MOLECULAR SIMULATIONS WITH HIGH-PERFORMANCE COMPUTERS

PIOTR ARŁUKOWICZ,[a] EWA BIERNAT, [a] JERZY CIARKOWSKI,[a] CEZARY CZAPLEWSKI,[a] MAŁGORZATA GROTH,[a] RAJMUND KAŹMIERKIEWICZ,[b] ADAM LIWO,[a] KATARZYNA MAKSIMIAK[a], MAŁGORZATA NOWACKA,[a] STANISŁAW OŁDZIEJ,[a] SYLWIA RODZIEWICZ[c] and EDYTA WOŹNIAK[a]

Faculty of Chemistry, University of Gdańsk, Sobieskiego 18, 80-952 Gdańsk, Poland

[a]http://chemik.chem.univ.gda.pl:8000/research/peptydy.html
[b]http://chemik.chem.univ.gda.pl:8000/research/hormony.html
[c]http://chemik.chem.univ.gda.pl:8000/research/bioch.html

Abstract: In this article, we present the work carried out in our group on various fields of theoretical biochemistry. Our main fields of research are as follows: i) design of an algorithm for de novo prediction of protein structure from amino-acid sequence using energetic criteria, ii) theoretical modeling of the structure and dynamics of neurohypophyseal-hormone receptors, iii) quantum-mechanical investigation of reactions in organic chemistry and biochemistry, and iv) theoretical conformational analysis of small peptides using experimental information. The performance of different machines, depending on the kind of calculations with special focus on the exploitation of parallelism as well as the applicability and performance of various commercial, free, and our home- made software (AMBER, GAUSSIAN, ECEPPAK, GAMESS, SYBYL, MSI/Biosym) available at TASK and the importance of graphical processing of the data are discussed.

"It must be emphasized that the machine does not think for itself. It may exercise some degree of judgement and discrimination, but the situations in which these are required, the criteria to be applied, and the actions to be taken according to the criteria, have all to be foreseen in the program of operating instructions furnished to the machine. Use of the machine is no substitute for thought on the basic organization of a computation, only for the labour of carrying out the details of the application of that thought."

Douglass R. Hartree, Moore School lecture, Univ. of Penn., 9 July 1946

1. Introduction

It was not long time ago that the experimental chemistry community considered theoretical branch of this field of science as something entirely unrelated and with no application to their work. This kind of thinking had the basis in the enormous complexity of systems that undergo chemical processes and must, unlike the processes in e.g. mechanics be treated at the microscopic level all the time. It takes billions arithmetic operations to solve the Schrödinger equation governing the motion of the electrons for even a small organic molecule and even after doing this one has only a faint idea, what is the electronic structure of the *isolated* molecule. So, how can one think of actually predicting the pathway of a chemical reaction in a condensed phase, which involves a practically uncountable number of chemical species? And, the more, what can be any other way, apart from experimental, of learning about the biochemical processes that involve ensembles of very complex macromolecules?

It can be said without reservation that the advent of high-speed computers opened a new era in the history of theoretical chemistry. While ten years ago it took eight hours on a R-32 to energy-minimize a single conformation of a nine-residue peptide hormone oxytocin, nowadays eight hours of computations on a single processor of the SP2 allows one, using the Monte Carlo methods to explore the conformational space of oxytocin fairly completely, obtaining around 1000 accepted energy-minimized conformations. Now experimental chemists ask their colleagues doing theoretical chemistry to predict the possible reaction pathways or the effect of point mutations on protein structure. In many cases, the accessible molecular-modelling software, such as HYPERCHEM, INSIGHT, SYBYL allows an experimentalist without particularly advanced knowledge in theory to carry out the computations. On the other hand, it must be borne in mind that the approximations involved in the methods of computational chemistry require that the results cannot be taken as they are, without any criticism. Nevertheless, a careful worker, with good background in chemistry and biochemistry, can do very well in his or her prediction of (bio)chemical processes, which can save expensive chemicals, test animals, and the time of the experimentalist, compared to situations where things need to be found out by a trial-and-error method. Moreover, the possibility of doing large-scale computations gave rise to entirely new branches of theoretical chemistry and biophysics, such as studying the mechanism of protein folding or solute-solvent interactions at the microscopic level with few approximations involved.

Our group's field of interest are the biochemical processes studied at various levels of theory, from atom-atom and mesoscopic empirical potentials to high-level quantum chemistry-depending on the kind of problem. We started about 15 years ago as a theoretical subgroup of the Division of Peptide Chemistry of our Faculty. From them on, our informal theory group has expanded to include people from the Division of Hormone Chemistry and Division of Bioorganic Chemistry and is still open to everybody in our Faculty who is interested in doing the work in varius aspects of computational chemistry or biochemistry. As more and more advanced computer

systems became accessible, the scope of our research gradually evolved from conformational analysis of small peptides to proteins and other macromolecules and from studying electronic properties of isolated organic molecules by semiempirical methods of quantum chemistry to the investigation of complex chemical reactions with the use of advanced *ab initio* methods. Our current main research topics are briefly described in the next section.

2. Description of research

2.1 Design of force fields for *de novo* prediction of protein structure and simulation of protein folding

Prediction of the spatial structures of proteins, which are key macromolecules in the biochemistry of all known living organisms, from amino-acid sequence still continues to be an unsolved and extremely challenging problem of theoretical molecular biology and biophysics. Unlike the case of ordinary polymers that are mixtures of random conformations, if the polymerization process occurs in an uncontrolled manner, each protein possesses a unique three-dimensional structure, the so-called *native structure*, stable within a range of conditions (the so-called *physiological conditions*) that determines its biological function(s). Knowledge of the three-dimensional structure of proteins is, in turn, the necessary condition to study their biological function and mechanism of action. Whereas tens of thousand of new amino-acid sequences of either wild-type or mutated protein are revealed each year, at the same time experimental methods, such as X-ray crystallography bring only about two hundred newly resolved three-dimensional structures.

One approach to predict the three-dimensional structure of proteins is to construct an appropriate (free) energy function that recognizes the native structure as the one distinctively lowest in (free) energy and then to carry out a search of the lowest-energy structure of a new protein. This approach utilizes Anfinsen's *thermodynamic hypothesis*,[1] according to which the native structure of a protein is the global minimum in its free energy surface. While simple in formulation, this approach faces two very difficult problems: finding the global minimum of a nonconvex function in hundreds of thousands of dimensions and constructing appropriate energy functions. Despite great progress that was made there during the recent years,[2-4] both problems are still far from solution and calculation of the global minimum of even a small protein ($\approx$ 40 - 50 amino-acid residues) is technically impossible, if all-atom representation of the molecule is implemented. A possible way out is the so-called hierarchical protocol of protein-structure prediction, in which extensive search of the conformational space of the polypeptide chain is carried out at a coarse-grained level, in which each amino-acid residue is represented as a few interaction sites (the so-called *united-residue* representation) and the whole atom chain is gradually reconstructed based on the global energy minimum structure of the simplified chain.[5-9] Even a longer history than the whole hierarchical protocol has the idea of

united-residue representation of polypeptide chain, which was initiated in 1975 by the pioneering work of Levitt and Warshell [10] and continues to receive great attention ever since.[5-19]

Work on united-residue model of the polypeptide chain and the associated force field, as well as the hierarchical protocol of protein-structure prediction was initiated in our group seven years ago when one of us (AL) was on a post-doctoral leave in Prof. H.A. Scheraga group in Cornell University, USA. The fruitful co-operation with Prof. Scheraga continues until the present and in addition to its scientific outcome, thanks to the massively parallel system of SP2 clusters installed at Cornell Theory Center and the parallel software, as well as virtual workshops on parallel programming it provides, we are able to increase and continuously update our knowledge of the art of massively parallel computations.

In our model,[8,9,17-20] a polypeptide chain is represented by a sequence of α-carbon (C^{α}) atoms linked by virtual bonds with attached united side chains (SC) and united peptide groups (p) located in the middle between the consecutive α-carbons. Only the united peptide groups and united side chains serve as interaction sites, the α-carbons assisting in the definition of the geometry, as illustrated in Fig. 1.

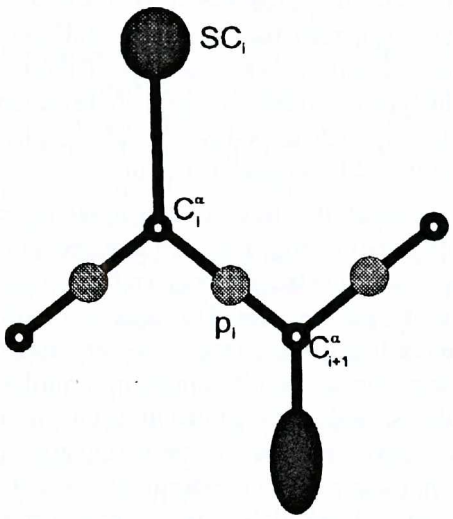


Figure 1. United residue model of the polypeptide chain

The energy of the virtual-bond chain consists of the pairwise interactions between united side chains ($U_{SC_iSC_j}$), between side chains and united peptide groups ($U_{SC_ip_j}$), between peptide groups ($U_{p_ip_j}$), local-interaction terms that describe the local energetic propensities of the united-residue chain (U_{loc}) and the correlation or multibody terms (U_{corr}) that arise from the fact that the energy of united-residue chain has the meaning of an average energy of the corresponding all-atom chain and does not, in general, retain the pairwise form, even if the parent all-atom energy function

is pairwise. The parameters of the force field were determined as statistical potentials of mean force based on an extensive analysis of all available entries in the Brookhaven Protein Data Bank (PDB). For detailed description of the model and the force field the reader is referred to original works.[8,9,17-20]

The derived force field and protocol for protein-structure prediction was successfully applied in the prediction of the native structures of simple helical proteins, such as the avian pancreatic polypeptide (APP),[9] galanin,[21] and the 10-58 fragment of staphylococcal protein A (Fig. 2).[19] The conformational space was searched using the Monte Carlo methods which enables to find the global energy minimum of simplified chains of up to 50-60 amino acid residue lengths. Our current work focuses on the improvement of the united-residue force field towards the treatment of more complicated structural motifs, including β-sheets. We are also working on the implementation of efficient global-optimization methods of smoothing the energy surface, such as the Diffusion Equation Method (DEM),[22,23] and the shift method[24] which will enable to extend the search of the conformational space to chains comprising several hundred amino-acid residues.

At the same time, we are working on the improvement of all-atom force fields, by parameterizing the interaction potential of proteins with water.[25] As mentioned above, these the so-called hydrophobic/hydrophilic interactions are largely responsible for the formation of characteristic three-dimensional structures of proteins.

In this part of our research, we use our own software for Monte Carlo simulation of the polypeptide chain in united-residue approximation. The program is massively parallelized with the use of the IBM SP2-based MPL message-passing library.

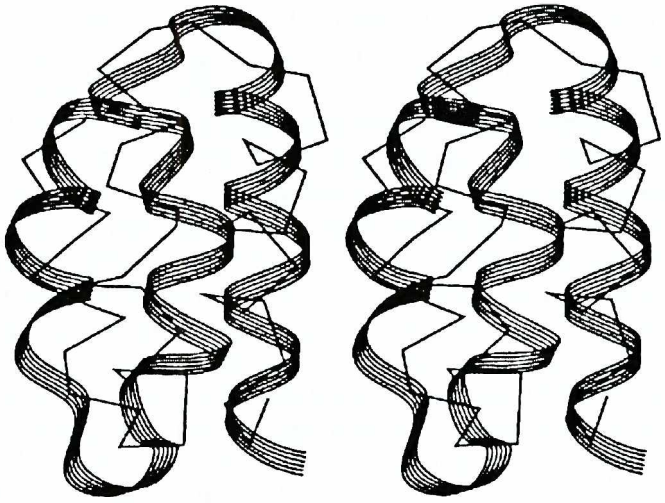


Figure 2. Predicted structure of the 10-58 fragment of protein A (sticks) superposed on the native structure (ribbon)

2.2 Modeling of the structure of neurohypophyseal hormone receptors and their interaction with ligands

Other topics under study have been neurophysin/bioligand interactions, related to the allosteric signal transmission. Neurophysins I and II (NPI and NPII) serve in the neurosecretory granules as carrier proteins for the neurophyseal hormones oxytocin (OT) and vasopressin (VP), respectively. The NPI2 and NPII2 homodimers and several (NP/ligand)2 heterotetramers were modelled from a low resolution structure, given by the C^{α}-carbon atom coordinates of the NPII/dipeptide complex (file IBN2 in the Brookhaven Protein Data Bank), by using standard modeling tools available within the SYBYL suite of programs (by Tripos, Inc.). The crude structures were refined by the use of a sophisticated protocol consisting, among other things, of a constrained simulated annealing and molecular dynamics in water, both implementing the AMBER 4.1 force field[26]. The protocol have been described in detail elsewhere[27,28] The MD simulations have indicated a presence of specific interaction paths, linking the ligand binding sites with the NP-NP interface, and thus pointing out at an allosteric communication between the ligand binding and the NPII dimerization[27-32] This area of research has also completed in a Ph.D. thesis by R. Kaźmierkiewicz[33]

Most recent area of our study concerns molecular modeling of the G protein-coupled receptors (GPCR)-bioligand complexes and interactions, underlying the most common way for conveying extracellular signals to the cytosol. Our interest has initially concentrated on the vasopressin V2 receptor (V2R)-agonist and V2R-antagonist interactions. Using the automated GPCR modeling facility available via Internet (http://expasy.hcuge.ch/swissmod/SWISSMODEL.html) for construction of the 7TM domain in accord with the bovine rhodopsin (RD) footprint, and the SYBYL software (Tripos, Inc, St.Louis) for addition of the intra- and extracellular domains, the human V2R was modelled. The structure was further refined and its conformational variability tested using a constrained simulated annealing (CSA) protocol developed in our laboratory. An inspection of the resulting structure has revealed that the V2R (likewise any GPCR modeled this way) is much thicker and accordingly forms a more spacious TM cavity than most of the hitherto GPCR constructs (typically based on the structure of bacteriophodopsin, BRD) do. Also, in our model the TM helices are arranged differently than they are in BRD-based models. Hence, the geometry of the TM cavity, potentially capable of receiving ligands, is in this model unlike those in earlier models. In the subsequent step, several ligands, including the native AVP and the non-peptide antagonist OPC-31260 were docked into the TM cavity in multiple ways and the resulting structures were equilibrated and their conformational variability tested using CSA as above. The best docking (Fig. 3) was selected and justified upon consideration of ligand-receptor interactions and structure-activity data, a number of amino acid residues were identified, mainly in TM helices 3-7, as potentially important in both AVP and antagonist docking. Most of them are invariant for either the GPCR superfamily or the neurophyseal (vasopressin V2R, VlaR and Vlbft and oxytocin OR) subfamily of receptors. Importantly,

some of the equivalent residues in VIaR have already been found critical for the ligand affinity.[34] Some of these results have been described.[35,36]

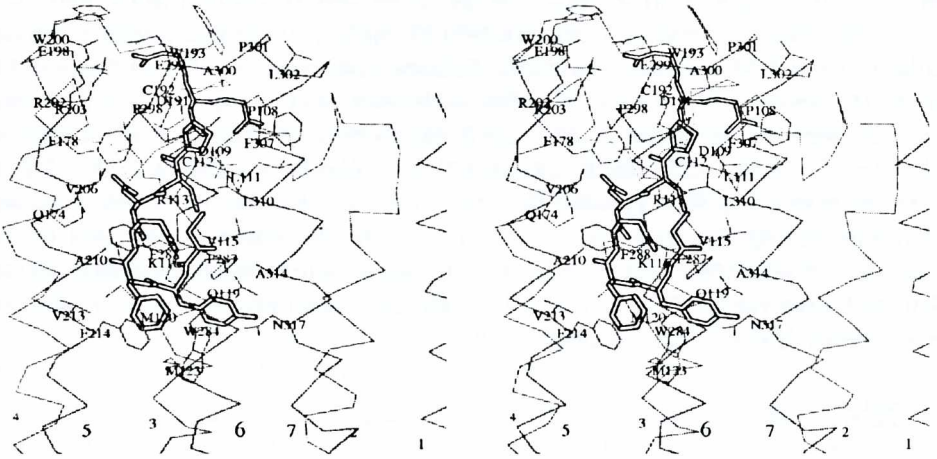


Figure 3. Extra-cellular part of the vasopressin receptor with docked vasopressin obtained in MD simulations

The complexity of the molecules analyzed requires the use of parallelism, in order to obtain the results in real time. The parallel code of the AMBER package was written for various machines, including CRAY, IBM-5P2, and Power Challenge SGI. It is interesting at this point to compare the dependence of real time on the number of processor for various architectures obtained in standard tests of AMBER (Fig. 4).

2.3 Theoretical studies of the conformation of bioactive peptides

In contrast to proteins, oligopeptides are relatively small molecules composed of a few (usually 2-30) amino-acid residues. These compounds fulfil various regulatory function in living organisms (e.g. hormones, neurotransmitters, analgesis). Oligopeptides are also the active components of plant and animal toxins, such as falloidin from death cap (*Amannita phalloides*) and apamin from honey bee and wasp. Instead of occurring in one well-defined conformation, most of oligopeptides have largely flexible structure in solution. Nevertheless, each of them is characterized by a spectrum of accessible conformations which, in turn, determines its biological activity. The small size of oligopeptides allows one to carry out a fairly complete conformational search with the computer power presently available.

We have carried out a number of extensive conformational studies, with the use of the ECEPP[37] and AMBER[26] force fields on a number of biologically active peptides, including the neurohypophyseal hormones oxytocin and vasopressin,[38] their analogs,[39-41] enkephalin analogs,[42] scyliorhinin,[43] and morphiceptin.[44] The conformational space was searched using the Electrostatically Driven Monte Carlo (EDMC)[45]

method implemented in the ECEPPAK package[46] and molecular dynamic methods implemented in the AMBER package. In all of these studies, we used the experimental information both to narrow down the number of possible conformations and to minimize the margin of error caused by the uncertainties inherent in the empirical force fields. The most common procedure is to implement the interproton distances available from NMR measurements as distance constraints in the conformational search. However, in the case of flexible molecules, such a procedure is generally incorrect, because the distances are only averages over all accessible conformation and not characteristic of a single conformation. Recently, we started to develop an algorithm that takes into account this fact. Instead of using the conformational constraints during the sampling procedure, we fit the statistical weights of the resulting conformations so as to obtain the best agreement between the measured and calculated observables. This procedure has already been implemented in conformational studies of enkephalin analogs.[47]

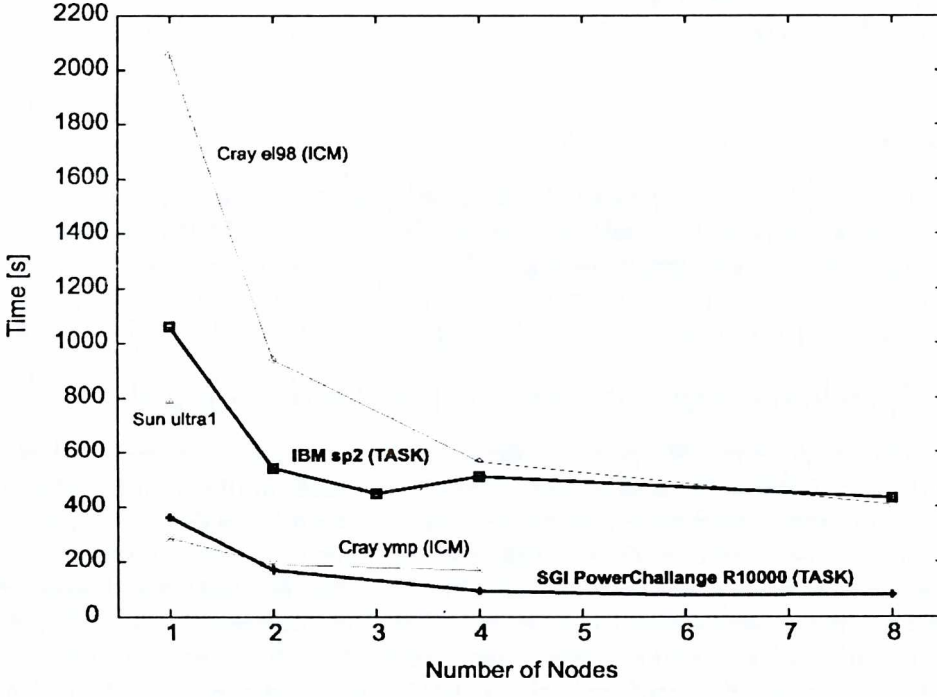


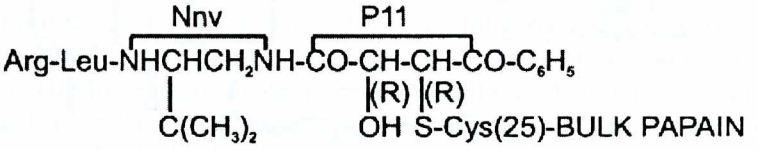
Figure 4. Dependence of the wall-clock time of a standard AMBER run (100 MD iterations at 300 K for plastocyanin; in water (a total of 11585 atoms)

2.4 Theoretical modeling of enzymatic processes

Using the AM1,[48] PM3[48] semiempirical methods, and the Density Functional Theory (DFT)[49] to the model of the protein catalytic site composed of ca. 160-190 atoms,

we have carried out studies aimed at the explanation of three aspects of the mechanism of action of asperity proteinases, concerning the site of dissociation within the catalytic diad COOH/COO^- (i) in the free enzyme, (ii) in the Michaelis complex and, (iii) concerning the energy changes associated with the catalytic paths. We have found that the state of dissociation within the catalytic diad is ligand-sensitive. In the free enzyme and in the intermediate complexes Asp33 prefers to be dissociated and the outer oxygen of Asp213 to be protonated, while in the Michaelis complex and the product complexes the opposite holds true. This is in agreement with a recent mechanistic hypotheses and with some experimental results by FTIR and NMR. The energy diagram for the catalysis indicates that electronic effects are most of all responsible for the relative reduction of energy of the intermediates (and possibly transition states) on the reaction path. The shape of the diagram qualitatively agrees with the Transition-State Analogue Theory for the enzymatic reactions. The work has been reported in several papers[50-53] and resulted with the Ph.D. thesis by St. Ołdziej.[54] We consider the subject closed for the time being.

A next subject of our interest concerns the studies on interactions between cystein proteinase papain and its inhibitors. Epoxide-type inhibitors covalently bind to the catalytic Cys25 in papain, with simultaneous opening of the -C-O-C- triangle. The inhibitor subject this work is drawn below:



Its initial 3D architecture was so modelled as to ensure the best overlap of the P11 part with the equivalent portions of the solid state structures of papain-E64c (two forms) and simultaneously the best overlap of the Leu-Nvv moiety with the equivalent portions of the papain-leupeptin (Leu-Leu) and papain-ZPA (Z-Phe) complexes (Protein Data Bank accession codes 1pe6/1ppp, 1pop and 6pad, respectively). This way, the Nnv isopropyl and the Leu residue would appear in the S2 and S3 enzymic pockets, respectively, just like successive Leu residues in leupeptin-papain and the Phe and Z residues, respectively, in ZPA- papain. The complex was submitted to constrained simulated annealing and subsequently to 200 ps molecular dynamics. The final structure, averaged over last 50 ps and minimized, did not drift far away of the starting one. Remarkably, the time-averaged fragmental protein mobility has correlated excellently with the thermal factors of the papain- leupeptin complex, which argues in favour of the force field used (AMBER 4.1). The inhibitors N-terminal Arg residue appears to be loosely fluctuating on the surface, away of the S1-S3 substrate pockets. Hence, Arg (potential P4) does not seem essential for nesting the inhibitor in the enzyme catalytic site, turning a role of an S4 site in papain elusive. This work has not been published yet.

In collaboration with Prof. Edward Borowski (Department of Pharmaceutical Technology and Biochemistry of the Technical University of Gdańsk), we also carry out theoretical studies of the mechanism of the action of glucosamine syntheses with its covalent inhibitors, mainly the derivatives of N^3-(4-methoxyfumaroyl)-L-2,3-diaminopropanoic acid (FMDP). Glucosamine syntheses is a very important target in anti-fungal drug design, because it is essential in the build-up of fungal cell walls, and therefore vital for the fungus organism. Recently, we have completed the theoretical studies of the inhibition mechanism, initiated a few years ago;[55] two papers on this subject will be published soon.[56,57]

2.5 Quantum-mechanical studies of oxygen interaction with unsaturated and aromatic compounds bearing the hydroxy groups

This part of our research is carried out collaboration with Dr. Danuta Jeziorek and Prof. Józef S. Kwiatkowski of the Institute of Physics of Nicholas Copernicus University in Toruń.

Interaction of reactive oxygen species with organic compounds play an important role in the biochemistry of all known living organisms.[58,59] The primary reactive oxygen species are singlet ($^1\Delta_g$) oxygen, superoxide anion radical ($O_2^{\bullet -}$), and the hydroperoxyl radical ($HOO^{\bullet}$). Our interest in the theoretical studies of the chemistry of reactive oxygen species was initiated by the still continued cooperation with Prof. Edward Borowski and is primarily focused on the production of superoxide anion radical on oxygen interaction with anthracenedione derivatives of that constitute an important class of antitumor drugs of anthracycline class [60,61] (Fig. 5). These species can substitute ubiquinone in quinone-reducing sites of NADH dehydrogenase, cytochrome P-450 oxidase and other redox enzymes and further mediate the transfer of one electron to the oxygen molecule, yielding the superoxide anion radical.[60,62,63] The superoxide anion radical is further converted to the hydroperoxyl and hydroxyl radicals which initiate a chain of radical reactions in lipid component of the cells, which is particularly damaging for the heart tissues, causing, in turn severe cardiotoxicity of anthracycline anti-tumour drugs.[60,61]

We investigated two possible mechanisms of superoxide production:

$$\begin{aligned} A + e^- &\rightarrow A^{\bullet -} && \text{(Ia)} \\ A^{\bullet -} + O_2 &\rightarrow AO_2^{\bullet -} && \text{(Ib)} \\ AO_2^{\bullet -} &\rightarrow A + O_2^{\bullet -} && \text{(Ic)} \end{aligned}$$

$$\begin{aligned} A + O_2\,(^1\Delta_g) &\rightarrow AO_2 && \text{(IIa)} \\ AO_2 + e^- &\rightarrow AO_2^{\bullet -} && \text{(IIb)} \\ AO_2^{\bullet -} &\rightarrow A + O_2^{\bullet -} && \text{(IIc)} \end{aligned}$$

where A denotes anthracycline derivative and AO_2 or $AO_2^{\bullet -}$ covalent or non-covalent adducts of the oxygen molecule to anthraquinone or its anion radical. Mechanism II

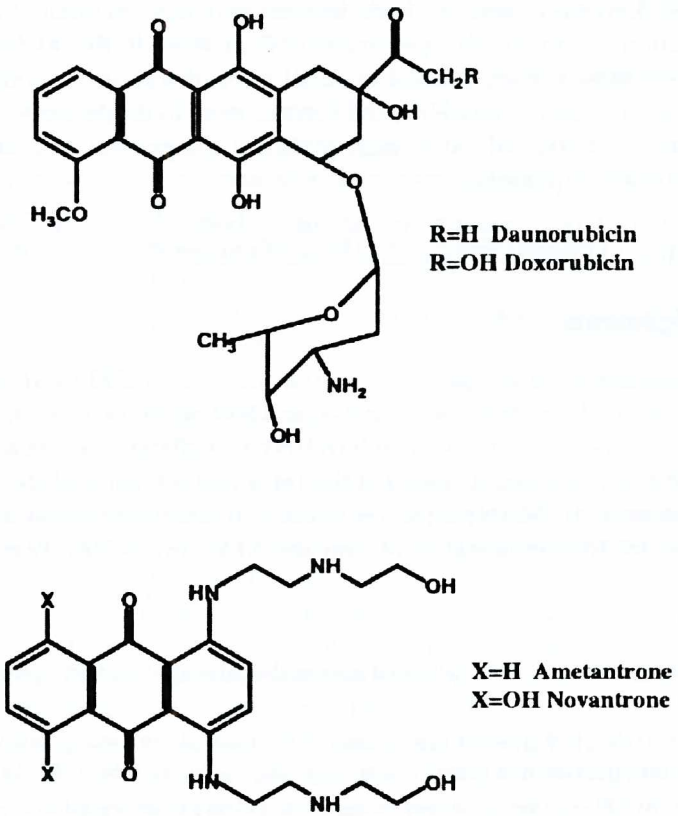


Figure 5: Structures of representative anthraquinone-based anti-cancer drugs

requires the excitation of ground-state $^3\Sigma_g$ oxygen molecule to a higher-energy $^1\Delta_g$ state; it was proposed for the first time by Tempczyk and co-workers.[64]

In a series of papers,[65-72] we described our research of the energetics of all stages of mechanism I and II and the structures of possible covalent intermediates. We concluded that the most probable way of electron-transfer mediation is mechanism I, provided that the anthraquinone derivative possesses a proton-donor group. In this case, the hydrogen bond between anthraquinone and superoxide anion radical is so strong that the conical crossing between the $A^{\bullet-} \cdots O_2$ and $A \cdots O_2^{\bullet-}$ states requires virtually no activation energy. This finding is in excellent agreement with electrochemical data on the reduction of various anthraquinone derivatives in the presence of oxygen,[65,69,70] as well as with the fact that the cardiotoxic activity of the anthraquinone-based anti-cancer drugs is reduced dramatically upon the removal of hydroxy groups.[60-63]

More recently, we extended the above-described research to the mechanism of oxygen addition to various unsaturated and aromatic compounds and the further fate of the adducts. The practical basis of this is that fact that while the amino and imino

anthraquinone derivatives apparently do interact with oxygen, their ability to superoxide generation is much less pronounced than that of the hydroxy derivatives.[65,69,70,73] We have already studied in detail the pathways of a model reaction of oxygen addition to ethene, involving the formation of hydroperoxide aldehyde and 1,2-dioxethane.[71] Studies of other than hydroxy compounds and other types of adducts are already in progress.

In this part of our research, we are using both *ab initio* and semiempirical methods, such as GAMESS,[74] Gaussian,[75] and MOPAC.[48]

Acknowledgements

This work was supported by grants No 3 T09A 027 11 , 3 T09A 036 10, and T09A 006 08, 3 T09A 112 12, BW/8000-5-0194-7, and BW/8000-5-0179-7 from the Polish State Committee for Scientific Research (KBN). Calculations were carried out with the use of the resources and software at the Informatics Center of the Metropolitan Academic Network (IC MAN) at the Technical University of Gdańsk and Interdisciplinary Center for Mathematical and Computer Modeling (ICM), Warsaw, Poland.

References

[1] Anfinsen, C.B. (1973) Principles that govern the folding of protein chains. *Science* **181**, 223-230.

[2] Scheraga, H.A. (1992) Some approaches to the multiple-minima problem in the calculation of polypeptide and protein structures. *Int. J. Quant. Chem.* **42**, 1529-1536.

[3] Vasquez, M., Nemethy, G. & Scheraga, H.A. (1994) Conformational energy calculations on polypeptides and proteins. *Chem. Rev.* **94**, 2183-2239.

[4] Scheraga, H.A. (1996) Recent developments in the theory of protein folding: searching for the global energy minimum. *Biophys. Chem.*, **59**, 329-339.

[5] Koliński, A., Godzik, A. & Skolnick, J. (1993) A general method for the prediction of the three-dimensional structure and folding pathway of globular proteins: Application to designed helical proteins. *J. Chem. Phys.* **98**, 7420-7433.

[6] Godzik, A., Koliński, A. & Skolnick, J. (1993) De novo and inverse folding predictions of protein structure and dynamics. *J. Comput.-Aided Mol. Design*, **7**, 397-438.

[7] Skolnick, J., Koliński, A., Brooks, C.L., Godzik, A. & Rey, A. (1993) A method for predicting protein structure from sequence. *Curr. Biol.* **3**, 414-424.

[8] Liwo, A., Pincus, M.R., Wawak, R.J., Rackovsky, S. & Scheraga, H.A. (1993) Calculation of protein backbone geometry from $\mathcal{L}$-carbon coordinates based on peptide-group dipole alignment. *Prot. Sci.* **2**, 1697-1714.

[9] Liwo, A., Pincus, M.R., Wawak, R.J., Rackovsky, S. & Scheraga, H.A. (1993) Prediction of protein conformation on the basis of a search for compact structures; test on avian pancreatic polypeptide. *Prot. Sci.* **2**, 1715-1731.

[10] Levitt, M. & Warshell, A. (1975) Computer simulation of protein folding. *Nature* **253**, 694-698.

[11] Levitt, M. (1976] A simplified representation of protein conformations for rapid simulation of protein folding. *J. Mol. Biol.* **104**, 59-107.

[12] Pincus, M.R. & Scheraga, H.A. (1977) An approximate treatment of long-range interactions in proteins. *J. Phys. Chem.* **81**,1579-1583.

[13] Miyazawa, S. & Jernigan, R.L. (1985] Estimation of effective interresidue contact energies from protein crystal structures: quasi-chemical approximation. *Macromolecules,* **18**, 534- 552.

[14] Seetharamulu, P. & Crippen, G.M. (1991) A potential function for protein folding. *J. Math. Chem.*, **6**, 91-110.

[15] Maiorov, V.N. & Crippen, G.M. (1992) Contact potential that recognizes the correct folding of globular proteins. *J. Mol. Biol.* **227**, 876-888.

[16] Sippl, M.J. (1993) Boltzmann's principle, knowledge-based mean fields and protein folding. An approach to the computational determination of protein structures. *J. Comput.-Aided Mol. Design.*, **7**, 473-501.

[17] Liwo, A., Ołdziej, St., Pincus, M.R., Wawak, R.J., Rackovsky, S. & Scheraga, H.A. (1997) A united-residue force field for off-lattice protein-structure simulations. I: Functional forms and parameters of long-range side-chain interaction potentials from protein crystal data. *J. Comput. Chem.*,**18**, 849-873.

[18] Liwo, A., Pincus, M.R., Wawak, R.J., Rackovsky, S., Ołdziej, St. & Scheraga, H.A. (1997) A united-residue force field for off-lattice protein-structure simulations. II: Parameterization of local interactions and determination of the weights of energy terms by Z-score optimization. *J. Comput. Chem.*,**18**, 874-887.

[19] Liwo, A., Ołdziej, St., Czaplewski, C., Groth, M., Kaźmierkiewicz, R., Pincus, M.R., Wawak, R.J., Rackovsky, S. & Scheraga, H.A. (1996) A knowledge-based united-residue force field for off-lattice calculations of protein structure that recognizes native folds. *Internationnal Symposium on Theoretical and Experimental Aspects of Protein Folding,* San Luis, Argentina, 17-21 June, 1996; Liwo, A., Kaźmierkiewicz, R., Czaplewski, C., Groth, M., Ołdziej, St., Wawak, R.J., Rackovsky, S., Pincus, M.R. & Scheraga, H.A. (1996) Origin of backbone hydrogen-bonding cooperativity in united-residue potentials. *Third Electronic Computational Chemistry Conference, ECCC-3*, Northern Illinois University,1-30 November 1996, poster No 19 (http://chemik.chem.univ.gda.pl:8000/ECCC3/19/poster.html).

[20] Liwo, A., Kaźmierkiewicz, R., Groth, M. & Czaplewski, C. (1997) Design of a knowledgebased force field for off-lattice simulations of protein structure. *Acta Biochim. Pol.*, accepted for publication.

[21] Liwo, A., Ołdziej, St., Ciarkowski, J., Kupryszewski, G., Pincus, M.R., Wawak, R.J., Rackovsky, S. & Scheraga, H.A. (1994) Prediction of conformation of rat galanin in the presence and absence of water with the use of Monte Carlo methods and the ECEPP/3 force field. *J. Prot. Chem.*, **13**, 375-380.

[22] Piela, L., Kostrowicki, J. & Schzeraga, H.A. (1989j The multiple-minima problem in thze conformational analysis of molecules. Deformation of the potential energy hypersurface by the diffusion equation method. *J. Phys. Chem.*, **93**, 3339-3346.

[23] Kostrowicki, J. & Scheraga, H.A. (1992) Application of the diffusion equation method for global optimization to oligopeptides. *J. Phys. Chem.*, **96**, 7442-7449.

[24] J. Pillardy, K.A. Olszewski, L. Piela, Theoretically predicted lowest-energy structures of water clusters. *J. Mol. Struct.*, **1992**, 270, 277-285.

[25] Liwo, A., Kim, S.-Y., Cho, K.-H., Gibson, K.D., Maksimiak, K., No, K.T., Rackovsky, S. & Scheraga, H.A. (1997) Performance of empirical energy functions in reproducing the crystal structures of ice - directions of improvement; manuscript in preparation.

[26] Weiner, S.J., Kolman, P.A., Nguyen, D.T. & Case, D.A. (1996) An all-atom force field for simulations of proteins and nucleic acid. *J. Comp. Chem.*, **7**, 230-252

[27] Kaźmierkiewicz, R., Czaplewski, C., Lammek, B., Ciarkowski, J. & Lesyng, B. (1995) Study of the interactions between neurophysin II and dipeptide ligand by means of molecular dynamics. *J.Molec.Model.*, **1**, 135-141.

[28] Kaźmierkiewicz, R., Czaplewski, C., Lammek, B., & Ciarkowski, J. (1997) Molecular modeling of the neurophysin I-oxytocin complex. *J.Comp.-Aided Molec. Design,*,**11**, 9-20.

[29] Kaźmierkiewicz, R., Czaplewski, C., Lammek, B., & Ciarkowski, J. (1997) Molecular modeling of the neurophysin II-vasopressin complex. QSAR,16,1-e.

[30] Czaplewski, C., Kaźmierkiewicz, R., Liwo, A., Ołdziej, S., & Ciarkowski, J. (1995) Symulacja struktury biocząsteczek z zastosowaniem komputerów dużej mocy. In *Zaawansowane Technologie Informatyczne w Nauce Polskiej,* Centrum Promocji Informatyki, Kraków 1995, pp. 97-105.

[31] Kaźmierkiewicz, R., Czaplewski, C., Lammek, B. & Ciarkowski, J. (1997) Modeling of the NPI/OT interactions. In *Peptides 1996,* (Proc 24 EPS, Edinbourgh, Sep. e-13 1996), Ramage, R. & Epton, R. Eds, European Peptide Society 1997, in press.

[32] Kaźmierkiewicz, R., Czaplewski, C. & Ciarkowski, J. (1997) Elucidation of neurophysin/bioligand interactions from molecular modeling. *Acta Biochimica Polonica,* ordered minireview, submitted.

[33] Kaźmierkiewicz R. (1997) Thesis, University of Gdańsk.

[34] Mouillac, B., Chini, B., Balestre, N.-N., Elands, J., Trumpp-Kallmeyer, S., Hoflack, J., Hibert, M., Jard, S. & Barberis, C. (1995) The binding site of neuropeptide vasopressin Vla receptor. *J. Biol. Chem*, **270**, 25771-25777.

[35] Czaplewski, C., Kaźmierkiewicz, R., Ciarkowski, J., Herzyk, P. & Hubbard, R. (1997) Molecular Modeling of the V2 receptor - bioligand interactions. In *Peptides 1996* (Proc 24 EPS, Edinbourgh, Sep. 8-13 1996), Ramage R & Epton R Eds, European Peptide Society 1997, in press.

[36] Czaplewski, C., Kaźmierkiewicz, R. & Ciarkowski, J. (1997) Molecular modeling of the human vasopressin V2 receptor/agonist complex. *J Comp.-Aided Molec Design*, submitted.

[37] Nemethy, G., Gibson K.D., Palmer, K.A., Yoon, C.N., Paterlini, G., Zagari, A., Rumsey, S. & Scheraga, H.A. (1992) Energy parameters in polypeptides. 10. Improved geometrical parameters and nonbonded interactions for use in the ECEP/3 algorothm with application to proline-containing peptides. *J. Phys. Chem.*, **96**, 6472-6484.

[38] Liwo, A., Tempczyk, A., Ołdziej, S., Shenderovich, M.D., Hruby, V.J., Talluri, S., Ciarkowski, J., Kasprzykowski, F., Łankiewicz, L. & Grzonka, Z. (1996) Exploration of the conformational space of oxytocin and arginine-vasopressin using the electrostatically-driven Monte Carlo and molecular dynamics methods. *Biopolymers*, **38**, 157-175.

[39] Kasprzykowski, F., Skurski, P., Liwo, A., Łankiewicz, L., Ołdziej, S. Łanoszka, J., Wiczk, W. & Grzonka, Z. (1994] Conformational studies of oxytocin analogues. Pol. *J. Chem.*, **68**, 987-995

[40] Kaźmierkiewicz, R., Liwo, A. & Lammek, B. (1995) Theoretical conformational analysis of three vasopressin antagonists with a modifled cyclohexyl ring in the flrst thioacid residue. *Int. J. Pept. Prot. Res.*, **45**, 451.

[41] Ołdziej, S., Ciarkowski, J., Liwo, A., Shenderovich, M.D. & Grzonka, Z. (1995) Conformational aspects of differences in requirements of oxytocin and vasopressin receptors. *Journal of Receptor & Signal Transduction Research*, **15**, 703-713.

[42] Kawińska, M., Łankiewicz, L., Skurski, P., Micewicz, B., Wiczk, W. Ołdziej, S. & Liwo, A. (1994] Fluorescence energy transfer in a series of leucine-enkephalin analogues. *Pol. J. Chem.*, **68**, 975-985.

[43] Qi, Xiao-Fei, Rodziewicz, S., Zboińska, J. & Rolka, K. (1997) Conformational Study of Scyliorhinin I by NMR and Molecular Dynamics Calculations. *Pol. J. Chem.*, in press.

[44] Woźniak, E., Nowacka, M. & Ciarkowski, J. (1997) manuscript in preparation.

[45] Ripoll, D.R. & Scheraga, H.A. (1988) On the multiple-minima problem in the conformational analysis of polypeptides. II. An electrostatically driven Monte Carlo method - tests on poly(L-alanine). *Biopolymers*, **27**, 1283-1303.

[46] Ripoll, D.R,., Pottle, M.S., Gibson, K.D., Liwo A., & Scheraga, H.A. (1995) Implementation of the ECEPP algorithm, the Monte Carlo Minimization method, and the Electrostatically Driven Monte Carlo method on the Kendall Square research KSR1 computer. *J. Comput. Chem.*, **16**, 1153-1163.

[47] Groth, M., Malicka, J., Czaplewski, C., Ołdziej, S., Łankiewicz, L., Wiczk, W. & Liwo, A. (1996) Computer modelling of the solution conformation of cyclic enkephalin analogues. Fourth World Congress of Theoretically Oriented Chemists WATOC '96, Jerusalem, Israel, July 7-12, 1996, Abstract PO-B-61, Book of Abstracts, p. 283; Groth, M., Malicka, J., Czaplewski, C., Ołdziej, S., Lankiewicz, L., Wiczk, W. & Liwo, A. (1996) Determination of solution conformation of Dns^1-c-[D-A_2bu^2,Tyr^4,D-Leu^5]enkephalin and Dns^1-c-(D-A_2bu^2,Tyr^4,D-Leu^5]enkephalin by Monte Carlo global conformational analysis, 2D NMR, and fluorescence spectroscopy, *J. Biomol.* NMR, submitted.

[48] Stewart, J.J.P. (1990). MOPAC - A semiempirical molecular orbital program. *J. Comput.Aided Mol. Design*, **4**, 1-104; Stewart, J.J.P. (1993). MOPAC 93.00 Manual. Fujitsu Limlted, Tokyo, Japan.

[49] Andzelm, J. (1993) DGauss: Density Functional Theory Approach. Implementation and applications. In *Density Functional Methods in Chemistry*, ed. J.K. Labanowski and J. W. Andzelm, Springer, New York, pp.155-174.

[50] Ciarkowski, J & Ołdziej, S. (1993) AM1 and PM3 study of hydrogen exchange within the catalytic center of aspartic proteases using a low-molecular-weight mimic. *Eur. Biophys. J.*, **22**, 207-212.

[51] Ciarkowski, J., Ołdziej, S., & Liwo, A. (1994) Mimicking the mechanism of aspartic proteases: l. Molecular mechanics study. *Polish J. Chem.*, **68**, 939-948.

[52] Ołdziej, S. & Ciarkowski, J. (1994) Mimicking the mechanism of aspartic proteases: 2. Molecular orbital study. *Polish J. Chem.*, **68**, 949-956.

[53] Ołdziej, S. & Ciarkowski, J. (1996) Mechanism of action of aspartic proteinases. Application of transition state analogue theory. *J. Comp-Aided Molec. Design*, **10**, 583-588.

[54] Ołdziej, S. (1995) Thesis, University of Gdańsk.

[55] Tarnowska M., Ołdziej S., Liwo A., Grzonka Z. & Borowski E. (1992) Investigation of the inhibition pathway of glucosamine synthase by N^3-(4-methoxyfumaroyl)-L-2,3-diaminopropanoic acid by semiempirical quantum mechanical and molecular mechanics methods. *Eur. Biophys. J.*, **21**, 273-280.

[56] Ołdziej, S., Liwo, A. & Borowski, E. (1997) Theoretical study on the mechanism of addition of mercaptans to ethylene and acrylic aldehyde. *J. Am. Chem. Soc.*, submitted.

[57] Ołdziej, S., Liwo, A. & Borowski, E. (1997) Theoretical studies of the mechanism of glucosamine-synthase inactivation by the derivatives of N^3-fumaroyl-L-2,3-diaminopropionic acid (FDP) *J. Computer-Aided Molec. Design*, submitted.

[58] Halliwell, B. & Gutteridge, J.M.C. (1985) Free radicals in biology and medicine. Clarendon Press, Oxford.

[59] Pryor, W.A. (1982) (ed.) Free radicals in biology, Vol. V, Academic Press, New York.

[60] Lown, J.W. (1988). (Ed.) Anthracycline and anthracenedione-based anticancer drugs, E1- sevier, Amsterdam.

[61] Lown, J.W. (1993) Discovery and development of anthracycline antitumor antibiotics. Chem. Soc. Rev.,165-17G.

[62] Davies, K.J.A. & Doroshow, J.H. (198G) Redox cycling of allthracyclines by cardiac mitochondria. *J. Biol. Chem.*, **261**, 3060-3067.

[63] Mian, M., Frata, D., Rainaldi, G., Sim, S., Mariani, T., Benetti, D. & Gervasi, P.G. (1991) Superoxide anion production and toxicity in V79 cells of six hydroxy-anthraquinones. *Anticancer Research,* **11**,1071-1076.

[64] Tempczyk, A., Tarasiuk, J., Ossowski, T. & Borowski, E. (1988) An alternative concept of the peroxidating ability of anthracycline anti-tumor antibiotics and anthracenodiones. *Anti-Cancer Drug Design*, **2**, 371-385.

[65] Tarasiuk, J., Liwo, A., Wojtkowiak, S., Dzieduszycka, M., Tempczyk, A., Garnier-Suillerot, A., Martelli, S. & Borowski, E. (1991) Molecular determinants of singlet oxygen binding by anthraquinones in relation to their redox cycling activity. *Anti-Cancer Drug Design,* **6**, 399-416.

[66] Jeziorek, D., Dyl, D., Liwo, A., Woźnicki, W., Tempczyk, A.& Borowski, E. (1992) A theoretical study of the mechanism of oxygen binding by model anthraquinones. Part I. Quantum mechanical evaluation of the oxygen-binding sites of 1,4-hydroquinone. *Anti-Cancer Drug Design,* **7**, 451-461.

[67] Jeziorek, D., Dyl, D., Liwo, A., Woźnicki, W., Tempczyk, A. & Borowski, E. (1993) A theoretical study of the mechanism of oxygen binding by model anthraquinones. Part II. Quantum-mechanical studies of the energetics of oxygen binding to model anthraquinones. *Anti-Cancer Drug Design,* **8**, 223-235.

[68] Jeziorek D., Dyl, D., Liwo A., Ossowski, T., & Woźnicki, W. (1994) Enlralpy of oxygen addition to anthraquinone derivatives determines their ability to mediating NADH oxidation. *Anti-Cancer Drug Design,* **9**, 435-448.

[69] Liwo, A., Jeziorek, D., Ossowski, T., Dyl, D., Tempczyk, A., Tarasiuk, J., Nowacka, M., Borowski, E. & Woźnicki, W. Molecular modeling of singlet-oxygen binding to anthraquinones in relation to the peroxidating activity of antitumor anthraquinone drugs. *Acta Biochim. Pol.*, **42**, 445-456.

[70] Jeziorek, D., Ossowski, T., Liwo, A., Dyl, D., Nowacka, M. & Woźnicki, W. (1997) Theoretical and electrochemical study of the mechanism of anthraquinone-mediated one-electron reduction of oxygen: the involvement of adducts of dioxygen species to anthraquinones. *J. Chem. Soc. Perkin Transactions II*, 229-236.

[71] Liwo, A., Dyl, D., Jeziorek, D., Nowacka, M., Ossowski, T. & Woźnicki, W. (1997) MCSCF study of singlet-oxygen addition to ethenol - a model of photooxidation reactions of unsaturated and aromatic compounds bearing hydroxy groups. *J. Comput. Chem.*, in press.

[72] Liwo, A., Jeziorek, D., Dyl, D., Ossowski, T. & Chmurzyński, L. (1997) Theoretical Study of the role of hydrogen bondiug and proton transfer in oxygen reduction by semiquinones. *J. Mol. Struct. THEOCHEM*, in press.

[73] Novak, R.F., Kharasch, E.D., Frank, P. & Runge-Morris, M. (1988). Anthracyclines, anthracenediones and anthrapyrazoles: comparison of redox cycling activity and effects on lipid peroxidation and prostaglandin production; in: Lown, J.W. (ed.) *Anthracycline and anthracenediorce-based anticancer drugs*, Elsevier, Amsterdam, pp. 475-526

[74] Schmidt, M.W., Baldridge, K.K., Boatz, J.A., Elbert, S.T.,Gordon, M.S., Jensen, J.J., Koseki, S., Matsunaga, N., Nguyen, K.A., Su, S., Windus, T.L., Dupuis, M., & Montgomery, J.A. (1993) General Atomic and Molecular Electronic Structure System. *J. Comput. Chem.*, **14**, 1347-1363.

[75] Gaussian 94, Revision D.4, M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. B. Stefanov, A. Nanayakkara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Pittsburgh PA,1995.