5-Aminolevulinic acid and derivatives thereof
“Erwin with his \( \psi \) can do
Calculations quite a few.
But one thing has not been seen:
Just what does \( \psi \) really mean?”

— Erich Hückel, translated by Felix Bloch
5-Aminolevulinic acid and derivatives thereof
Properties, lipid permeability and enzymatic reactions
Abstract


5-aminolevulinic acid (5-ALA) and derivatives thereof are widely used prodrugs in treatment of pre-malignant skin diseases of the cancer treatment method photodynamic therapy (PDT). The target molecule in 5-ALA-PDT is protoporphyrin IX (PpIX), which is synthesized endogenously from 5-ALA via the heme pathway in the cell. This thesis is focused on 5-ALA, which is studied in different perspectives and with a variety of computational methods. The structural and energetic properties of 5-ALA, its methyl-, ethyl- and hexyl esters, four different 5-ALA enols, and hydrated 5-ALA have been investigated using Quantum Mechanical (QM) first principles density functional theory (DFT) calculations. 5-ALA is found to be more stable than its isomers and the hydrolysations of the esters are more spontaneous for longer 5-ALA ester chains than shorter. The keto-enol tautomerization mechanism of 5-ALA has been studied, and a self-catalysis mechanism has been proposed to be the most probable. Molecular Dynamics (MD) simulations of a lipid bilayer have been performed to study the membrane permeability of 5-ALA and its esters. The methyl ester of 5-ALA was found to have the highest permeability constant ($P_{Me-5-ALA} = 52.8 \text{ cm/s}$).

The mechanism of the two heme pathway enzymes; Porphobilinogen synthase (PBGS) and Uroporphyrinogen III decarboxylase (UROD), have been studied by DFT calculations and QM/MM methodology. The rate-limiting step is found to have a barrier of 19.4 kcal/mol for PBGS and 13.7 kcal/mol for the first decarboxylation step in UROD. Generally, the results are in good agreement with experimental results available to date.

Keywords: 5-Aminolevulinic acid, tautomeration, PDT, DFT, MM, QM/MM, Porphobilinogen synthase, Uroporphyrinogen III decarboxylase, membrane penetration, enzyme mechanism.
Abstract


5-aminolevulinic acid (5-ALA) and derivatives thereof are widely used prodrugs in treatment of pre-malignant skin diseases of the cancer treatment method photodynamic therapy (PDT). The target molecule in 5-ALA-PDT is protoporphyrin IX (PpIX), which is synthesized endogenously from 5-ALA via the heme pathway in the cell. This thesis is focused on 5-ALA, which is studied in different perspectives and with a variety of computational methods. The structural and energetic properties of 5-ALA, its methyl-, ethyl- and hexyl esters, four different 5-ALA enols, and hydrated 5-ALA have been investigated using Quantum Mechanical (QM) first principles density functional theory (DFT) calculations. 5-ALA is found to be more stable than its isomers and the hydrolysations of the esters are more spontaneous for longer 5-ALA ester chains than shorter. The keto-enol tautomerization mechanism of 5-ALA has been studied, and a self-catalysis mechanism has been proposed to be the most probable. Molecular Dynamics (MD) simulations of a lipid bilayer have been performed to study the membrane permeability of 5-ALA and its esters. The methyl ester of 5-ALA was found to have the highest permeability constant (P$_{Me-5-ALA}$ = 52.8 cm/s). The mechanism of the two heme pathway enzymes; Porphobilinogen synthase (PBGS) and Uroporphyrinogen III decarboxylase (UROD), have been studied by DFT calculations and QM/MM methodology. The rate-limiting step is found to have a barrier of 19.4 kcal/mol for PBGS and 13.7 kcal/mol for the first decarboxylation step in UROD. Generally, the results are in good agreement with experimental results available to date.

Keywords: 5-Aminolevulinic acid, tautomerization, PDT, DFT, MM, QM/MM, Porphobilinogen synthase, Uroporphyrinogen III decarboxylase, membrane penetration, enzyme mechanism.
This thesis is based on following papers:


VI. Bushnell, Eric A. C.; Erdtman, Edvin; Llano, Jorge; Eriksson, Leif A.; Gauld, James W., A computational study into the first branching point in porphyrin biosynthesis; decarboxylation of ring D in URO-III by uroporphyrinogen-III decarboxylase (submitted to Biochemistry), 2010.

All published papers are printed with permission from the journals.
List of papers

This thesis is based on following papers:


VI. Bushnell, Eric A. C.; Erdtman, Edvin; Llano, Jorge; Eriksson, Leif A.; Gauld, James W., *A computational study into the first branching point in porphyrin biosynthesis; decarboxylation of ring D in URO-III by uroporphyrinogen-III decarboxylase* (submitted to Biochemistry), 2010.

All published papers are printed with permission from the journals.
My contributions to the papers are:

I-V  All calculations and analysis, writing of the first drafts and revisions of the papers.

VI  Supervision and discussions with the first author related to the docking, MM optimizations and MD simulations. Some of the MD simulations and the average distance calculations were performed by me.
I contributed to the papers as follows:

I-V All calculations and analysis, writing of the first drafts and revisions of the papers.

VI Supervision and discussions with the first author related to the docking, MM optimizations and MD simulations. Some of the MD simulations and the average distance calculations were performed by me.

Abbreviations, symbols and units

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-ALA</td>
<td>5-Aminolevulinic acid</td>
</tr>
<tr>
<td>5-ALA-hyd</td>
<td>5-Amino-4,4-dihydroxy-pentanoic acid (hydrated 5-aminolevulinic acid)</td>
</tr>
<tr>
<td>Me-5-ALA</td>
<td>5-Aminolevulinic acid methyl ester</td>
</tr>
<tr>
<td>Et-5-ALA</td>
<td>5-Aminolevulinic acid ethyl ester</td>
</tr>
<tr>
<td>He-5-ALA</td>
<td>5-Aminolevulinic acid hexyl ester</td>
</tr>
<tr>
<td>5-CLA</td>
<td>5-Clorolevulinic acid</td>
</tr>
<tr>
<td>ALAS</td>
<td>5-Aminolevulinic acid synthase</td>
</tr>
<tr>
<td>ALAD</td>
<td>5-Aminolevulinic acid dehydratase, synonym to PBGS</td>
</tr>
<tr>
<td>AO</td>
<td>Atomic orbital</td>
</tr>
<tr>
<td>B3LYP</td>
<td>Becke 3-Parameter (Exchange), Lee, Yang and Parr (Correlation)</td>
</tr>
<tr>
<td>CP-III</td>
<td>Coproporphyrinogen III</td>
</tr>
<tr>
<td>DFT</td>
<td>Density functional theory</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DPPC</td>
<td>Dipalmitoylphosphatidylcholine (a phospholipid)</td>
</tr>
<tr>
<td>EC</td>
<td>Enzyme commission number (classification number for enzymes)</td>
</tr>
<tr>
<td>FC</td>
<td>Ferrochelatase</td>
</tr>
<tr>
<td>GTO</td>
<td>Gaussian type orbital</td>
</tr>
<tr>
<td>HF</td>
<td>Hartree-Fock</td>
</tr>
<tr>
<td>IEFPCM</td>
<td>Integral equation formalism of the polarizable continuum model</td>
</tr>
<tr>
<td>LA</td>
<td>Levulinic acid</td>
</tr>
<tr>
<td>MD</td>
<td>Molecular dynamics</td>
</tr>
<tr>
<td>MM</td>
<td>Molecular mechanics</td>
</tr>
<tr>
<td>MO</td>
<td>Molecular orbital</td>
</tr>
<tr>
<td>MO-LCAO</td>
<td>Molecular orbital linear combination of atomic orbitals</td>
</tr>
<tr>
<td>PI, PII, …</td>
<td>Paper I, Paper II etc.</td>
</tr>
<tr>
<td>PA</td>
<td>Proton affinity</td>
</tr>
<tr>
<td>PBG</td>
<td>Porphobilinogen</td>
</tr>
<tr>
<td>PBGS</td>
<td>Porphobilinogen synthase, also called ALAD</td>
</tr>
<tr>
<td>PDT</td>
<td>Photodynamic therapy</td>
</tr>
<tr>
<td>PMB</td>
<td>Photodynamic molecular beacon</td>
</tr>
<tr>
<td>PpIX</td>
<td>Protoporphyrin IX</td>
</tr>
<tr>
<td>QM</td>
<td>Quantum Mechanics</td>
</tr>
<tr>
<td>QM/MM</td>
<td>Combined calculation method, with a QM and a MM part.</td>
</tr>
</tbody>
</table>
\textbf{RHF} \quad \text{Restricted Hartree-Fock}

\textbf{ROS} \quad \text{Reactive Oxygen Species}

\textbf{ROHF} \quad \text{Restricted Open-shell Hartree-Fock}

\textbf{SCF} \quad \text{Self consistent field}

\textbf{sCoA} \quad \text{succinyl-Coenzyme A}

\textbf{STO} \quad \text{Slater type orbital}

\textbf{TCA cycle} \quad \text{Tricarboxylic acid cycle or Citric acid cycle}

\textbf{UHF} \quad \text{Unrestricted Hartree-Fock}

\textbf{URO-III} \quad \text{Uroporphyrinogen-III}

\textbf{UROD} \quad \text{Uroporphyrinogen-III decarboxylase}

\textbf{ZPE} \quad \text{Zero-point vibrational energy}

\psi \quad \psi, \text{ wave function}

\lambda \quad \lambda, \text{ unit for wavelength}

\nu \quad \nu, \text{ unit for frequency. } \nu \text{ represents the energy of a photon.}

\hbar \quad \text{Planck constant } = 6.626 \times 10^{-34} \text{ Js}

S_0 \quad \text{Ground singlet state}

S_1 \quad \text{First excited singlet state}

S_n \quad \text{Higher (n th) excited singlet state}

T_1 \quad \text{First excited triplet state}

\AA \quad 1 \text{ Ångström } = 10^{-10} \text{ m}

\text{nm} \quad 1 \text{ nanometre } = 10^{-9} \text{ m}

\text{ns} \quad 1 \text{ nanosecond } = 10^{-9} \text{ s}

\text{ps} \quad 1 \text{ picosecond } = 10^{-12} \text{ s}

\text{fs} \quad 1 \text{ femtosecond } = 10^{-15} \text{ s}

\text{kcal/mol} \quad 1 \text{ kilo calorie per mol } = 0.239 \text{ kJ/mol}
## Contents

1 Introduction .................................................................................................................. 13

1.1 Photodynamic therapy .............................................................................................. 13

1.1.1 The PDT approach ................................................................................................. 14

1.1.1.1 Photosensitizers ............................................................................................... 14

1.1.1.2 Light .................................................................................................................. 16

1.1.1.3 Oxygen .............................................................................................................. 16

1.1.2 The PDT mechanism .............................................................................................. 17

1.1.3 Cellular mechanisms .............................................................................................. 18

1.1.4 PDT vs other treatments ....................................................................................... 19

1.2 5-Aminolevulinic acid ............................................................................................... 19

1.2.1 5-ALA metabolism ............................................................................................... 19

1.2.1.1 Porphobilinogen synthase ............................................................................... 21

1.2.1.2 Uroporphyrinogen III decarboxylase .............................................................. 23

1.2.2 5-ALA-PDT ........................................................................................................... 25

1.2.2.1 Fluorescence .................................................................................................... 26

1.2.2.2 Photobleaching ............................................................................................... 26

1.2.2.3 Limitations ....................................................................................................... 26

1.3 Tautomerism .............................................................................................................. 27

2 Computational Methods ............................................................................................. 29

2.1 Quantum Mechanics ................................................................................................. 29

2.1.1 Hartree-Fock ......................................................................................................... 30

2.1.2 Basis sets ................................................................................................................ 31

2.1.3 Density Functional Theory .................................................................................... 33

2.1.4 Hybrid methods ..................................................................................................... 34

2.2 Molecular Mechanics & Molecular Dynamics ....................................................... 34

2.3 QM/MM method ....................................................................................................... 36

2.4 Computational methods in the current studies ...................................................... 37

2.4.1 Paper I and II ........................................................................................................ 37

2.4.2 Paper III ................................................................................................................ 37

2.4.3 Paper IV and V ...................................................................................................... 39

2.4.4 Paper VI ................................................................................................................ 39

2.5 Computational facilities .......................................................................................... 40

3 Summary of results ..................................................................................................... 41
3.1 QM calculations of 5-ALA and its derivatives (P I & P II) ..... 41
3.1.1 Structural properties .................................................. 42
3.1.2 Free energies ........................................................... 43
3.1.3 Proton affinities ........................................................ 45
3.1.4 Tautomerization mechanism ....................................... 47
3.2 MD simulations of 5-ALA and its esters in membrane (P III) 50
3.3 Enzymatic reactions (P IV-VI) ........................................ 52
  3.3.1 The mechanism of PBGS (PIV-V) .............................. 52
    3.3.1.1 The Schiff base formation ................................. 52
    3.3.1.2 Schiff base transfer ......................................... 55
    3.3.1.3 Cyclization reaction mechanism ......................... 57
  3.3.2 The mechanism of UROD (PVI) ................................. 60

4 Conclusions and future perspectives ................................. 63

Acknowledgements .......................................................... 65

References ........................................................................... 67
CHAPTER 1

1 Introduction

The focus of this thesis is the drug 5-aminolevulinic acid (5-ALA), which is used to treat pre-malignant skin disorders with the treatment modality photodynamic therapy (PDT). The aims are to with various computational methods explore 5-ALA’s properties, understand how it behaves in cellular environments, and to get more insight into its metabolism. This knowledge can lead to further improvement of the treatment and the development of new drugs based on 5-ALA derivatives.

1.1 Photodynamic therapy

Photodynamic therapy (PDT) is a treatment modality for primary cancerous lesions, but also pre-malignant and non-malignant diseases. In the late 19th century Niels Ryberg Finsen began to take an interest in the healing effect of sunlight on the skin. Finsen, who received the Nobel Prize for his findings in 1903, was able to treat skin disorders such as smallpox and Lupus vulgaris. Inspired by Finsen’s publicity, a lot of research was started in this field in the beginning of the 20th century. Raab, Jesionek and von Tappeiner found that sunlight combined with a photosensitizer and oxygen could destroy cells, whereupon von Tappeiner coined the term photodynamic therapy. The dye eosin was then used to treat both epilepsy and cancer in conjunction with sunlight.

The curing capability of light was however not a new knowledge. The use of sun treatment was known thousands of years earlier, back in the ancient Egypt, India, China and Greece. They had found that by eating various plants in combination with exposure to sunlight, they could treat skin lesions as vitiligo, cancer, psoriasis and infections. Later these plants were found to contain psoralen compounds, which absorb the light of the sun.

In the last few decades, the number of scientific studies and the usage of PDT have enormously increased. Large steps have been taken in the improvement of photosensitizers and illumination techniques. Due to enhancement of the illumination there are now a variety of diseases that can be treated with PDT. In the beginning PDT was used mainly to treat skin disorders. However, endoscopes with lasers made it possible to bring the light to the lesion, even within the body. Therefore, is it now possible to treat cancer in for example: the bladder, lungs and the organs in the gastrointestinal tract (i.e. the part of the digestive system consisting of the stomach, small intestine and large intestine). In addition, in conjunction with surgery even brain cancer could be treated by PDT. Methods are under
development, where light is delivered via optical fibres through needles stuck into the tumour, which make it possible to treat larger tumours.\textsuperscript{5,6} Alongside cancerous treatment, PDT is used to treat some non-malignant and pre-malignant skin diseases, such as psoriasis, actinic keratosis, acne, age-related macular degeneration and blood sterilization.\textsuperscript{7,8}

### 1.1.1 The PDT approach

PDT is a three component method, where all of its three components need to be present simultaneously. These three essential components are a photosensitizer, light and oxygen. When all these components are present, the light excites the photosensitizer, which reacts with oxygen to form reactive oxygen species (ROS), e.g. singlet oxygen and/or peroxide radicals. The ROS are very reactive, and destroy the cancer cells by oxidation of cellular components. Normally there is oxygen present in tissue, whereas the other two components have to be added. Each of these components is discussed more in detail below.

#### 1.1.1.1 Photosensitizers

A photosensitizer is a compound that can be excited to a higher energy level upon illumination by a specific wavelength. It is preferably built up by a conjugated $\pi$-electron system.

There are a number of criteria for a good photosensitizer. First of all, it must be chemically and physically stable, and also chemically pure. It is an advantage if it is water soluble, but it must also be able to penetrate the lipophilic cell membrane. It should be nontoxic in the absence of light, and preferably become photoactive in the red to near IR region (i.e. it should have a high molar absorption coefficient at $\lambda = 600\text{-}900$ nm). Another important factor is that the photosensitizer should not become photoactive upon UV-radiation. Furthermore, the photosensitizer should accumulate more selectively in tumour cells than healthy tissue, and reach its max concentration there relatively fast. Finally, it should also leave the body rapidly to prevent sensitivity to light after the treatment.\textsuperscript{9,11}

Photosensitizers are divided into porphyrins and non-porphyrins. The porphyrins are in turn classified as first, second and third generation photosensitizers.\textsuperscript{8}

The first generation of photosensitizers are based on a compound called hematoporphyrin (Hp). Hp is a tetapyrrole extracted from blood, in which the iron has been removed.\textsuperscript{2} This compound was found to specifically accumulate in cancerous tissue. A lot of effort was taken into development of superior derivatives of Hp, since it was not effective enough and
a high dosage was needed. One of these derivatives is Photofrin – the most clinically used photosensitizer (also known as Porfimer sodium). Photofrin is a mixture of oligomers ranging from two to nine porphyrin units linked together by primarily ether bonds. Photofrin has been approved for use in treatment of many cancer diseases, such as lung-, oesophageal-, bladder- and cervical cancer as well as malignant and non-malignant skin diseases. However, Photofrin has serious disadvantages. First of all the clearance of Photofrin in the body is very slow. It stays photoactive for weeks after the treatment, and during this time the patient is very sensitive to light and is not able to stay in the sunlight for longer periods of time. Secondly Photofrin has a weak absorption peak above 600 nm (630 nm), why the dosage needs to be quite high instead.\textsuperscript{8,12}

To solve these problems, a number of second generation porphyrin based photosensitizers have been developed, which are more swiftly degraded in the body and absorbs at higher wavelengths. Various substituents are attached to the porphyrin ring to get a larger system of conjugated double bonds, which will red-shift the absorption maxima. Two examples are the drug Foscan, which is applied for clinical use in head and neck cancer, and Tookad, which is applied for prostate cancer treatment. A drawback of building these large molecules could be that the drugs become very lipophilic and may accumulate in the cell membrane.

Another approach is to apply a prodrug, which is metabolized \textit{in situ} to a photosensitizer that is naturally present in the body. 5-aminolevulinic acid (5-ALA) is the precursor to heme and other porphyrins in living organisms, and by excess of 5-ALA the photosensitizer protoporphyrin IX is accumulated. This will be discussed later in chapter 1.2.2 about 5-ALA-PDT.

A third generation of porphyrin photosensitizers has also begun to be examined. Beyond the second generation, these photosensitizers are designed to have more specific affinity to the tumour tissue, and are built up by second generation photosensitizers bound to carriers, such as antibodies or liposomes.\textsuperscript{8,13}

Besides the porphyrin derivatives there are a couple of other drugs in use and in development; such as metal complexes and dyes like the anthraquinone-derivative hypericin and Methylene Blue.\textsuperscript{10,12}

Another very recent method to get more specific treatment to the tumour tissue is the use of photodynamic molecular beacons (PMBs). A PMB consists of a photosensitizer which is combined with a linker to a ROS quencher. The linker is designed to bind to a cancer cell-specific biomarker, and when bound the quencher is cut off. This means that healthy tissue will not get injured, since the ROS produced of the photosensitizer are
16

I

EDVIN ERDTMAN

5-Aminolevulinic acid and derivatives thereof

However, in cancer cells the quencher will no longer be proximate to the photosensitizer and the ROS are free to destroy the cell.\(^8,^{14}\)

1.1.1.2 Light

The light used in PDT is photosensitizer specific, since each photosensitizer has its own absorption maxima. Even though it is possible to use white light, which consists of a wide spectrum of photons, better results have been found by using monochromatic coherent light.

The most relevant light used in PDT is roughly visible light (400-700 nm cf. Figure 1.1), but in practice the most used light ranges from 600 to 900 nm. Wavelengths shorter than 600 nm are not suitable, since there is an elevated risk for sunlight photosensitivity (sunlight contains radiation with wavelength \(\lambda < 600\) nm). Furthermore, hemoglobin absorbs most of the incoming photons at these wavelengths.\(^7,^{15}\) The penetration depth into tissue is also a limiting factor of illumination. Red light (630-710 nm) for example has a penetration depth of 2.0-4.5 mm in tumours, whereas near IR (1060 nm) light may penetrate up to 6.5 mm.\(^{16}\) Unfortunately, the photons in the IR region do not have enough energy to excite most photosensitizers and generate singlet oxygen.

![Figure 1.1 The radiation spectrum.\(^{17}\)](image)

1.1.1.3 Oxygen

Molecular oxygen is by definition mandatory for PDT. Moan et al. observed the reactive singlet oxygen during PDT and found a connection between low oxygen concentration and less PDT effect.\(^{18,19}\)
There are however non-oxygen dependent techniques, which are not strictly speaking PDT-methods (however, the wider concept photochemotherapy also involves these techniques). Without the involvement of oxygen the photosensitizer is excited to higher triplet states, and is quenched directly by the tissue. These techniques are useful in tissues with low oxygen levels, for example in the middle of larger tumours.

### 1.1.2 The PDT mechanism

The general mechanism of PDT can be explained as follows; a photosensitizer is excited by light, followed by the reaction of the excited photosensitizer with molecular oxygen to produce ROS, such as singlet oxygen, hydroxyl- or superoxide radicals. There are two types of photoreaction mechanisms; type I and type II. The first steps are the same in both mechanisms:

\[
P(S_0) \xrightarrow{h\nu} P(S_n) \xrightarrow{\text{IS}} P(T_1)
\] (1.1)

The photosensitizer \(P\) is excited by a photon \((h\nu)\) from its ground state \((S_0)\) to a singlet excited state \((S_n)\). The photosensitizer is then relaxed to the lowest singlet excited state \((S_1)\), followed by an intersystem crossing to the first excited triplet state \((T_1)\). Triplet states are relatively more stable than excited singlet states. Triplets have therefore more time to undergo further reactions. However, competing reactions from \(T_1\) are fluorescence (1.2) and radiation-less relaxation (1.3).

\[
P(T_1) \rightarrow P(S_0) + h\nu'
\] (1.2)

\[
P(T_1) \rightarrow P(S_0) + \text{heat}
\] (1.3)

Generally the dominating mechanism is determined by the concentration of oxygen. If the oxygen concentration is large, the type II mechanism is most probable, while the type I mechanism is predominant if there is a lower oxygen concentration. The definitions of how to distinguish between the Type I and Type II reaction types diverge. One definition is based on the primary interaction of the photosensitizer. If this first reacts with the solvent or a biological substrate, it is a Type I, but if the photosensitizer first reacts with oxygen it is a Type II process. Another classification is based on whether oxygen radicals are formed via electron-transfer or singlet oxygen via energy-transfer. A simplified scheme over Type I / Type II reactions are shown in Figure 1.2.
Figure 1.2 A simplified scheme describing the Type I and Type II photoreactions. P represents the photosensitizer and A the substrate; a molecule in the cancerous tissue, for example a membranal phospholipid or solvent.

A type I photoreaction is a hydrogen abstraction or an electron-transfer reaction between a photosensitizer and a substrate (A in Figure 1.2), which can either be the solvent, another photosensitizer or a biological molecule. Free radicals or radical ions are formed, which are very reactive and react with oxygen to produce superoxide radical anions or hydroxyl radicals. These radicals then cause oxidative damage to the cell.\textsuperscript{12,21} Type I reactions may also be independent of oxygen, as for psoralens reaction with DNA. These reactions are also sometimes called Type III reactions.\textsuperscript{10,24}

In a type II reaction, the photosensitizer transfers its excitation energy directly into the oxygen molecule. Singlet oxygen is generated via energy transfer from the excited photosensitizer to triplet oxygen when they collide. Singlet oxygen will then cause oxidative damage to the tissue (Figure 1.2).

The ROS do also oxidize and degrade the photosensitizer; a process called photobleaching. In average each photosensitizer molecule can catalyse the production of $10^{3}$-$10^{5}$ singlet oxygen molecules before it is destroyed by photobleaching or other processes.\textsuperscript{21}

1.1.3 Cellular mechanisms

The PDT treatment affects the cells in different ways, and causes cell death by either necrosis or apoptosis. Necrosis on one hand is a sudden cell death, where organelles and membranes are damaged, while apoptosis on
the other hand is a controlled cell death that is naturally taking place so that the organelles of the dead cell can be recycled.

Whether the cell death is caused by necrosis or apoptosis depends on the location of the photosensitizer when it is illuminated. It is found that if the photosensitizer is illuminated in the mitochondria, the cell predominantly undergoes apoptosis. However, if the photosensitizer is settled in the cell membrane necrosis is more predominant. Other factors which play an important role are the cell line and the dosage of light and photosensitizer. Generally low doses of PDT lead to apoptosis, while higher doses increase the possibility for necrosis.\textsuperscript{11,15,25}

1.1.4 PDT vs other treatments

Compared with other cancer treatment techniques PDT has several advantages. Besides killing the cancer cells directly, PDT can also damage the tumour’s associated vasculature. Thus, the blood transfer to the tumours is affected, which suffocates the tumour. Another important issue is the immune system’s response to the treatment. While surgery, ionizing radiation and chemotherapy suppress the immune system, PDT stimulates it. When these three mechanisms; cell death, vasculature destruction and immune response, can be controlled to act together, a long-term tumour regression is performed by PDT.\textsuperscript{11}

1.2 5-Aminolevulinic acid

5-ALA (Figure 1.3) represents a completely different aspect of PDT. It is not in itself photosensible but with an excess of 5-ALA, Protoporphyrin IX (PpIX) is produced in situ. In the following section the metabolism of 5-ALA will be discussed.

\begin{center}
\textbf{Figure 1.3} 5-Aminolevulinic acid in its zwitterionic form, with the numbering of the carbon atoms.
\end{center}

1.2.1 5-ALA metabolism

5-ALA is a delta amino acid that has a carbonyl group at the fourth carbon (systematic name: 5-amino-4-oxopentanoic acid, Figure 1.3). 5-ALA is present in all kinds of organisms. There are two distinct pathways in which 5-ALA is biosynthesized; from glutamate (the C\textsubscript{5} or Beale pathway) or
from succinyl-Coenzyme A (sCoA) and glycine (the C\textsubscript{4} or Shemin pathway). In plants, algae, cyanobacteria, most other bacteria and archaea the multistep C\textsubscript{5} pathway is used to synthesize 5-ALA, whereas the one-step C\textsubscript{4} pathway is found in humans, animals, yeasts, and a few bacteria.\textsuperscript{26}

Since we are more interested in the 5-ALA mechanisms in humans, we will not go into detail of the C\textsubscript{3} pathway. The C\textsubscript{4} pathway is in eukaryotes combined with the tricarboxylic acid (TCA cycle) by sCoA. sCoA is together with glycine the substrates of the mitochondria located enzyme aminolevulinic acid synthase (ALAS; EC: 2.3.1.37). ALAS is a homodimer with the active site located in the subunit interface, in which two pyridoxal 5'-phosphate cofactors are symmetrically bound. ALAS catalyses the de-carboxylative condensation of glycine and sCoA, where the release of 5-ALA is the rate-determining step. ALAS is considered as the first enzyme in the heme biosynthesis (see Figure 1.4).\textsuperscript{26,27}

![Figure 1.4](image)

**Figure 1.4** A simplified scheme of the heme biosynthesis, which is taking place in both the cytoplasm and in the mitochondria. The currently studied enzymes are marked in bold.
There are seven further enzymes involved in the formation of heme; porphobilinogen synthase, porphobilinogen deaminase, uroporphyrinogen III synthase, uroporphyrinogen III decarboxylase, coproporphyrinogen III oxidase, protoporphyrinogen IX oxidase and ferrochelatase. The latter three are located in the mitochondria and the others in the cytosol (see Figure 1.4). In the current study porphobilinogen synthase (PBGS) and uroporphyrinogen III decarboxylase (UROD) have been considered with special interest (marked in bold in Figure 1.4).

1.2.1.1 Porphobilinogen synthase

The second enzyme in the heme pathway is porphobilinogen synthase (PBGS), also called 5-ALA dehydratase (ALAD; EC 4.2.1.24). Two 5-ALA molecules are combined to the pyrrole porphobilinogen (PBG). PBGS is located in the cytosol in contrast to ALAS which is operational in the mitochondria. PBGS is a metalloenzyme, which is most active in a homooligomeric form. By the natural single mutation of phenylalanine to leucine (F12L) in human PBGS, a hexamer structure can be formed; however, with a much lower activity (~12% of the wild type enzyme). The active site of PBGS has been found to be highly conserved amongst different species. All residues of PBGS are hereafter identified according to their yeast numbering (PDB ID 1H7O). In particular, the active site contains two lysine residues, Lys210 and Lys263, in the A- and P-site respectively (Figure 1.5). The sites are named after the acid group (acetyl- and propionyl-) of the product PBG derived from the carboxylate moieties of the 5-ALA substrates. Experimental mutogenesis studies have suggested that the latter lysine is essential for enzyme catalysis, and the former is essential for the binding of the first substrate. Each of the two 5-ALA substrates is found to bind to a lysine with a so-called Schiff base. A Schiff base is an imine with a hydrocarbyl group on the nitrogen atom (R=C=NR').

Furthermore, the active site consists of several polar groups which form hydrogen bonds to the carboxylate moieties of the P- (Ser290 and Tyr329) and A-site (Gln236) bound 5-ALAs. Several residues (Ser179, Asp131 and Tyr 207) form a polar pocket around or hydrogen bond to the terminal amino group of the 5-ALA substrates. However, at least for the P-site; substrate analogs without the terminal amino group have been found to be good competitive inhibitors. Therefore it is suggested that the interactions between the 5-ALA amino group and the enzyme are not essential, at least not for the binding. A flexible segment of PBGS is also believed to seal the active site when the 5-ALAs are bound. When this ‘lid’ is closed there are
in for example humans and yeast two arginine residues (Arg220 and Arg232) of the lid, that form hydrogen bonds to the carboxylate of the A-site 5-ALA (see Figure 1.5).  

![Figure 1.5 Schematic illustration of the active site of PBGS with the two 5-ALA substrate molecules covalently bound at the A- (red) and P-site (blue) via Schiff-base linkages based on the yeast PBGS crystal structures PDB ID: 1H7O and 1OHL.](image)

There are at least two different sequences for metal binding in PBGS, one primary for zinc ions, and one for magnesium ions.

The first one is located in the active site. In archaea, some bacteria, metazoa (multicellular animals) and yeast organisms the sequence is very cysteine rich with the general sequence DXCX CX(Y/F)X;G(H/Q)CG, where the underlined cysteines coordinate a Zn$^{2+}$ ion (shown in Figure 1.5). In other organisms this sequence is instead aspartate rich (DXALDX(Y/F)X;G(H/Q)DG), which could bind either Mg$^{2+}$ or monovalent ions such as K$^+$ or Na$^+$.  

The Zn$^{2+}$ ion in human and yeast PBGS coordinates to the thiolates of the three cysteines in the sequence above. Since zinc can coordinate four or
five ligands, it is possibly involved in the reaction mechanism, forming bonds to either H₂O, or the substrates/product. Experimental pH, mutagenesis and kinetic studies have suggested that the zinc ion plays an important role in substrate binding at the A-site and in stabilizing intermediates and transition structures during the catalytic mechanism, but however not the binding of the first (P-site) 5-ALA molecule. This has been further supported by experimental NMR studies on the enzyme-bound product complex and crystal structures obtained from human and yeast PBGS with an ‘almost product’ intermediate bound within their active sites (PDB ID: 1E51 and 1OHL). In both these crystal structures, the terminal amino group corresponding to the A-site bound 5-ALA was found to be neutral and coordinated to the Zn²⁺ ion. It has also been proposed that the carbonyl of A-site 5-ALA also coordinates the zinc ion; however this has not been observed in any crystal structures.

The second sequence for metal binding includes a glutamine, two invariant aspartates and one arginine $RX_{-164} DX_{-65} EXXD$. This site coordinates an allosteric octahedral Mg²⁺ ion, with the glutamine and seven water molecules in the first coordination sphere. The aspartate and the arginine residues are ligated at the outer coordination sphere together with more water molecules. This metal binding site is found in all organisms except metazoa, fungi and a few bacteria. The PBGSs that have a Zn²⁺ binding site, but no Mg²⁺ site (metazoa and fungi), has a second Zn²⁺ bound in proximate position to the first. This ion is however not crucial for the reaction.

There are different proposed mechanisms of PBGS. The main differences lay in how many Schiff bases that are formed in the active site, and in which order the intersubstrate bonds are formed. Other differences are which roles the zinc ion and the basic residues play in the active site. By the proofs of the crystal structures, a consensus has now been built up that there are two Schiff bases formed in the active site; one to each 5-ALA molecule. A majority of recently studies also suggest that the C–C intersubstrate bond is formed before the C–N bond.1.2.1.2 Uroporphyrinogen III decarboxylase

Uroporphyrinogen III (URO-III) is the first cyclic compound in the heme biosynthesis. The enzyme uroporphyrinogen III decarboxylase (UROD; EC 4.1.1.37) catalyses the decarboxylation of the acetyl chains of URO-III to form coproporphyrinogen III (CP-III):
where P specifies the propionate side chains. Under physiological substrate concentrations the enzyme starts the decarboxylation of the acetyl chain of ring D, followed by the acetyls of the A, B, and C rings, whereas at higher concentration the order is random.\(^{47,48}\)

The overall structure revealed from X-ray structures of human, tobacco and Bacillus subtilis shows that UROD has a homodimer quaternary structure. The active site of UROD can generally be divided into three regions; one negative, one positive/polar, and one non-charged region.\(^{49}\) The non-charged region is believed to bind the relatively non-polar core of the tetrapyrrole. The negative region is consistent only by the invariant residue Asp86 (human UROD numbering), which coordinates to the –NH–groups in the tetrapyrrole, and is proposed to be deprotonated to stabilize the positively charged nitrogen atom in the intermediate structures. Indeed, through mutagenesis investigations it was found that a D86G mutation essentially killed the enzyme.\(^{50}\) In the positive/polar region there are four invariant residues; two arginines (Arg37 and 41), a histidine (His339) and a tyrosine (Tyr164). The arginines are found to interact with the carboxyl groups of the product.\(^{50,51}\) Several mutagenesis studies have suggested that His339 and Tyr164 are not in fact essential for catalysis.\(^{50,52}\) Interestingly, mutation of His339 was found to have little or no effect on the rate at which the initial decarboxylation, that of ring D, occurred. However, it resulted in accumulation of the first mechanistic intermediate; ring D of URO-III decarboxylated while the A, B and C rings remained unaltered.\(^{52}\) In contrast, it has been found that one or all arginine residues within the active site (human UROD: Arg37, 41 and 50) are essential for catalysis.\(^{49,51,53-55}\) However, the exact role of the residues remains unclear.

Barnard and Akhtar have proposed the general mechanism for UROD enzymes given in Figure 1.6.\(^{56-58}\) There are several proposed mechanisms, and research groups currently contend about which residues are involved in the catalysis.\(^{50,53,59,60}\) In addition, the existing proposed mechanisms argue the placement of URO-III within the catalytic site. Finally, it has been debated that the mechanism involves more than a single active site. It has been suggested that if catalysis occurs in a single site, the substrate needs to
make a 180° flip after the decarboxylation of the ring D acetyl, followed by 90° rotations after each the remaining decarboxylations. Since a 180° flip seems to be very unlikely to happen, it has been proposed that after the D-acetyl is decarboxylated, the substrate flips over to the other monomer’s active site, and the rest of the decarboxylations take place there.60 However, there are many results that indicate that a single unique catalysing site is used throughout the reaction, a mechanism which is now generally accepted.47,50,52,57,58,61-63

**Figure 1.6 Proposed general acid/base mechanism for decarboxylation of the pyrrole acetates in URO-III. HA and HB represent the general acids.**56-58

### 1.2.2 5-ALA-PDT

As already mentioned, 5-ALA-PDT is an elegant method of using the processes inside the cell to produce the photosensitizer in situ.

The formation of 5-ALA is the rate-determining step in the heme synthesis and it is regulated by heme with feedback inhibition. Addition of extracorporeal 5-ALA will bypass the feedback inhibition, and ferrochelatase (FC) instead becomes the rate-determining enzyme. PpIX, the substrate of FC, is then accumulated. To even more enhance the concentration of PpIX, a FC inhibitor or an iron chelator can be added. Fortunately, the activity of FC is found to be lower in tumour cells compared with other cells.64,65 Thereby the excess of 5-ALA will lead to PpIX accumulation especially in
cancer cells. This is one of the advantages with 5-ALA-PDT compared to other methods. Another advantage is that PpIX is primary located in the mitochondria during illumination, which causes a higher degree of apoptosis than other photosensitizers. Furthermore, the deliverance of 5-ALA is very flexible, since it is the only PDT-drug that can be administered both topically and systemically.

1.2.2.1 Fluorescence

PpIX is, like most of the other PDT photosensitizers fluorescent, which has been found to be very useful combined with its high accumulation in cancerous cells. The fluorescent light shows where the tumour is located, which is not always consistent with the shape of the lesion (e.g. on the skin). 5-ALA and other photosensitizers are approved for use in fluorescence photodiagnosis of cancer in dermatology, and the hexyl ester of 5-ALA is approved for use in fluorescent diagnosis of bladder cancer.

1.2.2.2 Photobleaching

PpIX is a very photolabile photosensitizer and undergoes photobleaching much easier compared with other synthetically made photosensitizes. PpIX reacts with singlet oxygen formed during the illumination and is converted into a form that is not photoactive anymore. However, photobleaching may actually have some important clinical effects.

Upon treatment of large tumours, PpIX in the outer regions of the tumour is filtering the radiation, so that the inner parts are not reached by the light, but when PpIX in the outer parts is photobleached the light can penetrate deeper.

A relatively small amount of PpIX is accumulated into healthy tissue compared to cancerous tissue. The PpIX molecules in normal tissue are then also more rapidly photobleached, before they can cause any serious damage. Because of this phenomenon light overdosage does not cause any problem. Increasing the light dosage will only increase the effect on the tumour, since the PpIX in normal tissue is already deactivated.

1.2.2.3 Limitations

There are two major limitations of 5-ALA-PDT. The first limitation is the penetration of both light and drug into the tissue. 5-ALA has an ability to stay superficial and not penetrate deep enough, due to its polarity. The light penetrates, as already mentioned, deeper into human tissue at longer wavelengths and since PpIX has its long-wavelength absorption peak at 635 nm, the light penetration is not optimal.
The second limitation relates to the two step functionality of 5-ALA-PDT; the uptake of 5-ALA and the PpIX formation, are not easily controlled compared to other photosensitizers, where only the uptake of the drug is needed to be taken into account.

These limitations challenge researchers to find out workarounds to get an increased effect of the 5-ALA-PDT modality. To solve the penetration problem, different derivatives of 5-ALA have been developed, such as the methyl- and hexyl- esters of 5-ALA, and 5-ALA containing dendrimers. Another possibility is to make more persisting modifications of the 5-ALA, to change the chemical structure of the photosensitizer, and hence red-shift the absorption band. This however can be problematic, since the enzymes in the biosynthesis of PpIX need to recognize them as substrates.

PDT has been started to be used in combination with other methods for cancer treatment. There are two approaches of combining other methods with PDT. i) A prior treatment which will increase the tumour cells susceptibility to the PDT treatment. ii) A second treatment after PDT, which is dependent on the prosurvival molecular responses to the PDT modality. 68

1.3 Tautomerism

Isomerism is a quite wide concept, including both structural isomers and stereoisomers. Tautomerism is a type of structural isomerism, where the chemical composition is the same, but the position of a double bond and a hydrogen atom differs. Unlike other structure isomers, tautomers interconvert spontaneously into each other, and there is always an equilibrium reaction between the tautomers. In most cases the equilibrium position is far to one side of the reaction. The most common types of tautomeric reactions are keto-enol-, amide-imidic acid-, amine-imine- and lactam-lactim tautomerism.

Keto-enol tautomerization is the equilibrium reaction between a ketone and one or two enol forms:

\[
\begin{align*}
\text{OH} & \quad \leftrightarrow \quad \text{O} \\
\text{H} & \quad \leftrightarrow \quad \text{H}
\end{align*}
\] (1.5)

The ketone is in most cases the most stable, and the equilibrium position lies far to the ketone. The enolic forms are stabilized if there is a stabilizing group next to the enolic hydroxyl group, as for example 2,4-pentanedione:
Here actually the enolic form is in excess at equilibrium (~78% of the reaction mixture are in the enol form).\cite{69} The enol form is stabilized by hydrogen bonding to the second carbonyl group, which also can be explained by resonance stabilization:

\[
\begin{array}{c}
\text{\includegraphics[width=0.5\textwidth]{reaction.png}}
\end{array}
\]

5-ALA has a more stable keto form than its two enolic forms (3-enol and 4-enol, Figure 1.7). Jaffe et al. studied 5-ALA with $^{13}$C- and $^1$H NMR and found the enolic forms not detectable, i.e. less than 0.3% of these were formed. Deuterium exchange rates indicate that the tautomerization is 4 times faster for the 4-enol than for the 3-enol in phosphate buffer at pH 6.8. Jaffe et al. also compared 5-ALA with the similar compounds 5-chlorolevulinic acid (5-CLA) and Levulinic acid (LA)\cite{70}. The rate of 5-CLA is three times slower than 5-ALA for 3-enol but in the same order for the 4-enol, whereas LA has more than 100 times slower rates than 5-ALA. The use of phosphate buffer resulted in the fastest rates, which indicates that phosphate is probably involved in the mechanism.

\[\text{5-ALA-hyd}\]
CHAPTER 2

2 Computational Methods

All work presented in this thesis is performed by computational methods to investigate the chemical and physical properties of 5-ALA and its derivatives. The three following calculation methods have been applied:

1. Density Functional Theory (DFT), which is derived from Quantum Mechanics (QM) have been performed for the optimization of the structures and calculation of the energetic properties of small systems (Paper I, II, IV and V).

2. Molecular Mechanics (MM) and Molecular Dynamics (MD) methods have been used to elucidate the movements of 5-ALA and its derivatives in a lipid bilayer (Paper III), and in parts of the UROD study (Paper VI).

3. The third one is a combined method of QM and MM simply called QM/MM, which is used to elucidate the enzyme mechanism of UROD (Paper VI).

In this chapter, these methods will be explained and discussed briefly.

2.1 Quantum Mechanics

Quantum chemistry, which is the application of QM in chemistry, is based on the Schrödinger equation. The most commonly used version of the Schrödinger equation is the time-independent, non-relativistic Schrödinger equation, which is generally expressed:

$$\hat{H}\Psi = E\Psi$$  \hspace{1cm} (2.1)

where $\hat{H}$ is the Hamiltonian operator, $\Psi$ the wave function, and $E$ the energy of the system. Mathematically, Eqn. (2.1) is an eigenvalue problem, which means that when the Hamiltonian operator acts upon $\Psi$, the outcome is the energy of the system (the scalar $E$) times the same wave function ($\Psi$). $\Psi$ is hence an eigenfunction to $\hat{H}$. The Hamiltonian operator is a sum of the kinetic $\hat{T}$ and potential energy $\hat{V}$ operators.

$$\hat{H} = \hat{T} + \hat{V}$$  \hspace{1cm} (2.2)

If $\Psi$ is known, the probability of finding particles at given positions in space, is given by the integral of the square of the wave function $\int|\Psi|^2\,d\tau$. An arbitrary property $A$ of the system can be calculated as an expectation value $\langle A \rangle$, by the use of the property-specific operator $\hat{A}$:
\[ \langle A \rangle = \frac{(\Psi^* | A | \Psi)}{(\Psi^* | \Psi)} \]  

(2.3)

The denominator represents the normalization of the wave function, and is equal to 1 if \( \Psi \) is already normalized.

### 2.1.1 Hartree-Fock

The Schrödinger equation can be solved analytically for systems with one electron and one nucleus, for example the hydrogen atom. However, approximations have to be made to solve larger systems. The Hartree-Fock (HF) method contains a number of approximations of the many-electron wave function (\( \Psi \)), from which many other computational methods are derived. The major approximations in HF are:

- **The Born-Oppenheimer approximation** is the decoupling of the motions of the electrons and the nuclei. Since the mass of the smallest nucleus (a proton) is 1836 times heavier than an electron, the electron movement is much faster than the movement of the nucleus. Hence, in relation to the electrons, the nucleus is approximately fixed in space. The Born-Oppenheimer approximation implies that the electronic wave function is only dependent on the positions of the nuclei, but independent of their momenta.\(^{72}\) The nucleus-nucleus interaction is now a constant, and could be added to the electronic energy (\( E_{\text{electronic}} \)) when it has been solved.\(^{73}\)

\[
\hat{H}_{\text{electronic}} \Psi_{\text{electronic}} = E_{\text{electronic}} \Psi_{\text{electronic}}
\]

\[
E_{\text{total}} = E_{\text{electronic}} + \sum_{A>B} \sum_{\text{nuclei}} \sum_{\text{nuclei}} \frac{Z_A Z_B}{R_{AB}}
\]

(2.4)

- In the HF method, the electronic wave function is expressed in a single Slater determinant:

\[
\Phi_{\text{SD}} = \frac{1}{\sqrt{N!}} \begin{vmatrix}
\psi_1(1) & \psi_2(1) & \cdots & \psi_N(1) \\
\psi_1(2) & \psi_2(2) & \cdots & \psi_N(2) \\
\vdots & \vdots & \ddots & \vdots \\
\psi_1(N) & \psi_2(N) & \cdots & \psi_N(N)
\end{vmatrix}
\]

(2.5)

where each column represents an electrons spin orbital; a product of a spatial single-electron wave function and a spin function (\( \psi_i = \phi_i \sigma_i \)). The rows of the determinant represent the electron co-
ordinates. By using a single determinant, the electron correlation is neglected, and the electron-electron interaction is only taken into account as an average effect.\textsuperscript{74}

- **MO-LCAO approximation.** For molecules, the one-electron molecular orbitals (MOs) are approximated to be a linear combination (LC) of atomic orbitals (AOs):

\[
\phi = \sum_i c_i \chi_i
\]  \hspace{1cm} (2.6)

where \(c_i\) are coefficients and \(\chi_i\) the AOs. The atomic orbitals are expressed as a linear combination of functions; a basis set (which is discussed in next section).

- **Self consistent field.** The variational principle states, that the energy calculated from an approximate wave function is higher than, or equal to the energy of the exact wave function. The equality holds only if it is the exact wave function. Consequently, the wave function could be determined by an iterative manner.\textsuperscript{74} To solve a HF problem, the wave function is guessed, followed by the minimization of the energy, to get a new and better wave function as input for the energy minimization. This iterative procedure is called self consistent field (SCF).

Even though approximations are made, HF (and HF derived) methods are defined as \textit{ab initio} (Latin for \textit{from the beginning}) calculations. \textit{Ab initio} indicates, in contrast to semi-empirical methods, that the calculations are performed without any input of experimental data. The \textit{ab initio} methods are derived from theoretical principles and the only parameters included are the initial coordinates of the nuclei and universal physical constants.

There are variations in the HF method depending on the electron configuration of the system. The restricted Hartree-Fock (RHF) method is designed for closed-shell systems, i.e. systems that only have paired electrons, while the Restricted Open-shell Hartree-Fock (ROHF) and Unrestricted Hartree-Fock (UHF) allow calculations of open-shell systems.

### 2.1.2 Basis sets

The functions which are combined to form an AO are called a \textit{basis set}. A \textit{complete} basis set is when an infinite number of functions are used. However, since the use of an infinite number of functions is impossible; finite basis sets always imply an approximation of the AOs. There are mainly two function types used for this purpose, the Slater type orbitals (STO) and the Gaussian type orbitals (GTO).
The STOs are functions that mimic an atomic orbital, and have the following mathematical form:

$$\chi_{\zeta n, \ell m}(r, \theta, \varphi) = N Y_{\ell m}(\theta, \phi)r^{n-1}e^{-\zeta r}$$  \hspace{1cm} (2.7)

where $N$ is a normalization constant, $Y_{\ell m}(\theta, \phi)$ the angular part (spherical harmonic functions), $r$ the radii, $\zeta$ (zeta) the orbital exponent, and $n$, $\ell$ and $m$ are quantum numbers. Integration of these functions are however not very practical. Therefore, GTOs have been proposed to be used to model the orbitals with the following mathematical form:

$$\chi_{\ell \xi \eta \zeta}(x, y, z) = N x^\xi y^\eta z^\zeta e^{-ar^2}$$  \hspace{1cm} (2.8)

The main difference between STO and GTO lies in that the GTO is expressed in cartesian instead of polar coordinates, and that the $r$ is squared in the power of the exponential function. Thus GTOs are easier to integrate, and the product of two GTOs gives another GTO.\(^{74}\)

A minimal basis set is the lowest number of basis functions that can be used for a system, where only enough functions are used to contain all the electrons in a neutral system. One example is the STO-3G basis set, which is a STO function, approximated by a linear combination (contraction) of three primitive GTOs per atomic orbital.

A double zeta basis set is the first improvement of the minimal basis set. For each STO in the minimal basis, another STO is included in the basis set. Similarly, a triple or quadruple zeta basis set contains three or four times as many functions as a minimal basis, respectively. Among double and triple zeta basis sets, the split valence basis sets are very popular. These basis sets are not pure double/triple zeta, since minimal basis are used on the core orbitals. For example the 6-31G basis set has the minimal basis in the core orbitals, each a contraction of six primitive GTOs. The valence orbitals however, are each described by two contractions; the first with three primitive GTOs and the second with only one primitive GTO.

To extend the basis set further, a set basis functions with angular momentum higher than the valence shell are added (e.g. $d$-orbitals are added to heavy atoms and $p$-orbitals to hydrogen atoms). These orbitals tend to characterize the polarity of the molecules, and are therefore called polarized basis sets. For example, the 6-31G(d) basis set contains $d$-orbital functions on the heavy atoms and the 6-31G(d,p) basis set contains both $d$-orbitals on the heavy atoms and $p$-orbital functions to the hydrogen atoms.\(^{75}\)

To describe systems where the electrons could be further away from a nucleus (for example: anions, systems with lone pairs and excited states), diffuse functions are added to basis set. Technically, these more diffuse
versions of $s$- and $p$-orbitals have a higher probability to find the electrons further away from the nuclei. The 6-31+G(d,p) basis set is the 6-31G(d,p) basis set with diffuse functions on the heavy atoms, and if another + is added, diffuse functions are included at the hydrogen atoms, as well.

There are several more extensions and improvements (like the correlation consistent (cc) basis sets), and basis sets designed for special molecules (for example LANL2DZ, which is designed for large nuclei), which are not considered in the current thesis.\textsuperscript{76}

### 2.1.3 Density Functional Theory

Density Functional Theory (DFT) is basically a very accurate method, which is derived from the Schrödinger equation of quantum chemistry. To make it more practical however, some parameters are added to the DFT computations. Therefore DFT methods are usually not termed \textit{ab initio} methods, but \textit{first principle} methods.

The DFT method is based on the electron density $\rho$ instead of the wave function $\Psi$. Compared to HF, the number of variables is reduced from $4N$ (three spatial and one spin variable per electron) to \textit{three} spatial variables in DFT, where $N$ is the number of electrons\textsuperscript{74}. This means that DFT calculations are much faster than \textit{ab initio} methods, and much larger systems can be handled. The energy of the system is in DFT a \textit{functional} of the electron density $E[\rho]$. The definition of a \textit{functional} is a function of a function. The energy ($E$) is a function of the electron density $\rho(r)$, which is a function of the positions ($r$).

Kohn and Sham introduced a formalism, which provides a practical way of calculating the electronic energy in an iterative SCF manner similar to the HF method (though more computation-expensive with $3N$ variables).\textsuperscript{77} The Kohn-Sham orbitals can be computed numerically, or be built up by a set of basis functions. There are special basis sets designed for DFT methods, however in comparison with conventional HF basis sets, very small or no improvements of the results were found. It is also found that further enlargement of the basis sets does not enhance the results either. Therefore, conventional basis sets, for example 6-31g(d) are the most widely used in DFT studies.\textsuperscript{75}

The functional of the electron density consists of four terms; the kinetic energy of the non-interacting electrons ($T[\rho]$), electron-nucleus attraction ($E_{ne}[\rho]$), Coulomb interactions ($f[\rho]$) and the exchange-correlation energy ($E_{xc}[\rho]$).

$$E[\rho] = T[\rho] + E_{ne}[\rho] + f[\rho] + E_{xc}[\rho] \quad (2.9)$$
The first three terms can be obtained analytically. However, the last term, which involves non-classical contributions, needs to be calculated approximately. Therefore, a number of different functionals in DFT have been developed to approximate the exchange-correlation term. The electron correlation refers to that the dynamic movement of one electron is dependent of all other electrons in the molecule. As a consequence of the Pauli principle, there is also an exchange interaction between the electrons that imply that the electrons of the same spin are avoiding each others. Usually the exchange-correlation functional is divided into an exchange and a correlation functional term.\textsuperscript{78}(p. 316-320)

2.1.4 Hybrid methods

A method to improve the exchange-correlation functional is to take parts of it from HF. The HF approach makes the approximation that each electron moves in a mean field created by the rest of the electrons. This approximation neglects correlation energy, but gives accurately calculated exchange energy. In DFT calculations however, which are based on the total electron density, the correlation energy is included from the beginning. It is often more accurate to use a hybrid HF-DFT method, where part of the exchange energy is taken from HF method and the electron correlation energy is calculated with DFT methodology.

The B3LYP functional\textsuperscript{79-81} is a hybrid method with three constants fitted to optimize the proton affinities, atomization energies and ionization energies of a number of compounds. B3LYP has shown very good accuracy and stability; the errors of B3LYP lie within 2 kcal/mol.\textsuperscript{82} Therefore, the QM calculations in this work were performed using the B3LYP functional.

2.2 Molecular Mechanics & Molecular Dynamics

Molecular mechanics (MM) is a less accurate method compared to QM methods, since it is based on classical physical laws. However, in studies of larger systems like lipid membranes and enzymes, this methodology is very useful. In MM the nuclear motions are modelled and the motion of the electrons is ignored. It is usually not possible to model any reactions with MM when the bonds between the atoms are defined in parameters. These parameters are based on experimental data, and are collected in a force field.\textsuperscript{83}

The potential energy in MM ($E_{pot}(r^N)$) – a function of the positions ($r$) of $N$ particles, is usually calculated as a sum of four terms related to: bond lengths ($E_b$), bond angles ($E_a$), torsion angles ($E_t$) and non-bonded interactions ($E_{nb}$), where the first three are considered as bonded interactions.
To make a good mathematical representation of a bond stretch, the Morse potential is the most accurate. However, since it is quite computational expensive, the approximate harmonic potential function is used in most force fields:

\[
E_\ell = \frac{1}{2} k_\ell (\ell - \ell_0)^2
\]  

(2.11)

where \( k_\ell \) is the spring constant, \( \ell \) the bond length and \( \ell_0 \) the reference bond length value.

The angular energy is also calculated by a harmonic potential similar to the bond stretch (Eqn. 2.11), while the torsion angles are calculated by one or several cosine functions:

\[
E_\omega = \sum_{\text{torsions}} \sum_{n=0} V_n \frac{1}{2} (1 + \cos (n\omega - \gamma))
\]  

(2.12)

where \( V_n \) controls the barrier height, \( n \) is the multiplicity (i.e. how many minima there are in 360°), \( \omega \) the torsion angle and \( \gamma \) the phase factor (i.e. the location of the minima). Out-of-plane bendings can be modelled as distances to the plane, angles to the plane or by improper torsions, where the latter is the most common.

The non-bonded interactions are divided into the long range electrostatic interactions between charged particles \( (E_{el}) \) and the van der Waals interactions between neutral particles \( (E_{vdw}) \):

\[
E_{nb} = E_{el} + E_{vdw}
\]  

(2.13)

where the electrostatic interactions are modelled by the Coulomb’s law:

\[
E_{el} = \sum_{i=1}^{N_A} \sum_{j=1}^{N_B} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}
\]  

(2.14)

and the van der Waals interactions are in most force fields modelled with the Lenard-Jones potential:

\[
E_{vdw} = \sum_{i=1}^{N} \sum_{j=i+1}^{N} 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]
\]  

(2.15)

In some force fields other so called cross terms are added, to model the couplings between the terms in Eqn. (2.10).

Molecular dynamics (MD) is a method to simulate a system over time. Newtonian mechanics and statistical physics are employed to calculate the
trajectory of all atoms (i.e. the movement of all atoms in the system under a period of time) and physical properties, such as radial distribution functions and diffusion coefficients. An ensemble is chosen for the calculation to decide which parameters to be conserved during the simulation. The most common are the NVT and NPT ensembles. In both cases the number of particle (N) and the temperature (T) are conserved. In NVT the volume is constant, while in NPT the pressure is conserved.

The forces acting on each atom and the energy of the system are calculated using a MM force field, and the positions of next time step are calculated with an appropriate finite difference method. The time step in the time difference method is chosen in the magnitude of fs (10^{-15} s), which is much shorter than the time between collisions and vibrations in the system.

MM methods are used to perform molecular docking. Docking predicts the orientation of one molecule to another molecule. It is widely used to suggest the binding of substrates or inhibitors in the enzymatic active site. The enzyme is usually rigid during the docking process, and the molecule to be docked could either be rigid or flexible. The structures are ranked using a scoring function, and there are several scoring functions available, taking various parameters into account, such as non-bonded interactions, solvation effects and the torsional angles degrees of freedom.

2.3 QM/MM method

When pure QM calculations are too heavy, but reactions with bond forming and bond breaking needs to be considered (which is usually the case in studies of enzyme mechanisms), at least two different methods can be used. First of all the system can be pruned to a size that is possible to handle with QM methods. Secondly a hybrid method can be utilized. In the QM/MM approach, the system is divided into (at least) two parts – the QM part and the MM part. The Hamiltonian (H_{tot}) is then divided into three parts:

$$H_{tot} = H_{QM} + H_{MM} + H_{QM/MM}$$  \hspace{1cm} (2.16)

where H_{QM} is the Hamiltonian of the QM part, H_{MM} of the MM part and the H_{QM/MM} is the Hamiltonian of the interactions between the QM and MM parts.

The interactions can be described in different levels in the QM/MM methods. The lowest is the mechanical embedding, where all the QM–MM interactions are treated at the MM level. In the electrostatic embedding method the electrostatic interactions of the MM part are taken into account in the QM calculation. The bonded interactions and the non-bonded
van der Waals interactions between the subsystems are treated at the MM level.\textsuperscript{85}

2.4 Computational methods in the current studies

2.4.1 Paper I and II

All calculations in of the work presented in paper I and II are performed in the GAUSSIAN03 program package.\textsuperscript{86} As mentioned above, the B3LYP\textsuperscript{79-81} functional has been used throughout in the DFT calculations. The basis set 6-31+G(d,p) was chosen, with diffuse functions (+), since anions were studied, and with polarized functions (d,p) for both hydrogen and heavier elements, since the protonation and hydrogen bonding was important.

The optimizations were carried out in gas phase and in bulk solvation using the integral equation formalism of the polarizable continuum model (IEFPCM).\textsuperscript{87,88} In these calculations, water was used as a solvent, with the value 78.4 of the dielectric constant ($\varepsilon$). To investigate how the different molecules behave in a highly non-polar environment such as lipid membranes, single point calculations with IEFPCM were also carried out with $\varepsilon = 4.0$, in both studies. Harmonic vibrational frequency calculations were performed in aqueous and lipid environments, to obtain zero-point vibrational energies (ZPE), thermal corrections to the Gibbs’ free energies and the enthalpy at the temperatures 298.15 K (P I & P II) and 310.15 K (P I).

Proton affinities were determined as the difference between the ZPE-corrected internal energies of the protonated and non-protonated forms. For proton affinity calculations, in reactions involving solvated protons, the estimated solvation energy of H\textsuperscript{+} of 267.68 kcal/mol previously established in our group,\textsuperscript{89} was employed to obtain reaction free energies.

The accuracy of gradient corrected DFT for the calculations of proton affinities is well studied, and generally lies within 1–7 kcal/mol in comparison with experimental data.\textsuperscript{90,91}

2.4.2 Paper III

In this study MD has been used to calculate the behaviour of the solutes in a membrane. For this, the GROMACS program package\textsuperscript{92} and its built-in GROMACS standard force field has been used. The membrane model used was downloaded from the webpage of P. Tieleman and his group.\textsuperscript{93,94} The model consists of 64 dipalmitoylphosphatidylcholine (DPPC) molecules and 3846 water molecules forming a $4.9 \times 4.4 \times 9.4$ nm periodic box. A short equilibration run of 1 ns was performed of the pure membrane for verification.
The simulations were performed with periodic boundary conditions with the NPT ensemble at 1 atm and 323 K, sufficiently above the DPPC melting point (mp = 305 K computationally\(^95\) and 315 K experimentally\(^96\)).

The Particle-Mesh Ewald (PME) electrostatics, Nose-Hoover temperature coupling and Parrinello-Rahman semi-isotropic pressure coupling has been used during the simulations. The time step was set to 2 fs, and all bond lengths were constrained during the simulation.

Simulations were performed on systems including one, two and four solute molecules; inserted in the middle of the bilayer, where the lipid density is low. After equilibration of 2-7 ns a production run of 20 ns was performed. The densities of the solutes were calculated based on these production simulations.

For each of the solutes; free energy profiles across the membrane, local diffusion constants and total permeability constants were calculated. However, the latter two could not be performed for the hexyl ester, due to problems in the calculations.

The free energy, diffusion and permeability calculations in this study is based on the methodology first applied by Marrink and Berendsen.\(^97,98\) Snapshots from the equilibration and production simulations were used as input structures for a number of 1 ns MD-simulations. The snapshots were chosen in order that the distance between the solute centre of mass and the bilayer centre of mass possessed discrete steps from 0 to 45 Å (i.e. 46 snapshots per solute). During the 1 ns simulation, this solute–bilayer distance was constrained along the bilayer normal (the z-direction). The force required to maintain the constraint of the solute was obtained in every MD time step.

The free energy difference (\(\Delta G(z)\)) of solute transfer between the bulk water and a specific distance (\(z\)) to the bilayer is calculated by integration of the average over time of force acting on the solute (\(\langle F(z') \rangle_t\)):\(^{2.17}\)

\[
\Delta G(z) = -\int_{z_{water}}^{z_{bulk}} \langle F(z') \rangle_t dz'
\]

(2.17)

The lower limit of the integral was set to 45 Å, as a reference point for bulk water, which corresponds to the furthest distance in the study.

The local 1-dimension diffusion constants (\(D(z)\)) along the bilayer normal were calculated from:

\[
D(z) = \frac{(RT)^2}{\int_0^\infty \bar{C}(t)dt}
\]

(2.18)
where $R$ is the gas constant, $T$ the temperature and $C(t)$ is the autocorrelation function of the constrained force.

The permeability constant $P$, was then obtained according to:

$$P = \frac{1}{\int_{-\infty}^{\infty} [D(z')]^{-1} \exp \left( \frac{\Delta G(z')}{RT} \right) dz'} \tag{2.19}$$

where the limits were approximated to $-45$ and $45$ Å, respectively.

### 2.4.3 Paper IV and V

In both studies the system of the enzyme PBGS was pruned to a reduced part of the enzyme; containing only the substrate and parts of the two covalent bonding residues of the active site.

The calculations were as in the studies presented in P I and P II, performed in GAUSSIAN03,\textsuperscript{86} with B3LYP level of theory.\textsuperscript{79-81} In these studies the 6-31G(d) basis set was used. In comparison with P I and P II, a smaller basis set was chosen since the systems are slightly larger, and the diffuse functions do not make any difference since no anions were studied.

Optimized geometries were obtained in gas phase, and their corresponding harmonic vibrational frequencies were calculated in gas phase and in water bulk (IEFPCM and $\varepsilon = 78.4$). The free energy of the systems in aqueous continuum was calculated by adding the thermal correction value to Gibbs’ free energy obtained from frequency calculations of the system in the gas phase. To validate the transition state optimizations, IRC (intrinsic reaction coordinate) calculations were performed.\textsuperscript{29,100}

### 2.4.4 Paper VI

In this study the enzyme UROD was pruned to the URO-III substrate and the surrounding active site (i.e. the first shell residues and water molecules of the active site), and applied to the QM/MM approach. Only parts of the substrate and the side chains of the residues suggested to be involved in the catalytic reaction was involved in the QM part.

To generate input structures for the QM/MM calculations MM dockings, minimizations and MD simulations were performed in the Molecular Operating Environment (MOE) program,\textsuperscript{101} CHARMM22 force field was used, since it is well parameterized for heme-like compounds.\textsuperscript{102}

For the QM/MM calculations the GAUSSIAN03 program\textsuperscript{86} was used in conjunction with the ONIOM code with mechanical embedding.\textsuperscript{103-109} The level of theory in the QM part was B3LYP and 6-31G(d) basis set,\textsuperscript{79-81} and in the MM part: AMBER94 force field.\textsuperscript{110}
Furthermore, calculations of the proton affinities were performed of small models of the substrate and involved residues in order to choose which reactant complex to start with. The B3LYP/6-31G(d) level of theory was used in these calculations and the IEFPCM\textsuperscript{87,88} solvating scheme was used in order to compare the proton affinities at three different dielectric constant values ($\varepsilon = 4, 10$ and 78.39). The latter constant is that of water, and the two former have earlier been used to model the polarity within enzymes.\textsuperscript{111}

### 2.5 Computational facilities

Several computer systems have been used for these studies; in house facilities in Örebro, resources from Swedish National Infrastructure for Computing (SNIC) at National Supercomputer Centre (NSC) in Linköping and clusters in Canada (see Table 2.1).

<table>
<thead>
<tr>
<th>System</th>
<th>Location</th>
<th># of nodes</th>
<th># of CPUs / node</th>
<th>RAM/node (GiB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydra</td>
<td>Örebro</td>
<td>30</td>
<td>1 x s Intel P IV 2.8 GHz</td>
<td>0.5</td>
</tr>
<tr>
<td>Albatross</td>
<td>Örebro</td>
<td>13</td>
<td>2 x d AMD 1 GHz</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>23</td>
<td>2 x q AMD 2.3 GHz</td>
<td></td>
</tr>
<tr>
<td>Phoenix</td>
<td>Karlskoga</td>
<td>66</td>
<td>1 x q Intel Xeon 2.33 GHz</td>
<td>4</td>
</tr>
<tr>
<td>Monolith</td>
<td>Linköping</td>
<td>198</td>
<td>2 x s Intel Xeon 2.2 GHz</td>
<td>2</td>
</tr>
<tr>
<td>Neolith</td>
<td>Linköping</td>
<td>6440</td>
<td>2 x q Intel Xeon E5345 2.3 GHz</td>
<td>16/32</td>
</tr>
<tr>
<td>Sharcnet</td>
<td>Canada</td>
<td>168-384</td>
<td>2-4 x s/d/q AMD Opteron/ Intel Xeon 2.2-2.5 GHz</td>
<td>4-32</td>
</tr>
</tbody>
</table>

\textsuperscript{s} = single, \textsuperscript{d} = double and \textsuperscript{q} = quadruple core.

**Table 2.1 The computational facilities used in the current studies.**

---

CHAPTER 3

### 3. Summary of results

Figure 3.1 shows an overview of the studies of 5-ALA presented in the current thesis. The first two papers present studies of the properties of 5-ALA in solution (\textsuperscript{1}). Further, 5-ALA has been studied in a cell membrane; how it and its esters behave and how high the barriers are to enter the cell (P III) (\textsuperscript{2}). Step number (\textsuperscript{3}) in the figure represents the metabolism of 5-ALA in the heme biosynthesis, which has been studied in part (P IV-VI).

In the following sections a summary of the results described in the papers will be presented and discussed.

---

Figure 3.1 A schematically overview of the outlines of 5-ALA-PDT and the studies presented in the current thesis. 1) Investigations of 5-ALA and its derivatives in aqueous solution. 2) The penetration of 5-ALA into the cell. 3) The metabolism of 5-ALA.
CHAPTER 3

3 Summary of results

Figure 3.1 shows an overview of the studies of 5-ALA presented the current thesis. The first two papers present studies of the properties of 5-ALA in solution (1). Further, 5-ALA has been studied in a cell membrane; how it and its esters behave and how high the barriers are to enter the cell (P III) (2). Step number (3) in the figure represents the metabolism of 5-ALA in the heme biosynthesis, which has been studied in part (P IV-VI).

In the following sections a summary of the results described in the papers will be presented and discussed.

Figure 3.1 A schematically overview of the outlines of 5-ALA-PDT and the studies presented in the current thesis. 1) Investigations of 5-ALA and its derivatives in aqueous solution. 2) The penetration of 5-ALA into the cell. 3) The metabolism of 5-ALA.

3.1 QM calculations of 5-ALA and its derivatives (P I & P II)

The aim of the study presented in P I was to theoretically investigate the properties of 5-ALA and three of its esters; the methyl-, ethyl- and the hexyl ester. Even if a lot of experimental and clinical studies have been
performed on 5-ALA and its esters, no theoretical study of the drugs has been performed before.

In the second paper the reactivity of 5-ALA is described. 5-ALA is a ketone that can undergo keto-enol tautomerization to its enolic forms. There are four different enolic tautomers of 5-ALA, as already mentioned in chapter 1.3. The double bond can be formed either between C3 and C4 or between C4 and C5 in 5-ALA (Figure 1.3). Furthermore, for each of them the two Z and E symmetries can be formed. In addition, a hydrated form of 5-ALA is formed in solution, with a mol fraction of ~0.5%.70 P II describes the theoretical study of these derivatives of 5-ALA and its interconventional reaction mechanisms (Figure 1.7).

### 3.1.1 Structural properties

The optimized structures of 5-ALA’s different protonation states are shown in Figure 3.2. The figure displays that all of them except the zwitterion have an all-trans orientation. The zwitterion (Figure 3.2c) instead has a ring structure with a hydrogen bond (1.497 Å) between the carboxylic acid and the amine moieties of the molecule.

![Figure 3.2 Optimized structures of a) 5-ALA, b) 5-ALA\(^{-}\), c) 5-ALA\(^{+}\) and d) 5-ALA\(^{+}\).](image)

The optimized structures of the neutral and cationic forms of the three 5-ALA esters studied are displayed in Figure 3.3. The methyl- and ethylcationic esters (Figure 3.3 d and f) also have an all-trans geometry, while
the longer hexyl ester cation is slightly twisted. The neutral esters however (Figure 3.3 a, c and e), have a 90 degrees twisted torsion angle between the carbonyl and the carboxylic acid group.

Figure 3.3 Optimized structures of the neutral (a, c, e) and cationic forms (b, d, f) of a-b) Hexyl-, c-d) Ethyl- and e-f) Methyl 5-ALA esters.

3.1.2 Free energies

According to calculated free energies, the order of the solutes’ stability in water (IEFPCM solvation) is (cf. Figure 3.4a):

$$5\text{-ALA}^+ > 5\text{-ALA} + H^+ > 5\text{-ALA}^\pm + H^+ > 5\text{-ALA}^- + 2 H^+$$

where the energy used for the proton is specified in method section. In lipid environment however, the neutral form is the most stable (Figure 3.4b).

The esters are most stable in their protonated form in aqueous solution, whereas in lipid environment deprotonation is spontaneous by 6.1 – 7.4 kcal/mol.
5-Aminolevulinic acid and derivatives thereof

Figure 3.4 Mass balanced relative free energies (kcal/mol) in A) water and B) lipid environment. The energy of 5-ALA$^+$ (aq) + H$_2$O (aq) is set to zero.
The hydrolysation free energy of the non-charged esters according to following reaction:

\[ X\text{-5-ALA} + \text{H}_2\text{O} \rightleftharpoons \text{5-ALA} + X^{-}\text{OH} \quad X= \text{Me, Et, He} \]

range from thermoneutral for the methyl ester, to be spontaneous by 5.5 kcal/mol for the hexyl ester. Hence, there is a clear relation between ester chain length and hydrolysation free energy.

As shown in Figure 3.4, all the 5-ALA enolic forms are thermodynamically less stable than the keto form in both water and lipid environment. The stability of the protonation state of the enols follows the same pattern as 5-ALA, with two exceptions; the zwitterion of the Z 3enol isomer is slightly more stable than the neutral form in aqueous solution and the cation of 5-ALA-hyd is more stable than its neutral form in lipid environment.

The enols are on the same energy level, differing with only a few kcal/mol, not following any clear pattern. The hydrated forms are less thermodynamically stable than all the enols in both lipid and aqueous environment, except the cation in aqueous solution, which is slightly more stable than the most stable enol (Figure 3.4a).

### 3.1.3 Proton affinities

The calculated proton affinities (PA) in vacuum are in very well agreement with theoretical and experimental data of the similar compounds: alanine, glycine and levulinate as shown in Table 3.1. The first column of data lists the protonation of the nitrogen atom, and the second data column, the protonation of the carboxylic oxygen atom. The difference between the PA of oxygen and nitrogen in gas phase is large, but in solution they are much closer (Table 3.2). This is expected since ions are much more stable in solution than in gas phase. The 5-ALA esters are stronger proton bases than 5-ALA itself in gas phase but weaker bases in aqueous and lipid environment (Table 3.2).

As seen in Table 3.2 and Figure 3.5, all solutes except 5-ALA-hyd\textsuperscript{−−} are more likely to be protonated to their cationic form in water than in lipid, while the anions take up a proton in lipid much easier than in water. This is a predicted result since ionized compounds are not stable in non-polar fluids. It is also found that the neutral and zwitterionic forms of 5-ALA-hyd are more likely protonated than any of their corresponding 5-ALA or 5-ALA enol compounds. Finally, the enols (in particular their anions) with \(E\) symmetry have slightly higher PAs than their corresponding \(Z\) enols, with a few exceptions. Thus, the \(E\) symmetry enols are generally stronger bases than the \(Z\) enols.
Table 3.1 Proton affinities (kcal/mol) of 5-ALA in comparison with similar compounds in vacuum.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Method</th>
<th>Nitrogen</th>
<th>Oxygen</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-ALA</td>
<td>Theoretical ‡</td>
<td>213.8</td>
<td>330.3</td>
</tr>
<tr>
<td>Glycine</td>
<td>Experimental</td>
<td>211.6 a</td>
<td>342.4 e</td>
</tr>
<tr>
<td></td>
<td></td>
<td>212 b</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>208.2 c</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>211.9 d</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Theoretical</td>
<td>211 f</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>211.5 g</td>
<td></td>
</tr>
<tr>
<td>Alanine</td>
<td>Experimental</td>
<td>214.8 a</td>
<td>340.7 e</td>
</tr>
<tr>
<td></td>
<td></td>
<td>212.3 b</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>212.2 c</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>215.5 d</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Theoretical</td>
<td>215.5 g</td>
<td></td>
</tr>
<tr>
<td>Levulinate</td>
<td>Experimental</td>
<td></td>
<td>340.7 h</td>
</tr>
</tbody>
</table>

‡ This work; also displayed in Table 3.2. References: a: 112, b: 113, c: 114, d: 115, e: 116, f: 117, g: 91, h: 118.

Figure 3.5 Proton affinities of the solutes in water and in lipid environment.
Table 3.1 Proton affinities (kcal/mol) of 5-ALA in comparison with similar compounds in vacuum.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Method</th>
<th>Nitrogen</th>
<th>Oxygen</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-ALA</td>
<td>Theoretical</td>
<td>213.8</td>
<td>330.3</td>
</tr>
<tr>
<td>Glycine</td>
<td>Experimental</td>
<td>211.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a</td>
<td>212</td>
<td></td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>208.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c</td>
<td>211.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>d</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>e</td>
<td>342.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>f</td>
<td>211.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>g</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>h</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

‡ This work; also displayed in Table 3.2. References: a: 112, b: 113, c: 114, d: 115, e: 116, f: 117, g: 91, h: 118.

Table 3.2 Calculated proton affinities (kcal/mol) of 5-ALA and its esters in gas phase, aqueous and lipid environment.

<table>
<thead>
<tr>
<th>Atom</th>
<th>Vacuum</th>
<th>Water</th>
<th>Lipid</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-ALA</td>
<td>N</td>
<td>213.8</td>
<td>279.4</td>
</tr>
<tr>
<td>5-ALA(^{-})</td>
<td>O</td>
<td>–</td>
<td>280.9</td>
</tr>
<tr>
<td>5-ALA(^{-})</td>
<td>O</td>
<td>330.3</td>
<td>279.6</td>
</tr>
<tr>
<td>5-ALA(^{-})</td>
<td>N</td>
<td>–</td>
<td>277.8</td>
</tr>
<tr>
<td>Me-5-ALA</td>
<td>N</td>
<td>217.8</td>
<td>276.8</td>
</tr>
<tr>
<td>Et-5-ALA</td>
<td>N</td>
<td>217.1</td>
<td>277.5</td>
</tr>
<tr>
<td>He-5-ALA</td>
<td>N</td>
<td>225.6</td>
<td>276.5</td>
</tr>
</tbody>
</table>

3.1.4 Tautomerization mechanism

The tautomerization process of the Z isomers of 5-ALA-3enol and 5-ALA-4enol to the ketone form (5-ALA) has been studied, as well as the hydration of 5-ALA to 5-ALA-hyd. The results show that the tautomerization of 5-ALA is a hydrogen transfer reaction; due to the fact that the Mulliken charge on the hydrogen atom hardly change during the process.

Three mechanisms were tested for both the 3enol and the 4enol; direct transfer, transfer via a bridging water molecule, and a self-catalysed mechanism. The direct transfer was the least favourable mechanism with an activation energy of 68.2 and 57.6 kcal/mol, respectively. The bridging water molecule did almost halve the barriers to 36.5 and 30.9 kcal/mol, respectively. Anyway, the most probable mechanism was the self-catalysed one (shown in Figure 3.6), with the rate-determining activation energy of 10.8 and 15.1 kcal/mol for the 3enol and the 4enol, respectively (Figure 3.7). The carboxylic acid oxygen picks up the enol hydrogen (TS1) and moves it like a ‘crane’ (TS2) to the carbon atom in question (C3 or C5).

The hydration process of 5-ALA has a high activation energy of 35 kcal/mol even with an explicit water molecule included in the model. This is not in agreement with experimental data that show that 5-ALA-hyd is more common than the enolic forms.\(^70\) An explanation could be the pH dependence of the reactions or the influence of the phosphate buffer used in the experiments.
Figure 3.6 The tautomerization mechanism a) 5-ALA-3enol$^{+−}$ to 5-ALA$^{+−}$ b) 5-ALA-4enol$^{+−}$ to 5-ALA$^{+−}$. (Distances are shown in Å).
Figure 3.6 The tautomerization mechanism a) 5-ALA-3enol+− to 5-ALA+− and b) 5-ALA-4enol+− to 5-ALA+−. (Distances are shown in Å).

Figure 3.7 Relative reaction free energy potential energy surface of the tautomerization of  A) 5-ALA-3enol+−, and B) 5-ALA-4enol+− (the energies are given in kcal/mol).
3.2 MD simulations of 5-ALA and its esters in membrane (P III)

To elucidate the permeability of 5-ALA and its methyl-, ethyl- and hexyl esters through cell membranes, MD simulations have been performed of them within a model membrane consistent of a bilayer of DPPC lipid molecules (Figure 3.8). It has previously been proved that various transfer processes are involved in the transport of 5-ALA, but passive diffusion may play a role in the transport of the 5-ALA esters. The current study focuses on the passive diffusion of the solutes over the membrane bilayer.

After equilibration, the molecules are located at different depths into the lipid bilayer, where they stay during the production simulation. The neutral 5-ALA is most of the time found in the interface, between the lipid-tail and the polar head-group regions (Figure 3.8a). The more hydrophilic zwitterion is predominately located in the head-group region, which is expected since the zwitterion is doubly charged. The esters are more distributed in the membrane, between the head-group and the lipid-tail regions. The two esters with the longest chains have their polar part predominantly in contact with the polar head groups and its apolar part directed to the middle of the membrane (cf. Figure 3.8b).

**Figure 3.8** A unit cell of the equilibrated membrane with A) one neutral 5-ALA molecule, and B) one neutral He-5-ALA molecule.

The free energy profile of the membrane penetration was calculated for the systems and is plotted in Figure 3.9 against the distance to the bilayer centre of mass. All solutes have similar curve shapes; a well at 1–2 nm from the bilayer centre of mass and a local energy maximum in the middle. This means that all solutes are attracted by the polar lipid head groups (with highest density at 1–2 nm), and have varying difficulties of passing the most apolar, but less dense, middle of the membrane.

The zwitterionic 5-ALA molecule has a much higher maximum in the middle than the neutral 5-ALA (66 and 33 kJ/mol, respectively). There is no apparent connection between ester length and barrier height, since Me-5-ALA has the lowest barrier of all compounds (26 kJ/mol) and Et-5-ALA has a higher barrier than He-5-ALA (49 and 39 kJ/mol, respectively).

The permeability constants of 5-ALA, the methyl- and the ethyl ester were calculated (Table 3.3). The zwitterion has the lowest permeability constant; six orders of magnitude lower than neutral 5-ALA. The permeability of 5-ALA and Me-5-ALA are in the same magnitude, with a slightly higher permeability of the ester. The ethyl ester is not penetrating as fast as the methyl ester, since it has both a broader well in the head-group region, and a higher maximum in the middle of the membrane.
To elucidate the permeability of 5-ALA and its methyl-, ethyl- and hexyl esters through cell membranes, MD simulations have been performed of them within a model membrane consistent of a bilayer of DPPC lipid molecules (Figure 3.8). It has previously been proved that various transfer processes are involved in the transport of 5-ALA, but passive diffusion may play a role in the transport of the 5-ALA esters. The current study focuses on the passive diffusion of the solutes over the membrane bilayer.

Figure 3.8 A unit cell of the equilibrated membrane with A) one neutral 5-ALA molecule, and B) one neutral He-5-ALA molecule. After equilibration, the molecules are located at different depths into the lipid bilayer, where they stay during the production simulation. The neutral 5-ALA is most of the time found in the interface, between the lipid-tail and the polar head-group regions (Figure 3.8a). The more hydrophilic zwitterion is predominately located in the head-group region, which is expected since the zwitterion is doubly charged. The esters are more distributed in the membrane, between the head-group and the lipid-tail regions. The two esters with the longest chains have their polar part predominantly in contact with the polar head groups and its apolar part directed to the middle of the membrane (cf. Figure 3.8b).

The free energy profile of the membrane penetration was calculated for the systems and is plotted in Figure 3.9 against the distance to the bilayer centre of mass. All solutes have similar curve shapes; a well at 1–2 nm from the bilayer centre of mass and a local energy maximum in the middle. This means that all solutes are attracted by the polar lipid head groups (with highest density at 1–2 nm), and have varying difficulties of passing the most apolar, but less dense, middle of the membrane.

The zwitterionic 5-ALA molecule has a much higher maximum in the middle than the neutral 5-ALA (66 and 33 kJ/mol, respectively). There is no apparent connection between ester length and barrier height, since Me-5-ALA has the lowest barrier of all compounds (26 kJ/mol) and Et-5-ALA has a higher barrier than He-5-ALA (49 and 39 kJ/mol, respectively).

The permeability constants of 5-ALA, the methyl- and the ethyl ester were calculated (Table 3.3). The zwitterion has the lowest permeability constant; six orders of magnitude lower than neutral 5-ALA. The permeability of 5-ALA and Me-5-ALA are in the same magnitude, with a slightly higher permeability of the ester. The ethyl ester is not penetrating as fast as the methyl ester, since it has both a broader well in the head-group region, and a higher maximum in the middle of the membrane.
Table 3.3 Calculated permeability coefficients (P).

<table>
<thead>
<tr>
<th>Solute</th>
<th>P / cm s$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-ALA$^{+}$</td>
<td>$6.44 \times 10^{-5}$</td>
</tr>
<tr>
<td>5-ALA</td>
<td>$1.89 \times 10^{1}$</td>
</tr>
<tr>
<td>Me-5-ALA</td>
<td>$5.28 \times 10^{1}$</td>
</tr>
<tr>
<td>Et-5-ALA</td>
<td>$7.45 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

To fully understand how the ester chain length affects the permeability, further investigations have to be performed with a wider spectrum of esters. On the basis of present results, it is not clear where the trend is broken. According to experimental studies, both the methyl- and the hexyl ester diverge from the linearity of increased ester chain length vs. penetration of the molecule. Cells in vitro produce less PpIX from the methyl ester than from 5-ALA, although the picture changes to show an increase from ethyl to pentyl ester in PpIX levels. For hexyl and longer ester chains the produced PpIX levels once again becomes lower in level.$^{124,125}$

3.3 Enzymatic reactions (P IV-VI)

The aim of the third part of the current work is to study the metabolism of 5-ALA. The story does not end when the drug molecules finally have found their way to the target cells and penetrated their cell membranes. To become photoactive, eight 5-ALA molecules are combined together to protoporphyrin IX (PpIX) within the cell. This process follows the heme biosynthesis and is catalysed by in total six enzymes in the cytosol and mitochondria (cf. Figure 1.4). Two of these enzymes were studied; Porphobilinogen synthase (PBGS) and Uroporphyrinogen-III decarboxylase (UROD).

3.3.1 The mechanism of PBGS (PIV-V)

The PBGS enzyme catalyses the asymmetrical condensation of two 5-ALA molecules. The active site contains, as already mentioned in chapter 1, two catalytic lysine residues. The two 5-ALA substrates bind to each of the two lysines by Schiff base linkages. The study of the formation of a Schiff base is described in P IV and the mechanism of the cyclization reaction of PBGS is described in P V. The energies are given as free energies in aqueous solution at 298.15 K and 1 atm ($\Delta G_{aq}^{298}$) unless otherwise noted.

3.3.1.1 The Schiff base formation

The Schiff base formation is divided into two steps; the carbinolamine formation and the Schiff base formation. To orient the carbinolamine for
the Schiff base formation step, some rearrangements have to be done, but their barriers are not high enough to affect the rate of the reaction.

![Figure 3.10](image) A schematic picture of how the system is pruned in the study of the Schiff base formation. The atoms within the white box are included in all systems, whereas the atoms in the light grey boxes are included in the listed systems beside each box and the atoms in the dark grey areas were not included in any system.

The calculations were performed on small model systems, where the lysine residues were pruned to a methylamine \([\text{CH}_3\text{NH}_2]\), and the substrate molecule (5-ALA) was either included as a whole or pruned to aminoacetone \([\text{NH}_2\text{CH}_2\text{C}(=\text{O})\text{CH}_3]\) or acetone \([\text{CH}_3\text{C}(=\text{O})\text{CH}_3]\). The content of each system is schematically shown in Figure 3.10. Both neutral and protonated systems were studied as well as the effect of an explicit water molecule. The names of the systems in Figure 3.10 are based on their content; 5-ALA (5-), aminoacetone \((\text{Aa})\), acetone \((\text{A})\), methylamine \((\text{M})\), water \((\text{W})\) and protonated \((\text{H}^+)\).

The results show that an explicit water molecule catalyses the carbinolamine formation step of the neutral system and halves the barrier from 36.6 to 18.1 kcal/mol (Figure 3.11).

In the protonated systems, the carbinolamine formation barrier is lowered to 22.3–26.3 kcal/mol when the amino group of the substrate is acting as a proton transporter \((5-\text{MH}^+, 5-2\text{MH}^+, \text{Aa2MH}^+)\). However, when a second lysine residue transfer the proton from the nucleophilic amine to the carbonyl oxygen atom \((\text{A2MH}^+ \text{ and A2MW}^+)\), the barrier is further reduced to 16.2–20.1 kcal/mol (Figure 3.12).
Figure 3.11 The neutral systems. The lowest barriers for the various mechanisms of (left) the carbinolamine formation (TS1) and (right) the Schiff base formation step (TS3). The free energy activation energies are given in kcal/mol.

Figure 3.12 The protonated systems. The lowest barriers for the various mechanisms of (upper row) the carbinolamine formation (TS1) and (bottom row) the Schiff base formation (TS3). The free energy activation energies are given in kcal/mol.
The Schiff base formation barrier in the neutral systems is halved by a bridging water molecule (from 48.7 kcal/mol of the AM system to 26.2 kcal/mol of the AMW system), but even more reduced by the catalysis of the carboxylic acid moiety of 5-ALA (5-M system) to 14.1 kcal/mol (Figure 3.11).

Water catalyse the Schiff base formation of the protonated system (AMWH+) with a barrier of 21.3 kcal/mol, while 5-ALA amine catalysis (5-MH+) reduces the barrier to 9.0 kcal/mol. However, the addition of a second lysine (5-2MH+, A2MWH+, Aa2MH+ and A2MH+) reduces the barrier even further to 6.0–7.5 kcal/mol (Figure 3.12).

Hence, we can conclude that the second lysine is the best catalyst for both the carbinolamine formation and the Schiff base formation steps.

3.3.1.2 Schiff base transfer

It has been proposed that the 5-ALA substrate is first bound to the A-site lysine, followed by a transfer of the Schiff base to the P-site lysine, so that a second 5-ALA molecule can bind to the A-site (cf. Figure 1.5). We therefore studied this transfer process of a neutral and a protonated model system. The system was pruned in the same way as in the Schiff base formation study (P IV). The input structures include 5-ALA bound with a Schiff base to a methylamine representing the A-site lysine, and a second methylamine as a model for the free P-site lysine.

In the neutral system the nucleophilic attack of the P-site lysine and the proton transfer from P-site amine to the deprotonated A-site imine is found to be a one-step reaction. This process is catalysed by the carboxylic acid of 5-ALA, with a barrier of 13.0 kcal/mol. From the now formed aminal (two amines bonded to the same carbon), the cleavage of the A-site C–NH bond is attained by the reversed reaction, as shown in Figure 3.13a as a mirror of TS1, with a barrier of 7.2 kcal/mol.

![Figure 3.13](image)

Figure 3.13 Relative potential energy surfaces in gas phase ($\Delta E_d$ in red) and in water ($\Delta E_{aq}^{298}$ in green) given in kcal/mol for the neutral (A) and protonated (B) system of the Schiff base transfer.
With a protonated A-site Schiff base the P-site lysine can attack the A-site without the need of a proton transfer. This process has a barrier of 7.4 kcal/mol. However, a proton needs to be transferred between the amines in order to break the A-site C–NH bond. This proton transfer can be catalysed by the amine of 5-ALA. The proton is transferred from A-site nitrogen to 5-ALA amine by a cost of 6.9 kcal/mol. The cleavage of A-site C–NH bond is as in the neutral system attained by the mirrored reaction (Figure 3.13b) which has the corresponding barriers 1.2 and 3.3 kcal/mol, respectively.

Figure 3.14 The Path 1 mechanism of PBGS, where the C–N inter-substrate bond is formed first.\(^{30,126}\)
3.3.1.3 Cyclization reaction mechanism

When the two substrates are bound to the active site, the cyclization reaction is catalysed. Three different mechanisms were compared in the study. The differences in these mechanisms lie in which order the intersubstrate bonds are formed. According to the first path, the C–N bond is formed prior to the C–C bond (Figure 3.14). The second and third paths suggest that the C–N bond is formed after the C–C bond. In the second path the C–N bond is formed subsequent to the C–C bond formation (Figure 3.15), whereas in the third path the enzyme substrate linkage is broken prior in between the two intersubstrate bond formations (Figure 3.16).

Figure 3.15 The Path 2 mechanism of PBGS, where the C–C inter-substrate bond is formed first. The last steps (1E-1H) follow the same mechanism as Path 1 in Figure 3.14.

The potential energy surfaces of the minimized local minima of the three mechanisms are plotted in Figure 3.17. The transition state calculations show that the highest barrier of path 1 is the first hydrogen transfer step (1C–1D) with 20.1 kcal/mol. The highest barrier in path 2 is the formation of the C–C bond (2B-2C), which is the first intersubstrate bond formed. The activation energy of this step is 19.4 kcal/mol. However, some of the proton transfer barriers are not investigated, and therefore we can not tell for sure if these are the highest barriers. Anyway, the calculated activation energies are in good agreement with experimental activation energy investigations, which found a barrier of 18.4 kcal/mol.
Figure 3.16 The Path 3 mechanism of PBGS, where the C–C inter-substrate bond is formed first, but the P-site enzyme-substrate bond is broken prior to the cyclization step.\textsuperscript{32,34,38} The first steps (1A–2C) follow the same mechanism as Path 2 (Figure 3.15) and the last step (1G–1H) the same as Path 1 (Figure 3.14).

The third path follows the same steps as path 2 with the C–C bond formation barrier mentioned above. The unique steps of the third path have very low calculated barriers. However, the activation energy of the 2C–3D step was not calculated, which is known to be a very endergonic step (10.9 and 17.0 kcal/mol for the formation of 3D with E and Z symmetry, respectively). In addition, the proton transfer of the 2C–3D step proceeds via a four-membered ring transition structure if there is not another residue that can catalyse the proton transfer. Maybe this would be possible with a catalysing active site water molecule present.

Since the highest barrier of path 1 is a hydrogen transfer, it is probable that this step can be catalysed by a base (e.g. the A-site lysine), which can reduce this barrier. Furthermore, many of the steps are less endergonic in path 1, compared to the other two. Therefore, we believe that this path is the most probable. However, the effect of the surrounding active site cavity and the zinc ion is not considered in this study, which can influence the barriers and the geometries.
The third path follows the same steps as path 2 with the C–C bond formation barrier mentioned above. The unique steps of the third path have very low calculated barriers. However, the activation energy of the $2C-3D$ step was not calculated, which is known to be a very endergonic step (10.9 and 17.0 kcal/mol for the formation of $3D$ with $E$ and $Z$ symmetry, respectively). In addition, the proton transfer of the $2C–3D$ step proceeds via a four-membered ring transition structure if there is not another residue that can catalyse the proton transfer. Maybe this would be possible with a catalysing active site water molecule present.

Since the highest barrier of path 1 is a hydrogen transfer, it is probable that this step can be catalysed by a base (e.g. the A-site lysine), which can reduce this barrier. Furthermore, many of the steps are less endergonic in path 1, compared to the other two. Therefore, we believe that this path is the most probable. However, the effect of the surrounding active site cavity and the zinc ion is not considered in this study, which can influence the barriers and the geometries.

**Figure 3.17** Potential energy surfaces in aqueous solution ($\Delta G_{aq}^{298}$) of Path 1 (—black), Path 2 ( - - -red) and Path 3 Z ( – – blue) and E symmetry (- – - green).
3.3.2 The mechanism of UROD (PVI)

The enzyme UROD catalyses the decarboxylation of the acetyl side chains of URO-III to CP-III. The mechanism of the first decarboxylation of the ring D acetyl has been investigated in the current study.

The initial docking study generated 30 structures of the enzyme—URO-III systems. The ones of these that were mechanistically relevant can be divided into three groups based on the acidic residue coordinated to the acetate on ring D. The three residues are: i) Arg50, ii) Tyr164 and iii) His339. The top scored structures in each group was chosen for further investigations as conformer I, II and III, respectively (Figure 3.18).

It was found that conformer I had the strongest interaction energy to the enzyme (~246.3 kcal/mol), while conformer II and III had weaker interactions by 14.5 and 22.8 kcal/mol, respectively.

![Figure 3.18 Schematic illustration of the three active site-bound substrate conformations: I, II and III. A and P denotes the substrates acetate and propionate groups, respectively and substrate—enzyme/solvent interactions are indicated by dashed lines.](image)

It has earlier been suggested that the solvent is acting as the acid in the first step of the mechanism (Figure 1.6). However, the distance between C2 of pyrrole D (Dc2) and the closest water molecule is in the three conformers between 5.82 and 7.31 Å. Therefore, we conclude that there must be another catalysing acid. The average distances between the a guanidinium proton of Arg37 and Dc2 during the MD simulations were calculated to 2.41, 2.40 and 2.97 Å for conformer I, II and III, respectively. This means that Arg37 is more likely to catalyse this step than water, and at least the former two distances are in the range of a hydrogen bond.

To examine if Arg37 is a good candidate to act as the acid HA in Figure 1.6, the active site environment effect on the proton affinities (PAs) of Arg37 and Dc2 was investigated. The PAs were calculated for models representing Arg37 and the pyrrole moiety, in a bulk solvation with the dielectric constants set to 78.39, 10.0 and 4.0. The results show that arginine has
higher proton affinity than the $D_{c2}$ before the decarboxylation in water ($\varepsilon = 78.39$), but lower values when the dielectric constant is reduced. If the pyrrole amine is coordinated by hydrogen bond to Asp86 the difference in PA is further increased in the calculations with lower dielectric constants. This means that the $D_{c2}$ is able to take the proton from Arg37 in the active site of UROD. In the last step in Figure 1.6, since the PA of $D_{c2}$ of the product is lower than the guanidinium at all dielectric constants $D_{c2}$ is also able to donate the proton back to Arg37.

The PA was calculated for the structure that corresponds to the protonation of $D_{c3}$ after the decarboxylation step (third structure in Figure 1.6). In comparison with the PA of arginine, it is found that that Arg50 is able to act as the acid HB and protonate $D_{c3}$; independent of the dielectric constant when the pyrrole ring is coordinated to the acetate.

Large system QM/MM calculations were performed on conformer I in order to calculate the energy barriers for the reaction. Conformer I was chosen on the following basis: a) conformer I is proven to have the strongest interaction energy, b) conformer II and III are questioned since experimental studies have found that Tyr164 and His339 are not essential for the enzymatic mechanism,\textsuperscript{49,50} c) the PA discussion above show that Arg37 and Arg50 are good candidates to be the catalysing residues.

The barriers of the reaction are shown in Figure 3.19. The TS1, which is found to be the rate-determining step with a barrier of 13.7 kcal/mol, re-
fers to the proton transfer from Arg37 to D$_{\text{C2}}$. TS2 refers to the decarboxylation with the concomitant proton transfer from Arg50 to D$_{\text{C3}}$, which is actually found to be spontaneous. TS3 refers to the proton transfer of the initially transferred proton back from D$_{\text{C2}}$ to Arg37. The rate-determining barrier agrees well with experimental results of various species, which are found in the range of 2.0–12.3 kcal/mol.$^{51,128}$
CHAPTER 4

4 Conclusions and future perspectives

5-ALA and Me-5-ALA are now widely used in photodynamic therapy as prodrugs in treatment of actinic keratosis and basal cell carcinoma. The aim of the studies presented in this thesis was to find out more about the properties of 5-ALA and its derivatives by the use of various computational techniques such as DFT, MM, MD and QM/MM. This research has led to more and deeper knowledge on how 5-ALA acts in aqueous solution and in a lipid membrane, and parts of its biosynthesis has been elucidated.

The main results are:

- The studies of the stability of the various protonation states of 5-ALA show that the protonated form is the most stable in aqueous solution (P I).
- The tautomerization reaction of 5-ALA follows a self-catalysed mechanism with a barrier of 15 kcal/mol (P II).
- The free energy profiles and permeability constants show that 5-ALA and its methyl ester are diffusing fastest through a lipid bilayer (P III).
- The reaction mechanisms of PBGS and UROD, where various pathways were compared. The activation energies were found to be in good agreement with experimental data (P IV-VI).

Of course, much more research can be done in this field to find new derivatives with enhanced effectiveness in their drug delivery. More specifically, the studies of the transport over lipid bilayers can be enhanced to involve a wider spectrum of molecules. Furthermore, transport mechanisms with for example transmembrane proteins can be studied.

Computational methods are powerful tools to elucidate the mechanisms of enzymes, and the results are found to be in good agreement with experiments. To evaluate the effects of the active site cavity and the zinc ion, larger models of PBGS with QM/MM methodology can be used. Furthermore, the following three decarboxylation steps in UROD are under further studies. Finally, there are still several other heme enzymes for which the enzymatic reaction is not fully understood. Their mechanisms could also be further investigated by computational methods.
Acknowledgements

First of all I would like to thank Professor Leif Eriksson, who has been my supervisor during these four years of PhD studies. Thank you for all your help and support and thanks for helping me to focus on the right things.

Thank you Dr James Gauld, for that you have generously shared your knowledge in the field of enzyme catalysis and have given me so much of your time and effort in supervision. Thanks also for your (and the rest of the group) hospitality and all your support during my stay at your research group in Windsor, (ON) Canada. A special thanks to Eric Bushnell and Dr Jorge Llano for introducing me to calculations with QM/MM.

Thank you MD Lennart Löfgren for your knowledge on the clinical field and for interesting and improving discussions about 5-ALA.

A special thanks to Dr Daniel dos Santos for supervision in the lipid membrane studies. Thanks for being available on e-mail and instant messaging, not hesitating to answer my questions.

I am very thankful to my workmates in the research group! I wish to give a special thank to my friends in PhD and MSc projects: Klefah, Emma, Li, Viarja, Magnus, Ann-Louise, Min, Boxue, Samuel and Ismael. Thank you postdocs and seniors for help and discussions about work: Yaoquan, Dragan, Patricia, Oles, Salama, David, Sofia and Rubo. Thanks all others in the department and especially you within the Life Science research centre and Modelling and Simulation centre. A special thanks to docent Jana Jass, the head of the Life science research school.

Thank you Lisa for your help in proof reading the thesis!

I would also like to acknowledge the faculty of Science and Technology of Örebro University, the Modelling and Simulation Centre (MoS), the Swedish Science Research Council (VR) and the Swedish Chemical Society for financial support. I also acknowledge the National Supercomputer Centre (NSC) in Linköping and SHARCNET for generous grants of computing time.

Jag skulle även vilja tacka min fru Lisa för din kärlek och att du stått ut med min arbetsbörda speciellt under de senaste månaderna. Tack mamma och pappa för all kärlek och uppskattning. Tack alla mina vänner här i Örebro, och ett speciellt tack till er i Brickebergskyrkan – tack för den goda gemenskapen och era böner! Slutligen, tack Gud för din nåd.
Acknowledgements

First of all I would like to thank Professor Leif Eriksson, who has been my supervisor during these four years of PhD studies. Thank you for all your help and support and thanks for helping me to focus on the right things.

Thank you Dr James Gauld, for that you have generously shared your knowledge in the field of enzyme catalysis and have given me so much of your time and effort in supervision. Thanks also for your (and the rest of the group) hospitality and all your support during my stay at your research group in Windsor, (ON) Canada. A special thanks to Eric Bushnell and Dr Jorge Llano for introducing me to calculations with QM/MM.

Thank you MD Lennart Löfgren for your knowledge on the clinical field and for interesting and improving discussions about 5-ALA.

A special thanks to Dr Daniel dos Santos for supervision in the lipid membrane studies. Thanks for being available on e-mail and instant messaging, not hesitating to answer my questions.

I am very thankful to my workmates in the research group! I wish to give a special thank to my friends in PhD and MSc projects: Klefah, Emma, Li, Viarja, Magnus, Ann-Louise, Min, Boxue, Samuel and Ismael. Thank you postdocs and seniors for help and discussions about work: Yaoquan, Dragan, Patricia, Oles, Salama, David, Sofía and Rubo. Thanks all others in the department and especially you within the Life Science research centre and Modelling and Simulation centre. A special thanks to docent Jana Jass, the head of the Life science research school.

Thanks Lisa, Ann, Klefah, Emma, Yaoquan, Eric and Dragan for your help in proof reading the thesis!

I would also like to acknowledge the faculty of Science and Technology of Örebro University, the Modelling and Simulation Centre (MoS), the Swedish Science Research Council (VR) and the Swedish Chemical Society for financial support. I also acknowledge the National Supercomputer Centre (NSC) in Linköping and SHARCNET for generous grants of computing time.

Jag skulle även vilja tacka min fru Lisa för din kärlek och att du stått ut med min arbetsbörda speciellt under de senaste månaderna. Tack mamma och pappa för all kärlek och uppskattning. Tack alla mina vänner här i Örebro, och ett speciellt tack till er i Brickebergskyrkan – tack för den goda gemenskapen och era böner! Slutfilen, tack Gud för din nåd.
I. 5-Aminolevulinic acid and derivatives thereof

References


References


43. Mills-Davies, NL, Thompson, D, Cooper, JB, Wood, SP and Shoolingin-Jordan, PM, PDB code: 1E5I; Crystal Structure of Native Human Erythrocyte 5-Aminolaevulinic Acid Dehydratase, 2001.


92. GROMACOS v. 3.3.1 / 3.3.2, Spoel, Dvd, Lindahl, E and Hess, B, 2006, University of Groningen, The Netherlands


95. Leekumjorn, S and Sum, AK, Molecular studies of the gel to liquid-crystalline phase transition for fully hydrated DPPC and DPPE bilayers, Biochimica et Biophysica Acta (BBA) - Biomembranes, 2007 1768(2):354-65.


119. Rud, E, Gederaas, O, Hogset, A and Berg, K, 5-Aminolevulinic Acid, but not 5-Aminolevulinic Acid Esters, is Transported into


