthesized OPs. hBChE is a structurally and functionally homologous enzyme that is present in the bloodstream (50 nM, in humans) and acts as a stoichiometric bioscavenger of OPs. [8] Irreversible inhibition of hAChE and hBChE by OPs generally proceeds in two steps; phosphylation of the catalytic serine with concomitant release of a leaving group, followed by dealkylation of the OP-enzyme adduct, a process also called aging. It is noteworthy that medical countermeasures to reactivate cholinesterases are only efficient against nonaged enzymes:[9] aged ChEs can indeed no longer be reactivated by oximes used as antidotes against OP poisoning.

hAChE and hBChE differ mostly by the amino acid distribution in their active site gorges, the former containing more aromatic residues than the latter. In particular, the replacement of two phenylalanine residues in the acyl-binding

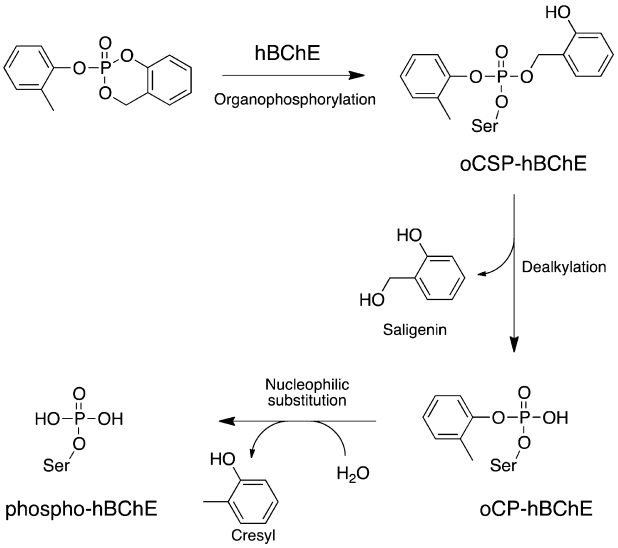
Received: November 9, 2012
Published: January 22, 2013

pocket of hAChE, by a leucine and a valine in BChE, results in an enlargement of the pocket, thereby allowing accommodation of bulkier substrates and inhibitors and expanding the substrate specificity of hBChE as compared to that of hAChE.[10] Accordingly, hAChE and hBChE have repeatedly been reported to preferentially react with different OP stereoisomers.[11]

The ensemble of neurological symptoms associated with aerotoxic syndrome includes memory impairment and uncontrollable shaking.[12,13] It has been shown that CBDP inhibits esterase enzymes, notably the neuropathy target esterase[14,15] but also hBChE and hAChE.[7,16−18] In addition, other serine esterases are likely inhibited by CBDP and tyrosine residues in several proteins have been shown to be reactive toward CBDP. Inhibition of such proteins may participate in the pathogenesis of aerotoxic syndrome.

We earlier investigated the inhibition mechanisms of hAChE and hBChE by CBDP using a combination of kinetic measurements, mass spectrometry, and X-ray crystallography.[7] As shown in Scheme 1, the inhibition of hBChE by CBDP occurs in three steps.

Scheme 1. Reaction of CBDP with hBChE



CBDP first forms a phosphorylserine conjugate of hBChE, with saligenin and *o*-cresyl as substituents (oCSP-hBChE). The first aging reaction corresponds to the release of the saligenin moiety, and results in the *o*-cresyl-phospho-BChE adduct (oCP-hBChE). A second aging reaction then occurs that allows the release of the *o*-cresyl moiety and yields phospho-hBChE, whose structure we reported already.[7] We also showed that the bimolecular rate constant of CBDP for hBChE is at least 10-fold higher than that for hAChE. CBDP is, in fact, one of the most potent BChE inhibitors reported to date (1.5×10^8 $M^{-1}\cdot min^{-1}$), suggesting that hBChE efficiently scavenges CBDP in the bloodstream. The reactivity of CBDP toward hAChE is in the same range as that exhibited by paraoxon, tabun, or soman (10^6−10^7 $M^{-1}\cdot min^{-1}$). It remains unclear, however, if hAChE and hBChE inhibition by CBDP proceeds in the same fashion, and whether the same intermediate states form.

Here, we provide a detailed description of mouse (m) AChE and human (h) BChE inhibition by CBDP, based on crystallographic snapshots of inhibition intermediate states and on mass spectrometry analyses. We show that mAChE inhibition by CBDP occurs in two steps and involves a transient enlargement of the acyl-binding pocket to allow accommodation of the bulky CBDP adduct. The resulting *o*-cresyl-phospho-mAChE is a stable OP-aged adduct that does not evolve further. In contrast, the *o*-cresyl-phospho-hBChE adduct represents an unstable, strained conformer of hBChE that undergoes a second aging reaction yielding a phospho-serine-hBChE adduct. CBDP inhibition of mAChE and hBChE involves a different stereoisomer of CBDP, occurs through different pathways, and yields a different end-conjugate.

EXPERIMENTAL PROCEDURES

Caution: *CBDP is a highly toxic organophosphorus compound. Handling requires suitable personal protection, training, and facilities. These requirements are the same as those for other poisonous organophosphorus compounds.*

Chemicals. CBDP was synthesized by Starks Associates Buffalo, NY, and provided by Dr. D. Lenz (USAMRICD, Aberdeen PG, MD) and Dr. Wolf Dettbarn (Vanderbilt University, Nashville, TN). A 0.1 M stock solution of CBDP in acetonitrile was stored at −20 °C and diluted in water on the day of the mass spectrometry or crystallography experiment. Pepsin (from porcine stomach mucosa) was purchased from Sigma (a member of the Sigma-Aldrich group, St. Louis, MO). ^{18}O-Water (99% ^{18}O) was purchased from ISOTEC (a member of the Sigma-Aldrich group). Chemical structures and reaction schemes were illustrated using the program ChemDraw (CambridgeSoft).

Production of Recombinant ChEs. Recombinant hBChE (L530stop) is a truncated monomer containing residues 1 to 529 but missing 45 C-terminal residues that include the tetramerization domain. Four of the nine carbohydrate attachment sites were deleted by site-directed mutagenesis. Mutagenesis of N486 resulted in glycosylation of N485, an asparagine that is not glycosylated in native hBChE so that the recombinant hBChE contains six N-linked glycans.[19] The recombinant hBChE gene was expressed in Chinese hamster ovary (CHO) cells, secreted into serum-free culture medium, and purified by affinity and ion-exchange chromatographies as described previously.[19]

The synthetic gene (GeneArt) coding for a truncated mutant of mAChE (L544Stop) was inserted into a pGS vector carrying the glutamine synthetase gene marker and expressed in Chinese hamster ovary (CHO)-K1 cells. The cells were maintained in serum-free Ultraculture Medium (BioWhittaker, Walkersville, MD) and transfected using DNA−calcium phosphate coprecipitation. Transfected clones were selected by incubation in media containing methionine sulfoximine. The mouse enzyme, secreted into the culture medium, was purified by affinity chromatography and ion-exchange chromatography using a protocol identical to that described for the recombinant human enzyme.[20] The enzyme was concentrated to 14 mg/mL using a Centricon-30 ultrafiltration microconcentrator (30000 MW cutoff, Amicon, Millipore, Billerica, MA) in 10 mM MES buffer at pH 6.5.

Crystallization of hBChE and Generation of CBDP-hBChE Conjugates. hBChE was concentrated to 9 mg/mL and crystallized using the hanging drop vapor-diffusion method as described previously.[19] Two different conjugates were obtained by soaking native crystals for either 2 min or 12 h at 4 °C in a mother liquor solution (0.1 M MES at pH 6.5 and 2.1 M ammonium sulfate) containing 1 mM CBDP. Crystals were then soaked for a few seconds in a cryoprotectant solution (0.1 M MES buffer at pH 6.5, 2.3 M ammonium sulfate, and 20% glycerol) before being flash-cooled in liquid nitrogen for data collection.

Crystallization of mAChE and Generation of CBDP-mAChE Conjugates. mAChE, concentrated to 14 mg/mL, was crystallized using the hanging drop vapor-diffusion method as described previously.[21] Two different conjugates were obtained by soaking native mAChE crystals for either 30 min or 12 h at 4 °C in a mother

liquor solution (0.1 M Tris HCl buffer at pH 7.4 and 1.6 M ammonium sulfate) containing 1 mM CBDP. Crystals were soaked for a few seconds in a cryoprotective solution (0.1 M Tris HCl buffer at pH 7.4, 1.8 M ammonium sulfate, and 18% glycerol) and flash-cooled in liquid nitrogen.

X-ray Data Collection and Processing, Structure Determination, and Refinement. Diffraction data were collected on the ID14-eh4 beamline[22] at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) with an ADSC Quantum Q315r detector. The beam was characterized by a wavelength of 0.9765 Å, and the crystals were held at 100 K during data collection. All data sets were processed with XDS,[23] intensities of integrated reflections were scaled using XSCALE, and structure factors were calculated using XDSCONV. The structures were solved by molecular replacement with the program MOLREP[24] of the CCP4 suite[25] using the recombinant hBChE structure (PDB entry 1P0I) and mAChE (PDB entry 4A16) as starting models. For all diffraction data sets, the initial models were refined as follows: a rigid-body refinement, carried out with REFMAC5,[26] was followed by iterative cycles of model building with Coot[27] and restrained TLS refinement was carried out with Phenix.[28] The bound ligands and their descriptions were built using the Dundee PRODRG2.5 server including energy minimization using GROMOS96.1 force field calculations. Successive alternation of refinement cycles and manual model building were performed until R_{cryst} and R_{free} did not decrease any further.

In addition, for mAChE, a refinement strategy involving TLS, Ramachandran, secondary structure restraints, and occupancy optimization was performed using Phenix. In the course of the refinement of oCSP-mAChE, it appeared that Phe297 adopts two alternative conformations and that occupancies of saligenin of oCSP-mAChE were partial. Therefore, the occupancies of Phe297 and saligenin in oCSP-mAChE were refined with Phenix as follows: *B*-factors and occupancy of Phe297 and saligenin were initially set to 30 Å^2 and 0.5, respectively. Then, coordinates and *B*-factors were refined resulting in a *B*-factor about 50 Å^2 for Ser198. To account for the higher disorder of oCSP, its *B*-factor was manually adjusted to the *B*-factor of Ser198 plus 20% (60 Å^2). Next, occupancies were refined, grouping the alternative conformation of Phe297 with saligenin. The resulting occupancies were rounded to the next digit, and a final refinement cycle of coordinate and *B*-factor was performed. Protein structures were illustrated using the program PyMOL (Schrödinger, LLC).

Mass Spectrometry of CBDP-mAChE and CBDP-hBChE after Reaction in $H_2{}^{18}O$. mAChE or hBChE (3.2 μM) was irreversibly inhibited by reaction with CBDP (93 μM) either in $H_2{}^{18}O$ (93% enriched) or $H_2{}^{16}O$ and 10 mM ammonium bicarbonate (pH 8) at 22 °C for 1 or 40 h. At the end of the inhibition period, there was no activity remaining for either mAChE or hBChE. Before proteolysis, water from both the $H_2{}^{18}O$ and $H_2{}^{16}O$ samples was removed by evaporation in a Savant SpeedVac (Thermo Fisher Scientific, Waltham, MA). The samples were then redissolved in $H_2{}^{16}O$ to a final concentration of 3.2 μM. Removal of the ^{18}O water simplified the mass spectra by preventing ^{18}O atoms from being incorporated into the carboxyl termini of the peptides during proteolysis. Proteolysis was performed by adding pepsin (5 mg/mL in 5% formic acid) to each tube to obtain a final cholinesterase to pepsin ratio of 1:2 (w/w) and a final pH of 2.5. Mixtures were incubated at 37 °C for 2 h. Aliquots of each digest were diluted 1 to 10 in 50% acetonitrile and 0.3% trifluoroacetic acid (TFA), spotted directly onto MALDI sample plates, and air-dried. Dried spots were overlaid with a 2,5-dihydroxy benzoic acid matrix (Acros Organics a part of the Thermo Fisher Scientific group, Geel, Belgium). Samples were analyzed in a MALDI TOF/TOF 4800 mass spectrometer (Applied Biosystems, Foster City, CA) in negative reflector mode with delayed extraction (625 ns) and a laser intensity of 5500 V.

Molecular Dynamics Simulations of oCP-mAChE. A monomer from the oCSP-mAChE X-ray structure was used as the initial model for molecular dynamics simulations. All crystallographic water molecules were conserved. Both alternate conformations of Phe297 being present in the experimental structure, we retained only the non-native one. The saligenin group located in the acyl-binding pocket was manually removed from the initial model in order to generate the oCP-mAChE conjugate. We developed a full force field (FF) for the o-CP serine. The modified serine was built using Chimera[29] with geometry optimization at the Molecular Mechanics level. RESP charges of the modified residue were calculated using the REDS server[30] following the recommended procedure, including a Quantum Mechanics optimization of the fragments achieved with the Gaussian09 package at the HF/6-31G* theoretical level. The new FF was included in the amber99sb FF.[31] When parameters (bond distances, angles, and dihedral) were not available in amber99sb, they were assigned from the General Amber Force Field[32] using Antechamber.[33] Molecular dynamics simulations were carried out using the GROMACS 4.5.4 package.[34] The protein system was immersed in a 10-Å layer truncated cubic periodic water box using the TIP3P solvation model.[35] Charge equilibration was made by adding 9 Na^+ ions. A 2-fs time step was used in all the simulations, and long-range electrostatic interactions were treated with the particle-mesh Ewald (PME) procedure[36] using a cubic B-spline interpolation and a 10^{-5} tolerance for the direct-space and with a 12-Å nonbonded cutoff. A 300 K temperature coupling scheme using a separated Berendsen thermostat for solvent and protein was applied. Pressure was kept constant at 1 bar by a Berendsen barostat considering a compressibility of 4.5×10^{-5}. Molecular dynamics was preceded by an energy minimization (500 cycles of Steepest Descent) and a 50 ps position-restrained MD-simulated soak. Six full MD simulations differing only by the seed used to generate initial velocities were performed for 10 ns at 300 K under these conditions. The resulting trajectories were visualized using VMD[37] and analyzed using the tools provided in the GROMACS package.

RESULTS

Crystallographic Structures of *o*-Cresyl-phospho-hBChE (oCP-hBChE) and Phospho-hBChE. The oCP-hBChE adduct was generated by soaking a hBChE crystal for 2 min in a mother liquor solution containing 1 mM CBDP. The crystal was then flash-cooled to 100 K, and diffraction data were collected to a resolution of 2.7 Å (Table 1).

The initial Fourier-difference electron density map ($F_o - F_c$) featured a strong positive peak (10.4 σ) at covalent bonding distance from catalytic Ser198Oγ (not shown), allowing to assign the position of the phosphorus atom of the CBDP adduct. The saligenin substituent was not present despite the short soaking time. Rather, the ligand was modeled as an *o*-cresyl phosphate, with its *o*-cresyl moiety bound in the acyl-binding pocket (Figure 1A); this binding scenario implies that the P(R) enantiomer of CBDP initially bound to the enzyme. In the refined model, the covalent bonds oCP600P-Ser198Oγ, Ser198Oγ-Ser198Cβ, and Ser198Cβ-Ser198Cα are elongated and characterized by distances of 2.0, 1.7, and 1.65 Å, respectively. These values can be compared to predicted values of 1.60, 1.45, and 1.55 Å (calculated *ab initio* using Gaussian09) and to observed values of 1.85, 1.45, and 1.55 Å, in the previously determined crystal structure of phospho-hBChE, the end product of the hBChE inhibition by CBDP (Figure 1B; pdb ID 2y1k). Restraining the length of these bonds to their ideal values led to the appearance, in the resulting Fourier-difference map, of a large negative peak (>5 σ) between the phosphorus atom and Ser198Oγ, and of equivalent positive peaks on the opposite side of the phosphorus atom and Ser198Oγ. We interpret this observation as an indication of strain in the 3 consecutive bonds between the phosphorus atom and the serine main chain, which we propose to stem from the steric exclusion of the *o*-cresyl group by Trp231 (closest distance between non-hydrogen atoms: 3.5 Å). The entire phosphoryl group is pushed-up the active site gorge by 1.3 Å

Table 1. Crystallographic and Refinement Statistics

	oCP-BChE	oCSP-AChE	oCP-AChE
PDB entry code	4bbz	4bc0	4bc1
space group	*I*422	$P2_12_12_1$	$P2_12_12_1$
unit cell (Å)			
a	154.7	135.5	136.9
b	154.7	173.3	174.0
c	127.0	224.9	225.6
resolution (Å)	54.7−2.7	48.3−2.95	48.6−3.35
completeness (%)[a]	98.1 (99.0)	99.7 (99.4)	96.6 (98.6)
R_{sym} (%)[a]	7.1 (46.5)	8.1 (52.8)	7.5 (58.4)
$I/\sigma(I)$[a]	28.1 (5.2)	17.1 (3.1)	16.2 (2.8)
unique reflections[a]	21 049	111 473	75 488
redundancy[a]	10.1 (9.2)	4.4 (4.5)	3.6 (3.4)
Wilson *B* factor (Å^2)	53.4	63.9	94.92
R_{fact} (%)	16.6	18.6	16.2
R_{free} (%)	22.5	23.8	20.7
non-hydrogen atoms	4547	17994	17432
protein	4222	16 899	16 806
ligands	186	260	215
solvent	139	835	411
RMS bond length (Å)	0.008	0.009	0.009
RMS bond angles (deg)	1.269	1.368	1.482
Ramachandran			
favored (%)	94	92	94
allowed (%)	6	7.8	5.7
outliers (%)	0	1.2	0.3
average *B* factor (Å^2)	41.3	46.8	47.2
protein	40	46.9	47.1
ligands	74.1	91.2	92.7
solvent	36.2	31.5	28.3

[a]Values in parentheses refer to the highest resolution shell.

and twisted counter-clockwise around the Ser198Oγ-P bond when compared to its conformation in the phospho-BChE structure (Figure 2, panel A). This results in the absence of hydrogen-bonding interaction between the phosphoryl oxygen (O3P) of oCP600 and Ala199N in the oxyanion hole, again in contrast to what is observed in the phospho-hBChE structure (4.5 Å interatomic distance vs 2.7 Å in the oCP- and phospho-hBChE structures, respectively). Likewise, the hydrogen bond interaction between O1P of oCP and His438Nε2 and Ser198Oγ are both weakened (3.3 Å and 3.2 Å, respectively) relative to phospho-hBChE (2.9 Å and 3.0 Å; Figure 1B).

Altogether, these observations indicate that the oCP-hBChE adduct is not as stable as a canonical aged-BChE-OP conjugate, which the phospho-hBChE structure resembles more.

Crystallographic Structure of *o*-Cresyl Saligenin Phospho-mAChE (oCSP-mAChE). The oCSP-mAChE adduct was trapped by flash cooling a mAChE crystal that had been soaked for 30 min in a mother liquor solution containing 1 mM CBDP. The structure was solved at 2.95 Å resolution (Table 1). In this space group, the asymmetric unit contains two mAChE biological dimers that are related by noncrystallographic symmetry, and in each, monomers associate through a four helix-bundle.[38] In each dimer, one monomer features a solvent accessible peripheral site, while the peripheral site of the other is occupied by the Cys257-Cys272 loop from another monomer in the asymmetric unit tetramer. The *o*-cresyl and saligenin substituents are located in the choline- and acyl-binding pockets, respectively, indicating a preferential binding of the P(S) enantiomer of CBDP in mAChE (Figure 1C). This is in contrast with hBChE, in which it is the P(R) enantiomer that preferentially binds and where the *o*-cresyl substituent is located in the acyl-binding pocket (Figure 1A). In the acyl-binding pocket of oCSP-mAChE, Phe297 adopts an alternate conformation with a χ1 rotation of about 145° (around the Cα-Cβ bond), as required for the accommodation of saligenin (Figure 1C). We used the direct correlation between the alternate conformation of Phe297 and the presence of the saligenin moiety in the acyl-binding site to evaluate the partial occupancy of saligenin in the oCSP-mAChE structure (see the Experimental Procedures section for details). The fraction of bound CBDP retaining saligenin ranges from 0.4 to 0.5, depending on which of the four monomers is inspected. Overall, the geometry of catalytic-Ser203 in oCSP-mAChE is not distorted compared to catalytic-Ser198 in oCP-hBChE (Figure 2, compare the slate colored structures in panels A and B). The refined length of the covalent oCSP600P-Ser203Oγ bond is 1.7 Å in the mAChE structure, compared to a predicted value of 1.59 Å (as calculated *ab initio*). Finally, the catalytic His447Nε2 forms a H-bond with Ser203Oγ (2.8 Å) but does not interact with O2P of the cresyl substituent (Figure 1C).

Crystallographic Structure of *o*-Cresyl-phospho-mAChE (oCP-mAChE). The oCP-mAChE adduct was obtained by extending the soaking time to 12 h before flash cooling the crystal. The resolution of this structure is relatively low (3.35 Å) in comparison to that of the analogous oCP-hBChE structure (Table 1). It is unclear whether it is the relatively long soaking time or the effect of enzymatically produced saligenin that affected the crystal quality and the resolution of the diffraction data. Electron density maps were yet of good quality and thus allowed the unambiguous modeling of the *o*-cresyl moiety into a strong peak of positive $F_o - F_c$ electron density (10 σ) in the vicinity of the catalytic-serine hydroxyl group (Figure 1D). No electron density is observed in the acyl-binding pocket, indicating that the saligenin moiety has exited the active site (Figure 1D). Accordingly, Phe297 fully occupies its native position as indicated by the electron density maps and further confirmed by the occupancy refinement of its side chain. The *o*-cresyl moiety is stabilized in the choline-binding pocket by a perpendicular π-stacking interaction with the aromatic rings of Trp86. No H-bond interaction is noticeable between the *o*-cresyl O2 atom and His447Nε2, the later being hydrogen-bonded to Ser203Oγ (2.9 Å). With the *o*-cresyl substituent in the choline-binding pocket, the oCP-mAChE adduct appears to be the mirror image of the oCP-hBChE adduct in which the substituent is found in the acyl-binding pocket (Figure 1A and D). X-ray data collected after soaking periods ≥12 h repeatedly yielded an oCP-mAChE adduct and not a phosphoserine adduct as observed in hBChE.

Mass Spectrometry Analysis. The repeatedly unproductive attempts to obtain the structure of a phospho-mAChE adduct suggested that CBDP inhibition of mAChE occurs in two steps, i.e. phosphorylation of the enzyme and a canonical aging reaction. From our crystallographic data alone, however, we could not determine whether the observation of an *o*-cresyl phosphoserine adduct as the end-product of mAChE inhibition by CBDP was a crystallographic artifact or, rather, a sign for CBDP inhibition of AChE and BChE occurring through different pathways.

In order to ascertain the number of steps involved in the *in vitro* mAChE inhibition by CBDP, MALDI-TOF mass-spectrometry of proteolyzed, CBDP-inhibited mAChE was

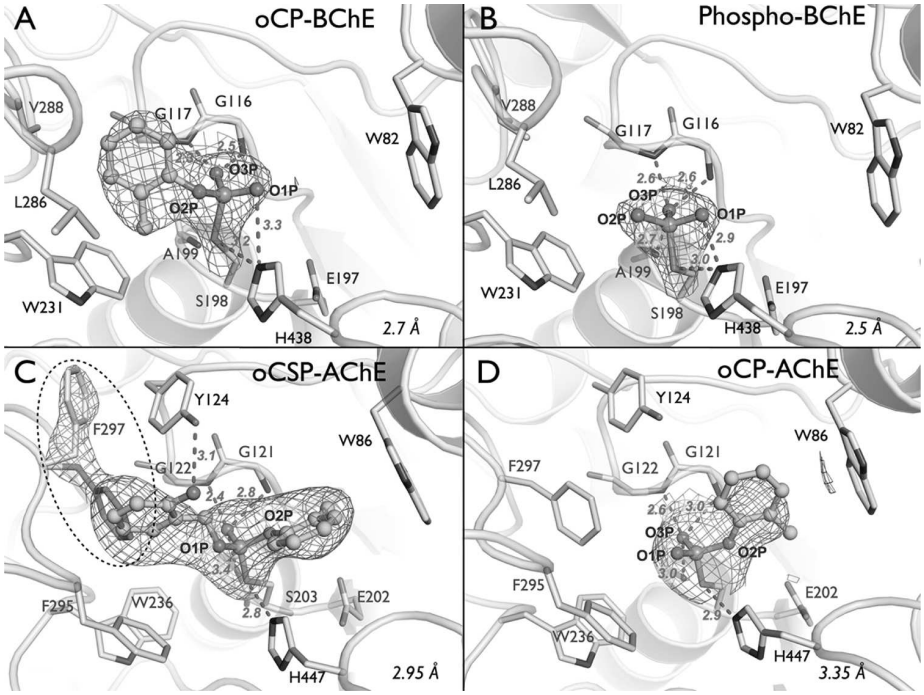


Figure 1. Sequential crystallographic snapshots of mAChE and BChE inhibition by CBDP. For BChE (upper panels), a soaking time of 2 min yielded an *o*-cresyl-phospho-BChE conjugate (A, current work). A soaking time of 12 h yielded the final phospho-BChE conjugate[7] (B). For mAChE (lower panels), soaking times of 30 min and 12 h yielded the *o*-cresyl saligenin phospho-AChE conjugate (C) and the *o*-cresyl-phospho-AChE conjugate (D), respectively. Key residues are represented as sticks with carbon atoms in green, nitrogen atoms in blue, phosphorus in orange, and oxygen atoms in red. Atoms and bonds of the adducts are represented as ball and sticks, respectively. Hydrogen bonds are represented by red dashes with distances in Å. A dashed ellipse in panel C highlights the native (magenta) and an alternate (green) position of Phe297. The electron density $2F_o - F_c$ is represented by a blue mesh contoured at 1.0 σ.

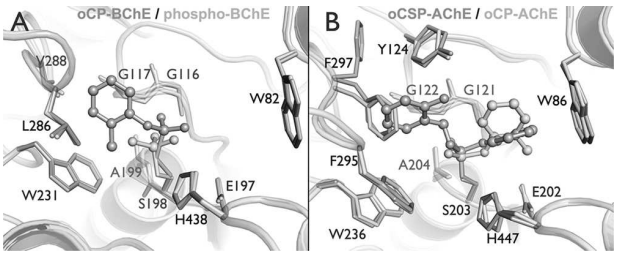


Figure 2. (A) Superimposition of oCP-BChE (slate) and phospho-BChE (green) structures. (B) Superimposition of oCSP-AChE (slate) and oCP-AChE (green) structures. Key residues are represented as sticks. Atoms and bonds of the adducts are represented as ball and sticks, respectively.

used, as previously described for hBChE.[7] We investigated the involvement of water molecules in the various steps of mAChE inhibition by CBDP by performing the inhibition in either $H_2{}^{16}O$ or in $H_2{}^{18}O$, which allowed us to characterize the step at which water is incorporated into the enzyme adducts. As a control, the same analyses were performed on CBDP-inhibited hBChE. Inhibition yields reached 100%, for both mAChE and hBChE, as evidenced by the absence of catalytic activity at the end of the reaction period and by the absence of unlabeled active-site peptides in the mass spectra.

From the hBChE-CBDP reaction, both *o*-cresyl phosphate (oCP) and phosphate adducts on the active-site peptic-peptide were detected.[18] Post source decay fragmentation confirmed that these adducts were on the active-site serine. Figure 3, panels A and B, show mass spectra (in negative mode) for *o*-cresyl phosphate adducts at 877 amu (for the peptic peptide FGESAGAA) and at 964 amu (for the miscleavage peptide FGESAGAAS). These adducts are formed following the release of saligenin from the initial CBDP-hBChE reaction product. The masses of these adducts are not affected by the presence of $H_2{}^{18}O$ indicating that no oxygen atoms from water are incorporated into the phosphorus moiety upon the release of saligenin. Figure 3, panels A and B, also show mass spectra (negative mode) for phosphate adducts. In $H_2{}^{16}O$, the phosphate adduct appears at 874 amu (for peptide FGESAGAAS). In $H_2{}^{18}O$, the 874 amu mass is shifted to 876 amu indicating the incorporation of an oxygen atom from water into the phosphate adduct upon the hydrolytic release of the *o*-cresyl moiety.

From the mAChE-CBDP reaction, only an *o*-cresyl-phosphate adduct was detected (Figure 3, panels C–D). No evidence for the phosphate adduct was obtained after the 1 h reaction period used for this experiment, nor was the phosphate adduct detected when the reaction was allowed to proceed for 40 h. Post-source decay fragmentation showed that the *o*-cresyl phosphate adduct was located on the active site serine. In mass spectra (negative mode) for both the $H_2{}^{16}O$ and $H_2{}^{18}O$ samples, masses for the oCP adducts appeared at 877 amu (for peptide FGESAGAA) and 964 amu (for peptide FGESAGAAS) indicating that no oxygen atom from water is incorporated into the adduct upon release of saligenin.

Molecular Dynamics Simulations. Phe297 is in an alternate conformation in the oCSP-mAChE adduct and moved back to its native position during aging as seen in the crystallographic structure of the oCP-mAChE adduct. We

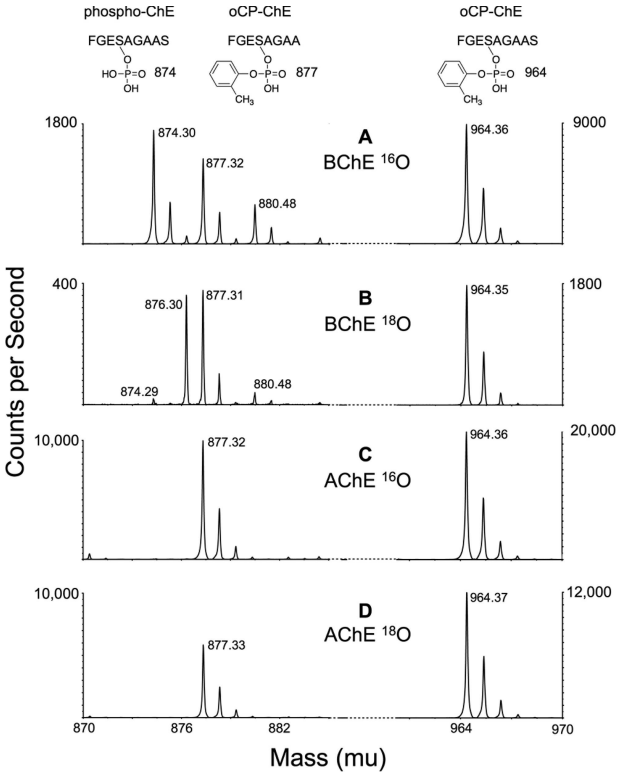


Figure 3. Reaction of hBChE and mAChE with CBDP in the presence of O^{16}-water (A and C) and O^{18}-water (B and D). The 874–877 amu region and 964 amu region of pepsin-digested 1-h CBDP-inhibited ChEs are shown. All spectra were taken in negative mode. The sequence of the phosphopeptide of mass 874 is FGES*AGAAS with PO_4 on serine. The sequences of the *o*-cresyl phosphate adducts of mass 877 and 964 are FGES*AGAA and FGES*AGAAS, where the asterisk indicates the labeled serine.

carried out molecular dynamics (MD) simulations in order to address whether the observed conformational change is part of equilibrium dynamics preexisting to ligand binding or a fit induced by ligand binding.[39] An initial model was created based on the oCSP-mAChE crystallographic structure, i.e., with the side chain of Phe297 in the alternate conformation but not with the saligenin substituent. In none of the six 10-ns MD simulations was a reversal of Phe297 to its native conformation observed.

■ DISCUSSION

Previously, we showed that hBChE inhibition by CBDP proceeds through a three-step reaction mechanism.[18] First, the catalytic serine is organophosphorylated to form the ring-opened CBDP-hBChE adduct. A dealkylation step then occurs viz. the first aging reaction, which results in the loss of the saligenin moiety. Last, the oCP-hBChE adduct undergoes hydrolysis with the concomitant departure of the *o*-cresyl moiety yielding the final phosphoserine hBChE adduct (Scheme 1).[7] That the latter is the end product of hBChE inhibition by CBDP was known from its X-ray structure at 2.7 Å resolution as well as from previous mass spectrometry experiments. It remained unclear, however, how these steps proceed at the atomic level of resolution. Also, it was unknown if mAChE follows the same inhibition pathway as hBChE and/or displays the same enantiomer selectivity.

In the current study, a combination of X-ray crystallography and mass-spectrometry was used to specifically address the above-mentioned unresolved issues. Atomic-resolution crystallographic snapshots of hBChE and mAChE were taken at different time points during their reaction with CBDP. Together with adduct masses obtained from the mass spectra, they allow one to develop detailed mechanistic models for the CBDP inhibitions of mAChE and hBChE, respectively.

Inhibition of mAChE by CBDP. In this study, mAChE was employed instead of hAChE because our preparation of the latter only crystallizes in the presence of the gorge-capping peptide fasciculin. Such capping prevents inhibitors from entering the active site precluding the use of soaking techniques to introduce inhibitors. mAChE is yet a reliable model for hAChE in that nearly all of the residues lining the active site gorge of mAChE and hAChE are identical. Noteworthy, Cheung et al. very recently reported a truncated form of hAChE able to crystallize without fasciculin and suitable for soaking techniques.[40] Use of this new form of hAChE will be preferred for future crystallographic studies.

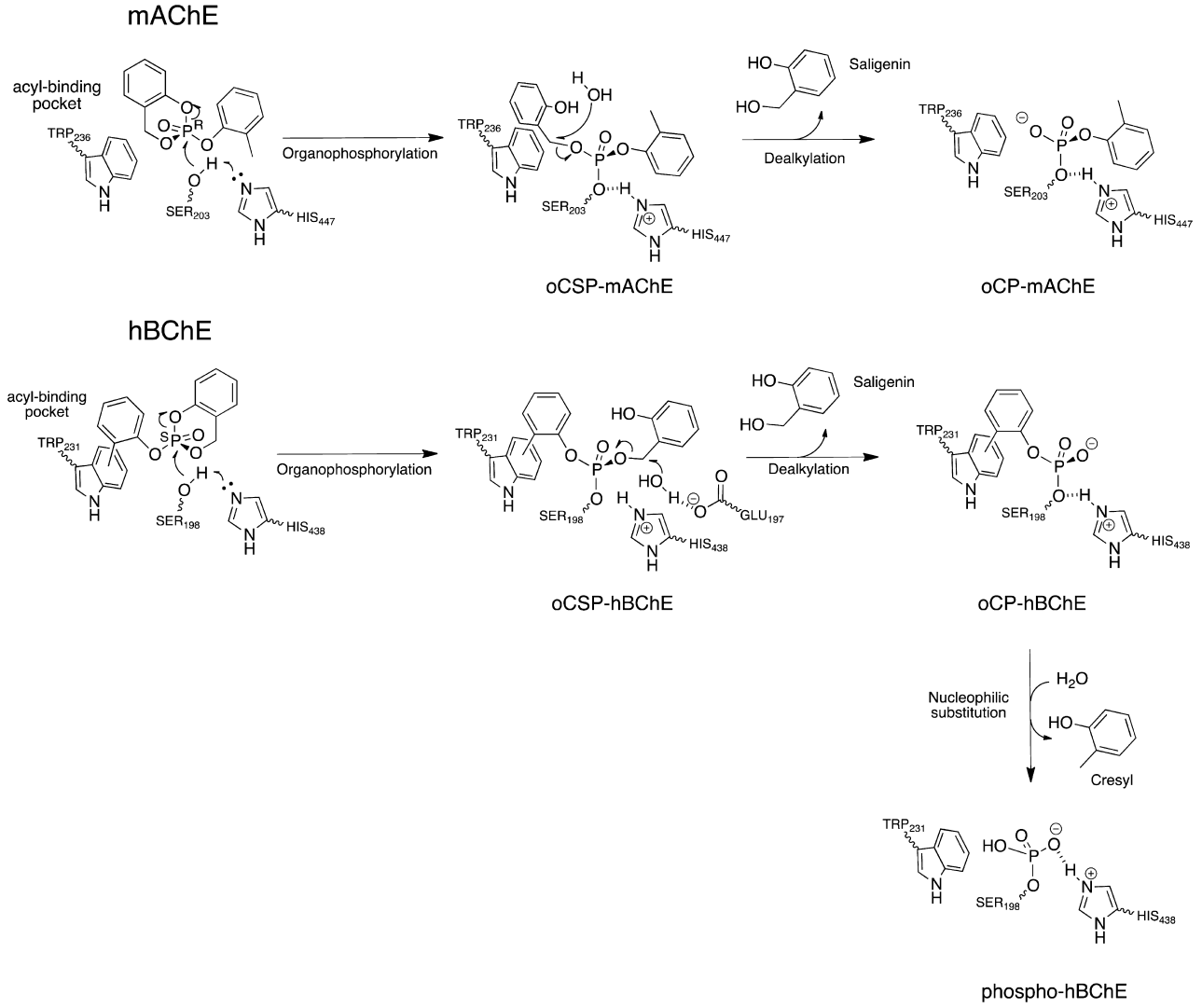
After soaking a crystal of mAChE for 30 min in a CBDP containing solution, the initial ring-opened oCSP-AChE adduct could be trapped by flash cooling (Figure 1C). The *o*-cresyl and saligenin moieties are located in the choline- and the acyl-binding pockets, respectively. This positioning indicates that crystalline mAChE preferentially binds and reacts with the P(S) enantiomer of CBDP.

Figure 1C also shows that the acyl-binding pocket is only partially occupied by saligenin. The partial occupancy (0.4–0.5) indicates that about half of the crystalline oCSP adducts have released saligenin after a 30 min soak (note that the *o*-cresyl site is fully occupied). Thus, the half-life for the first aging reaction is about 30 min *in crystallo*. Extending the soaking time to 12 h or more results in the complete elimination of saligenin and in the formation of an oCP-AChE conjugate (Figure 1D).

Mass spectral analysis after proteolysis, combined with the crystal structures, shows that mAChE inhibition by CBDP occurs in two steps, viz. organophosphorylation and dealkylation of saligenin in the acyl-binding pocket, yielding a final *o*-cresyl-phosphoserine adduct (Scheme 2). The fact that no water was incorporated in the *o*-cresyl-phosphoserine mAChE adduct indeed excludes the possibility that saligenin was released through its hydrolysis. Dealkylation is an unusual reaction for a substituent located in the acyl-binding pocket: the two residues known to catalyze dealkylation of OPs adducts, His447 and Glu202, are indeed part of the choline-binding pocket, on the opposite site of the active site gorge. No unusual strain is observed in the o-CSP adduct, which eliminates this source of enthalpy as the driving force for the dealkylation reaction. This implies that another driving force must be involved. A possibility is that the hydroxyl group of saligenin stabilizes a water molecule that could in turn react with the benzylic carbocation, thereby promoting the dealkylation reaction. The return to its native conformation of the Phe297 side chain, which was displaced from the acyl-binding pocket by the saligenin substituent, could constitute an additional driving force for the dealkylation.

Both the 40-h mass spectrometry and the 12-h crystallographic data indicate that the oCP adduct is the end product of mAChE inhibition by CBDP, in contrast to hBChE where the oCP adduct reacts further to yield a phosphoserine adduct (Figure 1B).[7] This difference is most likely to ascribe to the

Scheme 2. Detailed Chemical Mechanism for CBDP Inhibition of mAChE and hBChE



symmetric configuration of the *o*-cresyl substituent in mAChE and hBChE. In hBChE, the *o*-cresyl substituent is in the acyl-binding pocket; thus, the choline-binding pocket remains free for a water to approach the phosphorus atom, resulting in oCP adduct hydrolysis. Conversely in mAChE, hydrolysis does not occur because the *o*-cresyl substituent is located in the choline-binding pocket, meaning that the water would have to attack from the crowded acyl-binding pocket. While dearylation of the oCP-mAChE adduct could theoretically have occurred, considering that the *o*-cresyl substituent is located in proximity to His447/Glu202, our data unequivocally show that it does not. In this context, it is noteworthy mentioning that one of the strategies envisaged to counteract OP-intoxication is to realkykate aged phosphylconjugates;[41] by showing that dearylation does not occur on the time scale of our experiments (tens of hours), our data thus suggest that arylation may be more promising than realkylation.

In mAChE, organophosphorylation correlates with a rotation of the Phe297 side chain (Figure 1C), which enlarges the acyl-binding pocket and is thus required to accommodate the saligenin moiety. An identical conformational change was observed in the crystal structures of fenamiphos-inhibited mAChE in complex with the ortho-7 reactivator (pdb accession code 2wu4)[42] as well as in a *Tc*AChE/Bis(5)-tacrine complex (pdb accession code 2cmf).[43] In the ortho-7 complex, one pyridinium ring of ortho-7 pushes the isopropylamino substituent of fenamiphos toward the acyl-loop 285−298 and leads to its rearrangement. In the *Tc*AChE complex, non-covalent binding of the tacrine produces drastic rearrangements in the active site gorge, including a rotation of Phe290 (Phe297 in mAChE), that are induced by a tilt of Tyr121.[43] Since molecular dynamics simulations of native *Tc*AChE did not reveal alternate Phe290 conformations,[44] the rotation seen in the *Tc*AChE/Bis(5)-tacrine complex must be induced by the ligand rather than being selected by the ligand among preexisting native conformations. By extrapolation, the rotation of Phe297 observed in mAChE is also most likely the consequence of an induced fit mechanism, meaning that the conformational change is induced by ligand binding, rather than being part of preexisting equilibrium fluctuations. Although the acyl-loop determines the specificity of hAChE for the natural substrate acetylcholine,[45,46] the acyl-loop remains flexible enough to accommodate bulkier ligands at the cost of an induced fit rearrangement. This rearrangement is reversed once

the acyl-binding pocket becomes unoccupied, as after the aging of oCSP-mAChE. The conformational change was not reproduced in six 10 ns molecular dynamics simulations, indicating that the crystallographically observed conformational change is not a simple side chain flip but rather requires a larger rearrangement of the acyl-loop region that is not observed on the time scale of the simulations. These larger rearrangements of the acyl-loop are probably part of the proposed induced-fit mechanism.

Inhibition of hBChE by CBDP. The first species observed in the reaction of CBDP with hBChE is oCP-hBChE (Figure 1A), which corresponds to the second intermediate in the reaction mechanism (see Scheme 1). oCP-hBChE is created by the loss of saligenin from the initial intermediate, the ring-opened CBDP adduct. The latter could not be trapped, even by reducing soaking times to as short as 30 s, indicating that this adduct decays faster than it forms.

The *o*-cresyl moiety is located in the acyl-binding pocket of the oCP-hBChE adduct, which indicates that the saligenin moiety was bound in the choline-binding pocket prior to the first aging step and suggests a preferential reaction of the P(R) enantiomer of CBDP with crystalline hBChE. MD and QM/MM studies of hBChE reacting with CBDP, as well as titration experiments, support the proposed stereoselectivity (Lushchekina et al., unpublished results). Mass spectrometry showed that no oxygen atom from water is incorporated into the oCP-adduct, inferring that the release of saligenin likely stems from a His438/Glu197-catalyzed dealkylation.[7] This conclusion is in agreement with the proposal that aging of OP-hBChE proceeds through O-dealkylation of substituents bound in the choline-binding pocket.[47] However, an oxygen atom from water is incorporated during the second aging reaction. This observation supports the hypothesis of a nucleophilic attack, by a water molecule, on the phosphorus atom of oCP, resulting in the release of the o-cresyl moiety and the formation of the phospho-BChE adduct.[7]

From a mechanistic point of view, the most probable scenario is that this water attacks from the open face opposite the cresyl−O-P bond, vicinal to Glu197 (refer to Figure 1A), leading to the formation of a bipyramidal transition state and its subsequent collapse, with the release of *o*-cresyl (Scheme 2). The other possibility is that water approaches from the side opposite to the Ser198Oγ-P bond. However, we conjecture this scenario is unlikely considering the steric hindrance imposed by the presence of the *o*-cresyl moiety in the acyl-binding pocket. Also, a water attack from that side should lead to the scission of the Ser198Oγ-P bond and, therefore, to the release of the entire *o*-cresyl phosphate group, not just the *o*-cresyl moiety. Yet, neither was a native active site peptide identified in our mass spectrometry study, nor was self-reactivation observed in our previous kinetics studies. Thus, the attack of water is proposed to occur from the open face opposite the cresyl−O-P bond, vicinal to Glu197. A similar mechanism was proposed for the aging reaction of an analogue of the nerve agent tabun.[48]

A structural comparison between the oCP- and the phospho-hBChE adducts shows a release of the strain imposed on the enzyme, when *o*-cresyl is released and the phospho-conjugate is formed (Figures 2A). We conjecture that the release of this strain contributes to the driving force responsible for the scission of the P−O bond linking the *o*-cresyl to the phosphorus. It should be noted that unlike in mAChE, no side chain undergoes major conformational changes upon the binding of CBDP in hBChE.

CONCLUSIONS

The combination of crystallographic snapshots with mass spectrometric analysis shows that the inhibition of mAChE and hBChE by CBDP involves the binding of opposite enantiomers. Therefore, two different reaction pathways are undertaken by mAChE and hBChE for inhibition and aging. hBChE undergoes a three-step reaction process including (i) the organophosphorylation of the catalytic serine to form a CBDP ring-opened adduct, (ii) a dealkylation resulting in the loss of the saligenin substituent from the choline-binding pocket, and (iii) a nucleophilic hydrolysis of the *o*-cresyl substituent in the acyl-binding pocket, yielding a final phosphoserine-hBChE conjugate. In contrast, the inhibition of mAChE is a two-step reaction, resulting from of organophosphorylation and subsequent dealkylation, leading to the release of saligenin and the formation of the *o*-cresyl-phosphoserine mAChE adduct. Additionally, the crystallographic snapshots of the mAChE inhibition intermediates illustrate how a bulky inhibitor like CBDP is able to fit into the active site of mAChE through an induced fit rearrangement. It is worth noting that for both mAChE and hBChE, the final enzyme conjugate is aged, meaning that it cannot be reactivated by oximes.[47] In other words, there is currently no effective therapeutics to reactivate inhibited AChE in the case of severe TOCP intoxication.

ASSOCIATED CONTENT

Supporting Information

$2F_o - F_c$ and $F_o - F_c$ electron density maps. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*(F.N.) Tel: (+33) 476 636959. E-mail: florian@nachon.net. (M.W.) Tel: (+33) 438 789580. E-mail: martin.weik@ibs.fr.

Funding

Financial support by the DGA (project number DGA-REI 2009-34-0023 to M.W. and DGA/DSP/STTC 08co501 to F.N.), CEA, the CNRS, and the UJF is acknowledged, as well as grants from the Agence Nationale de la Recherche (ANR; project number ANR-09-BLAN-0192-04 to M.W. and F.N.) and the DTRA (HDTRA1-11-C-0047 to M.W. and F.N.). J.P.C. is a recipient of the Young International Scientists fellowship from the Chinese Academy of Science.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We are grateful to the ESRF for beam-time under long-term projects MX498, MX609, and MX722 (IBS BAG), and MX551 and MX 666 (radiation-damage BAG), and to the ESRF staff for providing efficient help during data collection. Mass spectra were obtained with the support of the Mass Spectrometry and Proteomics core facility at the University of Nebraska Medical Center.

ABBREVIATIONS

TOCP, tri-*o*-cresyl-phosphate; TCP, tricresyl phosphate; CBDP, 2-(*ortho*-cresyl)-4*H*-1,2,3-benzodioxaphosphoran-2-one, cresyl saligenin phosphate; OP, organophosphorus compound; hBChE, human butyrylcholinesterase; hAChE, human acetylcholinesterase; mAChE, mouse acetylcholinester-

ase; ChE, cholinesterase; oCP-hBChE, *o*-cresyl phosphate human butyrylcholinesterase conjugate; oCP, *o*-cresyl phosphate; oCSP-mAChE, *o*-cresyl saligenin phosphate mouse acetylcholinesterase conjugate; oCSP, *o*-cresyl saligenin phosphate; oCP-mAChE, *o*-cresyl phosphate mouse acetylcholinesterase conjugate

REFERENCES

(1) Winder, C. (2006) Air monitoring studies for aircraft cabin contamination. *Curr. Top. Toxicol. 3*, 33−48.

(2) Mackerer, C. R., Barth, M. L., Krueger, A. J., Chawla, B., and Roy, T. A. (1999) Comparison of neurotoxic effects and potential risks from oral administration or ingestion of tricresyl phosphate and jet engine oil containing tricresyl phosphate. *J. Toxicol. Environ. Health, Part A 57*, 293−328.

(3) Liyasova, M., Li, B., Schopfer, L. M., Nachon, F., Masson, P., Furlong, C. E., and Lockridge, O. (2011) Exposure to tri-o-cresyl phosphate detected in jet airplane passengers. *Toxicol. Appl. Pharmacol. 256*, 337−347.

(4) Casida, J. E., Eto, M., and Baron, R. L. (1961) Biological activity of a trio-cresyl phosphate metabolite. *Nature 191*, 1396−1397.

(5) Eto, M., Casida, J. E., and Eto, T. (1962) Hydroxylation and cyclization reactions involved in the metabolism of tri-o-cresyl phosphate. *Biochem. Pharmacol. 11*, 337−352.

(6) Eto, M., Oshima, Y., and Casida, J. E. (1967) Plasma albumin as a catalyst in cyclization of diaryl o-(alpha-hydroxy)tolyl phosphates. *Biochem. Pharmacol. 16*, 295−308.

(7) Carletti, E., Schopfer, L. M., Colletier, J. P., Froment, M. T., Nachon, F., Weik, M., Lockridge, O., and Masson, P. (2011) Reaction of cresyl saligenin phosphate, the organophosphorus agent implicated in aerotoxic syndrome, with human cholinesterases: mechanistic studies employing kinetics, mass spectrometry, and X-ray structure analysis. *Chem. Res. Toxicol. 24*, 797−808.

(8) Masson, P. (2011) Evolution of and perspectives on therapeutic approaches to nerve agent poisoning. *Toxicol. Lett. 206*, 5−13.

(9) Worek, F., Thiermann, H., Szinicz, L., and Eyer, P. (2004) Kinetic analysis of interactions between human acetylcholinesterase, structurally different organophosphorus compounds and oximes. *Biochem. Pharmacol. 68*, 2237−2248.

(10) Harel, M., Schalk, I., Ehret-Sabatier, L., Bouet, F., Goeldner, M., Hirth, C., Axelsen, P. H., Silman, I., and Sussman, J. L. (1993) Quaternary ligand binding to aromatic residues in the active-site gorge of acetylcholinesterase. *Proc. Natl. Acad. Sci. U.S.A. 90*, 9031−9035.

(11) Wandhammer, M., Carletti, E., Van der Schans, M., Gillon, E., Nicolet, Y., Masson, P., Goeldner, M., Noort, D., and Nachon, F. (2011) Structural study of the complex stereoselectivity of human butyrylcholinesterase for the neurotoxic V-agents. *J. Biol. Chem. 286*, 16783−16789.

(12) Winder, C., and Balouet, J. C. (2002) The toxicity of commercial jet oils. *Environ. Res. 89*, 146−164.

(13) van Netten, C. (2005) Aircraft air quality incidents: symptoms, exposures and possible solutions. *J. Occup. Health Saf. 21*, 460−468.

(14) Johnson, M. K. (1975) Structure-activity relationships for substrates and inhibitors of hen brain neurotoxic esterase. *Biochem. Pharmacol. 24*, 797−805.

(15) Glynn, P. (1999) Neuropathy target esterase. *Biochem. J. 344*, 625−631.

(16) Aldridge, W. N. (1954) Tricresyl phosphates and cholinesterase. *Biochem. J. 56*, 185−189.

(17) Earl, C. J., and Thompson, R. H. (1952) The inhibitory action of tri-ortho-cresyl phosphate on cholinesterases. *Br. J. Pharmacol. Chemother. 7*, 261−269.

(18) Schopfer, L. M., Furlong, C. E., and Lockridge, O. (2010) Development of diagnostics in the search for an explanation of aerotoxic syndrome. *Anal. Biochem. 404*, 64−74.

(19) Nachon, F., Nicolet, Y., Viguie, N., Masson, P., Fontecilla-Camps, J. C., and Lockridge, O. (2002) Engineering of a monomeric and low-glycosylated form of human butyrylcholinesterase: expression, purification, characterization and crystallization. *Eur. J. Biochem. 269*, 630−637.

(20) Carletti, E., Li, H., Li, B., Ekstrom, F., Nicolet, Y., Loiodice, M., Gillon, E., Froment, M. T., Lockridge, O., Schopfer, L. M., Masson, P., and Nachon, F. (2008) Aging of cholinesterases phosphylated by tabun proceeds through O-dealkylation. *J. Am. Chem. Soc. 130*, 16011−16020.

(21) Ronco, C., Carletti, E., Colletier, J. P., Weik, M., Nachon, F., Jean, L., and Renard, P. Y. (2011) Huprine derivatives as subnanomolar human acetylcholinesterase inhibitors: from rational design to validation by X-ray crystallography. *ChemMedChem 6*, 876−888.

(22) McCarthy, A. A., Brockhauser, S., Nurizzo, D., Theveneau, P., Mairs, T., Spruce, D., Guijarro, M., Lesourd, M., Ravelli, R. B., and McSweeney, S. (2009) A decade of user operation on the macromolecular crystallography MAD beamline ID14−4 at the ESRF. *J. Synchrotron Radiat. 16*, 803−812.

(23) Kabsch, W. (2010) XDS. *Acta Crystallogr., Sect. D 66*, 125−132.

(24) Vagin, A., and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. *J. Appl. Crystallogr. 30*, 1022−1025.

(25) Collaborative-Computational-Project-4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr., Sect. D 50*, 760−763.

(26) Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D 53*, 240−255.

(27) Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr., Sect. D 66*, 486−501.

(28) Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D 66*, 213−221.

(29) Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera: a visualization system for exploratory research and analysis. *J. Comput. Chem. 25*, 1605−1612.

(30) Vanquelef, E., Simon, S., Marquant, G., Garcia, E., Klimerak, G., Delepine, J. C., Cieplak, P., and Dupradeau, F. Y. (2011) R.E.D. Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments. *Nucleic Acids Res. 39*, W511−517.

(31) Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins 65*, 712−725.

(32) Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004) Development and testing of a general amber force field. *J. Comput. Chem. 25*, 1157−1174.

(33) Wang, J., Wang, W., Kollman, P. A., and Case, D. A. (2006) Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graphics Modell. 25*, 247−260.

(34) Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput. 4*, 435−447.

(35) Mahoney, M., and Jorgensen, W. (2000) A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys. 112*, 8910−8922.

(36) Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: an N.Log(N) method for Ewald sums in large systems. *J. Chem. Phys. 98*, 10089−10092.

(37) Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graphics 14*, 33−38.

(38) Bourne, Y., Taylor, P., Bougis, P. E., and Marchot, P. (1999) Crystal structure of mouse acetylcholinesterase. A peripheral site-

occluding loop in a tetrameric assembly. *J. Biol. Chem. 274*, 2963−2970.
(39) Xu, Y., Colletier, J. P., Jiang, H., Silman, I., Sussman, J. L., and Weik, M. (2008) Induced-fit or preexisting equilibrium dynamics? Lessons from protein crystallography and MD simulations on acetylcholinesterase and implications for structure-based drug design. *Protein Sci. 17*, 601−605.
(40) Cheung, J., Rudolph, M. J., Burshteyn, F., Cassidy, M. S., Gary, E. N., Love, J., Franklin, M. C., and Height, J. J. (2012) Structures of human acetylcholinesterase in complex with pharmacologically important ligands. *J. Med. Chem. 55*, 10282−10286.
(41) Wandhammer, M., de Koning, M., van Grol, M., Loiodice, M., Saurel, L., Noort, D., Goeldner, M., and Nachon, F. (2012) A step toward the reactivation of aged cholinesterases - Crystal structure of ligands binding to aged human butyrylcholinesterase. *Chem.-Biol. Interact.*, DOI: 10.1016/j.cbi.2012.1008.1005.
(42) Hornberg, A., Artursson, E., Warme, R., Pang, Y. P., and Ekstrom, F. (2010) Crystal structures of oxime-bound fenamiphos-acetylcholinesterases: reactivation involving flipping of the His447 ring to form a reactive Glu334-His447-oxime triad. *Biochem. Pharmacol. 79*, 507−515.
(43) Rydberg, E. H., Brumshtein, B., Greenblatt, H. M., Wong, D. M., Shaya, D., Williams, L. D., Carlier, P. R., Pang, Y. P., Silman, I., and Sussman, J. L. (2006) Complexes of alkylene-linked tacrine dimers with Torpedo californica acetylcholinesterase: Binding of Bis5-tacrine produces a dramatic rearrangement in the active-site gorge. *J. Med. Chem. 49*, 5491−5500.
(44) Xu, Y., Colletier, J. P., Weik, M., Jiang, H., Moult, J., Silman, I., and Sussman, J. L. (2008) Flexibility of aromatic residues in the active-site gorge of acetylcholinesterase: X-ray versus molecular dynamics. *Biophys. J. 95*, 2500−2511.
(45) Harel, M., Sussman, J. L., Krejci, E., Bon, S., Chanal, P., Massoulie, J., and Silman, I. (1992) Conversion of acetylcholinesterase to butyrylcholinesterase: modeling and mutagenesis. *Proc. Natl. Acad. Sci. U.S.A. 89*, 10827−10831.
(46) Ordentlich, A., Barak, D., Kronman, C., Flashner, Y., Leitner, M., Segall, Y., Ariel, N., Cohen, S., Velan, B., and Shafferman, A. (1993) Dissection of the human acetylcholinesterase active center determinants of substrate specificity. Identification of residues constituting the anionic site, the hydrophobic site, and the acyl pocket. *J. Biol. Chem. 268*, 17083−17095.
(47) Masson, P., Nachon, F., and Lockridge, O. (2010) Structural approach to the aging of phosphylated cholinesterases. *Chem.-Biol. Interact. 187*, 157−162.
(48) Nachon, F., Carletti, E., Worek, F., and Masson, P. (2010) Aging mechanism of butyrylcholinesterase inhibited by an N-methyl analogue of tabun: implications of the trigonal-bipyramidal transition state rearrangement for the phosphylation or reactivation of cholinesterases. *Chem.-Biol. Interact. 187*, 44−48.

Resumé de la thèse en français

L'acétylcholinestérase (AChE), une des enzymes les plus rapides dans la nature, est la cible d'un large nombre de toxiques, dont notamment les neurotoxiques organophosphorés. La première partie de ce manuscrit de thèse décrit le développement raisonné d'un nouveau réactivateur, qui présente des propriétés de réactivation supérieures aux molécules actuellement sur le marché. Les interactions entre cette molécule, KM297, et l'AChE ont été étudiées par dynamique moléculaire, docking et cristallographie aux rayons X. La connaissance des modes de liaison du KM297 dans l'AChE native ou inhibé par un OP ont permis de développer la molécule JDS207, qui se lie de façon exclusive au site périphérique de l'AChE. La deuxième partie de la thèse est dédiée à l'analyse des simulations de la AChE par dynamique moléculaire. On observe que la combinaison de multiples trajectoires générées avec des paramètres de vélocité initiale différents est une méthode fiable pour caractériser les conformations atteintes par les chaînes latérales des acides aminés. En comparant la distribution des rotamères pour l'AChE humaine et celle du poisson *Torpedo californica*, on montre que des différences importantes existent entre les enzymes des deux espèces. A partir de ces informations sur les conformations de résidus clés du site actif, une méthode a été développée pour générer des récepteurs utilisable pour des calcules de docking flexible, de façon à prendre en compte la dynamique propre à chaque résidu de l'enzyme. Cette méthode a été validé en comparent les résultats obtenues à des structures cristallographiques connues.

Resumé de la thèse en anglais

Acetylcholinesterase (AChE), one of nature fastest enzyme, is the target of multiple toxics, including organophosphate nerve agents (OP). In the first part of this thesis I present the structure-based development of a new uncharged reactivator, which showed characteristics better than any molecule commercially available to date. The molecule has been rationally designed to present both affinity to the inhibited enzyme and good reactivation capabilities. The interactions between the lead molecule KM297 and AChE has been characterized by means of flexible docking, molecular dynamics simulations and X-ray protein crystallography. The deeper understanding of its binding modes to both native and OP-inhibited AChE has helped in developing a derivative, JDS207, whose binding mode at the peripheral site of AChE is optimized. This derivative has also been studied by flexible docking and X-ray crystallography. The design of this family of reactivators taught us that a deep insight of the AChE dynamics is necessary to optimize ligands. The second part of the thesis is devoted to the analysis of molecular dynamics simulations of AChE. At first, we assessed that combining multiple short simulations is a fast and reliable method to characterize the dynamics of the amino-acids side-chains. By comparing dynamics of the side-chains from hAChE and *Tc*AChE, we confirm that some key dynamical differences exist between the two enzyme. The knowledge of the rotamers issued of MD simulation has lead us to develop a new method to generate flexible receptors for docking, which is specific to each single residue in the enzyme. This method has been validated by comparing its output structures with the ones found on the PDB database.