Blind docking: A novel approach of locating binding sites of potential drugs on macromolecules

(Unusual applications of AutoDock)

by

Csaba Hetényi

PhD thesis

Department of Medical Chemistry
University of Szeged, Hungary
2002

MAGNA EST VERITAS.
Nagy az igazság. Nem lehet vele megalkudni.

(Szent-Györgyi Albert, 1962)

To the memory of my GRANDPARENTS

ABBREVIATIONS

AD	Alzheimer's disease
AMBER	assisted model building and energy refinement (a force field)
AuD	AutoDock (the program package)
Aβ	beta amyloid peptide of Alzheimer's disease (any type)
BSB	beta sheet breaker
$\mathbf{E_{inter}}$	intermolecular interaction energy
FFVLDGn	Phe-Phe-Val-Leu-Asp-Gly-NH_2 (a hexapeptide amide)
GA	genetic algorithm
GPG®	Gly-Pro-Gly-NH_2 (a potential AIDS-drug)
GVVIAn	Gly-Val-Val-Ile-Ala-NH_2 (a pentapeptide amide)
LGA	Lamarckian genetic algorithm (in AutoDock)
MC	Monte-Carlo method (named originally by von Neumann)
MM	molecular mechanics
QM	quantum mechanics
SA	simulated annealing

TABLE OF CONTENTS

LIST OF PAPERS

(The papers included in this thesis are referred to as their Roman numbers in the text.)

I ***Hetényi, C.***, Paragi, G., Penke, B and Tímár, Z. Docking approach of inhibitor design for β-secretase of Alzheimer's disease. *Manuscript.*

II ***Hetényi, C.***, Körtvélyesi, T. and Penke, B. 2001.Computational studies on the binding of β-sheet breaker (BSB) peptides on amyloid βA(1-42). *Journal of Molecular Structure (THEOCHEM)* **542** 25-31.

III ***Hetényi, C.***, Körtvélyesi, T. and Penke, B. 2002. Mapping of the possible binding sequences of two beta-sheet breaker peptides on beta amyloid peptide of Alzheimer's disease. B*ioorganic & Medicinal Chemistry* **10**(5) 1587-1593.

IV ***Hetényi, C.***, Szabó, Z., Klement, É., Datki, Z., Körtvélyesi, T., Zarándi, M. and Penke, B. 2002. Pentapeptide amides interfere with the aggregation of beta amyloid peptide of Alzheimer's disease. *Biochemical and Biophysical Research Communications* **292** 931-936.

V ***Hetényi, C.*** and van der Spoel, D. 2002. Efficient docking of peptides to proteins without prior knowledge of the binding site. *Protein Science*, Accepted for publication.

Other papers (not included in this thesis)

Végvári Á., Földesi A., ***Hetényi C.***, Kochegarova, O., Schmid, M.G., Kudirkaite, V. and Hjertén, S. 2000. A new easy-to-prepare homogeneous continuous electrochromatographic bed for enantiomer recognition. *Electrophoresis* **21:** 3116-3125.

Tong, D., ***Hetényi, C.***, Bikádi, Z., Gao, J.-P. and Hjertén, S. 2001. Some studies of the chromatographic properties of gels ('artificial antibodies/receptors') for selective adsorption of proteins. *Chromatographia* **54:** 7-14.

Van der Spoel, D., ***Hetényi, C.***, McKenzie-Hose, A., Ågren, L., Goobar-Larsson, L., Höglund, S. and Vahlne, A. A tripeptide-amide interferes with self-assembly of the HIV capsid protein. Submitted.

INTRODUCTION

Computational chemistry and computer-aided molecular modeling are relatively young, but rapidly growing branches of chemistry. During the validation of the different modeling approaches, a - more or less - simplified model of the real chemical system is constructed and the results (outputs) of the calculations are compared to the experimental ones so as to evaluate the reliability of the theoretical considerations.

There are different levels of molecular modeling and different theories corresponding to them. The first - and maybe the most difficult - step of molecular modeling is the selection of the appropriate method for our problem and system. Several factors could influence our choice, such as the amount of computer time spent on the solution of the problem. the size of the studied system, the information requested from the calculations, the amount of available experimental data, etc.

Basically, computational chemistry can be divided into two fields due to the theory of methods: (1) quantum mechanics and (2) molecular mechanics. These main fields correspond to different level of approximations and degrees of freedom (van Gunsteren and Berendsen 1990). While a quantum chemical algorithm works with nuclei and electrons the MM programs use atom types for the simplification of the molecular systems. Furthermore, MM methods require the explicit distinction between the bonded and non-bonded interactions of a particular atom (connectivity lists or residue definitions). Fortunately, the latter information is available for numerous biologically important (macro) molecules and, therefore MM methods are real alternatives for the calculations of molecules of thousands of atoms. (QM methods are not able to handle very large molecules within acceptable calculation times.) The combination of the two approaches results QM/MM applications, which are used e.g. for the study of reaction mechanisms of enzyme catalysis (Wang et al. 2001) .

In most of the cases, biological processes involve interactions between two or more molecules. Many drugs are developed to block such interactions by binding to one of the interacting partners. Generally, the blocking molecule is the enzyme inhibitor, which fills the active site of the enzyme and sterically hinders the biocatalysis. Fortunately, sophisticated experimental methods, like X-ray crystallography are available to measure the exact atomic details of the complex of the enzyme and its substrate/inhibitor (ligand) molecule. However, crystallization of a protein and co-crystallization of the protein with the ligand are sometimes very time-

consuming and the success cannot be guaranteed in all cases. Additionally, if potential drugs (ligands) were selected by random screening (trial and error), drug research projects would require enormous amount of time. These problems legitimate the existence of alternative methods for identification of lead-compounds and structure-based drug design.

Structure-based design covers many strategies, including both manual and automated approaches. The core problem of these strategies is searching the very large conformational space of the ligand-macromolecule complexes. The searching methods can be divided roughly into two classes: the search for the new molecules with properties similar to a known ligand (intramolecular constraints), and the search for new molecules with properties complementary to the enzyme (intermolecular constraints). The latter type of methods seems to be more general, as in many cases there is no known structure of working inhibitor. There are at least two subtypes of this "intermolecular approach". One strategy often referred to as *de novo* drug design and builds molecules within a rigid enzyme (Moon and Howe 1991, Bohm 1994; Gillet et al. 1994, Murray et al. 1995, Wang et al. 2000, Budin et al. 2001). The other strategy is to dock a set of favorable conformations of a possible ligand to the active site of the protein (Miller et al. 1994, Ewing and Kuntz 1997). A refined version of this strategy is to dock conformationally *flexible* ligands, i.e. molecules with released torsional angles (Goodsell and Olson 1990, Oshiro et al. 1995). This approach has the potential to scan larger piece of the conformational space during the energy minimization.

Generally, the afore-mentioned docking algorithms involve the calculation of intermolecular interaction energy and use an MM approach with a simplified force field. (Details of the connection between MM and docking can be found in the theoretical section.)

In this thesis we present novel applications of a widely used docking program package AutoDock (Morris et al. 1996, Morris et al. 1998). The main goal of our studies is to explore the capabilities of the program package for an unusual task, namely *blind*[#] binding site search. Peptides were used in our investigations for ligand molecules, as they are flexible, versatile and important bioactive compounds. The target molecules of the calculations are selected proteins and the amyloid peptide of Alzheimer's disease see section "Some (molecular) aspects of Alzheimer's disease" for explanation of the role of Aβ.

[#] The term: *ab initio* docking is used in the literature if docking is performed without prior knowledge of location of the binding site (Morelli et al., 2001). Here, we prefer the use of an other attribute "blind" for this kind of dockings, as the term: *ab initio* is reserved for some QM methods

THEORY

1 Molecular mechanics

Molecular mechanics is no doubt an indispensable tool of calculation of macromolecular systems. The two most important features of an MM method are the *force field* and the *search method* used for minimization. Here, we sketch a general approach to MM and focus on the force field and algorithms used in our studies. There are several comprehensive studies available (e.g. Wang et al. 2001, Keserű et Náray-Szabó 1995) on this topic with more detailed information.

1.1 Force fields

Generally, all current force fields applied to biopolymers to date are two-body additive. This means, that the potential energy function (see a general formula below and the corresponding explanation in the footnote[#]) is a function of pairs of atoms.

$$V(\vec{r}_1,\vec{r}_2,...,\vec{r}_N) = \sum_{bonds} \tfrac{1}{2} K_b (b-b_0)^2 + \sum_{angles} \tfrac{1}{2} K_\Theta (\Theta-\Theta_0) + \sum_{dihedrals} K_\varphi \left[1+\cos(n\varphi-\delta)\right] +$$

$$+ \sum_{pairs(i,j)} \left[C_{12}(i,j)/r_{ij}^{12} - C_6(i,j)/r_{ij}^6 + q_i q_j /(4\pi\varepsilon_0\varepsilon_r r_{ij}) \right]$$

There are several force fields developed for calculation of biological molecules, like proteins: e.g., AMBER (Weiner et al. 1984, Cornell et al. 1995), CHARMm (MacKerell et al. 1998), GROMOS96 (Stocker and van Gunsteren 2000), OPLS-AA (Kaminski and Jorgensen 1996) and MMFF (Halgren 1996). All of these force fields has a similar potential as the above one and use empirical fits to liquid or solid-state small molecular systems in order to derive van der Waals parameters and charges (except of AMBER).

[#]The first term corresponds to the covalent bond stretching along bond b. The minimum energy bond length is b_0 and the force constant is K_b, which vary with the particular type of the bond. The second term describes the bond angle bending (three-body) interactions. Dihedral angle (four body) interactions are represented by the sinusoidal term for the φ dihedral angles. The last term includes all the non-bonded interactions: van der Waals and Coulombic ones.

The AMBER force field

In our studies the AMBER force field was used for the calculation of peptide structures and for docking simulations. In case of AMBER, ab initio Hartree Fock calculations are used with a 6-31G* basis set to derive the electrostatic potential for residues or molecules and fit the partial charges to optimally reproduce the QM potential. In this sense, AMBER differs from all the other force fields, mentioned in the last section. For modeling of hydrogen bonds Lennard-Jones 12-10 potential is included in the parameter set.

1.2 Search methods

There are several energy minimization methods for finding the best energy configurations with low V(r) in the configurational space. The type of the V(r) function, size of the system and the types of the freedom of the system influence the success of the particular search methods. Here, we focus on the algorithms, which are able to perform "global" minimum search, i.e. find the energy minima of the entire system. (During our work gradient search methods were also used for routine minimization of peptides. See section "Methods" for details of preparation of ligands for the docking calculations.)

Generally, the biologically important systems (proteins, nucleic acids, etc.) contain many degrees of freedom and straightforward (systematic) scanning of the complete configuration space is impossible. Hence, the generation of a representative set of configurations is necessary. In our studies, "step methods" were used to generate new configurations from the previous ones. These methods could be distinguished by the way the step direction and the step size are choosen.

1.2.1 Monte-Carlo methods

In the MC methods the step direction is chosen at random, the actual step size is determined by the change in energy gradient and a random element in case $\Delta E>0$. MC methods work with the random displacement of the configuration of a system. The available Cartesian space is uniformly sampled and the newly generated configurations is accepted if $\Delta E \leq 0$ or if $\Delta E>0$, when

$$e^{-\frac{\Delta E}{k_B T}} > R$$

where R is a random number taken from a uniform distribution over the interval (0,1). This Metropolis MC procedure generates a Boltzmann ensemble with a great number of low energy configurations. MC is often paired with simulated annealing (Kirkpatrick et al. 1983) which results a global→local search during the simulations.

1.2.2 Genetic algorithm

The GA procedure is often used to search for global minima in a larger parameter space. Basically, the language of natural genetics and biological evolution is used in the theory of GA. In molecular docking, there are state variables (translation, orientation and rotation), which exactly define the state of a ligand on a macromolecule. Each variable is referred to as a gene and the string containing all genes as a chromosome or genome (genotype). The atomic coordinates of the ligand correspond to the phenotype. Each chromosome in the population is evaluated using an appropriate scoring scheme (the total interaction energy function) to measure the fitness of the solution. In the beginning of the calculations, a random population of the chromosomes is generated. The population size (number of individuals in the population) is set as an input. The maximum number of generation is also an input and a termination criterion of the calculations. There are four procedures, which determine the composition of the new populations. *(1) Selection.* A portion of the better chromosomes is copied directly into the next generation. *(2) Elitism.* Some of the best performing chromosomes survive unchanged. *(3) Crossing-over.* New, offspring chromosomes are produced with the genes of the parents. *(4) Mutation.* Genes are changed in a random amount. The runs are terminated if the maximum number of generations has been reached.

1.2.3 Molecular dynamics

In a MD simulation both step size and direction are determined by the force and the velocity. In general a search algorithm will make a step, which is a linear combination of the gradient, previous steps and a random contribution. The trajectory of the molecular system is generated by simultaneous integration of Newton's equations (F obtained from the potential function) of motion for all the atoms in the system:

$$\vec{F}_i = -\left(\partial V(\vec{r}_i ... \vec{r}_N) / \partial \vec{r}_i \right)$$

$$d^2\vec{r}_i(t)/dt^2 = m_i^{-1}\vec{F}_i$$

The integration of the equation is performed in small time steps, typically 1-10 fs for molecular systems. State equilibrium quantities can be obtained by averaging over the trajectory, and dynamic information can also be extracted. Recent GROMACS (Lindahl et al. 2001) MD studies of phospholipids show that it is possible to simulate even the formation of bilayers of these amfipatic molecules *in silico* (Marrink et al. 2001a, Marrink et al. 2001b). These results highlight the promising perspective of MD methods.

2 Docking

Docking methods are developed to find binding pockets (sites) of smaller or larger molecules on macromolecules. Generally, there are two branches of such methods: (1) matching or directed methods, which create a model of the possible interacting points (e.g. hydrogen bridge donor/acceptor atoms) of the site and fit the ligand[#] to this scheme and (2) algorithms which use a (simple) molecular mechanics force field and a search method so as to optimize the position of the ligand on the target[#]. The program package: DOCK (Kuntz et al. 1982, Shoichet and Kuntz 1993) is an example of the first branch and AutoDock (Morris et al. 1996, Morris et al. 1998) is that of the second group. There are numerous docking programs available nowadays (for the names of some available packages see e.g. the article of Pang et al. 2001). Some programs are available only as parts of entire molecular modeling packages (e.g. SYBYL, Insight II, etc.). However, the cost of these program packages is enormous and their source code and theoretical background are sometimes not well documented. Hence, the use of the latter kind of program packages was omitted in our studies. In the beginning of our work AuD was selected out from the group of freely available programs.

[#]In the literature of docking, the term "ligand" is generally used for the small molecule, which is docked to the larger "target" molecule. In many cases, the ligand is a pro(drug) and the target is a protein or an other biopolymer.

There were three main causes of our choice: (1) the parameter set of AuD is based on the parameters of the AMBER force field (see previous sections), which is known to be adequate for both peptides and proteins; (2) the possibility of handling flexible ligands (peptides); (3) the ability of AuD to search the *entire* surface of the target molecules. Moreover, AuD has a growing number of applications, which proved and validated the methods of this program package. Here, a short description is given on the theoretical aspects and working philosophy of a type 2 docking program, with a focus on AuD.

2.1 Calculation of the grid maps

The first step of the docking procedure is the generation of the potential energy: V(r) function and the store of the numerical values of each atom types of the ligand molecule in text files. This approach was first used, as early as 1985 by Goodford. During the docking simulations the algorithm simply looks up the actual energy values corresponding to a ligand position in this numerical V (r) representation, which makes the navigation fast and effective. The stored energy values are generally called "grid map". The search algorithm uses the values of the grid map (the energy landscape of the protein) as "navigation points" for the ligand.

The molecular mechanics terms of van der Waals interactions, hydrogen bonding and electrostatic interactions of the general molecular mechanics force field (see Section 1.1 for the equation) are parts of the master equation of the free energy of binding (in the 3.0 version of AuD):

$$\Delta G = \Delta G_{vdw} \sum_{i,j} \left(A_{ij} / r_{ij}^{12} - B_{ij} / r_{ij}^{6} \right) + \Delta G_{hbond} \sum_{i,j} E(t) \left(C_{ij} / r_{ij}^{12} - D_{ij} / r_{ij}^{10} \right) + \Delta G_{elec} \sum_{i,j} q_j q_i / \varepsilon(r_{ij}) r_{ij} +$$

$$+ \Delta G_{tor} N_{tor} + \Delta G_{sol} \sum_{i,j} (S_i V_j + S_j V_i) e^{(-r_{ij}^2 / 2\sigma^2)}$$

The last two entropy terms were used in order to estimate the ΔG contribution from the restricted conformational degrees of freedom and the effect of desolvation. See the original paper of Morris et al. (1998) for details of the master equation.

2.2 Docking simulation

After obtaining the grid maps the ligand can be docked by stochastic methods, like: Monte Carlo simulated annealing or genetic algorithm (both of them are implanted in AuD, see Section 1.2 for some details). The docking simulation involves an energy minimization process along location, orientation and conformation search. Generally, a docking run is repeated several times to form a job. A job can be statistically evaluated at the last step of the work.

BACKGROUND

This thesis is based on the novel applications of the program package: AutoDock. Three of the papers (PAPER II-IV) concern with binding site search on the β-amyloid peptide of Alzheimer's disease. Here, a short background of the role of this peptide and also that of β-secretase (PAPER I) is given. Paper V deals with the methodological background of *blind* binding site search of peptides on proteins. Hence, selected applications in this filed are also mentioned in the next sections.

1 Some (molecular) aspects of Alzheimer's disease

Approximately 15 percent of the people who live to the age of 65 is affected by some form of dementia; by age 85, that portion increases to at least 35 percent. The most common of all the dementias is AD. Existing in two forms, early onset familial disease and late onset disease. Four million Americans currently suffer from the condition, and experts estimate that 22 million people around the world will be so afflicted by 2025. The peptide. 39-43 amino acid beta-amyloid, which is found in the plaques characteristic of AD, is now generally accepted marker and possibly etiological factors of AD (Selkoe, D.J. 1991, Hardy, J. 1997). The Aβ molecules readily form toxic fibrils that are thought to cause neuronal damage by altering calcium regulation, mitochondrial damage and/or immune stimulation.

There are several studies published on the structure of the Aβ peptide and its fragments. It was shown, that equilibrium exists between the helical, random coil and β-sheeted forms of this peptide (Terzi at al. 1994, Zhang and Rich 1997, Jarvet et al. 2000, Massi et al. 2001). The plaques consist of mainly the β-sheeted form of Aβ (Hilbich et al. 1991, Barrow et al. 1992, Lee et al. 1995, Benzinger et al. 1998). Although, both experimental (e.g. Sunde et al. 1997) and theoretical (e.g. Massi and Straub 2001) approaches were published, the exact structure of the fibrils and plaques of Aβ is not yet solved at atomic level. Reviews on this topic were published e.g. by Serpell (2000) and by Roher et al (2000).

In 1999 two groups independently developed atomic models of Aβ aggregates. Both of them based their models on the data available from the literature: the Aβ molecules adopted

antiparallel β-sheet conformation with a type I $G^{25}SNK^{28}$ turn in these studies. While Li et al. (1999) modeled the 12-42 part of the Aβ, George and Howlett (1999) used the $A\beta_{1-43}$ form of the peptide in their study (Fig. B1-1). As the crystallization of Aβ seems to be hopeless this time, such models are of great importance for drug (inhibitor) design.

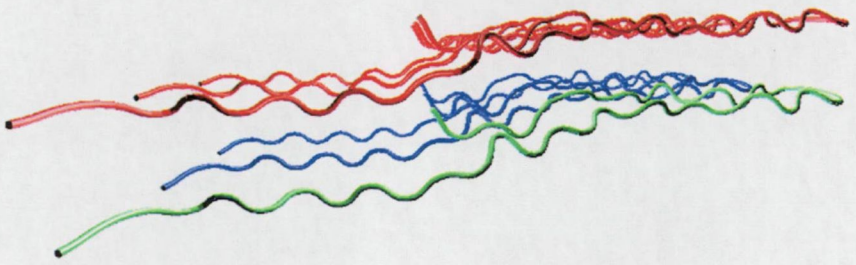

Fig. B1-1 *Two layers of the "solid state model" of amyloid fibrils of George and Howlett (1999). (One frontal hairpin-like Aβ molecule is highlighted in green)*

There are several approaches of the discovery of inhibitors of amyloid β-peptide polymerization (Findeis 2000). One branch of the aggregation (polymerization) inhibitors of Aβ is called "beta sheet breaker" peptides (Soto 2000). The *in vitro* effect of the first BSB was published by Tjernberg et al. (1996) and had the amino acid sequence: KLVFF. The original KLVFF sequence of Tjernberg is still applied as a template in recent studies (Gordon et al. 2001, Lowe et al. 2001). Soto et al. (1998) developed a similar (LPFFD) pentapeptide and demonstrated its effect *in vivo*, as well. [Interestingly, the molecular basis (change in molecular conformation) of AD and prion disease is partially common. Soto and his coworkers use the same BSB idea for the drug design against prion disease too (Soto et al. 1999, Soto et al. 2000).]

Aβ is indeed a cleavage product of the amyloid precursor protein. The enzyme β-secretase was shown to be responsible for the rate-limiting step of the cleavage of APP (Vassar et al. 1999). Thus, the design of inhibitors for β-secretase is also a real alternative of the drug development against AD (Ghosh et al. 2000, Turner et al. 2001).

2 Papers on flexible docking and binding site search

There are several applications of AuD for the reproduction of binding modes of smaller molecules of any type (e.g. haptens, HIV-1 integrase inhibitors, DNP-amino acids, etc.) in known binding pockets of proteins (e.g. Sotriffer et al. 1999, Sotriffer et al. 2000a, Sotriffer et al. 2000b, etc.). Moreover, both version of AuD was proved to be efficient in the docking of the highly flexible trisaccharides (Coutinho et al. 1998, Rockey et al. 2000), as well.

Peptides form a similar, computationally as problematic group of ligands as that of e.g. trisaccharides. A memorable and often cited paper of peptide docking is that of Stoddard and Koshland (1992). They published a novel "binary" approach for docking larger peptide fragments to proteins. The backbone conformation of the docked (ligand) peptide was the same as in the crystal structure and kept rigid during their simulation. Similar approach was used in the study of Neurath et al. (1996). However, in many cases the exact conformation of a ligand (peptide)-protein complex is not *blind* known, and therefore such "fixed conformation" simplifications cannot help in all situations.

Klepeis et al. (1998) pointed to the fact in their study, that rigid docking methods of peptides to proteins have many drawbacks. (There are algorithms, which use a complementary shape in the first step of docking.) However, the authors conclude that the use of an energy function-based algorithm is more straightforward for peptides as the former ones.

As peptides can be easily split into smaller fragments of amino acids, it is fairly plausible to dock the smaller fragments of the peptide to the pocket (Friedman et al. 1994). Convincing results of FFLD (Fragment-Based Flexible Ligand Docking) were published last year (Budin et al. 2001). However, the fragmentation is an intrinsic limitation of FFLD method since the number and fitness of the fragments depends on the definitions made by the user. Generally, one should be very careful by setting the size of the fragments as a too short fragment could lead to inadequate docking results (see PAPER V and Results for details).

Theoretically, the most exact approach of peptide docking is the one named "soft docking". During a soft docking simulation the rotations of the peptide are released (flexible ligand). Despite of the biological importance of peptide-protein complexes, relatively few soft docked peptide-protein complexes are published. Two examples are mentioned here to show the power of the flexible docking techniques. Monte Carlo soft docking of a heptapeptide was successfully done by a two-step method (Caflisch et al 1992). An octapeptide sequence was docked by Stigler et al. (1999) to mimic an interaction with an antibody.

However, the explicit handling of ligand flexibility is impossible for very large peptide or protein ligands. Soft docking is applied in an implicit manner in the case of e.g. protein-protein interactions (Palma et al. 2000, Morelli et al. 2001). This approach has the advantage of performing *ab initio* (*blind*) docking search of the binding sites on the macromolecular surface. *Blind* mapping of the binding sites of small molecules on proteins is not problematic. According to a recent paper (Dennis et al.) "consensus sites" of different organic solvents can be found with high accuracy on protein targets. Fortunately, the size of a synthetic drug-like compound to be modeled is usually much smaller than that of proteins or oligopeptides. Hence, AuD or its parallelized new version (Thormann and Pons 2001) could be an appropriate tool for detection of binding sites of flexible lead compounds on proteins.

METHODS

Molecular dynamics. Molecular dynamics studies were performed (PAPER II) to investigate the structure of two peptide amides in water. The TINKER (Pappu et al. 1998) program package was used for both MD simulations and MM optimizations. AMBER force field (Cornell et al. 1995) was applied during the calculations. The implicit GB/SA water model (Still et al. 1990) was used to estimate the solvent effects. One hundred simulated annealing MD runs were performed. Twenty picoseconds equilibration at 1050 K and 25 ps cooling till 50 K were applied by exponential protocol. The length of the time steps was 0.5 fs. The last sample was optimized with 0.001 kcal/mol tolerance. The best energy structure of 100 runs was chosen for evaluation.

Docking. *Preparation of the ligand molecules.* The peptide molecules were built and optimized by the modules of TINKER. In some cases, the original crystallographic conformation was used and decorated with polar hydrogen atoms. For the calculation of electrostatic interactions, the molecules were equipped with partial atomic charges on each atom. Either AMBER or Gasteiger-Marsili (Gasteiger and Marsili 1980)charges were used. (The latter method uses iterative partial equalization of orbital electronegativity for obtaining charges.) For the docking procedure the active torsions of the ligands were defined in the pdbq file. This file was created by the Autotors utility of AuD. The charges of the non-polar hydrogen atoms were united with that of the neighboring carbon atoms, and the coordinates of H-s were eliminated.

Preparation of the target molecules. The coordinates of the target protein molecules were downloaded from the Brookhaven Protein Databank. All of the water molecules, counter ions, ligands, etc. were removed (except of some cases, see Results for details). Polar (essential) hydrogen atoms were attached and Kollman united atom charges were assigned to the molecule. Solvation parameters and atomic fragmental volumes were attached to the end of the row of each atom in the pdbqs file of the target (in case of AutoDock ver. 3.0).

Grid map. The sizes of grid maps were set to cover either the binding pocket (PAPER I) or the whole surface of the targeted molecules (PAPER II-V). The grid spacing was either 0.375 Å or 0.550Å, depending on the map size. Lennard-Jones 12-6 and 12-10 parameters were applied for modeling non-bonded interactions and hydrogen bonds, respectively. The grid values of the electrostatic potential were calculated with the use the distance dependent dielectric approach of Mehler and Solmajer (1991).

Docking simulation. The minimization was performed either by the MC/SA (PAPER II) or by the LGA (PAPERS I, III-V) methods. The parameters of the algorithms were set as default, or modified and in one case systematically changed (PAPER V). The torsions of each ligands were released or partially fixed depending on the docking problem. The amide bonds were kept rigid in all cases. The starting position, orientation and conformation of the ligands were set randomly in all cases. Thus the ligands could find their positions either in the specific binding pocket (PAPER I) or on the entire surface (PAPERS II-V) of the proteins. Ten (PAPER I), one hundred (PAPERS II-IV) or several hundreds (PAPER V) of docking runs were performed during the docking jobs.

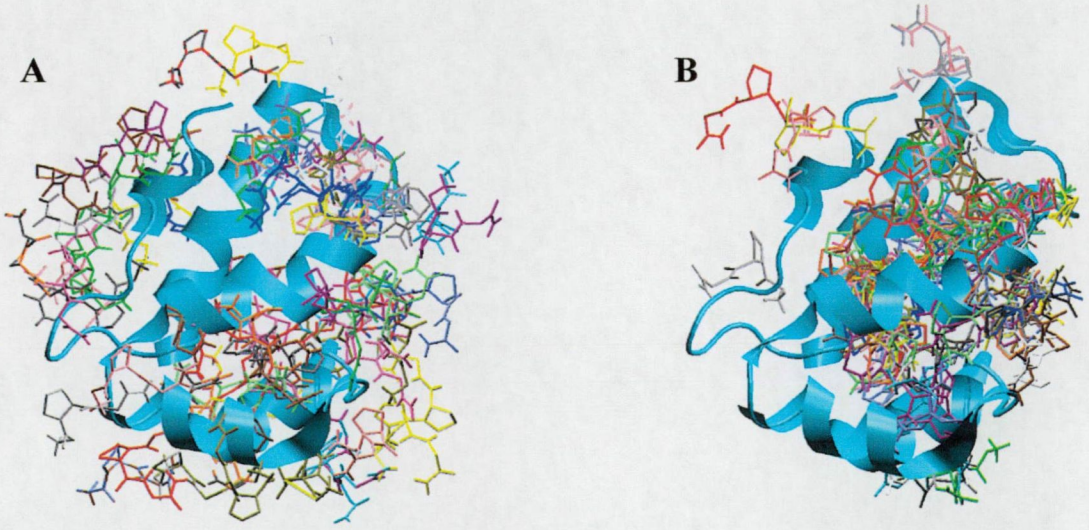

Fig. M-1 *The effect of the LGA-minimization during docking simulations.* ***A*** *The results of 100 very short docking runs (one docking job) of GPG® molecule (ligand) on the entire surface of the C-terminal part of p24 HIV capsid protein.* ***B*** *The results of 1000 times longer docking runs for the same system. [The capsid protein is a putative target of GPG® (van der Spoel et al. 2002)]*

The latter type of docking simulation is named "blind docking" in this thesis and in PAPER V. A typical example of *blind docking* simulations is shown on Fig. M-1. On part **A** of the figure, the "initial state" of a blind docking job is represented by a job of very short docking runs (200 LGA generations), whereas on part **B** the "final state" is depicted (200000 LGA generations). At the final state, the initially quasi randomly "disordered" ligands are ordered

into some branches on the surface of the protein. This figure is a graphical representation of the effect of the LGA minimization during docking runs.

Evaluation of the results. In some cases the energy minima of docking jobs were used for evaluation (PAPERS I, II). At PAPERS III-V home made C programs were used for evaluation of the docking results. The methodology is described in PAPER V in details. Briefly, after ordering the runs of a job by their energy, classification was made due to the position of the ligands. (The original clustering algorithm of AuD was not useful for our purpose, as it needs a reference structure of the ligand. Such reference ligand structure is not always available. Our *blind* approach produce classes of ligands which corresponds to the possible binding sites on the target.)

RESULTS AND DISCUSSION

PAPER I: Docking approach of inhibitor design for β-secretase of Alzheimer's disease.

Although, the main scope of this thesis is *blind* docking, in PAPER I a sample of "directed dockings" of a peptide is presented to show an example of the routine applications of AuD. Moreover, some difficulties and limitations of soft docking of peptides are outlined in this study.

In the BACKGROUND section a short summary of the role of β-secretase in AD was given. Here, we focus on the modeling of the inhibitor: Glu-Val-Asn-Leu-(Ψ)Ala-Ala-Glu-Phe (OM99-2) and its analogues. OM99-2 was designed using the SEVNL/DAEFR sequence of the cleavage site of Swedish mutant APP as a template. Asp of the site was changed to Ala and the peptide bond between P_1 and P_1' sites by a hydroxyethylene transition-state isostere in OM99-2 (Ghosh et al. 2000). Fortunately, the complex structure of β-secretase and OM99-2 is known (Hong et al. 2000) and provides a good basis for inhibitor design and modeling studies.

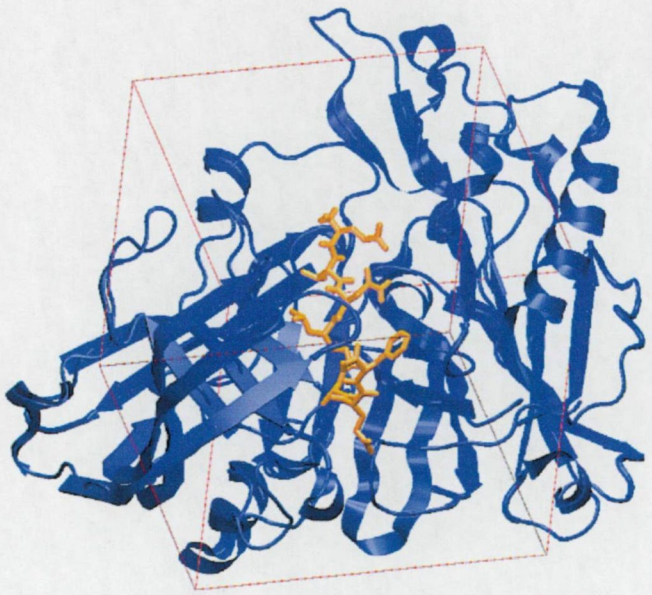

Fig. PI-1 *The complex of β-secretase and OM99-2 according to Hong et al. (2000). The border of the grid map used in our docking studies is presented with a red box.*

The octapeptide analogue OM99-2 molecule has several (31) possible torsional degrees of freedom (if that of amide bonds are not taken into account). Release of all of these torsions was proven to be impossible: no good quality docked structures were obtained. Hence, the backbone conformation of the molecule was kept rigid during the docking runs and good fit to the crystallographic position was obtained.

In the first step of our investigations the Glu(1), Glu(7) and Phe(8) amino acids of OM99-2 were removed and the Ace and NMe-blocked shorter core region was docked to the same binding pocket as the one depicted on Fig. PI-1. The docking procedure resulted a better fit to the crystallographic structure of OM99-2, even if all the torsions were released (Fig. 2 of PAPER I). These results pointed to the possibility of using AuD for the investigation of other inhibitor molecules with unknown crystal structures.

A series of inhibitor molecules of β-secretase and their K_i values had been published earlier (Ghosh et al. 2001). In the second step of the study three of the inhibitors were selected and docked to the same site as of OM99-2. Interestingly, docking with released torsions resulted in lower energy values then those with non-released backbones. The orientations of all the three molecules were the same as that of OM99-2. The lack of the C-terminal Glu and Phe resulted in the deficiency of these amino acids at their former positions (Figure 3 of PAPER I). These results are in good agreement with the rational considerations. The C-terminal part of the backbone of the inhibitors has good fit to the crystallographic structure, whereas the Boc-protected amino terminal part has (naturally) larger deviation due to the larger change of the structure of those parts of the inhibitors. Generally, the backbone H-bonding pattern and the hydrophobic interactions with the β-secretase binding pockets were conserved in these inhibitors, as well.

The studies of PAPER I focus on the reproduction and prediction of binding positions of ligands in their known binding pockets. In the case of peptides, even these routine applications of AuD are not self-evident because of the many degrees of torsional freedom. Due to the present results, it is realistic to model the binding of β-secretase inhibitors to their site. However, re-docking of present docking results and post-docking MD will be made soon so as to make a refinement on the structure of the calculated complexes and perform an exact ΔG calculation.

PAPER II: Computational studies on the binding of β-sheet breaker (BSB) peptides on amyloid βA(1-42).

MC/SA docking simulations were performed using simple target models of Aβ peptide. The backbone conformations of the target molecules were generated in βsheet by SYBYL. The docking of the GVVIAn peptide was done both for the $A\beta_{6-34}$ and $A\beta_{30-42}$ piece of the amyloid peptide, whereas the $A\beta_{6-34}$ fragment was used for the others (Tjernberg's KLVFF, Soto's LPFFD and our FFVLDGn). All peptides except of FFVLDGn showed good beta sheet breaker properties in the experiments (Tjernberg et al. 1996, Soto et al. 1998, PAPER IV). However, at that time only the experimental tests of KLVFF and LPFFD were published. FFVLDGn proved to be insoluble in our later experiments. This finding could be due to the strong intramolecular H-bridges (MD simulations and docking results, Figs. 5 and 6 of PAPER II) and the high hydrophobicity of this compound.

During the evaluation of the docking results the energy minima of the jobs were collected and featured. All the energy values of the docking jobs were listed in Table 1 of PAPER II. According to the energy values Tjernberg's KLVFF and its amide form KLVFFn showed the highest affinity to Aβ. (It should be remarked, that only the runs resulting in structures of intermolecular energy lower than 5 kcal/mol were used for average calculation and considered "successful". The average values of E_{inter} followed the trend of that of the minimum values of E_{inter}. Interestingly, Tjernberg's peptide was involved in the largest number (30-45/100) of successful runs, while the population was ca. 20/100 in the other cases.)

The ionic form of KLVFF docked to the region of E(11)...K(16) of the target, while the KLVFFn peptide amide was positioned by the F(19)...D(23) fragment of Aβ. Both associate are stabilized by salt bridges. In the KLVFFn complex hydrophobic interactions also occurs. Soto's LPFFD peptide was associated to the Y(10)...H(13) part of Aβ. Mostly ionic interactions (between the positively charged N terminal of the ligand and the carboxylate of the side-chains of the target) and hydrophobic interactions occurred.

The MD simulations resulted in a bend structure of the GVVIAn molecule. An intermolecular hydrogen bridge stabilized this conformation.

The E(22)...V(24) and the G(37)...V(39) regions of Aβ proved to be important in the binding of GVVIAn. The former complex had higher interaction energies. Besides the salt

bridges and H-bonds, the hydrophobic interactions of the side-chains stabilized the complexes. (In PAPER IV a more detailed discussion is given on the role of this peptide amide and the C-terminal fragment of Aβ.)

In conclusion, PAPER II was the first approach to find the binding sites of different potent peptides and peptide amides on the fragments of Aβ. Mostly, the hydrophobic cores of the amyloid peptide were identified as active regions. After finishing the studies of PAPER II, a new version (3.0) of AuD became available. Therefore, we decided to repeat all the docking simulation in our next papers using the new version of the program package.

PAPER III: Mapping of the possible binding sequences of two beta-sheet breaker peptides on beta amyloid peptide of Alzheimer's disease.

There are several regions of Aβ, which could be crucial for the aggregation of the amyloid peptide molecules into fibrils and plaques (Pike et al. 1995, Festy et al. 2001). The effect of hydrophobic sequences on aggregation was reported to be essential in independent papers (Hilbich et al. 1991, Roher et al. 2000).

The 3.0 version of AuD was used for the investigations of PAPER III. This version applies Lamarckian genetic algorithm (see THEORY for details) during the docking simulation. LGA was reported more efficient than MC/SA algorithm of the previous version (Morris et al., 1998). Additionally, the master equation of the scoring function was also changed. Entropy contributions were added to the ΔG function.

By the time of these studies a new model of the amyloid fibrils was available (George and Howlett, 1999). The model is discussed in details in the BACKGROUND section. The monomer Aβ of this model and also a dimeric piece of the structure of the fibril were used as target molecules in our docking experiments.

In PAPER III our investigations were focused on Tjernberg's KLVFF and Soto's LPFFD peptides. One hundred docking runs were performed in each docking jobs of the peptides. Grid maps with different grid spacings were used so as to test the reliability of the results. AAAAA and AAVFA (Tjernberg et al. 1996) "blank peptides" were also docked to the same targets to evaluate the specificity of the results.

Docking to the monomeric Aβ. Due to the docking results, three main sequences are important in the binding of both KLVFF and LPFFD. (1) The V(18)...D(23) sequence was

reported in several publications as a potent active region of the amyloid peptide. Additionally, it was shown, that this fragment of the amyloid peptide has high affinity to itself (Balbach et al. 2000, Watanabe et al. 2001). Good agreement was found between the NMR structure of Balbach et al. (2000) and the energetically favored structure of our study. (2) The R(5)...Y(10) sequence proved to be very important especially in the binding of the LPFFD peptide. Several H-bonds have stabilizing effect on this complex. (3)The G(38)...T(43) fragment of Aβ has also high affinity to the BSBs. The hydrophobic interactions between the side-chains have specific effect in this case. (In PAPER IV a more detailed discussion is given on the role of the end-sequences of Aβ.)

According to the comparison of the energy values of the BSBs and the blank peptides, our results proved to be reliable. The lack of side-chains always resulted in significantly higher energies. (Post-docking MD studies were also started for further verification of the results.)

Docking to the dimer Aβ. Dimeric unit of Aβ is sometimes considered as a basic building block of the fibrils (Roher et al. 2000). However, it is also a characteristic and – as a docking target – tractable unit of the model of George and Howlett. It is an interesting possibility to investigate whether the same or similar binding positions will be found by KLVFF and LPFFD on the amyloid molecule on the dimeric target. There could be several reasons why the affinity of the BSBs to a certain fragment of the target changes after dimerisation; e.g. sterical effect (elimination or creation of binding pockets), electronic effects (shielding of the charges of one another), just to name a few. The resulted complexes were divided into two groups according to the position of the BSBs on the plane of the dimer. In PAPER III a detailed description is given about the important sequences of Aβ found in these complexes. The D(7)...Y(10) and K(16)...E(22) parts of the amyloid remained important in the cases, where the BSBs associated to the Aβ in its plane (Group 1). Out-of-plane positions of KLVFF and LPFFD were found at the N- and C-terminal end of the target (Group 2). These findings point to the possible role of the BSBs in the breaking of inter-plane associations of the larger amyloid aggregates (sheets).

In conclusion, our *blind* docking simulations were able to find the active sequences on the amyloid target. The modeling results are in good agreement with the experimental findings and serve as realistic model of the peptide-peptide interactions. The time-evolution of the most important complexes in explicit solvent model is under investigation and the results of the detailed MD simulations will be published soon.

PAPER IV: Pentapeptide amides interfere with the aggregation of beta amyloid peptide of Alzheimer's disease.

The hydrophobic $A\beta_{29-42}$ segment of the amyloid peptide exists as an oligomeric β-sheet in solution regardless of differences in solvent, pH or temperature (Barrow and Zagorski 1991) and the last three amino acids have a particularly important role in the plaque formation (Barrow et al. 1992). Moreover, kinetic studies of fibril formation (Jarrett et al. 1993, Snyder et al. 1994) emphasize the role of the last residues in the rate-limiting step of aggregation of Aβ and the results were concluded in a "seeding hypothesis" (Jarrett and Lansbury 1993). Fortunately, a structural model was also published on the role of the C-terminal fragment of the amyloid peptide (Lansbury et al. 1995).

The above-mentioned facts of the literature encouraged us to develop small peptide inhibitors using the C-terminal part of the amyloid as a template. Two peptide amides GVVIAn and RVVIAn were constructed. Arg was inserted instead of Gly to strengthen the ionic interactions at the N-terminus and also to improve the solubility of the peptide amide. Biochemical and FT-IR measurements were used to prove the effect of the pentapeptide amides on the cells and the aggregates (PAPER IV).

During the docking simulations and the evaluations of the results of PAPER IV the same methods were used as in PAPER III. The results of the calculations, i.e. the properties of the most important complexes are summarized in Table PIV-1.

Table PIV-1 *The results of the docking simulations of GVVIAn on Aβ of AD*

Letter code of interaction	Interacting residues	Type of interaction*	Distance of N..O atoms (Å)	Angle (°)**
	GVVIAn – V(18)FFAED(23) complex			
a	G(1)...D(23)	SB	2.77	-
b	G(1)... E(22)	HB	3.22	13.02
c	V(3)... F(20)	HB	2.65	7.35
d	V(3)... F(20)	HB	2.90	46.60
e	A(5)...V(18)	HB	2.63	17.49
	GVVIAn – V(39)VIA(42) complex			
a	G(1)...I(41)	HB	2.88	28.06
b	V(2)...I(41)	HB	3.10	22.23
c	V(2)...I(41)	HB	2.66	14.51
d	A(5)...V(39)	HB	2.88	36.34

* SB: salt bridge (electrostatic interaction); HB: hydrogen bond; ** The angle (acceptor-donor-H atom) of the hydrogen bond

There are two fragments which proved to be important in the binding of GVVIAn: F(19)...D(23) and V(39)...A(42). Both sequences were also important in the binding of other BSB peptides (PAPERS II, III). Besides the salt bridges and hydrogen bondings (Table PIV-1) the hydrophobic interactions were also measurable at the present complexes.

Interactions between the hydrophobic side-chains of the same residues as in our modeling results were mentioned in the study of Lansbury et al. (1995).

The binding postions of RVVIAn on the Aβ were similar to those of GVVIAn. However, Arg was found to be involved in the interactions (Fig. PIV-1) considered during the design (see sections above).

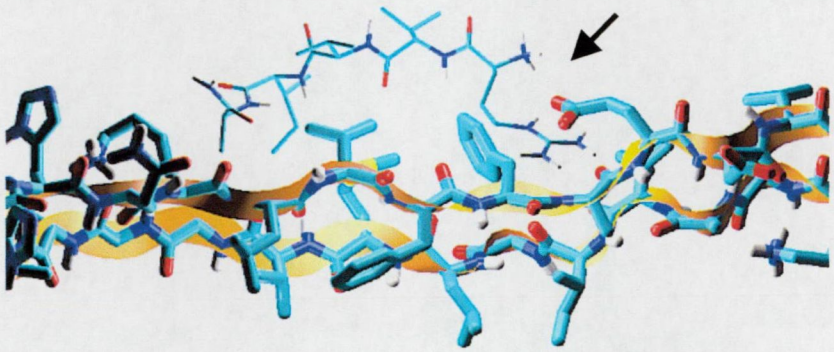

Fig. PIV-1 *The complex of RVVIAn (narrow lines) and the aromatic hydrophobic region of the amyloid peptide (marked with orange ribbon). The salt bridge between Arg, the positive N-termiuns of RVVIAn and Glu(22) of Aβ is marked with arrow*

The GVVIAn peptide was arranged in β-sheet at the amyloid peptide (Fig. 3 of PAPER IV). This finding is in good agreement with the FT-IR results.

In conclusion, the effects and possible mode of the action of two potent BSB pentapaptide amides were presented in PAPER IV.

PAPER V: Efficient docking of peptides to proteins without prior knowledge of the binding site.

In our daily practice the knowledge of the exact binding position of a ligand on a target molecule is often required. However, in many cases this practical wish cannot be satisfied

by X-ray crystallography or other sophisticated techniques or the solution requires enormous amount of time (see also INTRODUCTION). As our results (PAPER I-IV) and also the rising number of AuD (Homepage of AuD) and other docking applications (e.g. Homepage of Tripos Inc.) pointed to the usefulness of docking simulations, we decided to make a systematic study on the capabilities of docking. Our main interest focused on binding site search. Unfortunately, in our daily problems atomic details of the interactions of the ligand and the target were not available. Thus, we used the docking algorithm for the entire surface of our target proteins and checked if the original crystal structure of the control complexes could have been reproduced or not. Our present application of AuD is the first ***systematic*** binding site search of ***totally flexible*** small molecules on the ***entire surface*** of protein targets. We named our approach: ***blind docking*** so as to simplify the terminology of the future discussions about the method.

The following protein-peptide systems were chosen for the study (Table PV-1):

Table PV-1 *The protein-peptide systems used for the docking studies*

Letter code	PDB code	Protein	# residues	Ligand	# free torsions	# heavy atoms
A	3ptb	Trypsin	229	Benzamidine	0	9
B	1ak4	cyclophilin A	165	Ace-AGP-Nme	6	21
C	3tpi	Trypsinogen	281	IV	7	16
Cw	3tpi	Trypsinogen	281	IV	7	16
D	3cpa	carboxypeptidase A	307	GY	7	17
E	8gch	γ-chymotrypsin	237	GAW	8	24
F	5sga	SG protease A	181	Ace-APY	9	28
G	1ak4	cyclophilin A	165	Ace-HAGP-NMe	9	31
H	1lna	Thermolysin	316	VK	10	17
Hw	1lna	Thermolysin	316	VK	10	17
I	5sga	SG protease A	181	Ace-PAPY	11	35
J	1sua	subtilisin BPN'	262	ALAL	12	27

(The numbers of the free torsions of the ligands are increasing from the top to the bottom.)

In the first part of our study completely rigid ligands were *blind docked* to the proteins using their original conformation of the crystal complex. These dockings served as a test of the force field. If the FF would have not been accurate, the ligands had not docked to the correct positions. Excellent similarity between the docked structures and the measured ones were obtained (Table 2 of PAPER V).

Secondly, the same peptides were *docked blindly* with all their torsions released. This part of the investigation was a real challenge of the docking procedure as it is not obvious that a (quasi)stochastic method is able to find the right minima on such a complicated energy surface, which was given in the case of flexible ligands. Our results show (Table 2 of PAPER V), that it is not nonsense to find the exact position of flexible ligands *blindly* on the entire surface of the proteins. The best energy complexes (best classes) showed very good similarity to the reference structures. In many cases, the 1st class had an exact match to the reference (see RMSD values of the Tables).

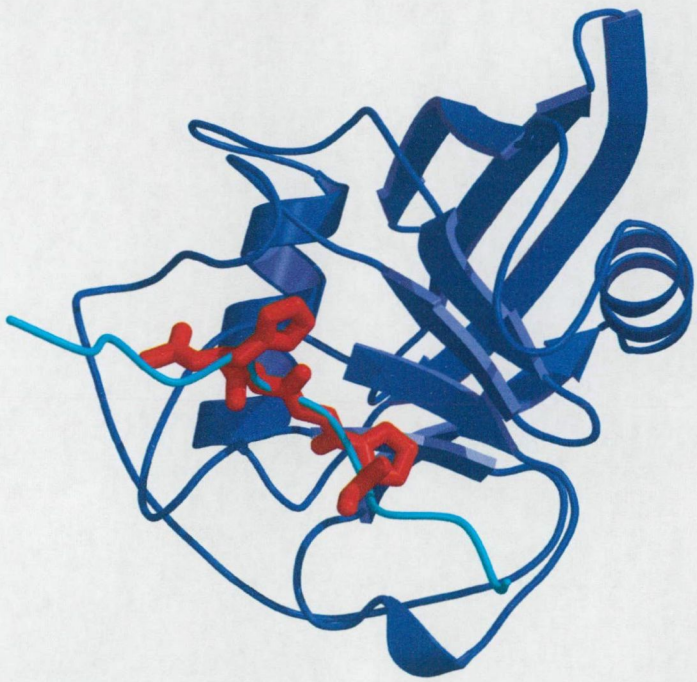

Fig. PV-1 *The complex of cypA (blue) and the N-terminal loop of p24 (green). The tetrapeptide (red) has a good match with the appropriate part of the loop, whereas the tripeptide ligand of system* ***B*** *proved to be too short to reproduce the crystal structure. (This figure was selected for cover illustration of the July, 2002 issue of Protein Science.)*

The robustness and limitations of the *blind docking* approach were also discussed in PAPER V. Docking of the problematic system **E** proved that the method is robust: if the most important part of the binding site is well-defined the method is able to maneuver the ligand to the right pocket. At the same time one should be careful while performing a fragment docking. Systems **B** and **G** is a good example for this statement since one missing residue

caused inappropriate docking results (Fig. PV-1). Very large ligands (system **I**) also cause trouble for blind docking (convergence failure).

The role of waters is a very important question of docking simulations. *Blind docking* is sometimes affected by the presence of crystal waters (systems **H** and **Hw**), but in some cases no effect of the water molecules can be concluded (systems **C** and **Cw**). Detailed investigations will be necessary on this topic to draw general conclusions.

The effect of three main LGA (docking) parameters on the docking results was investigated as a final step of the study. The ratio of successful runs of a job was almost constant with increasing number of runs. Increasing number of energy evaluations and population sizes resulted in higher efficiency (Fig. 4 of PAPER V).

As a conclusion, a section from the report of one of the referees of the manuscript of PAPER V is cited here (Anonymous 2002):

„ ...

While the scope of this paper is limited to the evaluation of a single docking code, it addresses issues that are becoming increasingly relevant as proteomics focuses on the mechanism of protein interactions and function. By evaluating the performance of one of the standard docking tools currently in use, the authors are able to explore the extent to which such methods can be utilized to answer questions that go beyond the common use of docking codes: in silico *screening of drug libraries. The results of the work are interesting and contribute to the field in two ways: 1) The authors evaluate a set of peptide receptor complexes and evaluate the docking results with regard to the nature of the complexes, and in some cases the role of bound waters; and 2) the exploration of the stochastic search parameters and their relationship to the convergence to a correct answer for a test system.*

..."

CONCLUSIONS AND PERSPECTIVES

Computational docking became one of the most powerful tools of calculation of macromolecule - ligand interactions in the last decades. The simple (MM level) approach of docking methods enables the search algorithms to find global maximums on the potential energy surface of such complexes. There are two types of search algorithms built in AuD named MC/SA and LGA. In the papers of this thesis both algorithm were used to investigate the binding of peptides on the Aβ and proteins.

In PAPER I a routine application of AuD is presented. A larger peptide was docked to a well-defined pocket. The sites of three new inhibitor molecules were also predicted within the same pocket. This paper pointed to the limitation of the LGA algorithm, as well. With 31 torsions, the OM99-2 peptide is well above the limit, which can be handled by the docking program within reasonable time limits. (It should be remarked, that the necessary time of convergence is not simply proportional, but after a limit value seems to change dramatically with the increasing number of the torsions of a ligand.) Fortunately, the real drug compounds (or lead compounds) are generally smaller than OM99-2 and therefore such a limit is rarely reached.

While PAPER I includes "simple" application of AuD, during the preparation of PAPERS II and III the main goal was the detection of the possible binding regions of known beta sheet breaker peptides on the *entire* surface of the amyloid target. As the Aβ molecule is a 39-43 AA long peptide, this is a relatively small target. In both papers *blind docking* was used (with different search methods) to select out the most important complexes. Good agreement with the available experimental data was achieved and reliable predictions were made. The results of these papers (and especially those of PAPER III) contribute to the clarification of the mode of action of the BSB peptides at atomic level and mark the attractive regions of Aβ for drug design.

Two potent BSB pentapeptide amides were presented in PAPER IV. Interestingly, similar regions were involved in the interactions of these molecules as for the other BSBs before. The possible reason of this finding is the importance of the hydrophobic regions at both the ligand and the target peptides. The positive results of PAPERS II-IV encouraged us to test our *blind docking* approach on several known protein-peptide systems.

In PAPER V, a set of peptides (ligands) with increasing size (up to tetrapeptides) and number of rotations (up to 12) were selected and proteins of moderate sizes (160-310 AAs) were used as targets. As absolutely no information was available on the capabilities of *blind docking*, these upper limits seemed to be realistic and computationally tractable. (Moreover, many compounds like the potential AIDS drug GPG® (Su et al. 2001a, Su et al. 2001b) and its possible target protein (van der Spoel 2002) fall into this range.) PAPER V includes the evaluation and detailed description of the *blind docking* approach based on the results of the above-mentioned set of test-systems.

In conclusion, our novel approach the *blind docking* was outlined in the papers of this thesis. It was shown, that the prior knowledge of the binding location of a ligand on the target is not obviously necessary in all cases. The limitations of our approach were also discussed in detail. However, these limitations are common in all applications of AuD-like docking programs, i.e. the intractable number of torsions of ligands, the rigid (and sometimes not realistic) target model, the uncertain positions of crystal waters or even the lack of crystal waters, etc. are common drawbacks of routine (PAPER I) and of blind (PAPER II-V) docking simulations. Possible solutions of these problems could be (1) the freezing of some of the torsional angles or narrowing the grid space; (2) the use of post-docking MD or QM/MM methods for refinement of the structure of the complex and involve induced effects; (3) use well-positioned crystal waters instead of the "dry" target molecule, etc. Though, an experienced modeling specialist can consider and solve the above-mentioned problems of docking simulations, the elimination of these difficulties is not yet trivial.

However, together with the afore-mentioned limitations *blind docking* is an excellent tool in selection of the most important possible binding regions of ligands on target molecules, according to our results. In combination with other techniques (e.g., MD and experimental methods: site-directed mutagenesis, affinity chromatographic techniques, etc.) it can be used for selecting the real (and only) binding site(s) of a drug on a macromolecule.

REFERENCES

Anonymous. (2002) 2nd referee's report on the manuscript of PAPER V (Protein Sci.; No. PS02-023; Hetényi and van der Spoel: Efficient docking of peptides to proteins without prior knowledge of the binding site.)

Balbach, J.J., Ishii, Y., Antzutkin, O.N., Leapman, R.D., Rizzo, N.W., Dyda, F., Reed, J. and Tycko, R. (2000) Amyloid fibril formation by $A\beta_{16-22}$, a seven-residue fragment of the Alzheimer's β-amyloid peptide, and structural characterization by solid state NMR. *Biochemistry* **39** 13748-13759.

Barrow, C.J. and Zagorski, M.G. (1991) Solution structures of β peptide and its constituent fragments: Relation to amyloid deposition. *Science* **253** 179-253.

Barrow, C.J., Yashuda, A., Kenny, P.T.M. and Zagorski, M.G. (1992) Solution conformations and aggregational properties of synthetic amyloid β-peptides of Alzheimer's disease. *J. Mol. Biol.* **225** 1075-1093.

Benzinger, T.L.S., Gregory, D.M., Burkoth, T.S., Miller-Auger, H., Lynn, D.G., Botto, R.E. and Meredith, S.C. (1998) Propagating structure of Alzheimer's β-amyloid$_{(10-35)}$ is parallel β-sheet with residues in exact register. *Proc. Natl. Acad. Sci. USA* **95** 13407-13412.

Bohm, H.J. (1994) On the use of ludi to search the fine chemicals directory for ligands of proteins of known 3-dimensional structure. *J. Comput.-Aided Mol. Des.* **8** 623-632.

Budin, N., Majeux, N. and Caflisch, A. (2001) Fragment-based flexible ligand docking by evolutionary optimization. *Biol. Chem.* **382** 1365-1372.

Budin, N., Majeux, N., Tenette-Souaille, C. and Caflisch, A. (2001) Structure-based ligand design by a build-up approach and genetic algorithm search in conformational space. *J. Comput. Chem.* **22** 1956-1970.

Caflisch, A., Niederer, P. and Anliker, M. (1992) Monte Carlo docking of oligopeptides to proteins. *PROTEINS* **13** 223-230.

Cornell, W. D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M. Jr., 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. *J. Am. Chem. Soc.* **117** 5179-5197.

Coutinho, P.M., Dowd, M. K. and Reilly, P.J. (1998) Automated docking of α-(1,4)- and α-(1,6)-linked glucosyl trisaccharides in the glucoamylase active site. *Ind. Eng. Chem. Res.* **37** 2148-2157.

Dennis, S., Körtvélyesi, T. and Vajda, S. (2002) Computational mapping identifies the binding sites of organic solvents on proteins. *Proc. Natl. Acad. Sci. USA* **99(7)** 4290-4295.

Ewing, T.J.A. and Kuntz, I.D. (1997) Critical evaluation of search algorithms for automated molecular docking and database screening. *J. Comput. Chem.* **18** 1175-1189.

Festy, F., Lins, L., Péranzi, G., Octave, J.N., Brasseur, R. and Thomas, A. (2001) Is aggregation of β-amyloid peptides a mis-functioning of a current interaction process? *Biochim. Biophys. Acta* **1546** 356-364.

Findeis, M.A. (2000) Approaches to discovery and characterization of inhibitors of amyloid β-peptide polymerization. *Biochim. Biophys. Acta* **1502** 76-84.

Friedman, A.R., Roberts, V.A. and Tainer, J.A. (1994) Predicting molecular interactions and inducible complementarity: fragment docking of Fab-peptide complexes. *PROTEINS* **20** 15-24.

Gasteiger, J. and Marsili, M. (1980) Iterative partial equalization of orbital electronegativity – A rapid access to atomic charges. *Tetrahedron* **36** 3219-3288.

George, A.R. and Howlett, D.R. (1999) Computationally derived structural models of the β-amyloid found in Alzheimer's disease plaques and the interaction with possible aggregation inhibitors. *Biopolymers* **50** 733-741.

Ghosh, A.K., Bilcer, G., Harwood, C., Kawahama, R., Shin, D., Hussain, K.A., Hong, L., Loy, J.A., Nguyen, C., Koelsch, G., Ermolieff, J. and Tang, J. (2001) Structure-based design: Potent inhibitors of human brain Memapsin 2 (β-secretase). *J. Med. Chem.* **44** 2865-2868.

Ghosh, A.K., Shin, D., Downs, D., Koelsch, G., Lin, X., Ermolieff, J. and Tang, J. (2000) Design of potent inhibitors for human brain memapsin 2 (β-secretase). *J. Am. Chem. Soc.* **122** 3522-3523.

Gillet, V.J., Newell, W. and Mata, P. (1994) SPROUT - Recent developments in the de-novo design of molecules. *J. Chem. Inf. Comput. Sci.* **34** 207-217.

Goodford, P.J. (1985) A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* **28** 849-857.

Goodsell, D.S. and Olson, A.J. Automated docking of substrates to proteins by simulated annealing. (1990) *Proteins* **8** 195-202.

Gordon, D.J., Sciarretta, K.L. and Meredith, S.C. (2001) Inhibition of β-amyloid(40) fibrillogenesis and disassembly of β-amyloid(40) fibrils by short β-amyloid congeners containing N-methyl amino acids at alternate residues. *Biochemistry* **40** 8237-8245.

Halgren, T.A. (1996) Merck molecular force field I. Basis, form, scope. parametrization, and performance of Mmff94. *J. Comp. Chem.* **17** 490-519.

Hardy, J. (1997) Amyloid, the presenillins and Alzheimer's disease. *Trends Neurosci.* **20** 154-159.

Hilbich, C., Kisters-Woike, B., Reed, J., Masters, C.L. and Beyreuther, K. (1991) Aggregation and secondary structure of synthetic amyloid βA4 peptides of Alzheimer's disease. *J. Mol. Biol.* **218** 149-163.

Homepage of AutoDock. http://www.scripps.edu/pub/olson-web/doc/autodock/references.html

Homepage of Tripos Inc. http://www.tripos.com/services/citations/index.html

Hong, L, Koelsch, G., Lin, X., Wu, S., Terzyan, S., Ghosh, A.K., Zhang, X.C. and Tang, J. (2000) Structure of the protease domain of Memapsin 2 (β-secretase) complexed with inhibitor. *Science* **290** 150-153.

Jarrett, J.T. and Lansbury, Jr., P.T. (1993) Seeding "one-dimensional crystallization" of amyloid: A pathogenic mechanism in Alzheimer's disease and scrapie? *Cell* **73** 1055-1058.

Jarrett, J.T., Berger, E.P. and Lansbury, Jr., P.T. (1993) The carboxy terminus of the β amyloid protein is critical for the seeding of amyloid formation: Implications for the pathogenesis of Alzheimer's disease. *Biochemistry* **32(18)** 4693-4697.

Jarvet, J., Damberg, P., Bodell, K., Eriksson, L.E.G. and Gräslund, A. (2000) Reversible random coil to β-sheet transition and the early stage of aggregation of the Aβ(1228) fragment from the Alzheimer peptide. *J. Am. Chem. Soc.* **122** 4261-4268.

Kaminski G. and Jorgensen, W.L. (1996) Performance of the Amber94, Mmff94 and Opls-Aa force fields for modeling organic liquids. *J. Phys. Chem.* **100** 18010-13.

Keserű György Miklós et Náray-Szabó Gábor. (1995) Molekulamechanika. (In: A kémia újabb eredményei) Akadémiai Kiadó, Budapest, Hungary.

Kirkpatrick, S., Gelatt, C.D. Jr. and Vecchi, M.P. (1983) Optimization by simulated annealing. *Science* **220** 671-680.

Klepeis, J.L., Ierapetritou, M.G. and Floudas, C.A. (1998) Protein folding and peptide docking: A molecular modeling and global optimization approach. *Comp. Chem. Engng.* **22** S3-S10.

Kuntz, I.D., Blaney, J.M., Oatley, S.J., Langridge, R. and Ferrin, T.E. (1982) A geometric approach to macromolecule-ligand interactions. *J. Mol. Biol.* **161** 269-288.

Lansbury Jr., P.T., Costa, P.R., Griffiths, J.M., Simon, E.J., Auger, M., Halverson, K.J., Kocisko, D.A., Hendsch, Z.S., Ashburn, T.T., Spencer, R.G.S., Tidor, B. and Griffin, R.G. (1995) Structural model for the β-amyloid fibril based on interstrand alignment of an antiparallel-sheet comprising a C-terminal peptide. *Nat. Struct. Biol.* **2(11)** 990-998.

Lee, J.P., Stimson, E.R., Ghilardi, J.R., Mantyh, P.W., Lu, Y.A., Felix, A.M., Llanos, W., Behbin, A., Cummings, M., Criekinge, M., Timms, W. and Maggio, J.E. (1995) ^{1}H NMR of Aβ amyloid peptide congeners in water solution. Conformational changes correlate with plaque competence. *Biochemistry* **34** 5191-5200.

Li, L., Darden, T.A., Bartolotti, L., Kominos, D. and Pedersen, L.G. (1999) An atomic model for the pleated β-sheet structure of Aβ amyloid protofilaments. *Biophys. J.* **76** 2871-2878.

Lindahl, E., Hess, B., van der Spoel, D. (2001) GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Model.* 7 306-317.

Lowe, T.L., Strzelec, A., Kiessling, L.L. and Murphy, R.R. (2001) Structure-function relationships for inhibitors of β-amyloid toxicity containing the recognition sequence KLVFF. *Biochemistry* **40** 7882-7889.

MacKerell, A.D., Bashford, D., Bellott, M., Dunbrack, R.L., Evanseck, J.D., Field, M.J., Fischer S., Gao J., Guo H., Ha S., Joseph-McCarthy D., Kuchnir L., Kuczera K., Lau F.T.K., Mattos C., Michnick S., Ngo T., Nguyen D.T., Prodhom B., Reiher W.E., Roux B., Schlenkrich M., Smith J.C., Stote R., Straub J., Watanabe M., Wiorkiewicz-Kuczera J., Yin D. and Karplus M (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem B* **102** 3586-3616.

Marrink, S.J. and Tieleman, D.P. (2001b) Molecular dynamics simulation of a lipid diamond cubic phase. *J. Am. Chem. Soc.* **123** 12383-12391.

Marrink, S.J., Lindahl, E., Edholm, O. and Mark, A.E. (2001a) Simulation of the spontaneous aggregation of phospholipids into bilayers. *J. Am. Chem. Soc.* **123** 8638-8639.

Massi, F. and Straub, J.E. (2001) Energy landscape theory for Alzheimer's amyloid β-peptide fibril elongation. *PROTEINS* **42** 217-229.

Massi, F., Peng, J.W., Lee, J.P. and Straub, J.E. (2001) Simulation study of the structure and dynamics of the Alzheimer's amyloid peptide congener in solution. *Biophys. J.* **80** 31-44.

Mehler, E.L. and Solmajer, T. (1991) Electrostatic effects in proteins: comparison of dielectric and charge models. *Protein Engng.* **4(8)** 903-910.

Miller, M.D., Kearsley, S.K., Underwood, D.J. and Sheridan, R.P. Flog - a system to select quasi-flexible ligands complementary to a receptor of known 3-dimensional structure. (1994) *J. Comput.-Aided Mol. Des.* **8** 153-174.

Moon, J.B. and Howe, J.W. (1991) Computer design of bioactive molecules – A method for receptor-based ligand design. *Proteins* **11** 314-328.

Morelli, X.J., Palma, P.N., Guerlesquin, F. and Rigby, A.C. (2001) A novel approach for assessing macromolecular complexes combining soft-docking calculations with NMR data. *Protein Sci.* **10** 2131-2137.

Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K. and Olson, A.J. (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comp. Chem.* **19**(14) 1639-1662.

Morris, G.M., Goodsell, D.S., Huey, R. and Olson, A.J. (1996) Distributed automated docking of flexible ligands to proteins: Parallel applications of AutoDock 2.4. *J. Comput.-Aided Mol. Des.* **10** 293-304.

Murray, C.W., Clark, D.E. and Byrne, D.G. (1995) PRO-LIGAND - an approach to de-novo molecular design .6. Flexible fitting in the design of peptides. *J. Comput.-Aided Mol. Des.* **9** 381-395.

Neurath, A.R., Jiang, S., Strick, N., Lin, K., Li, Y.-Y. and Debnath, A.K. (1996) Bovine β-lactoglobulin modified by 3-hydroxyphtalic anhydride blocks the CD4 cell receptor for HIV. *Nature Med.* **2**(2): 230-234.

Oshiro, C.M., Kuntz, I.D. and Dixon, J.S. (1995) Flexible ligand docking using a genetic algorithm. *J. Comput.-Aided Mol. Des.* **9** 113-130.

Palma, P.N., Krippahl, L., Wampler, J.E. and Moura, J.J.G. (2000) BIGGER: A new (soft) docking algorithm for predicting protein interactions. *PROTEINS* **39** 372-384.

Pang, Y.-P., Perola, E., Xu, K. and Prendergast, F.G. (2001) EUDOCK: A computer program for identification of drug interaction sites in macromolecules and drug leads from chemical databases. *J. Comp. Chem.* **22**(15) 1750-1771.

Pappu, R. V., Hart, R. K. and Ponder, J. W. (1998) Analysis and application of potential energy smoothing and search methods for global optimization. *J. Phys. Chem. B* **102** 9725-9742.

Pike, C.J., Walencewicz-Wasserman, A.J., Kosmoski, J., Cribbs, D.H., Glabe, C.G. and Cotman, C.W. (1995) Structure-activity analyses of β-amyloid peptides: Contributions of the β25-35 region to aggregation and neurotoxicity. *J. Neurochem.* **64(1)** 254-265.

Rockey, W.M., Laederach, A. and Reilly, P.J. (2000) Automated docking of α-(1→4)- and α-(1→6)-linked glucosyl trisaccharides and maltopentaoase into the soybean β-amylase active site. *PROTEINS* **40** 299-309.

Roher, A.E., Baudry, J., Chaney, M.O., Kuo, Y.-M., Stine, W.B. and Emmerling, M. (2000) Oligomerization and fibril assembly of the amyloid-β protein. *Biochim. Biophys. Acta* **1502** 31-43.

Roher, A.E., Baudry, J., Chaney, M.O., Kuo, Y.M., Stine, W.B. and Emmerling, M.R. (2000) Oligomerization and fibril assembly of the amyloid-β protein. *Biochim. Biophys. Acta* **1502** 31-43.

Selkoe, D.J. (1991) The molecular pathology of Alzheimer's disease. *Neuron* **6** 487-498.

Serpell, L.C. (2000) Alzheimer's amyloid fibrils: structure and assembly. (2000) *Biochim. Biophys. Acta* **1502** 16-30.

Shoichet, B.K. and Kuntz, I.D. (1993) Matching chemistry and shape in molecular docking. *Protein Eng.* **6** 723-732.

Snyder, S.W., Ladror, U.S., Wade, W.S., Wang, W.T., Barrett, L.W., Matayoshi, E.D., Huffaker, H.J., Krafft, G.A. and Holzman, T.F. (1994) Amyloid-beta aggregation: selective inhibition of aggregation in mixtures of amyloid with different chain lengths. *Biophys. J.* **67** 1216-1228.

Soto, C. (1999) Alzheimer's and prion disease as disorders of protein conformation: implications for the design of novel therapeutic approaches. *J. Mol. Med.* 77 412-418.

Soto, C., Kascsak, R. J., Saborio, G.P., Aucouturier, P., Wisniewski, T., Prelli, F., Kascsak, R., Mendez, E., Harris, D.A., Ironside, J., Tagliavini, F., Carp, R.I. and Frangione, B. (2000) *Lancet* **355** 192-197.

Soto, C., Saborio, G.P. and Permanne, B. (2000) Inhibiting the conversion of soluble amyloid-β peptide into abnormally folded amyloidogenic intermediates: relevance for Alzheimer's disease therapy. *Acta Neurol. Scand.* **Suppl. 176** 90-95.

Soto, C., Sigurdsson, E.M., Morelli, L., Kumar, R.A., Castaño, E.M. and Frangione, B. (1998) β-sheet breaker peptides inhibit fibrillogenesis in a rat brain model of amyloidosis: Implications for Alzheimer's therapy. *Nat. Med.* **4(7)** 822-826.

Sotriffer, C.A., Flader, W., Cooper, A., Rode, B.M., Linthicum, D.S.. Liedl, K.R. and Varga, J.M. (1999) Ligand binding by antibody IgE Lb4: Assessment of binding site preferences using microcalorimetry, docking and free energy simulations. *Biophys. J.* **76** 2966-2977.

Sotriffer, C.A., Flader, W., Winger, R.H., Rode, B.M., Liedl, K.R. and Varga. J.M. (2000a) Automated docking of ligands to antiobodies: Methods and applications. *Methods* **20** 280-291.

Sotriffer, C.A., Ni, H. and McCammon, J. A. (2000b) Active site binding modes of HIV-1 integrase inhibitors. *J. Med. Chem.* **43** 4109-4117.

Stigler, R.-D., Hoffmann, B., Abagyan, R. and Schneider-Mergener, J. 1999. Soft docking an L and a D peptide to an anticholera toxin antibody using internal coordinate mechanics. *Structure* 7(6) 663-670.

Still, W.C., Tempczyk, A., Hawley, R.C. and Henderickson, T. (1990) Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* **112** 6127-6129.

Stocker, U. and van Gunsteren, W.F. (2000) Molecular dynamics simulation of hen egg white lysozyme: a test of the GROMOS96 force field against nuclear magnetic resonance data. *Proteins* **40** 145-153.

Stoddard, B.L. and Koshland, D.E. (1992) Prediction of the structure of a receptor-protein complex using a binary docking method. *Nature* **358** 774-776.

Su, J., Andersson, E., Horal, P., Naghavi, M.H., Palm, A., Wu, Y.P., Eriksson, K., Jansson, M., Wigzell, H., Svennerholm, B. and Vahlne, A. (2001a) The nontoxic tripeptide glycyl-prolyl-glycine amide inhibits the replication of human immunodeficiency virus type 1. *J.Human. Virol.* **4** 1-7.

Su, J., Naghavi, M.H., Jejcic, A., Horal, P., Furuta, Y., Wu, Y.P., Li, S.L., Hall, W.W., Goobar-Larsson, L., Svennerholm, B. and Vahlne, A. (2001b) The tripeptide glycyl-prolyl-glycine amide does not affect the early steps of the human immunodeficiency virus type 1 replication. *J.Human. Virol.* **4** 8-15.

Sunde, M., Serpell, L.C., Bartlam, M., Fraser, P.E., Pepys, M.B. and Blake, C.C.F. (1997) Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J. Mol. Biol.* **273** 729-739.

Terzi, E., Hölzemann, G. and Seelig, J. (1994) Reversible random coil-β-sheet transition of the Alzheimer β-amyloid fragment (25-35). *Biochemistry* **33** 1345-1350.

Thormann, M. and Pons, M. (2001) Massive docking of flexible ligands using environmental niches in parallelized genetic algorithms. *J. Comput. Chem.* **22** 1971-1982.

Tjernberg, L.O., Näslund, J., Lindqvist, F., Johansson, J., Karlström, A.R., Thyberg, J., Terenius, L. and Nordstedt, C. (1996) Arrest of β-amyloid fibril formation by a pentapeptide ligand. *J. Biol. Chem.* **271(15)** 8545-8548.

Turner, R.T. III., Koelsch, G., Hong, L., Castenheira, P., Ghosh, A. and Tang, J. (2001) Subsite specificity of Memapsin 2 (β-secretase): Implications for inhibitor design. *Biochemistry* **40(34)** 10001-10006.

Van der Spoel, D., Hetényi, C., McKenzie-Hose, A., Ågren, L., Goobar-Larsson, L., Höglund, S. and Vahlne, A. (2002) A tripeptide-amide interferes with self-assembly of the HIV capsid protein. Submitted.

van Gunsteren, W.F. and Berendsen, H.J.C. (1990) Computer simulation of molecular dynamics: methodology, applications and perspectives in chemistry. *Angew. Chem. Int. Ed. Engl.* **29** 992-1023.

Vassar, R., Bennett, B.D., Babu-Khan, S., Kahn, S., Mendiaz, E.A., Denis, P., Teplow, D.B., Ross, S., Amarante, P., Loeloff, R., Luo, Y., Fisher, S., Fuller, J., Edenson, S., Lile, J., Jarosinski, M.A., Biere, A.L., Curran, E., Burgess, T., Louis, J.C., Collins, F., Treanor, J., Rogers, G. and Citron, M. (1999) Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. *Science* **286** 735-741.

Wang, R., Gao, Y. and Lai, L. (2000) LigBuilder: A multi-purpose program for structure-based drug design. *J. Mol. Model.* **6** 498-516.

Wang, W., Donini, O., Reyes and C.M., Kollman, P.A. (2001) Biomolecular simulations: recent developments in force fields, simulations of enzyme catalysis, protein-ligand, protein-protein, and protein-nucleic acid noncovalent interactions. *Annu. Rev. Biophys. Biomol. Struct.* **30** 211-243.

Watanabe, K., Segawa, T., Nakamura, K., Kodaka, M., Konakahara, T. and Okuno, H. (2001) Identification of the molecular interaction site of amyloid β peptide by using a fluorescence assay. *J. Peptide Res.* **58** 342-346.

Weiner, S.J., Kollman, P.A., Case, D.A., Chandra Singh, U., Ghio, C., Alagona, G., Profeta, S. Jr. and Weiner, P. (1984) A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* **106** 765-784.

Zhang, S. and Rich, A. (1997) Direct conversion of an oligopeptide from a β-sheet to an α-helix: A model for amyloid formation. *Proc. Natl. Acad. Sci. USA* **94** 23-28.

ACKNOWLEDGEMENTS

*TRIPEP AB (Huddinge, Sweden), the **Swedish Academy of Engineering Sciences**, TEMPUS Joint Research Project and the **Hungarian Academy of Sciences** are acknowledged for financial support.*

Professor Botond Penke, my supervisor helped me during my PhD studies in many ways. He introduced me to the interesting world of peptide science and pointed to the possible applications of peptides for important problems, like Alzheimer's disease and encouraged me to continue my modeling work. His friendly and human attitude was the guarantee of our pleasant and fruitful collaboration.

Dr. Tamás Körtvélyesi, my former supervisor introduced me to molecular mechanics and molecular modeling during our previous work. I am thankful to him for the suggestion of the use of highly interesting molecular docking techniques, especially AutoDock. He is the co-author of three papers included in this thesis. Tamás, I thank you also for your personal advice and help.

Dr. David van der Spoel, the best colleague, I have ever had... His positive, optimist attitude, friendship and help made my last stay in Uppsala pleasant and memorable. I hope I will have the opportunity to work together with you on our docking ideas in the future, as well. David, you are right: "***Experiments exist to verify calculations.***"

Professor Janos Hajdu is acknowledged for critically reading of the manuscript of Paper V and for his suggestions.

During my previous stays in Uppsala, I worked in the group of **Professor Stellan Hjertén**. I am thankful for the practical knowledge I obtained in his laboratory and the interesting discussions we had. **Dr. Ákos Végvári**, my friend and colleague gave me a lot of practical help during my stay in Uppsala. He introduced me in the preparation of CEC capillaries and living in Sweden. I am also thankful to **Professor Stefan Höglund**, who organized our work in the GPG project and encouraged me to continue my modeling work after the first months. **Anikó Takátsy**, my friend, helped me a lot with her presence in Uppsala and by sharing her shiny thoughts with me during our chats.

Dr. Miklós Hornyák is acknowledged for linguistic correction of the manuscript of this thesis.

My colleagues in the home department: **Zsolt Datki, Éva Klement, Gábor Paragi** and **Dr. Zoltán Tímár** are the co-authors of the papers of this thesis and hereby acknowledged for the common work.

This thesis would have never been finished without the love and help of **my parents, my brother and his family**.