Comprehensive Summaries of Uppsala Dissertations
from the Faculty of Science and Technology 965

Modern Computational Physical Chemistry

An Introduction to Biomolecular Radiation Damage and Phototoxicity

BY

JORGE LLANO

ACTA UNIVERSITATIS UPSALIENSIS
UPPSALA 2004

Dissertation presented at Uppsala University to be publicly examined in B42, BMC, Uppsala, Monday, May 3, 2004 at 09:15 for the degree of Doctor of Philosophy. The examination will be conducted in English.

Abstract
Llano, J. 2004. Modern Computational Physical Chemistry. An Introduction to Biomolecular Radiation Damage and Phototoxicity. (Modern fysikalisk-kemisk beräkningsmetodik. En introduktion till biomolekylära strålningsskador och fototoxicitet). Acta Universitatis Upsaliensis. *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 965. 80 pp. Uppsala. ISBN 91-554-5940-4

The realm of molecular physical chemistry ranges from the structure of matter and the fundamental atomic and molecular interactions to the macroscopic properties and processes arising from the average microscopic behaviour.

Herein, the conventional electrodic problem is recast into the simpler molecular problem of finding the electrochemical, real chemical, and chemical potentials of the species involved in redox half-reactions. This molecular approach is followed to define the three types of absolute chemical potentials of species in solution and to estimate their standard values. This is achieved by applying the scaling laws of statistical mechanics to the collective behaviour of atoms and molecules, whose motion, interactions, and properties are described by first principles quantum chemistry. For atomic and molecular species, calculation of these quantities is within the computational implementations of wave function, density functional, and self-consistent reaction field theories. Since electrons and nuclei are the elementary particles in the realm of chemistry, an internally consistent set of absolute standard values within chemical accuracy is supplied for all three chemical potentials of electrons and protons in aqueous solution. As a result, problems in referencing chemical data are circumvented, and a uniform thermochemical treatment of electron, proton, and proton-coupled electron transfer reactions in solution is enabled.

The formalism is applied to the primary and secondary radiation damage to DNA bases, e.g., absorption of UV light to yield electronically excited states, formation of radical ions, and transformation of nucleobases into mutagenic lesions as OH radical adducts and 8-oxoguanine. Based on serine phosphate as a model compound, some insight into the direct DNA strand break mechanism is given.

Psoralens, also called furocoumarins, are a family of sensitizers exhibiting cytostatic and photodynamic actions, and hence, they are used in photochemotherapy. Molecular design of more efficient photosensitizers can contribute to enhance the photophysical and photochemical properties of psoralens and to reduce the phototoxic reactions. The mechanisms of photosensitization of furocoumarins connected to their dark toxicity are examined quantum chemically.

Keywords: statistical mechanics, biophysical chemistry, interface, surface thermodynamics, bioelectrochemistry, ionizing radiation, radiation therapy, condensed matter, computational chemistry, nucleic acids, radiation damage, electrode potential, electronic transport, photochemistry, strand break, photodynamic action, cytostatic, solvation, solvated electron, absolute potential, chemical potential

Jorge Llano, Department of Cell and Molecular Biology, Box 596, BMC, Uppsala University, SE-75124 Uppsala, Sweden

ISSN 1104-232X
ISBN 91-554-5940-4
urn:nbn:se:uu:diva-4224 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-4224)

« *La logique qui peut seule donner la certitude est l' instrument de la démonstration : l' intuition est l' instrument de l' invention*. »

Henri Poincaré, *La valeur de la science*, Flammarion: Paris, 1905.

List of Papers

SELECTED ARTICLES INCLUDED IN THIS DISSERTATION:

I. J. Llano, L. A. Eriksson, **First Principles Electrochemistry: Electrons and Protons Reacting as Independent Ions**, *J. Chem. Phys*. **2002**, *117*, 10193.

II. J. Llano, L. A. Eriksson, **Mechanism of Photoinduced Formation of Radical Ions in DNA Bases and Base Pairs in Aqueous Solution**, 2004, manuscript.

III. J. Llano, L. A. Eriksson, **First Principles Electrochemical Study of Redox Events in DNA Bases and Chemical Repair in Aqueous Solution**, *Phys. Chem. Chem. Phys*. **2004**, in press.

IV. J. Llano, L. A. Eriksson, **Mechanism of Hydroxyl Radical Addition to Imidazole and Subsequent Water Elimination**, *J. Phys. Chem. B* **1999**, *103*, 5598.

V. J. Llano, L. A. Eriksson, **Oxidation Pathways in Adenine and Guanine in Aqueous Solution from First Principles Electrochemistry**, *Phys. Chem. Chem. Phys*. **2004**, submitted.

VI. J. Lipfert, J. Llano, L. A. Eriksson, **Radiation Induced Damage in Serine Phosphate—Insights into a Mechanism for Direct DNA Strand Breakage**, *J. Phys. Chem. B* **2004,** accepted.

VII. J. Llano, J. Raber, L. A. Eriksson, **Theoretical Study of Phototoxic Reactions of Psoralens***, J. Photochem. Photobiol. A* **2003**, *154*, 235.

RELATED ARTICLES NOT INCLUDED IN THIS DISSERTATION:

I. S. D. Wetmore, R. J. Boyd, and J. Llano, M. J. Lundqvist, L. A. Eriksson; **Hydroxyl Radical Reactions in Biological Media**. In *Recent Advances in Density Functional Methods*; Vol. 3; V. Barone, A. Bencini, P. Fantucci, Eds.; World Scientific: Singapore, 2000.

II. J. Raber, J. Llano, L. A. Eriksson; **Density Functional Theory in Drug Design: the Chemistry of the Anti-tumor Drug cis-Platin and Photoactive Psoralen Compounds**. In *Medical Quantum Chemistry*; L. P. Carloni, F. Alber, Eds.; Wiley-VCH: Weinheim, 2003.

Contents

List of Acronyms

CAS-SCF	Complete Active Space SCF [model]
CC	Coupled Cluster [model]
CCSD	CC Singles-and-Doubles [model]
CCSD(T)	CCSD [model] with approximate Triples correction
CIS	Configuration Interaction Singles [model]
DD	Dislocation Dynamics [model]
DFT	Density Functional Theory
ET	Electron Transfer
FCI	Full Configuration Interaction [model]
FEM	Finite Element Method
FF	Force Field
GGA	Generalized Gradient Approximation
GVB	Generalized Valence Bond [model]
HF	Hartree–Fock [model]
IR	Infrared [radiation]
KMC	Kinetic Monte Carlo [model]
KS	Kohn–Sham [theory]
LDA	Local Density Approximation
MC	Monte Carlo
MC-SCF	Multi-Configuration SCF [model]
MD	Molecular Dynamics
MM	Molecular Mechanics
MP*n*	*n*th-order Møller–Plesset [perturbation theory]
PCM	Polarized Continuum Model [of the solvent]
PES	Potential Energy Surface
PT	Proton Transfer
PT–ET	Proton-Coupled Electron Transfer
RHF	Restricted HF [model]
QC	Quantum Chemistry
QM	Quantum Mechanics

QM/MM	Hybrid QM–MM
SCE	Saturated Calomel Electrode
SCF	Self-Consistent Field [model]
SCRF	Self-Consistent Reaction Field [model]
SHE	Standard Hydrogen Electrode
SM	Statistical Mechanics
TD-DFT	Time-Dependent DFT
UHF	Unrestricted HF
UV	Ultraviolet [radiation]
VIS	Visible [radiation]
WFT	Wave Function Theory

Introduction

The ever-increasing complexity of scientific problems demands interdisciplinary approaches to solve them, and thus, the boundaries between many fields of physics, chemistry, and biology become diffuse. Today's problems concern phenomena whose descriptions coexist in different space and time scales. Scientific approaches are then challenged to deal simultaneously with atomistic and continuum systems and with ultrafast and slow phenomena.

This dissertation gives a brief overview of modern trends in chemical physics and physical chemistry. The discussion emphasizes the role and the extent of applicability of computational simulations in these fields. The material is thus organized according to the relevant space and time scales for chemical phenomena. Then, the presentation focuses on the realm of quantum chemistry to provide the basic details of the electronic and molecular structures needed to apply statistical mechanics to treat problems in the condensed phase. As a result, an accurate macroscopic description of redox reactions in solution arises from the molecular point of view. In particular, this enables a consistent treatment of electron, proton, and proton-coupled electron transfer reactions responsible for the radiation induced damage to DNA and many other crucial reactions in biology. Accordingly, these new models broaden the applications of theoretical and computational chemistry to material and biomedical sciences.

The contents are organized so that the applications to radiation chemistry and photochemistry discussed in the last two chapters refer to the theory discussed earlier. The chapters are rather brief and as self-contained as possible. Chapters 1–4 constitute an integrated overview of the combination of molecular, thermodynamic, and computational approaches to tackle basic physicochemical problems; Chapters 5 and 6 deal with the computational applications.

Chapter 1 reviews the space and time scales concerning atomic and molecular physics, chemistry, and other interdisciplinary fields. It discusses modelling of multiscale phenomena and the basic principles to apply the laws of physics consistently in computational simulations. Chapter 2 serves as a general introduction to molecular electronic structure theory. It should be regarded as concise revision notes on topics exhaustively reviewed in the key references cited therein. Chapter 3 contains the basics of interfacial electrochemistry. Even though this is a well-established field in physical

chemistry, the basic definitions found in the literature and textbooks can turn out to be contradictory. Accordingly, Chapter 3 follows IUPAC recommendations on the controversial definitions in order to recast the conventional electrodic problem into a tractable molecular problem enabling us to apply the Gibbs formalism. Chapter 4 deals with the definitions of absolute electrochemical, real chemical, and chemical potentials according to first principles quantum and statistical mechanics. Therein, the thermochemistry of electron, proton, and proton-coupled electron transfer reactions is defined in terms of the fundamental chemical potentials of the species in solution.

Chapter 5 describes the chemical effects of ionizing radiation on DNA. The discussion focuses on explaining and relating the reactions that several experimentally well-identified agents cause on DNA. This chapter includes the absorption of UV light to yield excited states in DNA, the formation of radical ions (charged molecules with unpaired electrons), and some rationale as to how the sugar–phosphate bonds directly break in the DNA strands. It also addresses the formation of hydroxyl radical adducts of purine bases and their oxidation to 8-oxoadenine and 8-oxoguanine, which are chemical lesions causing a particular type of mutation called transversions.

Finally, Chapter 6 deals with the application of the photosensitization to chemotherapy. This chapter describes the photodynamic action of psoralens, which is responsible for the dark toxicity of these compounds.

1. Modelling, Simulation, and Theory

Modelling refers to the construction of scale models, as well as to the devising or use of abstract or mathematical models. A *model* is thus a simplified or idealized description or conception of a particular system, in a sufficiently simple form to enable a certain problem to be solved. Models are put forward as a basis for theoretical or empirical understanding, or for calculations, predictions, etc. A *model* is a conceptual or mental representation of a system or process. Models are also physical three-dimensional representations of real systems and processes, especially on a smaller scale.[1]

Simulations are techniques of reproducing the behaviour of some existing or intended system, or some aspect of that behaviour, by means of models. They replicate the mathematical formulations of complex physical processes computationally. They are extensively used in the scientific prediction ranging from economy and meteorology to chemistry and biology.[2,3]

Modelling and simulation are based on the theories that describe natural phenomena. A simple and concise definition of a *physical theory* was given by Pierre Duhem in 1906: [4]

> "A physical theory is a system of mathematical propositions, deduced from a small number of principles, which has the object of representing a set of experimental laws as simply, as completely, and as exactly as possible".

The description of systems, processes, models, simulations, and so forth can be classified in terms of various criteria. According to the level of the physical theory, the description can be *quantum* or *classical*. According to the pattern of the evolution, the description can be *deterministic* or *stochastic*. Deterministic denotes a process whose resulting behaviour can be predicted exactly and is uniquely determined by the initial state of the system. In contrast, stochastic denotes a process whose resulting behaviour has a randomly determined distribution or pattern, which can be analyzed statistically, but cannot be predicted exactly.[5] According to the internal structure, an *atomistic* description is interpretable through analysis into distinct, separable, and independent elementary components; whereas a *continuum* description neglects any internal, possibly discrete, structure.

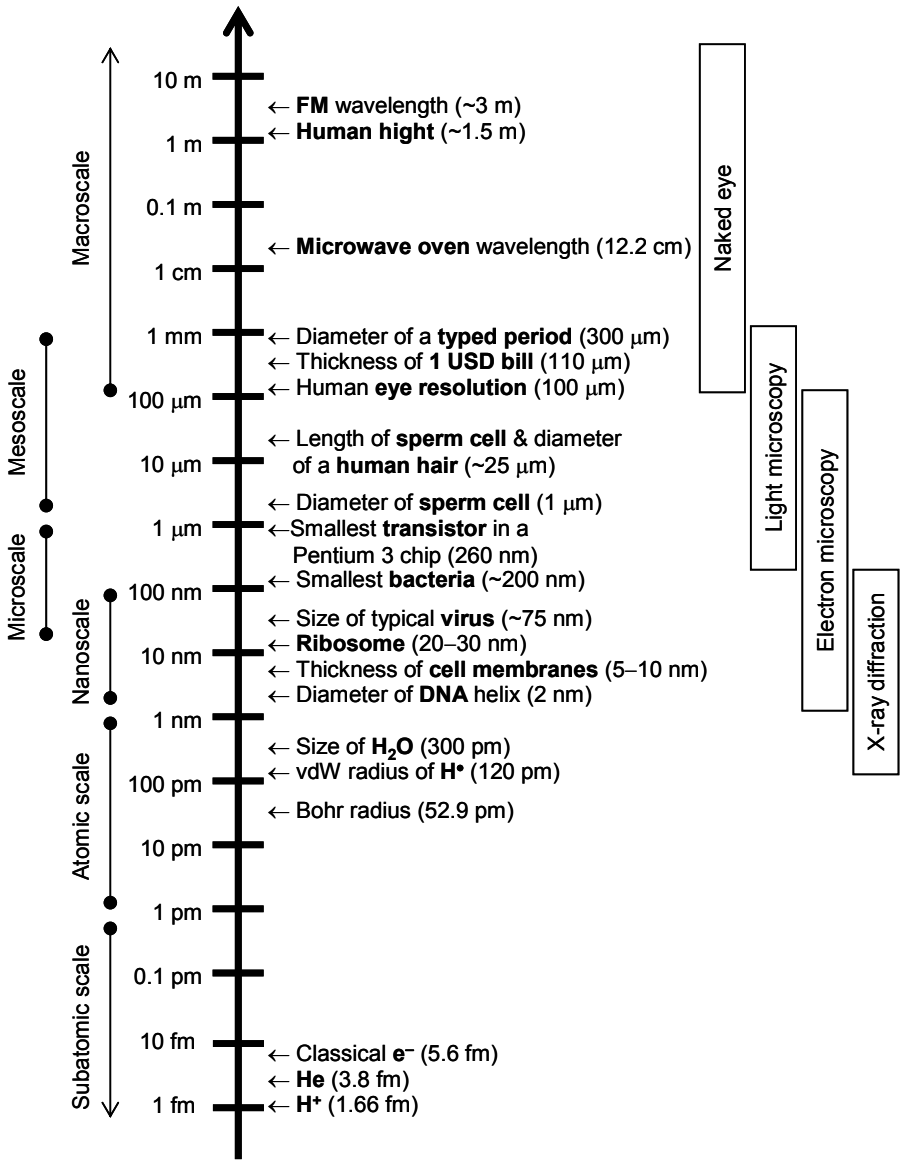


Figure 1.1. The relevant size scales.

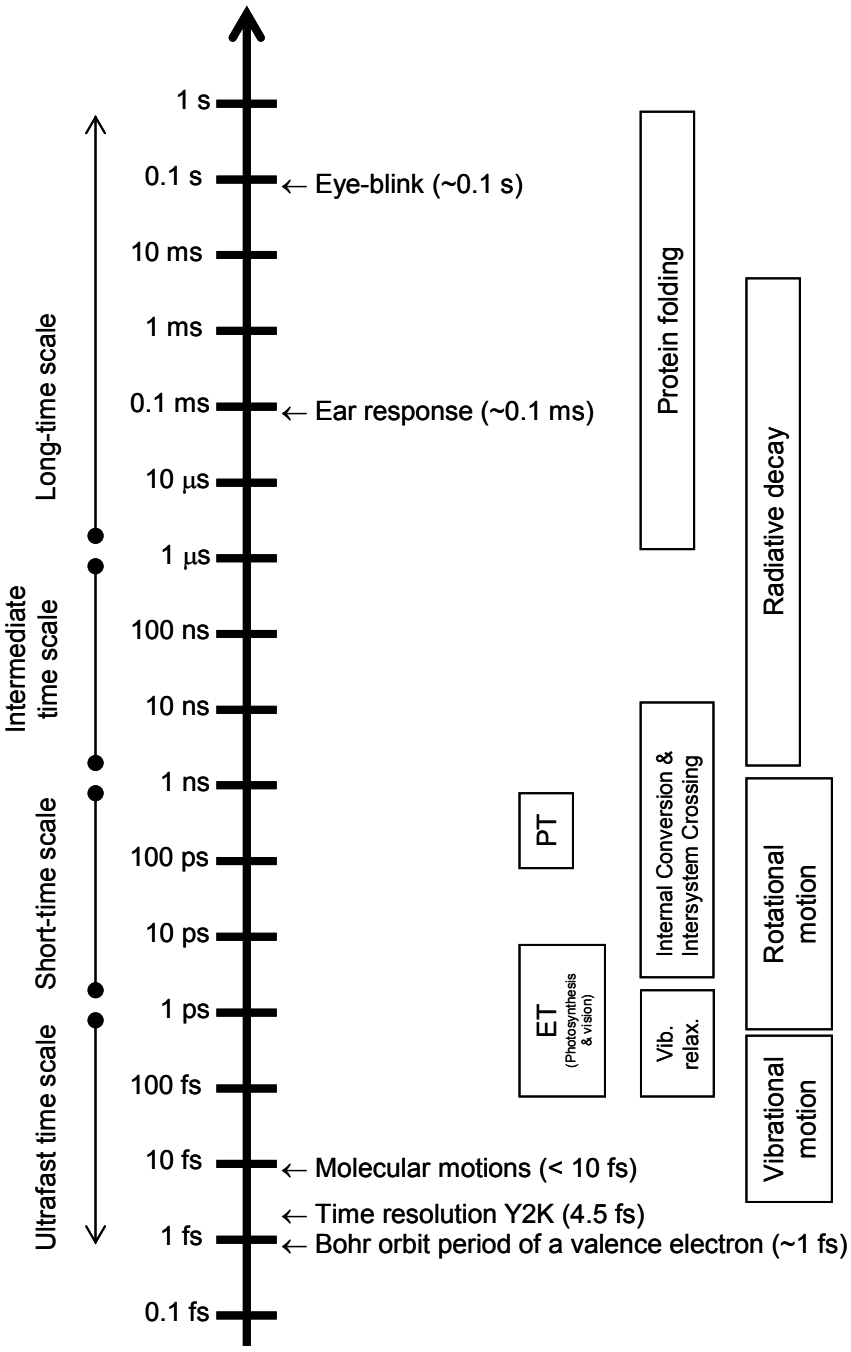


Figure 1.2. The relevant time scales.

In the design of models, theories, and simulations, one should notice that the appearance of objects and phenomena changes in the real world according to the scale of observation. For instance, the concept of branch of a tree makes sense on a scale of few centimeters to a few meters, while it is meaningless on the nanometer or kilometer level. Likewise, the concept of total solar eclipse makes sense on a scale of minutes—its maximum duration is seven minutes and eight seconds only—while it is meaningless on the millisecond or hour level. Moreover, one can observe many phenomena taking place within a single entity depending on the size and time scales the observations are made. Zooming into the space and time scales on entities and processes reveals their internal nature and complexity.

1.1. Space and time scales

Most nonliving and all living matter is characterized by its sheer complexity. Processes in complex systems (e.g. fracture in real materials, phase transitions, diffusion, surface and bulk phenomena, chemical reactions, metabolic pathways, protein folding, self-organization and self-assembly far from equilibrium, etc.) involve structures on many different space and time scales. Many occur on a macroscopic scale, but an accurate and detailed description of such phenomena often requires atomic level resolution. The relevant space and time scales for physics, chemistry, and biology are illustrated in Figures 1.1 and 1.2.

The size of entities can be conveniently grouped into six length ranges. Entities larger than $\sim 10^{-4}$ m can be regarded to be on the *macroscale*. The *mesoscale* (10^{-6} to $\sim 10^{-3}$ m) designates a size scale intermediate between the microscale and the macroscale, in which objects can usually be resolved by light microscopy. The *microscale* ($\sim 10^{-8}$ to 10^{-6} m) represents the size level typical of structures that can be made visible and examined by electron microscopy. Complex macromolecules and self-organized macromolecular structures can be found on the *nanoscale* (10^{-9} to $\sim 10^{-7}$ m). Atoms and small molecules lie on the *atomic scale* (10^{-12} to 10^{-9} m). The *subatomic scale* ($< 10^{-12}$ m) includes the elementary particles in physics (i.e., those that are the basic constituents of matter and cannot be subdivided) and in chemistry (i.e., electrons and nuclei).

In terms of modelling, continuum descriptions are usually applied at the macroscopic level. Mesoscopic systems cannot be successfully described neither with atomistic models alone nor with continuum models alone. In contrast, full atomistic models are possible from the microscopic level downwards.

The duration of physical and chemical events can be grouped into five significant periods of time. The realm of chemical kinetics (i.e., rates of reactions, long-lived intermediates) lies on a *long-time scale* (>10^{-6} s). Short-lived intermediates and free radicals are typical of an *intermediate-time scale* (10^{-9} to 10^{-6} s). Non-radiative photophysical processes (e.g. internal conversion, intersystem crossing) occur within the intermediate-time scale and the *short-time scale* (10^{-12} to 10^{-9} s), whereas the bond breaking/making dynamics arises on the *ultrafast-time scale* (10^{-15} to 10^{-12} s).[6] The frontier in ultrafast spectroscopy is the motion of electrons bound inside the atom, in orbits close to the nucleus; this motion lies on the *attosecond-time scale* (10^{-18} to 10^{-15} s).[7,8]

The overwhelming desire for manipulating, controlling, and observing atoms in motion opens up new fields in science and technology. *Nanoscience* deals with the observation and description of phenomena at the nanoscale, while *nanotechnology* is the popular term for the construction and utilization of functional structures with at least one characteristic dimension measured in nanometers.[9] The resolution of the dynamics of atomic motion has brought forth a new branch termed *femtoscience* (e.g., femtophysics, femtochemistry, and femtobiology). Given that the speed of the atomic motion is ~1 km/s, the average time required to record atomic scale dynamics over a distance of 1 Å is ~100 fs.[6] Furthermore, *attoscience* is at the cutting edge of femtoscience and aims at resolving the dynamics of the electronic motion. The time scale in attophysics is defined by the time it takes the electron in the innermost orbit of Bohr's hydrogen atom to complete one turn around the proton nucleus. The period of this orbit is 24×10^{-18} s or 24 attoseconds.[7]

The theoretical descriptions of the macro- and microworld must be complementary. On the one hand, as one zooms into the space scale, the continuous descriptions that work on the macroscale start failing on the mesoscale and break down on the nanoscale.[10,11] At this level, atomistic descriptions are thus essential. The gaps in between the extremes of the size scale also pose logical gaps in the time scale.

On the other hand, the complex appearance of a natural process is due to the fact that the observations are done on an extended time scale, during which many elementary steps in the process are integrated.[6] These elementary steps usually occur on smaller size scales. The increase in resolution of the time scale enables us to observe the slow-motion picture of what actually happens within the microworld.

These two problems reflect in the simulations. The more detailed the model is, the more the limitations occur in size and structural complexity of the model, as well as in computing time.

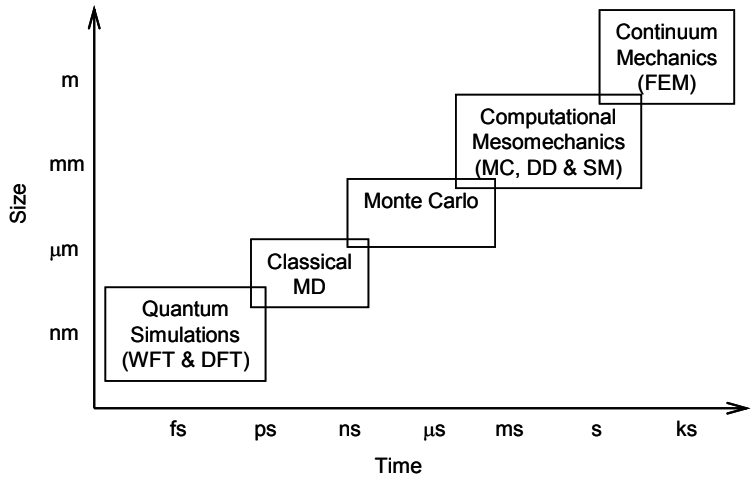


Figure 1.3. Multiscale hierarchy.

1.2. Multiscale modelling

A *multiscale model* is a composite mathematical model encompassing a wide range of length and time scales,[12] i.e., it combines partial models that describe phenomena at different characteristic space and time scales (Figure 1.3).[13] Nowadays, linking up several space and time scales is absolutely required to solve many questions in science and technology. The relevant scales to solve a problem are usually unknown a priori and should be identified during the process of solution itself. Moreover, the simultaneous resolution of the problem at all scales is likely to be unattainable. Hence, scale separation should be carried out so that the problem gets decomposed into a hierarchy of manageable models. The crucial question is to identify the level of theory on the finer scale needed to complement an effective theory on the coarser scale.

An ordinary strategy of multiscale modelling is to build a finite set of parameters for the theory holding on a coarser scale so that the parameters may condense and embody as much physical information as possible from a reduction of the many degrees of freedom of the model on a finer scale into fewer generalized degrees of freedom on a coarser scale.

1.2.1. Levels of theory and simulation

The level of theory and simulation required to explain or predict a phenomenon is decided by the particular features of the process of interest, the complexity of the system in which it takes place, and the accuracy pursued. At this point, one should take into account that an increase in

accuracy implies a surge in the computational cost of the simulation. Accordingly, time scale, space scale, and accuracy are the three dimensions involved in the multiscale hierarchy depicted in Figure 1.3.[9]

The behaviour of light entities—i.e., those in the subatomic, atomic, and nano- scales—is governed by the laws of *quantum mechanics* (QM). In particular, chemical reactions of atomic and molecular species are the most relevant events happening at this domain and gather together on the ultrafast-time scale. Hence, the description of phenomena at this level on the space and time scales requires electronic structure theories (Chapter 2). The formal time scaling of simulations with electronic structure methods is at least proportional to $M^2 - M^3$, being M the molecular size. The accurate description of the electronic structure with quantum chemical methods ranges from the order of 10 atoms (i.e., small molecules) with wave function methods to the order of 1000 atoms (i.e., in the unit cells of periodic systems) with density functional methods.[14]

As most problems in quantum chemistry can be solved approximately but not exactly, and as the size and complexity of systems increase on the nanoscale and the microscale, quantum chemical descriptions of the microworld become computationally prohibitive. Atomistic descriptions of the microscopic structure, in which whole atoms are represented explicitly, are an alternative to quantum and continuum structural pictures. *Force field* (FF) methods, also known as *molecular mechanics* (MM), ignore the electronic motions and use the laws of classical Newtonian physics to describe the interactions within the system (bond stretching, bond angle bending, torsional distortions, electrostatic and Van der Waals interactions, etc.)

Force field methods enable us to study the classical dynamics of motion of atoms within a system using the Newton's laws. The atomic positions can be followed over time by applying the Newton's equations of motion in a simulation technique known as *molecular dynamics* (MD). Molecular dynamics is a deterministic method, since any state of the system in the past or in the future can be predicted from its current state. Macroscopic properties can be computed as time averages over the dynamic evolution.[15] MD simulations have limitations on the time and space scales. As they are based on an atomistic model of the microscopic structure, atoms are bound to move in time and length scales of the order of 10^{-15} s (1 fs) and 10^{-10} m (1 Å).[10] MD simulations are typically run for tens or hundreds of picoseconds (a 100-ps simulation using a 1-fs time step requires 100 000 steps).[15] Accordingly, atomic collective motions in a wide range of length and duration pose a challenge to MD simulations, since the system relaxation can be longer than the MD simulation time-scale. The case of protein folding is a typical instance of a process whose relaxation may go beyond the MD time scale, since it occur in a long-time scale (10^{-6} to 10^{-1} s) with a broad amplitude (10 to 100 Å).[16]

The time scale of the atomic motion is of the order of femtoseconds, and this impose two fundamental limitations to atomistic MD simulations over the long periods of time required on the experimental time scales (from 10^{-15} to 10^{1} s and beyond). First, a numerical simulation uses finite time steps, and the step size is required to be small enough (of the order of the atomic motion) to keep the simulation stable. Second, simulating over long periods of time (microseconds and longer) relative to the short time scale of the atomic motion (femtoseconds) would require more than billion time steps.[10,17] Spanning the space scale of atomistic MD simulations is a computational problem that can be solved with parallel computing. Yet spanning the time scale is more a physical than a computational problem, which cannot be solved even with completely efficient implementations on the world's fastest parallel supercomputers.[18]

MD simulations connect successive configurations of the system in time.[15] The time-scale problem is partly overcome by *Monte Carlo* (MC) simulations. The MC method refers to any stochastic technique that involves random sampling for the mathematical simulation of a physical system. It is a statistical method for solving deterministic or probabilistic problems. MC methods generate configurations randomly, which allow for transitions between configurations that would be possible in MD only by extraordinarily long time trajectories.[18] MC is a stochastic method to explore the phase space. In molecular modelling, it is usually implemented with the Metropolis procedure so as to ensure that the configurations in the ensemble obey a Boltzman distribution. *Kinetic Monte Carlo* (KMC) is an alternative approach to circumvent the atomic time-scale problem. The KMC method is used for simulations of slow thermally-activated processes, e.g., atomic diffusion in solids. In a KMC simulation the system evolves from one configuration to another neglecting the thermal motion of atoms but moving atoms with rates defined by the activation barriers. The main drawback of the KMC method is that the activation barriers must be supplied as input parameters, and in turn, as the complexity of the process increases, the accuracy may decrease if information about some steps, e.g., activation barriers, is incorrect or missing.[10]

Last of all, computer simulation methods (i.e., MD and MC) tend to sample the region of the phase space close to the starting condition.[14] Even though computing power has been rapidly increasing, brute force simulations using atomistic modelling methods cannot describe systems much larger than 1 μm (billions of atoms) or longer than 1 ms (billions of femtosecond time steps).[10]

At the mesoscopic level, the problems grow more complex and computationally expensive to use atomistic simulations, and at the same time, they become too small for macroscopic continuum simulations. Problems in this realm are plastic deformations in solids,[10,11] diffusion of polymers in liquids and interfaces (polymer welding),[13] liquid crystalline

polymers,[13] composite materials, etc. Mesoscopic simulation methods are currently under development to bridge the gap between the extremes of the size scale. Computational *mesomechanics* include Monte Carlo (MC), dislocation dynamics (DD), and statistical mechanics (SM) techniques.[10,11]

At the macroscopic level, the *finite-element method* (FEM) is a broadly used computational technique for the continuum description of materials. The macroscopic object is broken down into finite elements, i.e., into simpler, disjoint components—like Lego pieces. Each element is specified by its nodes and is assigned a behaviour identified by physical quantities. The mathematical model for the whole system describes the behaviour of the individual elements plus their interaction.

1.2.2. Multiscale integration

The integration of different levels in a multiscale model structure can be suitably classified into serial, simultaneous, hierarchical, parallel, and multidomain.[12] As the boundaries between the approaches to integration may be diffuse, the combination of several approaches can be used for solving a single multiscale problem. Serial and hierarchical integrations seem so far to be thriving in the realm of molecular modelling.

In the *serial integration* approach, the solutions of the problem on different scales are decoupled; accordingly, part of the simulation outcome on a lower level of organization can be recast into income parameters for simulations on a higher level of organization. For instance, MD (atomistic description) can be used to calculate diffusion coefficients that are parameters of the parabolic partial differential equations of diffusion (continuum description).

In the *hierarchical integration* approach, the microscale model is embedded in the macroscale model, and hence, the theories and the solutions on different scales are coupled. In this class we can find some QM/MM methods and *ab initio* MD. Even in formal quantum chemistry, certain parts of the system may be treated fully quantum mechanically, e.g., the electron motion; other parts can be approximated semiclassically, e.g., the nuclear motion; and the surrounding environment can be described using a macroscopic theory, e.g., the continuum dielectric theory.

Among the QM/MM techniques, the ONIOM method (Morokuma and collaborators' ***o****wn* ***n****-layered* ***i****ntegrated molecular* ***o****rbital + molecular* ***m****echanics method*) is a multilayer approach to describe a molecular system by simultaneously using different levels of theory, which include quantum mechanical descriptions, classical mechanics, and continuum representations of the solvent.[19] Each level describes a particular region within the molecular system, and all must be properly coupled to reflect the interactions among the subsystems. In the case of *ab initio* MD, the potential energy surface is

computed by means of a quantum description, whereas the motion of atoms on that surface is treated classically.[20,21]

In summary, a multiscale model should ideally follow some basic principles: [10,12,22]

1. *Hierarchy in space and time scales*: Variables defining the state of the system at each level should be identified to make a rigorous reduction of degrees of freedom from the finer level to the coarser following level of the space and time scales.
2. *Compatibility of the theories verifying at each level*: Physical theories describing the behaviour of the system at each level should be hierarchically compatible in the length and time scales. Coupling of the physical laws of the finer descriptions with those of the coarser ones provides the means of determining the parameters essential for the description of the latter. This modelling procedure acts as a filter removing the complexity of smaller scales and restricting attention to the processes directly relevant to the scale of interest.
3. *Self-consistency within the multiscale model*: The accuracy of predictions with the laws verifying at a given level should show reasonable agreement with the predictions done with the laws of the preceding level in the space and time scales.
4. *Consistency with the experimental data*: A multiscale model should provide data that is consistent with the experimental data at every level of the space and time scales.

1.2.3. Pitfalls in handling complexity

Complexity is a term difficult to define, and thus, to quantify and measure. Complex qualifies the way in which the whole is different from its constituent parts. In natural sciences, complex phenomena can be approached from two opposed methodological points of view: holism and reductionism. Methodological holism seeks to understand a certain kind of complex systems at the level of principles governing the behaviour of the whole system, and not at the level of the structure and behaviour of its component parts. In contrast, methodological reductionism seeks to understand a complex system at the level of the structure and behaviour of its component parts.[23] In practice, some systems have qualitatively new, holistic, emergent properties arising from the strong and lasting interaction of their parts, whereas other systems have properties arising from the parts as adding up to the whole, usually when the components are weakly interacting.

Multiscale modelling is a reductionist approach to analyze the levels of organization that emerge in complex systems, and hence, it is debatable its application to predictions about emerging properties originating on a coarser scale from the underlying finer scales. In the case of emergent phenomena,

multiscale modelling can be a good approach for explaining but not for predicting. Conversely, in some cooperative phenomena and in those processes occurring within systems consisting of weakly-interacting subsystems, it is effective and accurate to reduce the large number of degrees of freedom on a finer scale into a minimum number of degrees of freedom on a coarser scale. This is, for instance, what is done with statistical mechanics: the collective behaviour of a huge number of atoms and molecules of infinite degrees of freedom can be described by simple scaling laws. These laws are obtained by averaging techniques, and the relevant information contained in the statistical mechanical equations represents some average behaviour of many basic entities.[10]

According to the principle of *causality*, effect cannot precede cause; nevertheless, we find two basic forms of simultaneous causality: upward causation and downward causation.[24] *Upward causation* means that entities acting at a lower level of organization simultaneously bring about effects on the entities at a higher level of organization. It accounts for the fact that the same physical interactions among the huge variety and number of simpler building blocks determine the formation of qualitatively different and more complex assembled entities. *Downward causation* is the reverse, i.e., entities acting at a higher level of organization simultaneously bring about effects on the entities in a lower level of organization. This type explains the selectivity and specificity for the building blocks in the process of assembling complex entities.

In natural phenomena, we find both forms of causality, which should also be reflected in the models, theories, and simulations of our reality. Causation from bottom to top in a multiscale problem indicates that phenomena on a finer scale, to some extent, enable and account for new phenomena on subsequent coarser scales. Nonetheless, we also find causation from top to bottom, i.e., different new properties emerging at the top level determine which building blocks and which interactions among them are compatible with the complex superstructure. This sort of selectivity at the top level does not interfere with or constrain the laws at the bottom level, but it introduces new causal interactions in the complexity of natural phenomena.

In line with the views of the Aristotelian causality and the Cartesian reductionism, a unidirectional bottom-up conception of causality has reinforced radical reductionism as one solution to the problem of the relationship between different sciences. According to this point of view, biology could be reduced to chemistry, supposing that no distinctive biological facts exist, and chemistry to physics, supposing that no distinctive chemical facts exist. This has lead to the logical positivist version of reductionism, the doctrine of unity of science, which claims that particular sciences, e.g., biology and chemistry, can eventually be reduced to one embracing science, usually thought of as physics. The physicalist ideology regards the real world as nothing more than the physical world. The doctrine

may, but need not, include the view that everything that can truly be said in the language of physics, i.e., the laws of those sciences lower in hierarchy should be explained by those of the more basic science. Accordingly, the theoretical entities of biology and chemistry could be redefined in terms of physical observables constituting the common basis of all scientific laws. Even though the question of the reducibility of one science to another may be controversial, the top-down forms of causality reflected in the emergent phenomena have to be always taken into account properly for a correct theoretical description of complexity at or within any level of the hierarchy of sciences.[25,26]

Multiscale modelling faces several challenges that will not be sorted out with improvement in computer hardware and architecture alone. From the physical point of view, the accurate theoretical description of emergent properties and the interactions among different space and time scales. From the computational point of view, the development of efficient algorithms enabling us to use multiple resolutions in both space and time, as well as allowing the extension of simulations to larger spatial and temporal scales.

2. Molecular Electronic Structure Theory

The electronic structure theories can be grouped into two complementary approaches: *wave function theory* (WFT) and *density functional theory* (DFT), according to the function chosen to represent the electron distribution in space for an atomic or molecular system composed of *M* atoms.

WFT methods are applicable to small molecules, with *M* of the order of tens of atoms, i.e., $M \sim 10$, and the theory gives a detailed insight into the nature of the electron correlation. On the contrary, DFT supplies little information about the nature of electron correlation, but the methods can be applied to larger systems. For polyatomic molecules, the upper bound is *M* of the order of one to two hundred atoms, i.e., $M \sim 100-200$. For periodic structures (solids), the upper bound is *M* of the order of thousands of atoms, i.e., $M \sim 10^3$.[27,28,40]

In regard to the computational cost, the computing time (t_{cp}) in DFT grows with the number of atoms *M* as $t_{cp} \sim M^{\alpha}$. KS-DFT methods formally scale as the forth power of the molecular size, but currently show $\alpha \approx 2-3$ with ongoing efforts to make the computer time grow only linearly in *M*, i.e., $\alpha \approx 1$, usually referred to as $O(M)$ or linearly-scaling methods. WFT are computationally more expensive: HF (KS-DFT may be viewed essentially identically), MP2, CCSD, and CCSD(T) for a given size basis set show computational requirements scaling as M^4, M^5, M^6, M^7, respectively. FCI, regarded as "exact" for a given basis set, shows computing time growing as $t_{cp} \sim \exp(\alpha M)$ with $\alpha \approx 1$.[40,27,28]

This chapter provides a comprehensive overview of the molecular problem, i.e., the simultaneous motion of electrons and nuclei in a molecular system. The starting point is the separation of the nuclear and electronic motions. Second, the nuclear motion is addressed by means of a semiclassical approach to describe the molecular vibrations. Third, a summary of the electron correlations are outlined to tackle the problem of the electronic motion by way of WFT and DFT.

2.1. The molecular problem

The puzzle of describing the simultaneous motion of nuclei and electrons in a molecule could be formulated as the molecular problem. The molecule can be pictured as a system consisting of *N* electrons and *M* nuclei, specified by their position vectors, collectively denoted as $\{\mathbf{r}_i\}$ and $\{\mathbf{R}_A\}$.

The Born–Oppenheimer approximation separates motions from two different time scales: the (slow) nuclear motion in the femtosecond domain from the (fast) electronic motion in the attosecond domain; therefore, one can regard the electrons in a molecule moving in the field of fixed nuclei. In this approximation, the *full hamiltonian* $\hat{H}$ for the system of *N* nonrelativistic interacting electrons and *M* nuclei consists of the kinetic energy of the nuclei $\hat{T}_{\mathrm{n}}$ and the electronic hamiltonian $\hat{H}_{\mathrm{el}}$. The *electronic hamiltonian* contains the kinetic energy of the electrons $\hat{T}_{\mathrm{e}}$, the Coulomb attraction between electrons and nuclei $\hat{V}_{\mathrm{ne}}$, the electrostatic repulsion between electrons $\hat{V}_{\mathrm{ee}}$, and the repulsion between nuclei $\hat{V}_{\mathrm{nn}}$:

$$\hat{H} = \hat{T}_{\mathrm{n}}(\{\mathbf{R}_A\}) + \hat{H}_{\mathrm{el}}(\{\mathbf{r}_i\};\{\mathbf{R}_A\}), \tag{2.1}$$

$$\hat{H}_{\mathrm{el}}(\{\mathbf{r}_i\};\{\mathbf{R}_A\}) = \hat{T}_{\mathrm{e}}(\{\mathbf{r}_i\}) + \hat{V}_{\mathrm{ne}}(\{\mathbf{r}_i\};\{\mathbf{R}_A\}) + \hat{V}_{\mathrm{ee}}(\{\mathbf{r}_i\}) + \hat{V}_{\mathrm{nn}}(\{\mathbf{R}_A\}); \tag{2.2}$$

$\hat{H}_{\mathrm{el}}$ depends *explicitly* on the electronic coordinates and *parametrically* on the nuclear coordinates. As electrons move much faster than nuclei, the average value of $\hat{H}_{\mathrm{el}}$ over the electronic coordinates can be split into the contributions of the *electronic energy* E_{el} for the motion of electrons in the field of *M* point charges and the *nuclear repulsion energy* V_{nn}:

$$E_{\mathrm{tot}}(\{\mathbf{R}_A\}) = \left\langle \hat{H}_{\mathrm{el}}(\{\mathbf{r}_i\};\{\mathbf{R}_A\}) \right\rangle = E_{\mathrm{el}}(\{\mathbf{R}_A\}) + V_{\mathrm{nn}}(\{\mathbf{R}_A\}) \equiv V(\{\mathbf{R}_A\}). \tag{2.3}$$

In the Born–Oppenheimer approximation, the *potential energy surface* (PES) is the total energy of the system as a function of the coordinates of the nuclei, $E_{\mathrm{tot}}(\{\mathbf{R}_A\})$. The total energy constitutes an effective potential $V(\{\mathbf{R}_A\})$ for the nuclear motion.[29] As a result, the motion of nuclei on the PES defines the possible chemical reactions, the molecular vibrations, etc. For a nonlinear molecule consisting of *M* atoms, the PES depends on $3M-6$ independent coordinates, and depicts how the potential energy (PE) changes as relative coordinates of the atomic nuclei involved in the chemical reaction are varied. An analytic function $V(\{\mathbf{R}_A\})$ that represents a PES is called *potential energy function*. The points of chemical interest in a PES are the stationary points. The saddle points define transition states and the minima define equilibrium geometries.[30,31]

2.1.1. The problem of nuclear motion

Once solved the electronic problem, the nuclear problem can be solved under the same assumptions. One should take into account that the complete specification of the molecular system in space requires $3M$ Cartesian coordinates, and hence, the molecule has $3M$ *degrees of freedom*. Out of these $3M$ coordinates, three can be used to specify the center of mass of the molecule. Motion along these three coordinates corresponds to translational motion of the center of mass, and so these three coordinates are called *translational degrees of freedom*. Moreover, two/three coordinates are required to specify the orientation of a linear/nonlinear molecule about its center of mass. Since motion along these coordinates corresponds to rotational motion, it is said that a linear/nonlinear molecule has two/three *degrees of rotational freedom*. The remaining coordinates ($3M-5$ for a linear molecule and $3M-6$ for a nonlinear molecule) specify the relative positions of the M nuclei. The motion along these coordinates correspond to vibrational motion, and hence, a linear molecule has $3M-5$ *vibrational degrees of freedom*, while a nonlinear molecule has $3M-6$.[32] Taking into account this separation of variables, the full hamiltonian $\hat{H}$ can be recast into the contributions of the translational, $\hat{H}_{\text{tr}}$, rotational, $\hat{H}_{\text{rot}}$, vibrational, $\hat{H}_{\text{vib}}$, and electronic, $\hat{H}_{\text{e}}$, motions:

$$\hat{H} = \hat{H}_{\text{tr}} + \hat{H}_{\text{vib}} + \hat{H}_{\text{rot}} + \hat{H}_{\text{e}} . \tag{2.4}$$

At absolute zero temperature, only vibrational and electronic motions are permitted. Provided the Born–Oppenheimer approximation holds, vibrations can be described by the *nuclear Hamiltonian* $\hat{H}_{\text{nu}}$ for the motion of nuclei in the average field of electrons:

$$\hat{H}_{\text{nu}} = \hat{T}_{\text{n}}(\{\mathbf{R}_A\}) + V(\{\mathbf{R}_A\}) . \tag{2.5}$$

The *vibrational hamiltonian* $\hat{H}_{\text{vib}}$ is constructed by assuming that the total energy $V(\{\mathbf{R}_A\})$ can be approximated by a second-order Taylor expansion around the stationary geometry $\{\mathbf{R}_{A0}\}$; therefore, the first derivative terms are zero. $V(\{\mathbf{R}_{A0}\})$ is an arbitrary potential energy, and hence, its value can be conveniently shifted to zero. The problem is then simplified with the introduction of the *mass-weighted coordinates* $\{\mathbf{q}_A\}$ ($\mathbf{q}_A = m_A^{1/2}\mathbf{R}_A$, where A runs from 1 to M, and m_A is the Ath atomic mass). There exist linear combinations $\{\mathbf{Q}_A\}$ of the mass-weighted coordinates $\{\mathbf{q}_A\}$ in which $\hat{H}_{\text{vib}}$ has a separable form. The linear combinations $\{\mathbf{Q}_A\}$ that achieve this separation of modes of vibration are called *normal coordinates*. The vibrational motions that correspond to displacements along these normal coordinates are called the *normal modes of vibration*.[33]

The normal modes of vibration display the following properties: (*i*) the normal coordinates form an orthogonal basis, and in turn, they represent linearly independent motions of the nuclei about the molecular center of mass—i.e., normal modes do not interact; (*ii*) the molecular center of mass remains motionless; (*iii*) the normal modes represent concerted motions of atoms in phase—i.e., at the same time, all atoms go past their equilibrium positions or reach the turning points of motion; (*iv*) the vibrational motion of polyatomic molecules can be pictured as the motion of n_{vib} independent and generalized harmonic oscillators of unit reduced mass and force constants κ_i with energies E_i,

$$E_j = (v_j + \tfrac{1}{2})\hbar\omega_j \qquad \omega_j = \kappa_j^{1/2} \qquad v_j = 0, 1, 2, \ldots, \tag{2.6}$$

where n_{vib} is the number of vibrational degrees of freedom, v_j is the vibrational quantum number, and ω_j is the vibrational angular frequency ($\omega_j = 2\pi\nu_j$) of the *j*th normal mode.[14,29,32,33]

The vibrational energy E_{vib} of a polyatomic molecule in the harmonic approximation is

$$E_{\text{vib}} = \sum_{j=1}^{n_{\text{vib}}} (v_j + \tfrac{1}{2})\hbar\omega_j \,. \tag{2.7}$$

A general vibrational state is $\left|v_1, v_2 \ldots v_{n_{\text{vib}}}\right\rangle$, with v_1, v_1, …, $v_{n_{\text{vib}}}$, the vibrational quantum numbers of the modes 1, 2, …, n_{vib}. The energy of the vibrational ground state $\left|0_1, 0_2 \ldots 0_{n_{\text{vib}}}\right\rangle$ is not zero, as the lowest classical energy is. Owing to the quantum mechanical behaviour, the energy of the lowest level is the *zero-point energy*, E_{ZP},[33]

$$E_{\text{ZP}} = \tfrac{1}{2}\sum_{j=1}^{n_{\text{vib}}} \hbar\omega_j = \tfrac{1}{2}\sum_{j=1}^{n_{\text{vib}}} h\nu_j \,. \tag{2.8}$$

Since the molecular energy—neglecting rigid translations and rotations—is $E_{\text{tot}} + E_{\text{ZP}}$ in the Born–Oppenheimer approximation, it follows that the *absolute molar internal energy* at zero absolute temperature, $E_{0\text{K}}$, is

$$E_{0\text{K}} = E_{\text{tot}} + E_{\text{ZP}} \,. \tag{2.9}$$

2.2. Electron correlations

Electron correlation refers to the adjustment of electron motion to the instantaneous—instead of spatial averaged—positions of all the electrons in a many-electron system. It expresses the tendency of electrons to correlate their motions in order to keep as far apart as possible because of the restrictions set by the Pauli exclusion principle (exchange correlation) and because of the electrostatic repulsions (Coulomb correlation).[34]

↑↑ Electrons with *parallel spins* must keep a certain separation because they cannot occupy the same space volume (antisymmetry principle for fermions).

↑↓ Electrons with *anti-parallel spins* must keep apart in order to lower their Coulomb repulsion—however, ↑↓ can be arbitrarily close in HF).

Accordingly, a many-electron system is stabilized:

↑↑ By moving *parallel spin* electrons apart to lower the *exchange energy* E_{X}.

↑↓ By moving *anti-parallel spin* electrons apart to lower the *correlation energy* E_{C}.

Hartree–Fock (HF) theory treats interacting electrons as being independent particles within a *self-consistent field* (SCF). The SCF is the interaction field an electron experiences when we take spatial average over the positions of all other electrons.[35] The HF-SCF method recovers ~99% of the total energy of a molecule but the theory neglects the instantaneous interactions (correlations) between electrons.[36] The *electron correlation energy* E_{C} is defined as the difference of the exact nonrelativistic energy and the HF total energy of the system:[37]

$$E_{\mathrm{C}} = E_{\mathrm{tot}}^{\mathrm{Exact}} - E_{\mathrm{tot}}^{\mathrm{HF}}. \tag{2.10}$$

The absolute magnitude of the correlation energy per electron pair is about 10–25 kcal/mol, and thus, is in a sensitive range for the accurate quantitative description of the bond-breaking/forming energetics.[28]

As an electron move, it tries to avoid the repulsive interaction from other electrons by creating a void around into which other electrons cannot penetrate, owing to the Pauli exclusion principle and the Coulomb repulsion. The void is called the *exchange–correlation hole* and gives rise to the *exchange–correlation energy* E_{XC}. The hole is assumed to be positively charged and to have a radius corresponding to exactly one electron being excluded. The electron and its exchange–correlation hole form a quasiparticle whose charge is zero.[38]

Electron correlation can be divided into two main types whose most important features are outlined below:[14,35,36,39]

(*i*) Long-range or non-dynamic or static or "left–right" correlation:

- Arises from large separation of electrons in a pair.
- Associated with the interaction of electrons at long distance, and hence, it accounts for a correct molecular dissociation (distribution of electrons on atoms at infinite distance).
- Non-local: multicentered, present in molecules, not in atoms.
- Long-range correlation makes HF strongly disfavour bond breaking and favour overbinding (bond lengths too short, frequencies too large).
- Multireference wave function usually mandatory to correct for long-range correlation. Static correlation accounts for the resonance stabilization of wave functions owing to favourable interactions between different electronic configurations.
- Included in UHF, MC-SCF, CAS-SCF, GVB, CC, FCI.

(*ii*) Short-range or dynamic or "in–out" correlation:

- Arises from short separation of electrons in a pair.
- Associated with the absence of interelectronic cusp in HF (i.e., electrons with anti-parallel spins can be arbitrarily near in HF), and hence, anti-parallel electrons $\uparrow\downarrow$ tend to be closer than they should be in reality.
- Local: monocentered, identifiable in atoms, not in molecules.
- Short-range correlation makes HF overestimate bond lengths (longer bonds) and underestimate binding.
- Single-reference wave functions correctly represent the electronic state of the system.
- Included in RHF, DFT, CC, FCI.

2.3. Electronic structure theories

Electronic structure theories aim at solving the electronic problem, i.e., the question of describing the motion of interacting electrons in the field of fixed nuclear point charges. The electronic hamiltonian $\hat{H}_{\text{el}}$ [Eq. (1.2)] describes this problem, and it can be tackled from two complementary approaches: WFT and DFT.

The electronic structure of the system consisting of N electrons and M nuclei can be represented by the wave function $\Psi(\mathbf{r}_1,...,\mathbf{r}_N)$, in WFT, or by the electron density $n(\mathbf{r})$, in DFT. The wave function belongs to the complex $3N$-dimensional space, depends on the physical representation, and is a non-visualizable concept, it does not represent any observable. On the contrary,

the electron density belongs to the real three-dimensional space, is independent of the physical representation, and is a visualizable concept.[40]

The most important property of an electronic ground state is its energy. WFT and DFT can estimate it by means of variational principles. Variational ground-state energies computed at any level give upper bounds to the exact energy E_0^{Exact}.

WFT follows the *Rayleigh–Ritz minimal principle*:[29,40] given a normalized trial (approximate) wave function $\widetilde{\Psi}$ depending on certain parameters, the best approximation to the ground-state energy E_0 of a hamiltonian $\hat{H}$ is the expectation value

$$E_0^{\text{WFT}} \equiv E\left[\Psi_0(\mathbf{r}_1 ... \mathbf{r}_N)\right] = \min_{\widetilde{\Psi}} \left\langle \widetilde{\Psi} \middle| \hat{H} \middle| \widetilde{\Psi} \right\rangle \geq E_0^{\text{Exact}} . \tag{2.11}$$

DFT follows the *Hohenberg–Kohn minimal principle*:[40,41] given a trial (approximate) electron density $\widetilde{n}$ depending on certain parameters, the best approximation to the ground-state energy as a functional of the density of interacting electrons in some external potential $v(\mathbf{r})$ (e.g., that created by the fixed positions of the nuclei, V_{ne}) is

$$E_0^{\text{DFT}} \equiv E\left[n_0(\mathbf{r}); v(\mathbf{r})\right] = \min_{\widetilde{\rho}} E\left[\widetilde{n}(\mathbf{r}); v(\mathbf{r})\right] \geq E_0^{\text{Exact}} . \tag{2.12}$$

In the Born–Oppenheimer approximation, the ground-state energy functional depends explicitly on $\widetilde{n}$ and parametrically on $v(\mathbf{r})$. The trial electron density $\widetilde{n}$ should be positive-definite and contain N electrons in that region of the space:

$$\widetilde{n}(\mathbf{r}) \geq 0 , \qquad \int \widetilde{n}(\mathbf{r}) d^3\mathbf{r} = N . \tag{2.13}$$

2.3.1. Wave function theory

WFT methods can be classified into *single-* or *multi-configuration* formalisms whether the electronic wave function is represented by a single Slater determinant or by a linear combination of Slater determinants. Single Slater determinants can describe the electronic state near equilibrium geometries in the PES. However, more than one Slater determinant is necessary to describe the electronic state of compounds containing multivalent elements and the electronic state specifying the dissociation fragments of a molecule. The reminder of this section addresses single-configuration approaches.

In WFT, the exact electronic energy [from Eq. (1.2)] should contain the contributions of the kinetic energy E_{K} of the interacting electrons, the classical Coulomb electron–nuclear attraction E_{ne}, and the electron–electron repulsion energy E_{ee}:

$$E_{\mathrm{el}}^{\mathrm{WFT}} = E_{\mathrm{K}} + E_{\mathrm{ne}} + E_{\mathrm{ee}} \,. \tag{2.14}$$

As WFT are referred to HF theory, the term E_{ee} consists of the contributions of the Coulomb, exchange, and electron correlation energies.[14]

2.3.1.1. Hartree–Fock theory

The simplest form of the trial many-electron wave function obeying the Pauli exclusion principle is the antisymmetrized product of one-electron spin functions (spin orbitals χ_i), i.e., a *Slater determinant* $\Psi = |\chi_1 \ldots \chi_N\rangle$. As a result of applying the variational principle [Eq. (1.11)], one can derive the Hartree–Fock (HF) equations, which determine the optimal spin orbitals by minimizing the value of the total energy with respect to the choice of spin orbitals.[29,42]

The HF method is a mean field approach applied to the many-electron problem, and thus, the HF solution is called the self-consistent field (SCF). The SCF averages out the instantaneous electron correlations; consequently, electrons approach each other more than they should, making the HF wave function become "too ionic" in character.[28]

The HF equations can be recast into a problem of linear algebra by expanding the unknown molecular orbital functions in terms of a fixed and finite set of functions. These functions are usually called the *atomic orbital basis*, because they are atom-centered and resemble solutions to the HF problem for the constituent atoms of a molecule.[28]

In principle, the main sources of error in an *ab initio* electronic structure calculation are, first, the approximate—instead of the exact—solution of the electronic Schrödinger equation, and second, the use of an incomplete set of expansion functions for the molecular orbitals.[28]

2.3.1.2. Many-body perturbation theory

Møller–Plesset many-body perturbation theory is a standard formulation of Rayleigh–Schrödinger perturbation theory using the Fock operator as the zeroth-order Hamiltonian and a Slater determinant of spin orbitals as the zeroth-order wave function. The total electronic energy is given by the sum:

$$E_{\mathrm{el}}^{\mathrm{MP}n} = E^{(0)} + E^{(1)} + E^{(2)} + E^{(3)} + E^{(4)} + \ldots + E^{(n)} + \ldots \tag{2.15}$$

The zeroth-order energy is the sum of the eigenvalues of the occupied spin orbitals, and the sum of the zeroth- and first-order energies is the HF energy. The sum of the remaining terms in the series is the correlation energy.[43] MP*n* energies neither give an upper nor a lower bound to the exact energy, because the formalism is not variational. They could even be lower than the exact energy, but the underestimation of the correlation energy usually avoids it, owing to the basis set truncation.

MP2 is the simplest method to treat electron correlation within WFT. Only single and double excitations contribute to the different correlation terms arising from the interaction of the excited-state wave functions with the HF wave function, because the hamiltonian contains one- and two-electron terms only. The self-consistent optimization of the HF wave function prevents its mixing with the single-excitation wave functions. Accordingly, second- and third-order energies have contributions from double excitations only. In higher orders, there is an indirect coupling through double excitations, and thus, the fourth- and fifth-order energies contain contributions from single, double, triple, and quadruple excitations.[36]

The second-order energy (MP2) is always negative and recovers ~80% of the correlation energy, while the forth-order energy (MP4) may recover up to ~95% of the correlation energy with large basis sets.[28]

2.3.2. Density functional theory

DFT is based on two fundamental theorems. The basic Hohenberg–Kohn theorem ensures the existence of the exact ground state energy of a molecular system as a functional only of the electron density and the fixed positions of the nuclei. Accordingly, for a given nuclear distribution, the electron density uniquely determines the energy and all properties of the ground state, including the electronic hamiltonian and its ground state wave function. Yet the theorem does not state the explicit formula for this functional dependence of the energy on the electron density. The second Hohenberg–Kohn theorem, in turn, ensures that the exact electron density function is the one that minimizes the energy, in that way giving a variational principle to find the density [Eq. (1.12)].[28,40,41]

Kohn–Sham (KS) theory proves that the electronic energy $E_{\mathrm{el}}^{\mathrm{DFT}}$ for a system of N interacting electrons with a hamiltonian $\hat{H}_{\mathrm{el}}$ [Eq. (1.2)] can be written as the contributions of the kinetic energy E_{KS} of the ground state of non-interacting electrons with density distribution $n(\mathbf{r})$, the classical Coulomb electron–nuclear attraction E_{ne}, the classical Coulomb electron–electron repulsion J, and a non-classical term accounting for correlation and exchange effects included in the exchange–correlation (XC) energy E_{XC}:

$$E_{\mathrm{el}}^{\mathrm{DFT}}[n] = E_{KS}[n] + E_{\mathrm{ne}}[n] + J[n] + E_{\mathrm{XC}}[n]. \tag{2.16}$$

As WFT and DFT are complementary formalisms, by equating the exact WFT and DFT energies [Eqs. (1.14) and (1.16)] a "definition" of E_{XC} can be written as:

$$E_{\mathrm{XC}}[n] = (E_{\mathrm{K}} - E_{KS}) + (E_{\mathrm{ee}} - J). \tag{2.17}$$

The first term in parentheses is the kinetic correlation energy including many-body effects. The second term in parentheses contains exchange and potential correlation energies.[14]

DFT is a mathematical framework for viewing electronic structure from the perspective of the electron density.[40] Physical models and approximations of the electronic structure should come from outside KS-DFT to formulate E_{XC} and to enable concrete applications.

In the *local-density approximation* of DFT,[41,44] the exchange–correlation energy of an inhomogeneous system, such as an atom or a molecule, is estimated by an integral whose integrand samples the local density $n(\mathbf{r})$ at each integration point **r**:

$$E_{\mathrm{XC}}^{\mathrm{LDA}} = \int \varepsilon_{\mathrm{XC}}\left[n(\mathbf{r})\right] n(\mathbf{r})\, d^3\mathbf{r} \,. \tag{2.18}$$

In the *generalized gradient approximation* (GGA) of DFT,[41,44] the exchange–correlation energy of the inhomogeneous system is estimated by an integral whose integrand $\varepsilon_{\mathrm{XC}}^{(1)}$ depends on the density and its gradient:

$$E_{\mathrm{XC}}^{\mathrm{GGA}} = \int \varepsilon_{\mathrm{XC}}^{(1)}\left[n(\mathbf{r}), \left|\nabla n(\mathbf{r})\right|\right] d^3\mathbf{r} \,. \tag{2.19}$$

Becke successfully introduced semiempirical fitting parameters in the GGA based on the atomization energies of a standard set of molecules. In a subsequent approach founded on the adiabatic connection formalism, Becke further introduced a hybrid method, with an exchange–correlation expression that contains a parameter a_0 to include non-locality in the real exchange–correlation hole:[28,40,41]

$$E_{\mathrm{XC}}^{\mathrm{Hyb}} = E_{\mathrm{XC}}^{\mathrm{DFT}} + a_0\left(E_{\mathrm{X}}^{\mathrm{Exact}} - E_{\mathrm{X}}^{\mathrm{DFT}}\right) . \tag{2.20}$$

This was shortly thereafter combined with the LYP correlation functional to yield the hybrid B3LYP functional. Hybrid approximations reduce average bond energy errors from about 6 kcal/mol for pure GGA's to roughly 2 kcal/mol. Reaction barrier heights are also improved by exact exchange mixing.[27] For these reasons, DFT qualifies as a powerful and efficient approach to solve problems in electronic structure theory, with accuracy comparable to the MP2 method and lower basis set requirements.

2.4. Electronic structure and environment interaction

The energy exchange of molecules with their surroundings modifies their electronic structures. In particular, this section addresses two types of interactions: those with the bulk solvent and those with certain electromagnetic waves.

2.4.1. Continuum solvation theory

Polarizable continuum solvation models (e.g., Tomasi's Polarized Continuum Model (PCM) and the Solvation Model (SM*x*) of Cramer and Truhlar) provide rather accurate values of the bulk Gibbs energy of solvation $\Delta_s G^\theta$ and a good physical description of the components involved in the bulk solvation process, as reviewed elsewhere.[45,46,47] In the models, the Gibbs energy of solvation $\Delta_s G^\theta$ can be split up into three main contributions: ΔG_{cav}, relating to the creation of an empty cavity in a dielectric medium where the solute molecule is placed; ΔG_{el}, describing the interaction of the solute charge distribution within the cavity with the continuum medium of dielectric constant ε; and ΔG_{vdW}, describing the solute–solvent van der Waals interactions:

$$\Delta_s G^\theta = \Delta G_{cav} + \Delta G_{el} + \Delta G_{vdW} . \tag{2.21}$$

The electrostatic component ΔG_{el} directly depends on the solvent inner potential. For a given charge distribution of the solute inside a cavity, the solvent inner potential consists of the electrostatic potential in the vacuum inside the cavity and the electrostatic potential generated by the solvent charge distribution outside. In the polarizable models, the electric field exerted on the solute by the solvent it has polarized is called *reaction field*. This electric field modifies the electronic structure of the solute, which brings about a subsequent change in polarization of the solvent. Iterating to self-consistency is called the *self-consistent reaction field* (SCRF) method.[48]

2.4.2. Molecular electric response theory

An absorption electronic spectrum is a record of spectral intensity as a function of frequency (ν) of the radiation absorbed by molecules. The frequency assigned to a band is related to the electron excitation energy following the law of conservation of energy in a quantum transition (Fermi's golden rule). In the *dipole approximation*, the spatial variation of the radiation field across the quantum system can be neglected, because the

radiation wavelength λ is large compared with the size a of the system ($\lambda >> a$). Assuming this approximation, the electronic transition takes place provided the spectrum of the incident electromagnetic radiation contains frequencies satisfying the Bohr frequency condition ($\varepsilon_f - \varepsilon_i = h\nu$). For such resonant transitions, the intensity of a band is directly proportional to the transition dipole moment.[33,49]

2.4.2.1. Static molecular response to an electric field

Neighboring molecules or an external apparatus can create an electrostatic field around a molecule. The simplest perturbation can be generated, for instance, with a uniform electrostatic field applied in an arbitrary direction, say, the z-direction; $\mathbf{F} = (0,0,F_z)$, where F_z is the field strength. The total electronic energy may be expanded in the field **F** as follows:[50]

$$E_{\text{tot}}(\mathbf{F}) = E_0 - \mathbf{F}^{\text{T}}\boldsymbol{\mu}_0 - \tfrac{1}{2}\mathbf{F}^{\text{T}}\boldsymbol{\alpha}\mathbf{F} + O(\mathbf{F}^3)\,, \tag{2.22}$$

where

$$E_0 = E_{\text{tot}}(\mathbf{0}) \qquad \boldsymbol{\mu}_0 = -\left.\frac{dE_{\text{tot}}}{d\,\mathbf{F}}\right|_{\mathbf{F}=\mathbf{0}} \qquad \boldsymbol{\alpha} = -\left.\frac{d^2E_{\text{tot}}}{d\,\mathbf{F}^2}\right|_{\mathbf{F}=\mathbf{0}}.$$

The *permanent molecular dipole moment* $\boldsymbol{\mu}_0$ and the *dipole-polarizability tensor* $\boldsymbol{\alpha}$ at zero-field describe the response of the molecular system to the external electric perturbation **F**. Expanding the energy to higher orders in the field **F**, we obtain molecular hyperpolarizability tensors of different orders (e.g., the first-order hyperpolarizability $\boldsymbol{\beta}$ and so on), needed for an accurate description of the system in strong fields.

Assuming that the molecular charge distribution is fixed, the only influence of the external field **F** is a change in the total energy. If the charge distribution is not fixed, it gets polarized in the presence of the external field **F**, thereby reorganizing and relaxing itself so as to minimize the energy. The dipole moment of the molecular charge distribution $\boldsymbol{\mu}$ changes; a field-dependent dipole moment arises $\boldsymbol{\mu}_{\text{ind}}$, induced by the external field, in addition to the permanent field-independent dipole moment $\boldsymbol{\mu}_0$:

$$\begin{aligned}\boldsymbol{\mu}(\mathbf{F}) = -\frac{dE_{\text{tot}}}{d\mathbf{F}} &= \boldsymbol{\mu}_0 + \boldsymbol{\alpha}\mathbf{F} + \tfrac{1}{2}\mathbf{F}^{\text{T}}\boldsymbol{\beta}\mathbf{F} + O(\mathbf{F}^3) \\ &= \boldsymbol{\mu}_0 + \boldsymbol{\mu}_{\text{ind}}(\mathbf{F})\end{aligned}\,. \tag{2.23}$$

2.2.4.2. Dynamic molecular response to an electric field

The interaction of the molecule with light can be modelled as the interaction with an electric field varying sinusoidally in time. If the molecule is subject to a time-dependent electric field strength, $\mathbf{F}(t) = (0,0,F_z(t))$ —e.g., $F_z(t) = F(\omega)\cos\omega t$ —the dipole moment of the molecular charge distribution $\boldsymbol{\mu}$, transformed to the frequency domain ($\omega = 2\pi\nu$), is

$$\boldsymbol{\mu}(\omega) = \boldsymbol{\mu}_0 + \boldsymbol{\mu}_{\text{ind}}(\omega) = \boldsymbol{\mu}_0 + \boldsymbol{\alpha}(\omega)\mathbf{F}(\omega) + \ldots , \tag{2.24}$$

where the frequency-dependent $\boldsymbol{\alpha}(\omega)$ is the *dynamic dipole-polarizability*. Applying perturbation theory, it can be shown that the *mean dynamic polarizability* $\overline{\alpha}(\omega)$ is

$$\overline{\alpha}(\omega) = \tfrac{1}{3}\,\mathrm{tr}\,\boldsymbol{\alpha}(\omega) = \sum_I \frac{f_I}{\omega_I^2 - \omega^2} , \tag{2.25}$$

with *vertical excitation energies* ($I \leftarrow 0$) from the ground state to the excited state I :

$$\hbar\omega_I = E_I - E_0 . \tag{2.26}$$

The electronic excitations can be described as quantized oscillators and the corresponding transition probabilities as oscillator strengths. The poles of $\overline{\alpha}(\omega)$ occur at the excitation energies with *oscillator strengths* defined as

$$f_I = \tfrac{2}{3}\hbar\omega_I \left| \left\langle \Psi_0 \left| \hat{\boldsymbol{\mu}}_0 \right| \Psi_I \right\rangle \right|^2 . \tag{2.27}$$

The oscillator strengths depend on the excitation energies and the square of the *transition dipole moments* $\left\langle \Psi_0 \left| \hat{\boldsymbol{\mu}}_0 \right| \Psi_I \right\rangle$. Oscillator strengths and experimentally measurable *molar extinction coefficients* ε ($\mathrm{M^{-1}cm^{-1}}$) of the absorption bands are related by the expression

$$f_I = 4.319\times10^{-9}\,\frac{9n_0}{\left(n_0^2+2\right)^2}\int_{\overline{\nu}_1}^{\overline{\nu}_2}\varepsilon(\overline{\nu})\,d\overline{\nu} \cong 4.319\times10^{-9}\,\frac{9n_0}{\left(n_0^2+2\right)^2}\,\Delta\overline{\nu}\times\varepsilon_{\max} , \tag{2.28}$$

where n_0 is the refractive index of the medium, $\overline{\nu}$ the wavenumber in $\mathrm{cm^{-1}}$, $\Delta\overline{\nu}$ the bandwidth, and $\varepsilon_{\max}$ the maximum value of the molar extinction coefficient for that absorption band.[51]

2.4.2.3. Time-dependent density functional response theory (TD-DFT)

TD-DFT is applied to the calculation of dynamic polarizabilities, hyperpolarizabilities, and electronic excitation spectra. The classic formulation of KS-DFT is restricted to the time-independent case. The formalism of TD-DFT generalizes KS theory to include the case of a time-dependent, local external potential. A practical computational formulation of TD-DFT can be developed using time-dependent linear response theory.[52]

The interaction of a harmonically time-dependent electric field with a molecule brings about electronic excitations that can be associated with spin-orbital permutations in the KS ground-state wave function $\Psi_0^{\mathrm{KS}} = \left| \chi_1 \ldots \chi_N \right\rangle$. These permutations form a point group of symmetry isomorphic with that of the molecular orbital rotations. The permutations are given by a unitary transformation $\mathbf{U}$ (i.e., $\mathbf{U}^{+} = \mathbf{U}^{-1}$), which can be written as an exponential transformation of molecular orbitals by the response matrix $\mathbf{P}$. The response matrix is anti-hermitian (i.e., $\mathbf{P}^{+} = -\mathbf{P}$) and contains the orbital rotation parameters. Using the exponential ansatz,[53] the time-dependent perturbed KS wave function can then be written as

$$\Psi^{\mathrm{KS}} = \exp(\mathbf{P})\Psi_0^{\mathrm{KS}}. \tag{2.29}$$

The interaction with a time-dependent electric field can be expressed as a first-order external perturbation to the total energy for small values of the electric field strength of the electromagnetic wave. We can write the KS-hamiltonian as $\hat{H}^{\mathrm{KS}} = \hat{H}_0^{\mathrm{KS}} + \lambda \hat{V}$ and the total energy of the perturbed molecular system as

$$E_{\mathrm{tot}}^{\mathrm{DFT}} = \left\langle \Psi_0^{\mathrm{KS}} \left| \exp(-\mathbf{P}) \hat{H}^{\mathrm{KS}} \exp(\mathbf{P}) \right| \Psi_0^{\mathrm{KS}} \right\rangle. \tag{2.30}$$

Using the Hausdorff commutator expansion, the KS-DFT total energy turns out to be Eq. (2.31), which can further be recast into the matrix form[54] of Eq. (2.32):

$$E_{\mathrm{tot}}^{\mathrm{DFT}} = E_0^{\mathrm{DFT}} + \lambda \sum_i V_{ii} + \left[\mathbf{P}, \lambda \mathbf{V}\right] + \left[\mathbf{P}, \left[\mathbf{P}, \lambda \mathbf{V}\right]\right] + O(\mathbf{P}^3), \tag{2.31}$$

$$E_{\mathrm{tot}}^{\mathrm{DFT}} = E_0^{\mathrm{DFT}} + \lambda \left[\sum_i V_{ii} + \begin{pmatrix} \mathbf{P} \\ \mathbf{P}^* \end{pmatrix}^{+} \begin{pmatrix} \mathbf{V}^* \\ \mathbf{V} \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \mathbf{P} \\ \mathbf{P}^* \end{pmatrix}^{+} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix} \begin{pmatrix} \mathbf{P} \\ \mathbf{P}^* \end{pmatrix} \right] + O(\mathrm{P}^3). \tag{2.32}$$

In these equations, the first term is the KS-DFT ground-state energy. The second term contains the trace of $\hat{V}$, and the third term, which is related to the dynamic dipole-polarizability, consists of contributions of the non-diagonal elements of $\hat{V}$. The forth term contains the frequency-dependent stability matrices $\mathbf{A}(\omega)$ and $\mathbf{B}(\omega)$, whose elements consist of the applied potential in the basis of the perturbed KS molecular orbitals. The optimal response matrix **P** makes the total energy stationary, and the excitation energies are the poles of $\boldsymbol{\alpha}(\omega)$:

$$\left.\frac{dE_{\text{tot}}^{\text{DFT}}}{d\,\mathbf{P}}\right|_{\lambda=1} = \begin{pmatrix}\mathbf{V}^*\\ \mathbf{V}\end{pmatrix} + \begin{pmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{B} & \mathbf{A}\end{pmatrix}\begin{pmatrix}\mathbf{P}\\ \mathbf{P}^*\end{pmatrix} = \mathbf{0}\,, \quad \boldsymbol{\alpha}(\omega) = \begin{pmatrix}\mathbf{V}^*\\ \mathbf{V}\end{pmatrix}^{+}\begin{pmatrix}\mathbf{P}\\ \mathbf{P}^*\end{pmatrix}. \tag{2.33}$$

In the dipole approximation, the electric field depends only on time, since the spatial variation of the field across the quantum system can be neglected. The subsequent approximations deal with this explicit dependence of the perturbation on time.[52,55,56,57,58,59,60,61] The interaction is usually represented by a sinusoidally varying time-dependent potential referred to as a harmonic perturbation, i.e., $V(t) = V(\omega)\exp(i\omega t) + V^*(\omega)\exp(-i\omega t)$.

The accuracy of TD-DFT excitation energies is ~0.3 eV for the lowest valence excited states.[62] For a given basis set, TD-DFT shows computational requirements scaling as $M^3 - M^4$, whereas CIS and CCSD have accuracies of ~1 eV and ~0.15 eV, and computational requirements scaling as M^5 and M^6, respectively.

2.5. Theoretical model chemistries

A *theoretical model chemistry* is a specification of an approximate method to solve the electronic problem by WFT or DFT and a representation of the atomic orbital basis. A theoretical model for any complex process is approximate, but it must be a well-defined mathematical procedure of simulation.[63,64] Accordingly, it should have some desirable properties such as unbiased model, correct size scaling, calculation of continuous PES, high accuracy, low computational cost, and wide application.[28,64]

In particular, a critical issue of theoretical model chemistries is their size-scaling properties, i.e., accuracy of a calculation with approximate methods should not degrade as the molecular system gets larger. Therefore, the total energy has to scale correctly with the number of independent subsystems and the number of electrons. The property of a theoretical method termed *size-consistency* guarantees that the total energy of a system composed of two molecules far apart, in the non-interacting limit, is equal to the sum of total energies of the isolated molecules, i.e., $E_{\text{AB}}(r_{\text{AB}} \to \infty) = E_{\text{A}} + E_{\text{B}}$.[65,66] Hence, size-consistency accounts for a correct dissociation into fragments. The

property of a theoretical method termed *size-extensivity* ensures that the total energy is a linear function of the number of electrons in a homogeneous electron gas.[67] In general, size-extensivity accounts for the correct growth of the total energy with increasing number of electrons in systems with non-homogeneous electron densities.[28,29,64,68]

3. Basics of Interfacial Electrochemistry

Electrochemistry deals with the processes of charge separation and flow in chemical systems. Those processes usually take place at interphases, in which one of the phases carries electrons, and the other ions. Electrochemical reactions are the electron-transfer reactions that take place on the interface.

Interfacial electrochemistry is a field that has received the attention of both physicists and chemists, and as a consequence, the same phenomena have been described in two different scientific languages. That has produced a variety of conventions and equivalent physical quantities, bringing about plenty of confusion in both scientific communities. This chapter summarizes the very essential definitions and our theoretical assumptions to characterize the electrochemical behaviour of electrons in metal–solution interphases. Herein the conventional electrodic and thermodynamic descriptions are recast into a tractable form enabling us to consistently separate out the molecular problem from the interfacial problem. Details are included in Ref. **S-I**.

3.1. Thermodynamics of a metal–solution interphase

Our model of interphase consists of a single M|Red,Ox electrode in electrochemical equilibrium under standard temperature T and pressure p (M|S in shorthand notation). This single electrode consists of an inert metal M dipped into an electrolyte solution (S) of the oxidized (Ox) and reduced (Red) forms, as depicted in Figure 3.1. It should be noted that part of the surfaces of M and S are in direct contact with each other and the remainders are surrounded by a field-free vacuum. In addition, the solution is protected with a rigid and impermeable wall to the passage of substance from solution to vacuum so as to avoid the irreversible boiling.

In our simplifications of the perfectly polarized interphase, we assume that the metal surface is negatively charged and the excess charge is spread over the surface in contact with both the solution and the vacuum. Unlike the metal, in which the surface altogether can be negatively charged owing to the highly delocalized band structure of the electronic states, the solution phase is an ionic conductor and most of the excess positive charge is most likely to be spread over the solution side of the metal–solution interface. Far

enough from the metal–solution interface, the solution surface in contact with the vacuum can be assumed to be uncharged.

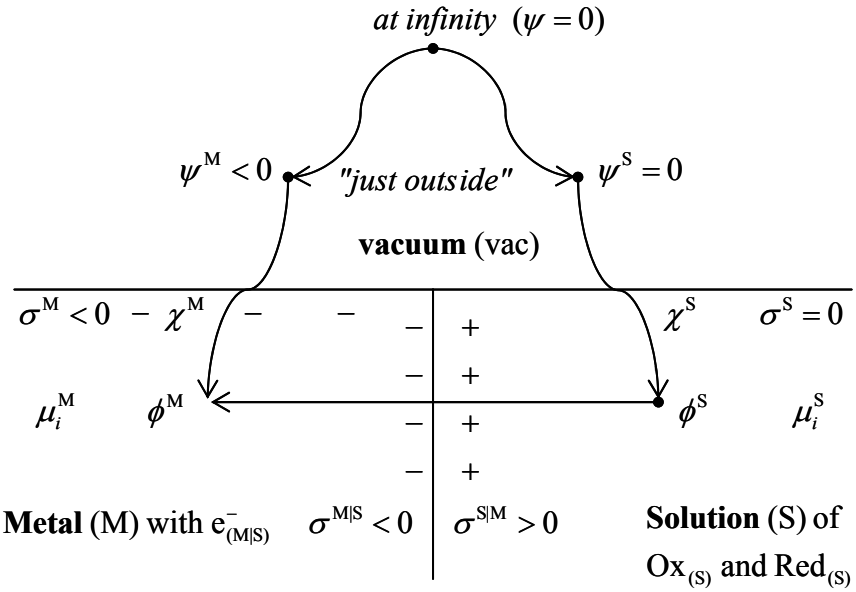


Figure 3.1. A single metal–solution electrode.

3.1.1. Electrostatic potentials

The electric field arising from the internal charge distribution of phase β ($\beta = \mathrm{M,S}$) defines an electrostatic potential within β, termed the *inner potential* ϕ^{β}, the gradient of which is equal to the negative of the electric field strength $-\nabla\phi^{\beta} = \mathbf{F}^{\beta}$.[69] If the phase has a total electric charge, the internal charge distribution also defines an electrostatic potential outside β, termed the *outer potential* ψ^{β}.[69] Concerning experimental measurement, the behaviour of the outer potential defines two relevant regions in the outside vacuum:

- *Vacuum "just outside" the phase*: This region is located at distances larger than ~100 Å, far enough from the surface so that a test point charge is out of the range of the image forces from the surface and does not disturb the charge distribution within the phase. The value of the outer potential is mainly defined by the surface charge distribution σ,[70] as illustrated in Figure 3.1.
- *Vacuum at infinity*: This region is located very far away from the M|S interphase where it is seen as a dipole, and the outer potential vanishes asymptotically at infinite distance, given that $\psi \propto 1/r^2$.

Particles in the surface region are subject to orienting forces owing to the absence of charge balance on the surface. The array of polar or polarized molecules or both on the surface of a phase in contact with the vacuum is called the *surface dipole layer* and gives rise to the *surface potential* χ^β, which is independent of the surface charge.[70] The surface potential is the work to make a positive point charge traverse the surface dipole layer from the bulk of the phase into the vacuum "just outside" it.[71] The inner, surface, and outer potentials are related by:

$$\phi^\beta = \chi^\beta + \psi^\beta \tag{3.1}$$

3.1.2. Chemical potentials

The *electrochemical potential* $\overline{\mu}_i^\beta$ and the *real chemical potential* α_i^β are state functions defined in terms of the electrostatic potentials within phase β:[69,72]

$$\overline{\mu}_i^\beta \equiv \mu_i^\beta + z_i e \phi^\beta , \tag{3.2}$$

$$\alpha_i^\beta \equiv \mu_i^\beta + z_i e \chi^\beta . \tag{3.3}$$

If the total charge of the phase is zero, its outer potential is also zero, $\psi^\beta = 0$. The real chemical potential may thus be regarded as the value of the electrochemical potential of the uncharged phase ($\overline{\mu}_i^\beta = \alpha_i^\beta$),[69] since $\overline{\mu}_i^\beta$ can alternatively be written as

$$\overline{\mu}_i^\beta = \alpha_i^\beta + z_i e \psi^\beta . \tag{3.4}$$

The chemical potential μ_i^β of a component of a mixture in phase β, is the partial molar Gibbs energy of that component, i.e., the state function $\mu_i^\beta = \mu_i^\beta(T, p, \{x_j^\beta\})$ of temperature T, pressure p, and composition of the mixture in terms of mole fractions $\{x_j^\beta\}$. The chemical potential μ_i^β can be expressed as an explicit function of a concentration variable (mole fraction x_i^β or activity a_i^β) of component i only, i.e., $\mu_i^\beta = \mu_i^\beta(T, p, x_j^\beta)$ for the *ideal mixture* or $\mu_i^\beta = \mu_i^\beta(T, p, x_j^\beta)$ for the *real mixture*. The chemical potential μ_i^β can be split up into the contributions of a standard term $\mu_i^{0,\beta}$, which depends on T and p for unit concentration, and a mixing term, $RT \ln x_i^\beta$ or $RT \ln a_i^\beta$, which depends on the composition variable of constituent i in the mole fraction scale:[73]

$$\mu_i^\beta \equiv \mu_i^{0,\beta} + RT \ln x_i^\beta , \tag{3.5}$$

$$\mu_i^\beta \equiv \mu_i^{0,\beta} + RT \ln a_i^\beta . \tag{3.6}$$

For practical purposes, chemical potentials in solution are given in three concentration scales: mole fraction x, molality m, and molarity c.[73,74] These concentration scales in a dilute solution S are related by

$$x_i = M_1 \times m_i = \frac{M_1}{\rho_1} \times c_i , \tag{3.7}$$

where M_1 and ρ_1 are the molar mass (in g/mol) and the density (in g/mL) of the solvent. Accordingly, standard chemical potentials in the molality and molarity scales, $\mu_{m,i}^{0,\mathrm{S}}$ and $\mu_{c,i}^{0,\mathrm{S}}$ can respectively be obtained from standard chemical potentials defined in the mole fraction scale $\mu_i^{0,\mathrm{S}}$ [Eqs. (3.5) and (3.6)] by adding a constant dependent on the solvent:

$$\mu_{m,i}^{0,\mathrm{S}} = \mu_i^{0,\mathrm{S}} + RT \ln M_1 \times m^0 , \tag{3.8}$$

$$\mu_{c,i}^{0,\mathrm{S}} = \mu_i^{0,\mathrm{S}} + RT \ln \frac{M_1}{\rho_1} \times c^0 . \tag{3.9}$$

These expressions hold for aqueous and non-aqueous solutions. In the especial case of dilute aqueous solutions, the difference between the molality and molarity scales can be ignored, because it depends on the density of liquid water, which is ~1 g/mL. Accordingly, $\mu_i^{0,\mathrm{S}}$ must be shifted by –2.38 kcal/mol to get $\mu_{m,i}^{0,\mathrm{S}}$ and $\mu_{c,i}^{0,\mathrm{S}}$ at 298 K and 1 atm, taking into account the standard state conditions $m^0 = 1\,\mathrm{mol/kg}$ and $c^0 = 1\,\mathrm{mol/L}$.

Standard thermodynamic quantities of solutes in liquid solutions are referred to a standard state of a standard reference system, which is a hypothetical ideal solution with $m_i = m^0 = 1\,\mathrm{mol/kg}$ or $c_i = c^0 = 1\,\mathrm{mol/L}$ and the properties of an *ideal-dilute solution*. Solute–solute interactions are present in nonideal-dilute solutions, but they vanish in the limit of infinite dilution. Only solute–solvent and solvent–solvent interactions remain in an ideal-dilute solution.

3.2. Electrochemical potential of electrons

In a condensed phase, the experimental and theoretical location of the electronic energy levels requires the specification of explicit relationships of the electrochemical potential of electrons with the measurable bulk thermodynamic properties (i.e., electron work function and electrode potential), as well as with computable quantum statistical mechanical properties (i.e., Fermi energies).

3.2.1. Work function and Fermi level

3.2.1.1. In a simple phase with an uncharged surface

The real chemical potential of electrons $\alpha^{\beta}(\mathrm{e}^{-})$ can be measured as the negative of the work function for the electron emission Φ^{β}, provided the simple phase β with an uncharged surface in contact with the vacuum has $\psi^{\beta} = 0$ from the region "just outside" the phase to infinity:[70]

$$\overline{\mu}^{\beta}(\mathrm{e}^{-}) = \alpha^{\beta}(\mathrm{e}^{-}) = -\Phi^{\beta} \tag{3.10}$$

Since the real chemical potential depends on the surface potential, the work function Φ^{β} can vary from face to face in a single crystal.[70,72] According to the free-electron ideal gas model for conductors (e.g., for metals), at absolute zero the most energetic electrons in the uncharged phase are at the highest occupied energy level, the Fermi level ε_{F}, which coincides with the chemical potential of electrons $\mu^{\beta}(\mathrm{e}^{-})$.[75] It is also known that for typical metallic densities the chemical potential remains equal to the Fermi energy from absolute zero to room temperature. Therefore, combining $\mu^{\beta}(\mathrm{e}^{-}) = \varepsilon_{\mathrm{F}}$ and Eqs. (3.3) and (3.10), we can write:

$$\overline{\mu}^{\beta}(\mathrm{e}^{-}) = \alpha^{\beta}(\mathrm{e}^{-}) = \mu^{\beta}(e^{-}) - e\chi^{\beta} = \varepsilon_{\mathrm{F}} - e\chi^{\beta} = -\Phi^{\beta} \tag{3.11}$$

Consequently, the electron work function of a solid whose surface is in contact with a vacuum is defined as the energy required to raise an electron in the bulk of the solid from the Fermi level ε_{F} to the energy level corresponding to an electron at rest in the vacuum "just outside" it across the uncharged surface.[76,77,78]

It should be noted that the assumption of an uncharged surface guarantees $\psi^{\beta} = 0$ from "just outside" the phase to infinity and enables us to relate the experimental electron work function with the theoretically calculated Fermi level. Φ^{β} is defined relative to the field-free vacuum "just outside" the phase. In contrast, the Fermi level is defined relative to the state where the electron and the phase are at rest in a field-free vacuum at infinite separation. This is so because the Fermi energy ε_{F} is an eigenvalue of the one-electron hamiltonian containing an effective single-particle mean-field potential in which the electrons move under the influence of the nuclear lattice.

3.2.1.2. In a single electrode

An electron can be brought from infinity to the metal of a single electrode following the two pathways depicted in Figure 3.1, i.e., it can be brought directly from vacuum to the metal (path on the left side) or from vacuum to the metal traversing the solution (path on the right side). The work to reversibly bring one mole of electrons from the vacuum at infinity to the

bulk metal at constant T and p is equal to the electrochemical potential of electrons in the metal. To make a clear distinction between the pathways the electrons take from vacuum to the bulk metal, the electrochemical potential is denoted $\overline{\mu}^{\mathrm{M}}(\mathrm{e}^-)$ and $\overline{\mu}^{\mathrm{M|S}}(\mathrm{e}^-)$ for the paths on the left and right sides.

It can be demonstrated (Ref. **S-I**), first, that the electrochemical potential of electrons in the metal is the same regardless the path chosen to bring them from infinity into the metal of the single electrode, and second, that the electrochemical potential of electrons in the metal is equal to the negative of the electron work function of the solution Φ^{S}, since $\overline{\mu}^{\mathrm{M}}(\mathrm{e}^-) = \overline{\mu}^{\mathrm{S}}(\mathrm{e}^-)$ at electrochemical equilibrium.[72,79] In short, since the electrochemical potential is a state function and provided $\psi^{\mathrm{S}} = 0$, one can show that:

$$\overline{\mu}^{\mathrm{M|S}}(\mathrm{e}^-) = \overline{\mu}^{\mathrm{M}}(\mathrm{e}^-) = \overline{\mu}^{\mathrm{S}}(\mathrm{e}^-) = -\Phi^{\mathrm{S}}. \tag{3.12}$$

Moreover, if $\mathrm{Ox}_{(\mathrm{S})} + \nu\mathrm{e}^-_{(\mathrm{M|S})} = \mathrm{Red}_{(\mathrm{S})}$ is the half-reaction for the single electrode, then $\overline{\mu}^{\mathrm{M|S}}(\mathrm{e}^-)$ coincides with the Fermi level of the redox couple, since the Fermi–Dirac distribution function describing the equilibrium concentration of Red leads to the correct equilibrium condition.[72,76,77]

3.2.2. Electrode potentials

The half-reaction holding at our single electrode can in general be written as

$$\mathrm{Ox}_{(\mathrm{S})} + \nu\mathrm{e}^-_{(\mathrm{M|S})} = \mathrm{Red}_{(\mathrm{S})}. \tag{3.13}$$

The redox potential, either in the reduction $E(\mathrm{Ox/Red})$ or the oxidation $E(\mathrm{Red/Ox})$ conventions, is equivalent to the absolute electrochemical potential of electrons $\overline{\mu}^{\mathrm{M|S}}(\mathrm{e}^-)$ in the metal–solution interphase, according to the linear relationship in terms of the Gibbs energy:[76]

$$\tfrac{1}{\nu}\Delta_{\mathrm{ET}}G(\mathrm{Ox/Red}) = -eE(\mathrm{Ox/Red}) = \overline{\mu}^{\mathrm{M|S}}(\mathrm{e}^-) + C. \tag{3.14}$$

The constant C defines the energy level of a conventional reference state. The vacuum at infinity, where $C = -\mu^{\mathrm{vac}}_{0\mathrm{K}}(\mathrm{e}^-) = 0$, defines the *absolute electrode potential*, $E_{\mathrm{abs}}(\mathrm{Ox/Red})$. Single reference electrodes, e.g., the saturated calomel electrode (SCE) or the standard hydrogen electrode (SHE), define electrode potentials in the *SCE scale* $E_{\mathrm{SCE}}(\mathrm{Ox/Red})$, where $C = -\overline{\mu}^{\theta,\mathrm{SCE}}(\mathrm{e}^-)$, or in the *standard hydrogen scale* $E_{\mathrm{SHE}}(\mathrm{Ox/Red})$, where $C = -\overline{\mu}^{\theta,\mathrm{SHE}}(\mathrm{e}^-)$.

The *relative electrode potential* of a redox couple in an interphase is defined with respect to another reference interphase. The term *absolute* potential stands for an electrode potential not based on another reference electrode but on a given reference electronic energy taken as zero. The

definition of a common absolute scale makes it possible to directly locate electronic energy levels of the solid–liquid, solid–gas, and solid–vacuum interphases, as well as the energy levels from the Schrödinger equation in a common frame of reference for the electronic energy.

For cells without liquid junctions, and for those from which the liquid junction potentials have been eliminated in the conventional manner, one can assume that both $\psi^{\mathrm{S}} = 0$ and $\chi^{\mathrm{S}} = 0$. It thus turns out that $\overline{\mu}^{\mathrm{M|S}}(\mathrm{e}^-)$ becomes equal to $\mu^{\mathrm{S}}(\mathrm{e}^-)$, and a more general form of the absolute potential may be defined, i.e., the *reduced absolute potential*:

$$-eE_{\mathrm{abs,r}}(\mathrm{Ox/Red}) = \mu^{\mathrm{S}}(\mathrm{e}^-)\,. \tag{3.15}$$

Assuming electrochemical equilibrium and standard conditions of temperature, pressure, and composition in a given basis, electrode potentials can be expressed in terms of the electrochemical potentials of the species involved in the half-reaction:

$$-eE^{\theta}_{\mathrm{abs}}(\mathrm{Ox/Red}) = \Delta_{\mathrm{ET}}G^{\theta} = \overline{\mu}^{\theta,\mathrm{aq}}(\mathrm{Red}) - \overline{\mu}^{\theta,\mathrm{aq}}(\mathrm{Ox}) - \mu^{\mathrm{vac}}_{0\mathrm{K}}(\mathrm{e}^-)\,, \tag{3.16}$$

$$-eE^{\theta}_{\mathrm{SHE}}(\mathrm{Ox/Red}) = \Delta_{\mathrm{ET}}G^{\theta} = \overline{\mu}^{\theta,\mathrm{aq}}(\mathrm{Red}) - \overline{\mu}^{\theta,\mathrm{aq}}(\mathrm{Ox}) - \overline{\mu}^{\theta,\mathrm{SHE}}(\mathrm{e}^-)\,.$$

If we further assume that $\psi^{\mathrm{S}} = 0$ and $\chi^{\mathrm{S}} = 0$ for cells without liquid junction potentials, electrode potentials can be written in terms of chemical potentials:

$$-eE^{\theta}_{\mathrm{abs,r}}(\mathrm{Ox/Red}) = \Delta_{\mathrm{ET}}G^{\theta} = \mu^{\theta,\mathrm{aq}}(\mathrm{Red}) - \mu^{\theta,\mathrm{aq}}(\mathrm{Ox}) - \mu^{\mathrm{vac}}_{0\mathrm{K}}(\mathrm{e}^-)\,, \tag{3.17}$$

$$-eE^{\theta}_{\mathrm{SHE}}(\mathrm{Ox/Red}) = \Delta_{\mathrm{ET}}G^{\theta} = \mu^{\theta,\mathrm{aq}}(\mathrm{Red}) - \mu^{\theta,\mathrm{aq}}(\mathrm{Ox}) - \mu^{\theta,\mathrm{SHE}}(\mathrm{e}^-)\,.$$

Therefore, the interfacial problem is simplified to a general and accurate definition of electrochemical, real chemical, and chemical potentials according to quantum and statistical mechanics. A suitable definition should be at the same time compatible with the experimental absolute scale and the electrostatic convention, as it is presented in Chapter 4.

3.3. Concluding summary

Our model for the zeroth-class single electrode assumes that the negative charge density can be spread over the whole metal surface, because of the highly delocalized band structure of the electronic states. In contrast, the excess positive charge is more likely to be spread over the solution side of the metal–solution interface, being the solution phase an ionic conductor. Far

enough from the metal–solution interface, the solution surface in contact with the vacuum is assumed to be uncharged.

The absolute electrochemical scale uses the energy level corresponding to an electron at rest in a field-free vacuum as its zero (referred to as the reference vacuum level), and hence, $\mu_{0K}^{vac}(e^-) = 0$.[70,79] Electrons in the reference vacuum level are assumed to behave as classical particles so that they can be at rest in a zero electrostatic potential vacuum, implying $T = 0$ and $p = 0$ in the macrostate. This is in agreement with the electrostatic convention.

We hold that the reference vacuum level "just outside" the phase is in accord with the experimental determination of the electron work function, but the reference vacuum level "at infinity" is in agreement with the electrostatic convention, and in consequence, a more general choice.

We assume that the solution outer potential should be zero, $\psi^S = 0$, so that both reference vacuum levels yield the same electrostatic work to bring the electron from the reference vacuum level to a point within S or M going through S. As a result, $\psi^S = 0$ enables that an experimental absolute electrode potential, relative to the vacuum "just outside", can be directly referenced to the vacuum at infinity, in line with the electrostatic convention.

Under these approximations, we recast the conventional electrodic problem into the simpler molecular problem of finding out the electrochemical, real chemical, and chemical potentials of all the species involved in the half-reaction.

4. Molecular Physical Chemistry

Molecular physical chemistry concerns the study of physical changes associated with chemistry, which ranges from the structure of matter and the fundamental atomic and molecular interactions to the macroscopic properties and processes arising from the average microscopic behaviour.

We herein follow this molecular approach to define absolute electrochemical, real chemical, and chemical potentials of species in solution and estimate their standard values. This is achieved by applying the scaling laws of statistical mechanics to the collective behaviour of atoms and molecules, whose motion, interactions, and properties are described by first principles quantum chemistry. As chemical species are composed of electrons and nuclei, we should find an internally consistent set of absolute standard values for all three chemical potentials of electrons and protons. As a result, we can get rid of problems in referencing chemical data and enable a uniform thermochemical treatment of electron, proton, and proton-coupled electron transfer reactions in solution. Details are mainly included in Ref. **S-I**, but also in Refs. **S-II**, and **S-III**.

4.1. Absolute electrochemical potentials

The electrochemical potential $\overline{\mu}_i^{\beta}$ is said to be *absolute* if we define $\overline{\mu}_i^{\beta}(T=0, p=0, c_i=0) \equiv 0$ for a hypothetical reference state. If $\overline{\mu}_i^{\beta}$ is absolute, then $\overline{\mu}_i^{0,\beta} \equiv \overline{\mu}_i^{\beta}(T=298\,\mathrm{K}, p=1\,\mathrm{bar}, c_i=1\,\mathrm{mol/L})$ is the *absolute standard* value of the electrochemical potential (cf. §3.1.2).

Absolute standard electrochemical potentials $\overline{\mu}_i^{0,\beta}$ have a simple and primary physical meaning: they represent the Gibbs energy change to assemble one mole of any chemical species *i* from nuclei and electrons in the vacuum state and bring it to the condensed phase *β*, at a predefined standard macrostate of temperature, pressure, and concentration (e.g., $T=298\,\mathrm{K}$, $p=1\,\mathrm{bar}$, and $c_i=1\,\mathrm{mol/L}$). The assembled species (e.g., atoms, molecules, as well as activated complexes and atomic or molecular ions of any spin multiplicity) and the elementary species (i.e., nuclei and electrons) are assumed to be at rest and at infinite separation from each other in the field-free vacuum—i.e., in a hypothetical reference macrostate with $T=0$, $p=0$, and $c_i=0$.

From the macroscopic standpoint, owing to the electroneutrality condition, the outer potential of a single solution is zero ($\psi^{\mathrm{S}} = 0$; therefore, $\phi^{\mathrm{S}} = \chi^{\mathrm{S}}$) and the electrochemical potential of species *i* is their real chemical potential, according to Eq. (3.4). Since the electrochemical potential can be calculated from the properties of the individual molecules composing the system, it can be further split up into two equivalent ways that allow us to define the absolute standard chemical potentials in the gas phase and solution in the suitable concentration scale ($\mu_i^{0,\mathrm{g}}$ and $\mu_i^{0,\mathrm{S}}$, cf. §4.1.1) as well as the bulk and real Gibbs energies of solvation ($\Delta_{\mathrm{s}}G_i^0$ and $\Delta_{\mathrm{rs}}G_i^0$, cf. §4.1.2):

$$\begin{aligned}\overline{\mu}_i^{0,\mathrm{S}} = \alpha_i^{0,\mathrm{S}} &\equiv \mu_i^{0,\mathrm{S}} + z_i e\chi^{\mathrm{S}} \\ &= \left(\mu_i^{0,\mathrm{g}} + \Delta_{\mathrm{s}}G_i^0\right) + z_i e\chi^{\mathrm{S}} \\ &= \mu_i^{0,\mathrm{g}} + \left(\Delta_{\mathrm{s}}G_i^0 + z_i e\chi^{\mathrm{S}}\right) \\ &= \mu_i^{0,\mathrm{g}} + \Delta_{\mathrm{rs}}G_i^0\end{aligned}. \tag{4.1}$$

4.1.1. Bulk and real Gibbs energies of solvation

The macroscopic standard Gibbs energy of solvation, defined as the Gibbs energy change at 298 K to transfer 1 mol of solute X from an ideal gas at $p^0 = 1\,\mathrm{bar}$ to a hypothetical ideal solution of $m^0 = 1\,\mathrm{mol/kg}$ $\left(\mathrm{X}_{(\mathrm{g})} \rightarrow \mathrm{X}_{(\mathrm{aq})}\right)$ having the properties of an ideal-dilute solution, is usually identified with the microscopic work of transferring one particle of solute (atom, molecule, ion, etc.) at rest from a vacuum into the solvent $\left(\text{i.e., } \mathrm{X}_{(\mathrm{vac})} \rightarrow \mathrm{X}_{(\mathrm{aq})}\right)$ corrected for the standard state conditions of composition (i.e., $p^0 = 1\,\mathrm{bar}$ and $m^0 = 1\,\mathrm{mol/kg}$). Moreover, in order to equate macroscopic and microscopic Gibbs energies of solvation, one should assume then that the standard entropy and the mean kinetic energy of the solute particles are the same in the gas phase and in solution at a given temperature. This microscopic standard Gibbs energy of solvation is thus in agreement with the macroscopic reference state for solutes, i.e., the hypothetical ideal-dilute 1 mol/kg solution in which only solute–solvent and solvent–solvent interactions are present.

Two kinds of Gibbs energies of solvation relative to the reference vacuum level can be defined. The *bulk Gibbs energy of solvation* $\Delta_{\mathrm{s}}G_i^0$ is the Gibbs energy change to place a molecule at rest from the vacuum into the bulk solvent, whereas the *real Gibbs energy of solvation* $\Delta_{\mathrm{rs}}G_i^0$ is the Gibbs energy change to transfer the molecule at rest from a vacuum into solvent across the surface of the solvent, and it turns out from Eq. (4.1) that

$$\Delta_{\mathrm{rs}}G_i^0 = \Delta_{\mathrm{s}}G_i^0 + z_i e\chi^{\mathrm{S}}. \tag{4.2}$$

Self-consistent reaction field models of the solvent (§2.4.1) are based on calculating the work of placing a molecule from a vacuum into the bulk solvent, assuming $\chi^{\mathrm{S}} = 0$. Furthermore, χ^{S} is usually unknown because it cannot be measured directly from experiment. SCRF methods can yield rather accurate values of $\Delta_{\mathrm{s}}G_i^0$. It is worth noting that the standard state conditions of composition for calculated $\Delta_{\mathrm{s}}G_i^0$ with solvation models often correspond to $c^0 = 1\,\mathrm{mol/L}$ for both ideal gas and solution.

4.1.2. Absolute chemical potentials

The absolute standard chemical potential of species *i* in an ideal-dilute solution in the mole fraction basis $\mu_i^{0,\mathrm{S}}$ [i.e., as defined by Eqs. (3.5) and (3.6)] can under standard state conditions of temperature, pressure, and composition be estimated from the contributions of the chemical potential in the gas phase $\mu_i^{0,\mathrm{g}}$ and its bulk Gibbs energy of solvation $\Delta_{\mathrm{s}}G_i^0$. The quantity $\mu_i^{0,\mathrm{g}}$ is the Gibbs energy change to assemble one mole of molecules from the nuclei and electrons in the vacuum, plus the Gibbs energy change to bring the mole of molecules from the vacuum to an ideal gas under standard temperature and pressure:

$$\mu_i^{0,\mathrm{S}} = \mu_i^{0,\mathrm{g}} + \Delta_{\mathrm{s}}G_i^0 \,. \tag{4.3}$$

The assumptions of ideal gas behaviour and *reference state consisting of nuclei and electrons at infinite separation and at rest* (i.e., a macrostate with $T = 0$, $p = 0$, and $c_i = 0$) enable us to get $\mu_i^{0,\mathrm{g}}$ at given *T* referenced to the level of zero electrostatic potential, in agreement with both the absolute electrochemical scale and the electrostatic convention. The standard chemical potential relative to the reference vacuum level at infinity, i.e., the *absolute standard chemical potential*, can be obtained by applying quantum and statistical mechanics:[64,75]

$$\mu_{T,i}^{0,\mathrm{g}} = h_{T,i}^{0,\mathrm{g}} - TS_{T,i}^{0,\mathrm{g}} \,, \tag{4.4}$$

$$h_{T,i}^{0,\mathrm{g}} = u_{T,i}^{0,\mathrm{g}} + RT \,, \tag{4.5}$$

$$\begin{aligned} u_{T,i}^{0,\mathrm{g}} &= E_{\mathrm{tr}} + E_{\mathrm{vib},i}^{(v=0)} + E_{\mathrm{rot},i} + E_{\mathrm{0K},i} \\ &= E_{\mathrm{tr}} + E_{\mathrm{vib},i}^{(\mathrm{BOT})} + E_{\mathrm{rot},i} + E_{\mathrm{tot},i} \end{aligned} \,, \tag{4.6}$$

where $h_{T,i}^{0,\mathrm{g}}$, $u_{T,i}^{0,\mathrm{g}}$, and $S_{T,i}^{0,\mathrm{g}}$ are the absolute standard enthalpy, absolute standard internal energy and absolute standard entropy, respectively, in the gas phase at temperature *T* and the standard unit pressure. The thermal correction to $u_{T,i}^{0,\mathrm{g}}$ in the classical limit contains: (*i*) a translational contribution $E_{\mathrm{tr}} = \frac{3}{2}RT$; (*ii*) a vibrational contribution $E_{\mathrm{vib},i}^{(v=0)}$ relative to the

lowest molecular vibrational state $v = 0$, $E_{\mathrm{vib},i}^{(\mathrm{BOT})}$ is the vibrational contribution relative to the bottom of the potential energy surface, $E_{\mathrm{vib},i}^{(\mathrm{BOT})} = E_{\mathrm{vib},i}^{(v=0)} + E_{\mathrm{ZP},i}$; and (*iii*) a rotational contribution $E_{\mathrm{rot},i} = \frac{3}{2} RT$ (*RT* for a linear molecule). The internal energy $E_{0\mathrm{K},i}$ at 0K, relative to the state in which nuclei and electrons are infinitely separated and at rest, is the sum of the total electronic energy $E_{\mathrm{tot},i}$ and the zero-point vibrational energy $E_{\mathrm{ZP},i}$ (cf. §2.1.1).

Absolute chemical potentials in solution can in general be defined for transition states, relaxed electronic ground states, relaxed electronically excited states, and vertically excited electronic states. In principle, Eqs. (4.3)–(4.6) are valid for relaxed electronic states (ground and excited states) with spin multiplicity $2S+1$. Adiabatic electronic transitions can therefore be treated within this formalism properly.

A vertical electronically excited state is, however, a nonequilibrium electronic state, since a molecular species in its ground state, e.g., $^1\mathrm{M}$, undergoes a vertical electronic excitation, $^1\mathrm{M}_{(\mathrm{S})} \rightarrow {}^{2S+1}\mathrm{M}^*_{(\mathrm{S})}$, which allows for a relaxation of the electron density keeping the nuclear positions fixed. The absolute standard chemical potential of the excited molecule M^* can be computed from that of the molecule M in its ground electronic state and the energy difference between the lowest vibrational levels of the ground and the electronically excited states ΔE_{0-0}:

$$\mu^{0,\mathrm{S}}(\mathrm{M}^*) = \mu^{0,\mathrm{S}}(\mathrm{M}) + \Delta E_{0-0} \tag{4.7}$$

The energy of the 0–0 transition, ΔE_{0-0}, is calculated using correlated wave function or density functional methods in conjunction with an SCRF model, without allowing for nuclear motions.

Computed first principles standard chemical potentials in the mole fraction basis $\mu_i^{0,\mathrm{S}}$ can thereafter be shifted into the molality and molarity bases by adding constants dependent on the solvent [cf., Eqs. (3.8) and (3.9)] to give $\mu_{m,i}^{0,\mathrm{S}}$ and $\mu_{c,i}^{0,\mathrm{S}}$.

4.2. Absolute electrochemical, real chemical, and chemical potentials of elementary species

Electrons and nuclei are the elementary particles of chemistry. They turn up to behave as structureless and indivisible entities in most of the phenomena regarded to be in the realm of chemistry. Furthermore, electrons and protons, in particular, behave as abnormal ions in the condensed phase, and not surprisingly, their description as individual ions brings about several

fundamental problems of referencing chemical data in both gas and condensed phases, as pointed out in Ref. **S-I**.

Following the prescription outlined in §4.1.2 in combination with Fermi–Dirac and Boltzmann statistics, experimental solvation data, and the absolute potential of the SHE, we have estimated an internally consistent set of absolute standard values for the chemical potentials of the elementary species in aqueous solutions in the molality and molarity bases (cf. Table I) within an accuracy of 0.5 kcal/mol. These values can be directly used in the simultaneous estimate of standard Gibbs energies of electron, proton, and proton-coupled electron transfer reactions.

Table 4.1. Estimated absolute standard values of chemical potentials for elementary species.

Potential	(eV)	(kcal/mol)	Ref.
$\overline{\mu}^{\theta,\mathrm{aq}}(\mathrm{H}^{\bullet}) = \mu^{\theta,\mathrm{aq}}(\mathrm{H}^{\bullet})$	–13.862 [a]	–319.67 [a]	**S-I**
$\overline{\mu}^{\theta,\mathrm{aq}}(\mathrm{H}^{+}) = \alpha^{\theta,\mathrm{aq}}(\mathrm{H}^{+})$	-11.5488 ± 0.02	-266.32 ± 0.5	**S-I**
$\mu^{\theta,\mathrm{aq}}(\mathrm{H}^{+})$	-11.6511 ± 0.009	-268.68 ± 0.2	**S-I**
$\mu_{0\mathrm{K}}^{\mathrm{vac}}(\mathrm{e}^{-})$	0	0	70
$\overline{\mu}^{\theta,\mathrm{SHE}}(\mathrm{e}^{-})$	-4.44 ± 0.02	-102.39 ± 0.5	79
$\alpha^{\theta,\mathrm{aq}}(\mathrm{e}^{-}, p_{\mathrm{H}_2} = 1\,\mathrm{atm}, \mathrm{pH} = 0)$	-4.44 ± 0.02	-102.39 ± 0.5	**S-I**
$\mu^{\theta,\mathrm{SHE}}(\mathrm{e}^{-})$	-4.34 ± 0.02	-100.03 ± 0.5	**S-I**
$\overline{\mu}^{\theta,\mathrm{aq}}(\mathrm{e}^{-}) = \alpha^{\theta,\mathrm{aq}}(\mathrm{e}^{-})$	-1.6079 ± 0.02	-37.08 ± 0.5	**S-I**
$\mu^{\theta,\mathrm{aq}}(\mathrm{e}^{-})$	-1.6638 ± 0.04	-38.37 ± 1	**S-1**

[a] Estimated from $E_{0\mathrm{K}}(\mathrm{H}^{\bullet}) = -0.5$ a.u., i.e., the exact ground state energy of $\mathrm{H}^{\bullet}$, and $\Delta_{\mathrm{s}}G^{\theta}(\mathrm{H}^{\bullet}) = 3.11$ kcal/mol, calculated at the IEF-PCM/B3LYP/cc-pVTZ level using a spherical cavity of radius 1.908 Å (~3.6 a_0).

4.3. First principles ion thermochemistry

A heterogeneous reaction can be conveniently represented in the thermochemical convention as an algebraic equation where $\mathrm{R}_{i(\beta)}$ is the ith reactant or product in phase β depending on whether the stoichiometric coefficient ν_i is negative or positive:

$$\sum_{\beta}\sum_{i}\nu_i \mathrm{R}_{i(\beta)} = 0 \tag{4.8}$$

If we allow $\mathrm{R}_{i(\beta)}$ to be ions, the state of that system can be described by temperature T, pressure p, composition in terms of the number of particles N_i^{β}, and net charges q_i^{β} of component i within the phase β; $q_i^{\beta} = N_i^{\beta}(z_i^{\beta}e)$. The composition of the system varies linearly with the extent of reaction ξ, i.e., $dN_i^{\beta} = \nu_i d\xi$, and thus, the Gibbs energy of the system can be formulated as $G = G(T, p, \{N_i^{\beta}(\xi)\}, \{q_i^{\beta}(\xi)\})$. Under standard conditions of temperature, pressure, and composition in a given concentration scale, it can be easily shown, from the definition of the reaction affinity[18] **A**, that:

$$-\mathbf{A}^0 = \left(\frac{\partial G}{\partial \xi}\right)^0_{T,p} = \sum_{\beta}\sum_{i}\nu_i \overline{\mu}_i^{0,\beta} \equiv \Delta_{\mathrm{r}} G^0_{\overline{\mu}}. \tag{4.9}$$

Accordingly, the Gibbs energy of reaction $\Delta_{\mathrm{r}} G^0_{\overline{\mu}}$ is the "products minus reactants" difference in electrochemical potentials. When the overall electroneutrality condition for a heterogeneous system results from electroneutrality in every phase β, then the Gibbs energies of reaction in the three bases of chemical potentials are equal: $\Delta_{\mathrm{r}} G^0_{\overline{\mu}} = \Delta_{\mathrm{r}} G^0_{\alpha} = \Delta_{\mathrm{r}} G^0_{\mu}$. The equality holds for homogeneous reactions and for heterogeneous reactions in which there is no transfer of charge from one phase to another, i.e., when all the ionic components participating in the reaction remain in the same phase β. Hence, we can simplify the calculation of the Gibbs energies of reaction from electrochemical and real chemical potentials to chemical potentials in a given concentration basis, which can be estimated accurately from first principles (cf. §4.1.2) given that they are independent of the unknown surface potentials χ^{β}.

Particularly relevant to chemistry and biology are the reactions in which electrons and protons occur as independent ions in the stoichiometry, inasmuch as the Gibbs energy changes are directly proportional to experimentally measurable electrode potentials and pK_{a} values. These reactions can be classed into three groups; for the sake of simplicity, we assume that all $|\nu_i| = 1$:

- Proton-coupled electron transfer (PT–ET) reactions

  (i) Reduction by electrons from the reference vacuum level:

$$\mathrm{Ox}_{(\mathrm{aq})} + \mathrm{H}^+_{(\mathrm{aq})} + \mathrm{e}^-_{(\mathrm{vac})} = \mathrm{Red}_{(\mathrm{aq})}$$

$$\begin{aligned} -eE^{\theta}_{\mathrm{abs,r}}(\mathrm{Ox/Red}) &= \Delta_{\mathrm{PT-ET}} G^{\theta} \\ &= \mu^{\theta,\mathrm{aq}}(\mathrm{Red}) - \mu^{\theta,\mathrm{aq}}(\mathrm{Ox}) - \mu^{\theta,\mathrm{aq}}(\mathrm{H}^+) - \mu^{\mathrm{vac}}_{0\mathrm{K}}(\mathrm{e}^-) \end{aligned}. \tag{4.10}$$

(ii) Reduction by electrons from the SHE:

$$\mathrm{Ox}_{(aq)} + \mathrm{H}^{+}_{(aq)} + \mathrm{e}^{-}_{(SHE)} = \mathrm{Red}_{(aq)}$$

$$\begin{aligned} -eE^{\theta}_{\mathrm{SHE}}(\mathrm{Ox/Red}) &= \Delta_{\mathrm{PT-ET}}G^{\theta} \\ &= \mu^{\theta,aq}(\mathrm{Red}) - \mu^{\theta,aq}(\mathrm{Ox}) - \mu^{\theta,aq}(\mathrm{H}^{+}) - \mu^{\theta,\mathrm{SHE}}(\mathrm{e}^{-}) \end{aligned} . \quad (4.11)$$

(iii) Reduction by hydrated electrons:

$$\mathrm{Ox}_{(aq)} + \mathrm{H}^{+}_{(aq)} + \mathrm{e}^{-}_{(aq)} = \mathrm{Red}_{(aq)}$$

$$\Delta_{\mathrm{PT-ET}}G^{\theta} = \mu^{\theta,aq}(\mathrm{Red}) - \mu^{\theta,aq}(\mathrm{Ox}) - \mu^{\theta,aq}(\mathrm{H}^{+}) - \mu^{\theta,aq}(\mathrm{e}^{-}) . \quad (4.12)$$

- Electron transfer (ET) reactions

Reduction by $\mathrm{e}^{-} = \mathrm{e}^{-}_{(vac)}, \mathrm{e}^{-}_{(SHE)}, \mathrm{e}^{-}_{(aq)}$

$$\mathrm{Ox}_{(aq)} + \mathrm{e}^{-} = \mathrm{Red}_{(aq)}$$

Pure ET reactions fall into the same three cases of reduction as PT–ET reactions. The equations for ET can be written down using the above equations for PT–ET, [i.e., Eqs. (4.10)–(4.12)] just neglecting all the terms concerning H^{+}.

- Proton transfer (PT) reactions

Deprotonation equilibrium:

$$\mathrm{HX}_{(aq)} = \mathrm{X}^{-}_{(aq)} + \mathrm{H}^{+}_{(aq)}$$

$$2.303RT \times \mathrm{p}K_{\mathrm{a}} = \Delta_{\mathrm{PT}}G^{0}_{T} = \mu^{0,aq}_{T}(\mathrm{X}^{-}) + \mu^{0,aq}_{T}(\mathrm{H}^{+}) - \mu^{0,aq}_{T}(\mathrm{HX}) \quad (4.13)$$

The Gibbs energy changes of these reactions must be shifted to the chemical dissociation energy scale. Herein, electron and proton transfer from a bound state to a state in which they behave as free particles; therefore, a contribution from the binding energy must be added to the absolute standard electrochemical potential of the free particle. Accordingly, we correct the standard absolute potentials of electrons and protons with the difference in zero-point vibrational energies of the reduced and oxidized forms for ET and PT-ET or of the conjugate base and the acid for PT.

4.4. Concluding summary

We here present the first principles definitions of absolute electrochemical, real chemical, and chemical potentials of any species in a given phase, i.e., atoms, molecules, activated complexes, atomic and molecular ions of any spin multiplicity, as well as electronically excited species. Our definitions are fully compatible with the concentration scales, the electrostatic convention, and the absolute and relative electrochemical scales. A general prescription to estimate the absolute standard values of the three types of chemical potentials is provided, departing from the molecular standpoint of quantum chemistry. The scaling laws of classical and quantum statistical mechanics are thereafter applied to average out the quantum behaviour of these microscopic species within their three types of chemical potentials.

We also emphasize the importance of defining and making clear distinction between bulk and real Gibbs energies of solvation in other to make compatible theoretical approaches and experimental measurements. The selection and use of the appropriate physical quantities can help to sort out part of the current problems in the accuracy of the theoretical predictions.

We apply the Gibbs formalism in its most general form by the straightforward microscopic definition and quantification of the electrochemical potentials. The calculation of Gibbs energies of reaction involving ions in solution can be reduced to the basis of chemical potentials for homogeneous reactions and for heterogeneous reactions in which all the ionic components involved in reaction remain in the same phase.

Last of all, we separate the problem of calculation of Gibbs energies of reactions involving electrons and protons behaving as ions out of the interfacial problem. For electrons and protons within an ideal-dilute aqueous solution and the SHE, we estimate an internally consistent set of electrochemical, real chemical, and chemical potentials in the molality and molarity bases within the chemical accuracy, i.e., ~1 kcal/mol—cf. Table 4.1. The values of these chemical potentials are bound by Born–Haber-type cycles, thereby forming a unique set of numbers for given solvent, temperature, and pressure in the molality and molarity scales of composition. Accordingly, our formalism ensures an internally consistent treatment of the Gibbs energetics of ET, PT, and PT–ET reactions in solution.

5. Radiation Chemistry of DNA

Radiation chemistry deals with the chemical effects of ionizing radiation. Radiation-induced damage is hazardous to living organisms because it brings about chemical modifications in the structure of macromolecular building blocks. Chemical lesions in DNA are highly risky since they can replicate and lead to permanent changes in the DNA sequence. Enzymes usually repair isolated lesions rather efficiently. Damaging agents, however, tend to give clustered lesions that can be more difficult to recognize by the enzymatic repair systems. In particular, when enzymes fail to detect and fix chemical lesions in DNA, mutations arise.[80,81,82,83]

DNA-damaging agents are of both intracellular and environmental origin. To assess the extent of the damage, one should note that the human genome contains $\sim 3 \times 10^9$ base pairs and that the DNA polymerase error rate typically is of the order of one out of 10^9 base pairs replicated. Estimates of the number of damaging events in Mammalian cells are shown in Table 5.1.[82,83]

Table 5.1. Estimated number of damaging events in Mammalian cells.

Agent	No. nucleobases per cell per day
Basal metabolism	10^4
Ionizing radiation (α,β,γ,e^-)	10^3–10^4
Alkylating agents	10^4–10^6
UV light	10^5–10^6
H_2O_2	10^6–10^7

Radiation damage can be conveniently classified into two main types. Primary damage includes the absorption of radiation to yield electronically excited species, radical ions, and their direct fragmentation products. Secondary damage groups the cascade of reactions of the highly reactive species formed in the primary ionization events with the building blocks of carbohydrates, lipids, proteins, and nucleic acids.

In the living cell, water absorbs about 70% of the energy carried in the ionizing radiation, whereas organic compounds absorb the remaining 30%. Radiolysis of water yields hydroxyl radicals, solvated electrons, and

hydrogen atoms, which are the major reactive species initiating the cascade of secondary events:[80]

$$H_2O_{(aq)} + h\nu \rightarrow H_2O^*_{(aq)}$$
$$H_2O^*_{(aq)} \rightarrow H^\bullet_{(aq)} + OH^\bullet_{(aq)}$$
$$H_2O_{(aq)} + h\nu \rightarrow H_2O^{\bullet+}_{(aq)} + e^-_{(aq)}$$
$$H_2O^{\bullet+}_{(aq)} \rightarrow OH^\bullet_{(aq)} + H^+_{(aq)}$$

This chapter outlines the feasibility of a sequence of primary and secondary damaging events to DNA nucleobases. First, we address the formation of electronically excited states on base pairs and their subsequent conversion into radical ions in solution (Ref. **S-II**). Second, we examine the redox reactions undergone by the radical ions and the repairing action of thiols on these chemical lesions (Ref. **S-III**). Radical cations can further hydrolyze to yield OH radical adducts. Nonetheless, the main yield of OH radical adducts is by the direct reaction of the OH radicals produced by water radiolysis with nucleobases. The hydroxylation mechanisms of nucleobases are reviewed in Ref. **R-I**, and the role of stereoelectronic effects ruling the OH radical selectivity for the addition sites is unravelled in Ref. **S-IV**. In particular, 8-hydroxy-purine radical adducts (A8OH$^\bullet$, G8OH$^\bullet$) are prone to undergo further reactions as ring-fragmentation and oxidation to yield 6-amino-5-N-formamidopyridine (FAPyA, FaPyG) and 8-oxopurine (8-oxoA, 8oxoG) derivatives. These compounds are well known to cause transversions in DNA (A:T→T:A, G:C→T:A). The ring-opening mechanism of the 8-hydroxypurine radical can be found in Ref. **R-I** and the oxidation pathways of adenine and guanine are examined in Ref. **S-V**. Insights into the mechanism of DNA strand breakage are presented in Ref. **S-VI**.

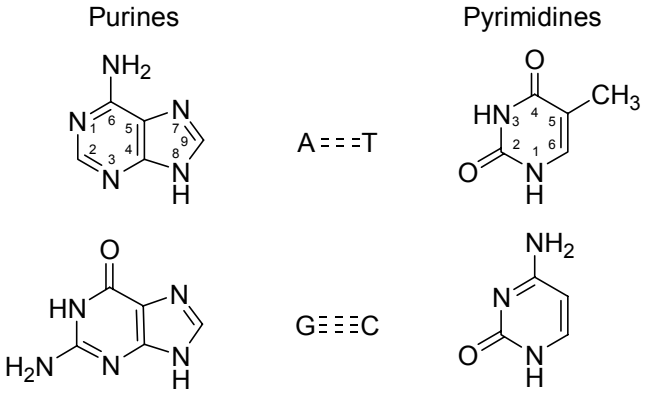


Figure 5.1. Nucleobases.

5.1. Photochemistry of nucleobases

In DNA, electronically excited states are formed by direct absorption of photons and by photosensitization. DNA acts as a cromophore itself by directly absorbing UV-C and UV-B light (UV-C: 190–290 nm, 6.525–4.275 eV; UV-B 290–320 nm, 4.275–3.875 eV). In these wavelength regions, the action spectra for biological responses of cells exposed to UV light overlap with the absorption spectrum of DNA.[84] Although the absorption of DNA in the UV-A region (320–380 nm, 3.875–3.263 eV) is negligible, electronic excitations can arise as a result of photosensitization. Porphyrins and their derivatives, reduced nicotinamide coenzymes (NADH and NADPH), and flavins are known sensitizers that enhance DNA damage *in vitro* if they are present during exposure to UV-A photons.[85]

Photoelectrochemical evidence in condensed phase suggests that formation or decay of electronically excited species can also occur as electron transfer reactions involving solvated electrons. Even though UV light is not generally considered as ionizing radiation, photoionization experiments in aqueous and nonaqueous solutions indicate that the ionization energy of the solute lowers with increasing polarity of the solvent, and hence, the ionization of certain aromatic compounds takes place to yield radical cations and solvated electrons ($B_{(aq)} = B^{\bullet+}_{(aq)} + e^{-}_{(aq)}$).[80,86,87,88] In addition, electrogenerated chemiluminescence experiments indicate that some aromatic radical cations and neutral radicals can form excited states through uptake of solvated electrons,[89] whereas anions formed in redox reactions between organic compounds and solvated electrons point to the fact that electron transfer reactions can proceed by excited states.[90]

It is well known that the UV-light absorption by DNA in solution brings about photoionization.[80] Even though it is difficult to track down detail experimental data on the mechanistic fate of excited states in DNA and its building blocks, it is plausible that the formation and decay of electronically excited nucleobases may proceed by very fast electron transfer reactions as observed for model aromatic compounds.[80,86,87,88,89,90] The pathways in DNA are however unknown because experimental measurements (e.g., EPR, absorption spectroscopy, conductimetry, redox titrations) are usually unable to follow the fate of short-lived species generated in the photoionization and radiolysis experiments.

The question as to whether radical ion and excited electronic states are localized in a DNA molecule is essential to gain insight into the primary ionization events. Two extreme models have been applied to describe the electronic structure of DNA. In the tight-binding approach, π-electrons in a base pair are delocalized into molecular orbitals that spread along the DNA helix. In the discrete molecular orbital approach, the π molecular orbitals do not overlap significantly between adjacent base pairs.

Experiments and calculations on long-range charge migration in DNA show that a hopping mechanism prevails over a bandlike charge transport mechanism, although some features of the latter are found, as the local structural distortions of DNA required in charge migration.[91] Evidence thus indicates that radical ions are rather localized on the nucleobases or the base pairs.

Likewise, the absorption spectra of *E. coli* DNA in the UV-C and UV-B regions[92] suggest that a discrete molecular orbital description could be appropriate for the lowest excitations. The spectra show that, first, the UV absorptions of nucleotides and nucleosides are very similar to those of the nucleobases; second, the absorption of DNA at 260 nm (4.769 eV) to form the lowest excited singlet is about 60% of the weighted average spectrum for the component nucleobases; and third, below 200 nm (higher excitations) nucleobase–nucleobase interactions dominate the features of the absorption.[87] Hence, the formation of the lowest singlet excited states in DNA seems to be also rather localized on the nucleobases.

Since experiments point to a considerable localization of radical ions and the lowest excited electronic states in DNA, we applied our first principles approach to photoelectrochemistry and ion thermochemistry (Ref. **S-I**), in the framework of TD-DFT and PCM, to study the photoinduced ionization of nucleobases B (B = A, G, T, C) and base pairs (AT and GC) in aqueous solution by way of electronically excited species B^* of multiplicity $2S+1$ and solvated electrons $e^-_{(aq)}$. This process follows the general sequence of steps:

$$^1B_{(aq)} \rightarrow {}^{2S+1}B^*_{(aq)} \qquad \Delta_{exc}G^\theta(B) = \mu^{\theta,aq}(B^*) - \mu^{\theta,aq}(B)$$

$$B^*_{(aq)} \rightarrow B^{\bullet+}_{(aq)} + e^-_{(aq)} \qquad \Delta_{ox}G^\theta(B^*) = \mu^{\theta,aq}(B^{\bullet+}) + \mu^{\theta,aq}(e^-) - \mu^{\theta,aq}(B^*)$$

$$B^*_{(aq)} + e^-_{(aq)} \rightarrow B^{\bullet-}_{(aq)} \qquad \Delta_{red}G^\theta(B^*) = \mu^{\theta,aq}(B^{\bullet-}) - \mu^{\theta,aq}(B^*) - \mu^{\theta,aq}(e^-)$$

The standard Gibbs energy diagrams (Figure 5.2) illustrate the processes of excitation and ionization of base pairs in aqueous solution. The first singlet of the AT basepair lies ~1 eV higher in energy than that of the GC basepair. However, the first triplet states and the radical anions lie at about the same standard Gibbs energies relative to the ground states. The formation of radical cations from the first triplet state is ~0.5 eV more endergonic for AT than for GC. The diagrams show that the formation of radical cations from the first excited states is endergonic, whereas the uptake of solvated electrons by the excited states is very exergonic. It should be noted that in the traditional theoretical treatment of ion formation, the energy level of electrons in aqueous solution is unknown, and hence, the radical ions can be incorrectly located on a standard Gibbs energy diagram of such processes in the condensed phase (Ref. **S-II**).

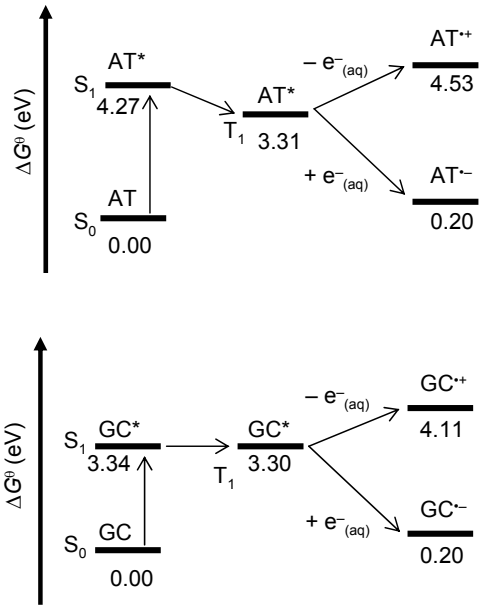


Figure 5.2. IEF-PCM/(TD-)B3LYP/6-311+G(2df,p)//B3LYP/6-31(+)G(d,p) UV excitation and ionization of the AT and GC base pairs in aqueous solution.

5.2. Redox reactions of radical ions

Once the radical ions are formed, they undergo a cascade of redox reactions depending on the effective proton concentration in the medium. The radical cations readily deprotonate to yield more stable neutral radicals (Figures 5.3 and 5.4). The nucleobases are restored to their native forms through chemical or enzymatic repair. A simple mechanism of chemical repair of the acidic radical cations $B^{\bullet+}$ consists of regenerating the nucleobase B by reaction with a reducing agent, e.g., reduced glutathione (about 10^{-2} mol/L in living cells) or other thiols present in the cells. Thiols function by donating H atoms to the deprotonated radical cation $B(-H)^{\bullet}$, thereby repairing the lesion:

$$B_{(aq)} \rightarrow B^{\bullet+}_{(aq)} + e^{-}_{(aq)}$$

$$B^{\bullet+}_{(aq)} \rightarrow B(-H)^{\bullet}_{(aq)} + H^{+}_{(aq)}$$

$$B(-H)^{\bullet}_{(aq)} + RSH_{(aq)} \rightarrow B_{(aq)} + RS^{\bullet}_{(aq)}$$

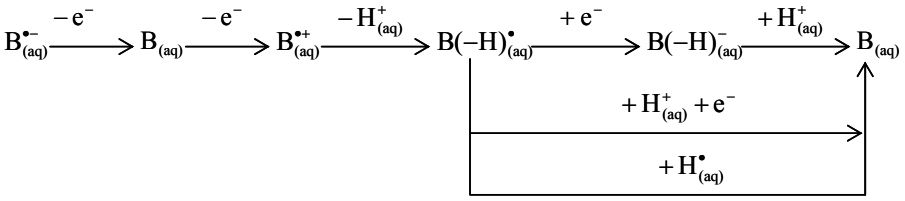


Figure 5.3. Sequence of primary ionization, decay of radical cations, and regeneration of nucleobases.

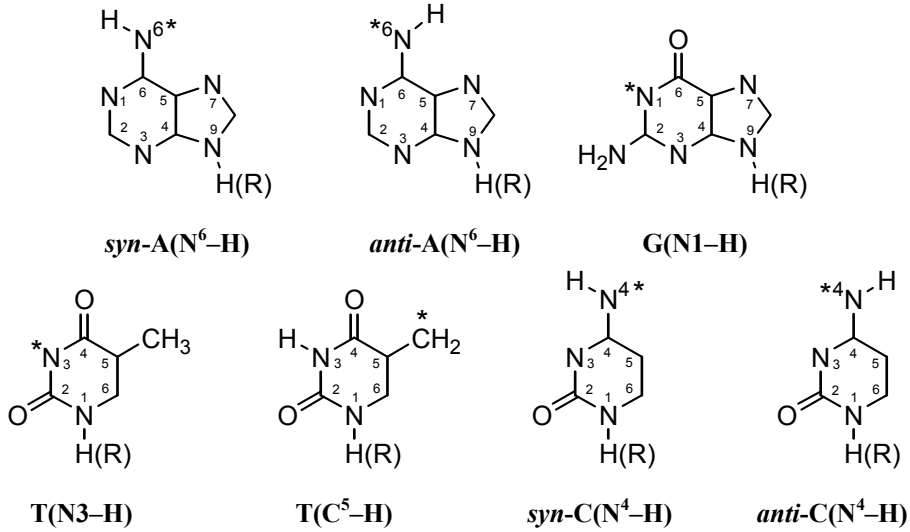


Figure 5.4. Dehydrogenated nucleobases B(–H) radicals and anions. An asterisk indicates the position of the missing hydrogen.

From experimental and computed Gibbs energies of reaction, we are able to gain insight into three relevant issues for the redox events in nucleobases: the nature of the ET process, the influence of PT on ET and *vice versa*. Calculated standard Gibbs energy diagrams of primary ionization, formation of dehydrogenated nucleobase radicals and chemical repair by thiols allow us to compare with experimentally derived data (Ref. **S-III**). Unfortunately, no experimental thermochemical data is available for assessing computed Gibbs energetics in the redox processes involving PT–ET or species like $H^{\bullet}_{(aq)}$ and $e^{-}_{(aq)}$, albeit this information would help unravelling some mechanistic details of the reaction kinetics.

In agreement with the experimental evidence, our calculations suggest that, in aqueous solution, the ease of reduction of the nucleobases B to yield the radical anions decreases in the order T > C > A > G whereas the ease of reduction of the radical cations $B^{\bullet+}$ to yield the nuclebases B is C > T > A > G. Radical cations $B^{\bullet+}$ and their deprotonated forms $B(-H)^{\bullet}$ can gain electrons via ET or PT–ET, respectively, to regenerate the nucleobases B or their deprotonated forms $B(-H)^{-}$, depending on the effective concentration

of hydrogen ions. Both pathways leading to formation of B are essentially equally favoured. Reduction of $B(-H)^{\bullet}$ by PT–ET to yield B is however more favorable than the reduction by ET to give $B(-H)^{-}$.

Our correct reference level for the electronic energy reveals that the standard Gibbs energies of reduction of $B(-H)^{\bullet}$ by hydrogen atoms are halfway in between those for pure ET and PT–ET reductions by hydrated electrons. At neutral or acidic pH the PT–ET reduction by hydrated electrons should dominate over the reduction by hydrogen atoms, whereas both reactions will compete around pH 9.

In the case that the reduced form is an acid, we have shown that the higher the acidity of the reduced form, the more favourable is the proton-coupled reduction of the corresponding oxidized form. Accordingly, the deprotonated forms $B(-H)^{\bullet}$ of the nucleobase radical cations $B^{\bullet+}$ are better oxidants when reducing to the neutral nucleobase B than when reducing to the anions $B(-H)^{-}$, since the nucleobases B are very weak acids with pK_a's larger than 9. Moreover, the oxidizing powers of the nucleobase radical cations $B^{\bullet+}$ by one-electron reduction and the deprotonated forms $B(-H)^{\bullet}$ by one-electron proton-coupled reduction are very similar.

The deprotonated forms $B(-H)^{\bullet}$ of the nucleobase radical cations are lesions that can be chemically repaired by thiols. We have shown that a PT–ET pathway should be favoured for A and C at any pH whereas for G and T, a PT–ET pathway could be preferred at low and neutral pH but in the pH range from 9 to 11, the ET pathway should dominate.

5.3. Hydroxylation mechanism

In cells, OH radicals are formed as by-products of respiration, through exposure to ionizing radiation or other agents that generate free radicals. About half of the DNA damage caused by OH radicals occurs on nucleobases, where these add to the double bonds of purine bases to yield C4, C5, and C8 radical adducts in adenine (A4OH$^{\bullet}$, A5OH$^{\bullet}$, and A8OH$^{\bullet}$) and guanine (G4OH$^{\bullet}$, G5OH$^{\bullet}$, and G8OH$^{\bullet}$).[80,82]

Assuming quasiequilibrium, the mechanism of hydroxylation starts with the diffusion of the OH radical to the nucleobase to form a π-complex $B:OH^{\bullet}$, in which the OH radical lies parallel to the purine ring. Once this reactant complex is formed, the OH radical adds to a specific position of purine:

$$B_{(aq)} + OH^{\bullet}_{(aq)} = B:OH^{\bullet}_{(aq)} \qquad K^{\theta}_{ass} = \frac{c_e(B:OH^{\bullet})}{c_e(B)c_e(OH^{\bullet})}$$

$$B:OH^{\bullet}_{(aq)} = BOH^{\bullet\neq}_{(aq)} \qquad k,\ K^{\theta\neq} = \frac{c_e(BOH^{\bullet\neq})}{c_e(B:OH^{\bullet})}$$

$$BOH^{\bullet\neq}_{(aq)} = BOH^{\bullet}_{(aq)} \qquad k^{\neq}$$

The observed rate constant is thus $k_{obs} = kK^{\theta}_{ass}$; therefore, the activation energy is the sum of the barrier $\Delta^{\neq}E^{\theta}$ and the association energy $\Delta_{ass}E^{\theta}$ to form the π-complex $B:OH^{\bullet}$, i.e., $E_a = \Delta^{\neq}E^{\theta} + \Delta_{ass}E^{\theta}$. Figure 5.5 shows that the activation energies to hydroxylation of the 4- and 8-positions of purines are lower than 9 kcal/mol. The attempts to find the transition state at the 5-position always lead to the structure of transition state at the 4-position. Accordingly, the SCI-PCM/B3LYP energetics indicate that hydroxylation of nucleobases should be diffusion-controlled (Ref. **R-I**). The energetics partly explain that A8OH$^{\bullet}$ and G8OH$^{\bullet}$ are major adducts over A4OH$^{\bullet}$ and G4OH$^{\bullet}$. Nonetheless, the energetics alone do not allow to find a general rule of thumb to account for the site specificity of hydroxylation.

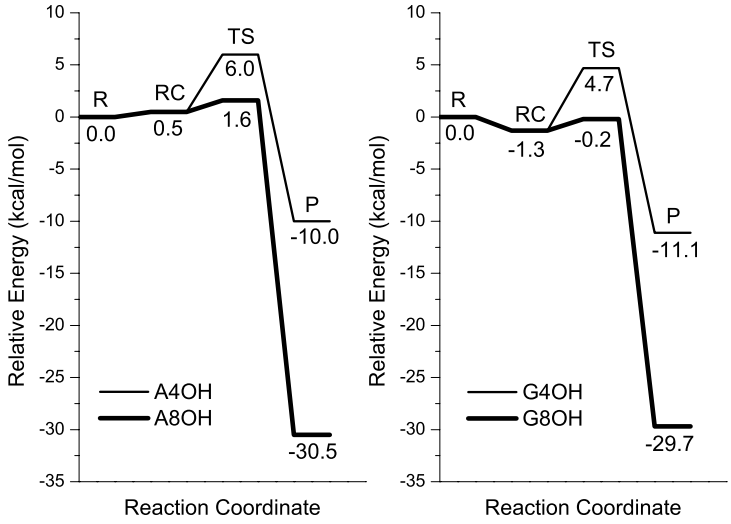
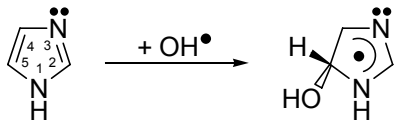


Figure 5.5. SCI-PCM/B3LYP/6-311G(2df,p)//MP2/6-31G(d) hydroxylation profiles for adenine and guanine. **R**, free reactants B + OH$^{\bullet}$; **RC**, π reactant complex B:OH$^{\bullet}$; **TS**, transition state BOH$^{\bullet\neq}$; **P**, product BOH$^{\bullet}$.

5.3.1. Role of the generalized anomeric effect in hydroxylation

The site specificity of hydroxylation was investigated using imidazole as a model compound (Ref. **S-IV**). Experimental evidence demonstrates that the OH radical reacts with pyrroles and imidazoles by addition at a carbon adjacent to the nitrogen.[93,94] In particular, the reaction of imidazole (I) with OH radicals could yield three possible adducts, i.e., the 2-, 4-, and 5-hydroxyimidazolyl radicals (I2OH$^{\bullet}$, I4OH$^{\bullet}$, and I5OH$^{\bullet}$). However, both experimental and electron spin resonance (ESR, EPR) measurements as well as theoretical studies have shown that, in neutral and alkaline (pH = 9–10) aqueous solution, the OH radical specifically adds to the 5-position of

imidazole and, in acidic media (pH = 2), to the same position in the side chain of histidine leading to the 5-adduct.[93,94,95] Hydroxylation is site-specific even though I2OH$^{\bullet}$ is energetically more favoured by ~4 kcal/mol with respect to I5OH$^{\bullet}$.



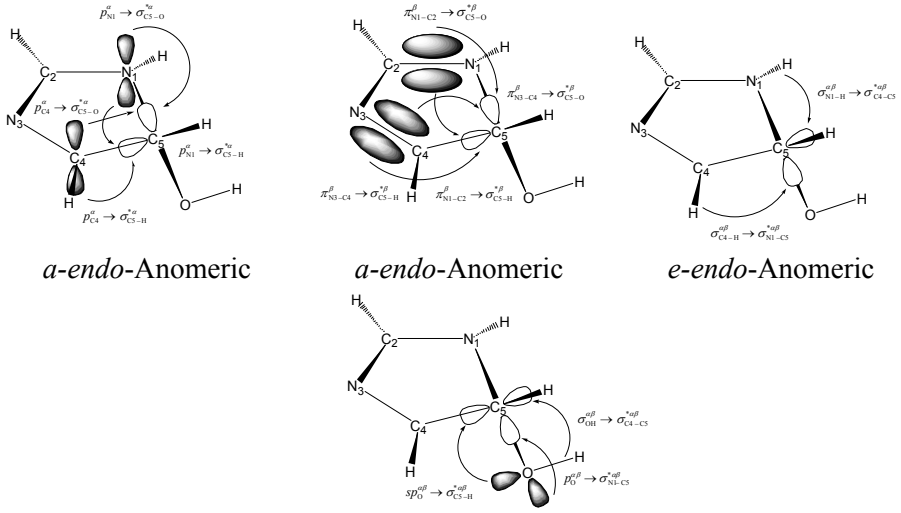
$$I_{(aq)} + OH^{\bullet}_{(aq)} = I5OH^{\bullet}_{(aq)}$$

Two stereoelectronic effects are present in the hydroxyimidazolyl radicals: conjugation and generalized anomeric effect. For instance, in I5OH$^{\bullet}$, conjugation is associated with the N1–C2–N3–C4 fragment while the generalized anomeric effect is related to the C2–N1–C5–OH moiety. Accordingly, the OH$^{\bullet}$ addition to the 2-, 4- or 5-positions in imidazole should be ruled by the generalized anomeric effect and the influence of the conjugated fragment thereon. The generalized anomeric effect is the preference for the synclinal (*sc*) or *gauche* arrangement over the antiperiplanar (*ap*) or *trans* arrangement in cyclic or aliphatic compounds containing the *R*–*X*–*T*–*Y* moiety.[96,97,98,99,100,101] *R* can be H or C, *X* can be a nonmetallic element or a metalloid which binds leaving lone electron pairs (*X* = N, O, S, Se, or Te), *T* is a tetrahedral center of intermediate electronegativity (*T* = C, Si, P), and *Y* denotes a more electronegative element than *T* (*Y* = N, O, S, Se, Te, F, Cl, or Br). Moreover, if *Y* is also bonded to a substituent *R*' (*R*' = H, C), the anomeric effect in the fragment *R*–*X*–*T*–*Y*–*R*' turns to be the contribution of two components: the endo-anomeric effect (*endo*-An.), connected with the *R*–*X*–*T*–*Y* fragment and the exo-anomeric effect (*exo*-An.) related to the *X*–*T*–*Y*–*R*' fragment.

The generalized anomeric effect has been rationalized in terms of electrostatic[96b,97,98,99,102] and hyperconjugative[98,100,102,103,104,105,106] interactions. According to the hyperconjugative hypothesis, the anomeric effect could be qualitatively explained as an $n_X \rightarrow \sigma^*_{T-Y}$ overlap, which is the strongest interaction among other stabilizing hyperconjugative ones. In the presence of a continuum solvent model, the $n_X \rightarrow \sigma^*_{T-Y}$ interaction should weaken with respect to that in the gas phase so as to reproduce the experimental evidence that the anomeric effect is reduced as the solvent polarity increases.[102,107]

The anion-like and cation-like Lewis structures were chosen to tailor the α- and β-electron densities of I2OH$^{\bullet}$, I4OH$^{\bullet}$, and I5OH$^{\bullet}$, respectively, as well as for their transition states. These Lewis structures consist of more than 97% of the α- and β-electron densities. As depicted in Figure 5.6, the energy stabilization due to the generalized anomeric effect can be split into three components for a cyclic compound: the *exo*-anomeric effect and the two

components of the *endo*-anomeric effect, i.e., the axial (*a-endo*-An.) and equatorial (*e-endo*-An.) contributions.



$$\text{Generalized Anomeric Effect} = \textit{endo}\text{ - An.} + \textit{exo}\text{ - An.}$$
$$= (\textit{a}\text{ - }\textit{endo}\text{ - An.} + \textit{e}\text{ - }\textit{endo}\text{ - An.}) + \textit{exo}\text{ - An.}$$

Figure 5.6. Components of the generalized anomeric effect in I5OH$^{\bullet}$.

For the three possible hydroxyimidazolyl radical adducts, the decreasing order of stability I2OH$^{\bullet}$ > I5OH$^{\bullet}$ > I4OH$^{\bullet}$ can be accounted for in terms of the generalized anomeric effect, since the generalized anomeric stabilization linearly decreases in the same order (Figure 5.7). Nonetheless, the adduct stability cannot explain the experimentally observed specificity of 5-hydroxylation of imidazole and histidine. In contrast, the calculated barriers increase in the order I5OH$^{\bullet}$ < I2OH$^{\bullet}$ < I4OH$^{\bullet}$, indicating that I5OH$^{\bullet}$ is kinetically favoured. The kinetic preference can be rationalized from the mechanistic standpoint as the contribution of two factors. First, the transition state at the 5-position is the most favoured by the *exo*-anomeric interaction among the non-bonding spin orbitals of OH$^{\bullet}$ and the spin antibonds at C5 in imidazole. The *exo*-anomeric interaction decreases in the order I5OH$^{\bullet}$ > I2OH$^{\bullet}$ > I4OH$^{\bullet}$ for the transition states. Second, the interaction of the attacking sp^n_{O} non-bonding spin orbital of OH$^{\bullet}$ with the π-cloud is the lowest at the 5-position; hence, the interaction with the Rydberg orbitals at carbon to form the $\sigma_{\mathrm{C-O}}$ spin orbital would be less hindered.

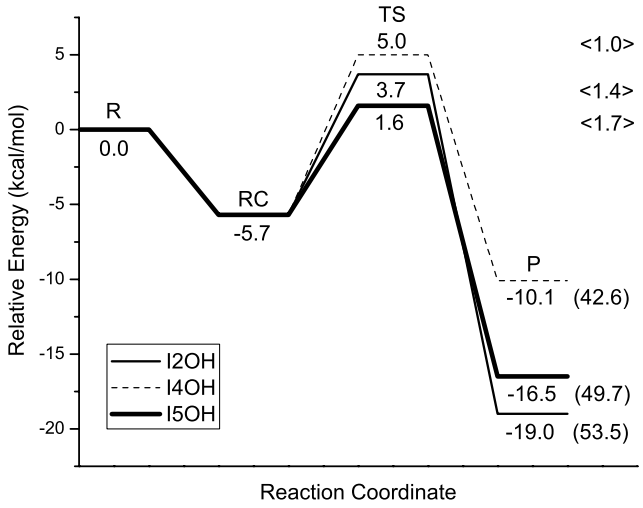


Figure 5.7. Hydroxylation profile of imidazole calculated at the SCI-PCM/PMP2/6-311G(2df,p)//MP2/6-31G(d,p) level. ZPE-corrected total energies relative to the isolated reactants **R**. **RC** stands for reactant complex, **TS** for transition state, and **P** for product, i.e., the OH radical adducts. The stabilization energies of the *exo*-anomeric hyperconjugative interactions (in kcal/mol) in the **TS** are in angular brackets. The stabilization energies of the generalized anomeric hyperconjugative interactions (in kcal/mol) in **P** are in parenthesis.

5.4. Oxidation of 8-hydroxy-purine radical adducts

Under oxidative conditions, the 8-hydroxyl radical adducts (A8OH$^{\bullet}$ and G8OH$^{\bullet}$) will react further to yield the 8-hydroxypurine derivatives (8-OHA and 8-OHG) and their 8-oxo-7,8-dehydropurine tautomers (8-oxoA and 8-oxoG).[80,82,108] Addition of water to G$^{\bullet+}$ is an alternative and feasible pathway to yield 8-oxoG. The reaction seems to be important in DNA, but experimental evidence suggests that it does not occur in solutions of nucleotides.[109]

8-oxoA and 8-oxoG are the major stable purine products produced by oxidation and radiolysis.[110,111,112,113] The 8-oxopurine derivatives are highly mutagenic lesions, as mispairing of 8-hydroxyguanine and 8-oxoguanine with adenine brings about G:C to T:A transversions.[83,114,115]

Two pathways for oxidation have been proposed (Figure 5.8). The first pathway (Figure 5.8, PT branch) is initiated by deprotonation of the 8-hydroxypurine radical at C8 followed by subsequent electron loss and tautomerization,[116] whereas the other (Figure 5.8, PT–ET branch) is initiated by a proton-coupled one-electron oxidation of the 8-hydroxypurine radical

(A8OH$^{\bullet}$ and G8OH$^{\bullet}$) to yield the 8-hydroxypurine derivative (8-OHA and 8-OHG) followed by tautomerizarion.[109,117,118]

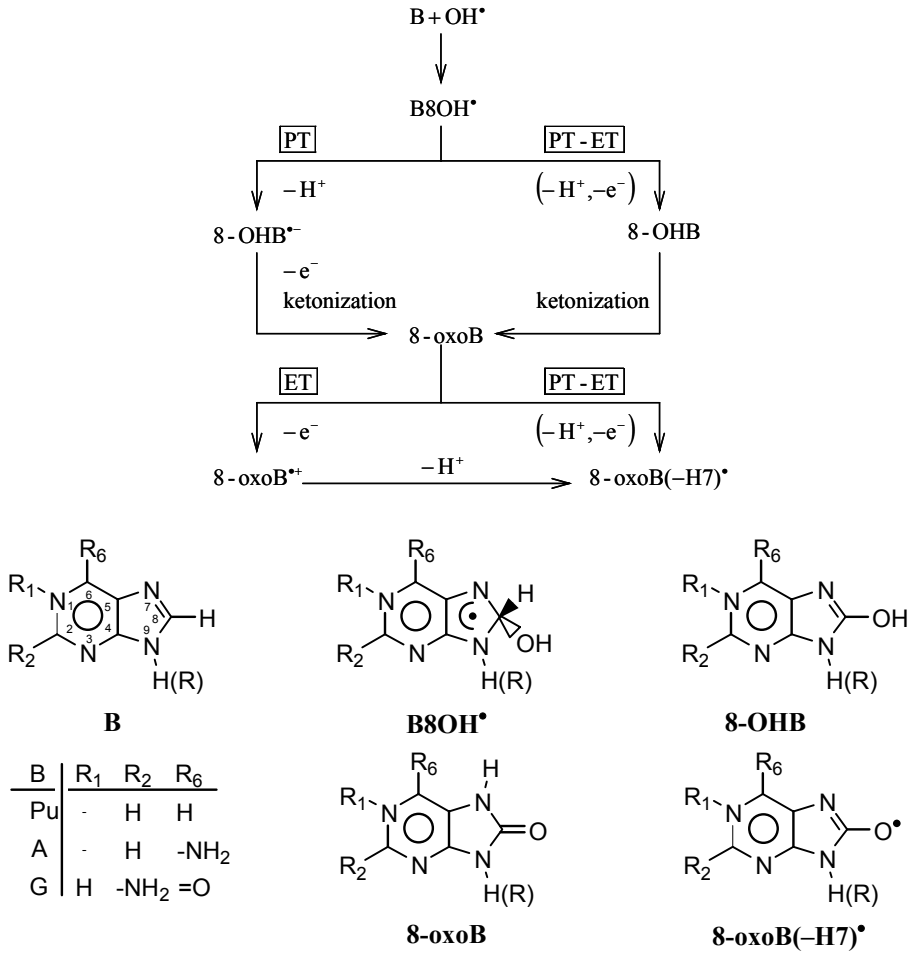


Figure 5.8. Oxidation of purine nucleobases.

The 8-oxopurine derivatives (8-oxoA and 8-oxoG) can be subjected to further oxidation. 8-OHG and 8-oxoG may undergo one-electron oxidation, tautomerization, and deprotonation, depending on the reaction conditions.[119]

The PT-initiated oxidation pathway of the 8-hydroxypurine radicals (A8OH$^{\bullet}$ and G8OH$^{\bullet}$) followed by electron loss seems to be very unlikely in aqueous solution because of the very low acidity of the hydrogen at the 8-position. Instead, the PT–ET-initiated oxidation pathway of A8OH$^{\bullet}$ and G8OH$^{\bullet}$ seems to be the favoured route in aqueous solution.

In the case of the PT–ET-initiated oxidation, G8OH$^{\bullet}$ is only slightly more sensitive than A8OH$^{\bullet}$ to proton-coupled one-electron oxidation followed by tautomerization to yield 8-oxoG and 8-oxoA. Likewise, 8-oxoG and 8-oxoA

behave similarly towards one-electron oxidation to yield the radical cations 8-oxoG$^{\bullet+}$ and 8-oxoA$^{\bullet+}$. The difference in reactivity does not appear to be related to the 8-hydroxylation step either, since the barriers of adenine and guanine are close to the diffusion-controlled limit, and both reactions are exergonic by about 15 kcal/mol.

However, the proton-coupled one-electron oxidation of 8-oxoG is preferred over that of 8-oxoA by about 5 kcal/mol. Hence, a higher yield of 8-oxoG(–H7)$^{\bullet}$ should be expected over 8-oxoA(–H7)$^{\bullet}$ in DNA.

Our calculations furthermore show that the ease of ionization of nucleobases in their native and oxidized forms to yield B$^{\bullet+}$ (via ET) or B(–H)$^{\bullet}$ (via PT–ET) is 8-oxoG > G > 8-oxoA > A > C > T in aqueous solution. Accordingly, 8-oxoG can reduce any nucleobase radical cation B$^{\bullet+}$ (via ET) or their deprotonation products B(–H)$^{\bullet}$ (via PT–ET) to the nucleobase native forms. The smallest driving force thus corresponds to the repair of lesions in G (G$^{\bullet+}$ or G(–H1)$^{\bullet}$):

ET: $\mathrm{G}^{\bullet+}_{(aq)} + 8\text{-}\mathrm{oxoG}_{(aq)} = \mathrm{G}_{(aq)} + 8\text{-}\mathrm{oxoG}^{\bullet+}_{(aq)}$

$\Delta_r G^{\theta} = -4.4$ kcal/mol (expl: −12.7 kcal/mol)

PT–ET: $\mathrm{G}(-\mathrm{H1})^{\bullet}_{(aq)} + 8\text{-}\mathrm{oxoG}_{(aq)} = \mathrm{G}_{(aq)} + 8\text{-}\mathrm{oxoG}(-\mathrm{H7})^{\bullet}_{(aq)}$

$\Delta_r G^{\theta} = -9.3$ kcal/mol (expl: −12.7 kcal/mol)

Since the driving force for the PT–ET repair is about twice as strong as that for the ET repair, the presence of 8-oxoG could also contribute to shift the equilibrium at neutral pH towards the deprotonation of the remaining G$^{\bullet+}$, which behaves as a weak acid.

Likewise, 8-oxoG can reduce both 8-oxoA$^{\bullet+}$ (via ET) and 8-oxoA(–H7)$^{\bullet}$ (via PT–ET). Hence, albeit the conditions for damages to occur on A and G are essentially identical, 8-oxoG is more prone to reduce/repair damages on other sites (including G$^{\bullet+}$ and 8-oxoA) and to function as a sink for oxidative damage occurring in the base stack. Thus, the detection of 8-oxo-7,8-dehydropurine derived radicals would yield higher amounts of 8-oxoG than 8-oxoA. This provides an explanation as to why the observed ratio 8-oxoA:8-oxoG is 1:3 when both naked DNA and chromatin are irradiated, and could be a likely way in which 8-oxoG acts as a trap of radical cations and neutral radicals in DNA.

5.5. Radiation induced damage in serine phosphate

A DNA strand break involves the cleavage of the phosphate–ester bond. It can be single (on one of the strands) or double (on both strands). From a chemical standpoint, two types of strand breaks are identified: (*i*) the direct strand break resulting from straight cleavage of the phosphate bond, and (*ii*) the strand break following a chemical modification of the nucleobase[108] or deoxyribose.[120]

Apart from the cleavage of the phosphate–ester bond, no alterations of the phosphate moiety as such have been reported in experimental studies on full DNA samples or DNA oligomers.[121,122,123] Phosphate centered radicals in DNA ($ROPO_2^{\bullet -}$) are in yields of less than 0.2%.[124] They are formed by local backbone uptake of non-solvated low-energy electrons. Irradiation of DNA primarily leads to ionization of the aromatic bases,[82] and there is experimental evidence that primary ionic sites in DNA can trigger strand breaks.[125] However, the mechanisms by which the base damages are transferred to the sugar–phosphate region and how subsequent breakages of the sugar–phosphate bonds occur are still largely unknown.

Small model systems have been used to get insight into the strand breaks, because of both experimental difficulties in performing and interpreting the results of work on full DNA, and computational limitations on the size of a molecule that can be investigated at high accuracy. In radiation chemistry, experiments on L-*O*-serine phosphate (Ser-P) have enabled some pathway proposals.[126] Ser-P displays key features also found in the DNA strand: the phosphate group, the phosphate–ester bond, and the electron-withdrawing carboxyl group acting as the electron-withdrawal nucleobases in DNA.

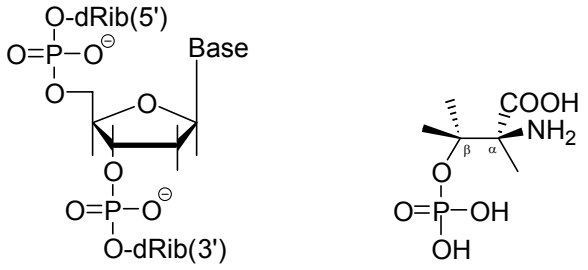


Figure 5.9. Structures of a DNA strand and L-*O*-serine phosphate.

The experimentally identified radicals[126] are formed by ionization of Ser-P followed by deamination, decarboxylation with subsequent radical exchange, or dephosphorylation (radical **I**, **II**, or **III** in Ref. **S-VI**, respectively; cf. Figure 5.10).

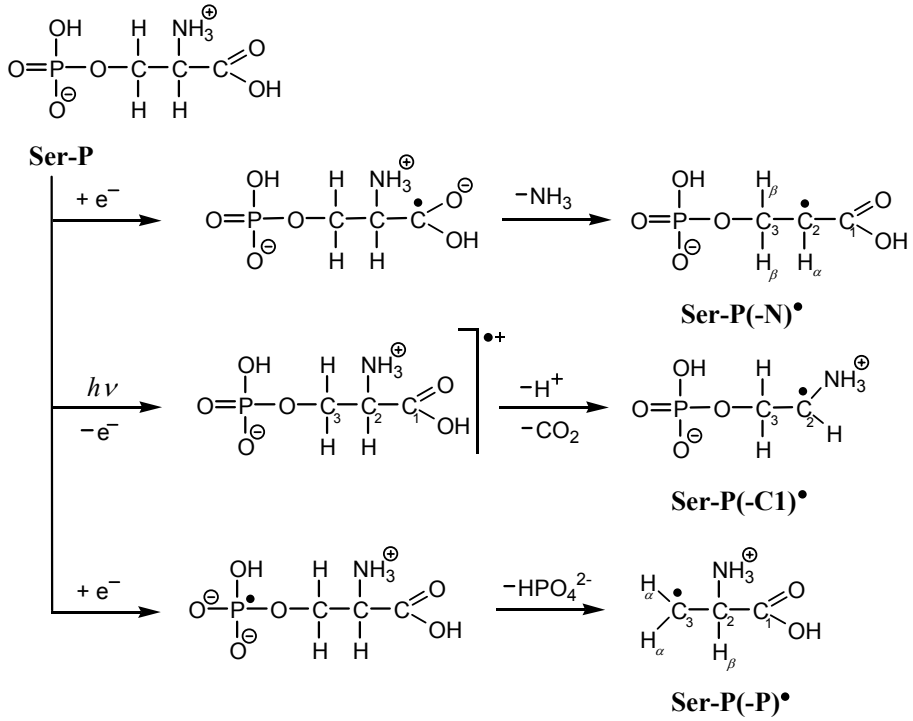
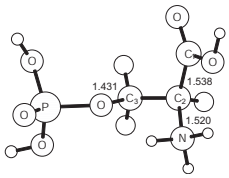
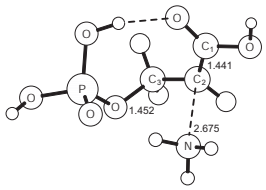


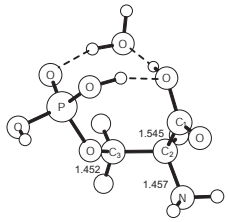
Figure 5.10. Proposed experimental pathways yielding deaminated, decarboxylated, and dephosphated radicals of Ser-P.

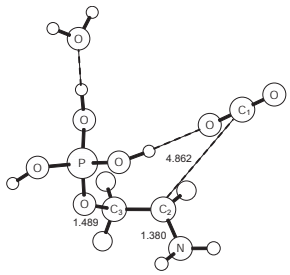
Calculations on our models (Ref. **S-VI**) show that after electron capture, the deamination of Ser-P is exothemic and barrierless to yield Ser-P(–N)$^{\bullet}$ (Figures 5.11a). The energetics supports that Ser-P(–N)$^{\bullet}$ is the major product in irradiated Ser-P. Likewise, after electron loss, the decarboxylation of Ser-P follows the same trend to give Ser-P(–C1)$^{\bullet}$ (Figures 5.11b). The dissociative electron capture to yield Ser-P(–P)$^{\bullet}$ competes with the deamination reaction. Electron uptake followed by dephosphorylation (Figure 5.12) involves an initial accumulation of the unpaired spin on the phosphorous atom, and the formation of a trigonal bipyramidal phosphoranyl radical intermediate. The system then passes over a transition barrier for the dephosphorylation reaction and a back-transfer of the unpaired spin to the adjacent carbon atom.





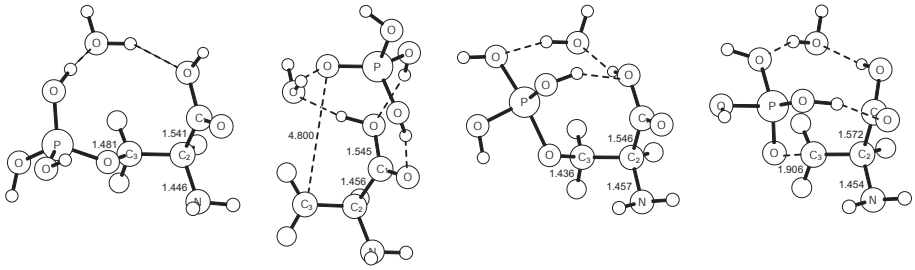
(a) Ser-$PH^{\bullet+}$ + e^- → Ser-$P(-N)^{\bullet}$ + NH_3





(b) Ser-P + H_2O – e^- → Ser-$P(-C1)H^{\bullet+}$ + CO_2 + H_2O

Figure 5.11. B3LYP/6-31G(d) optimized structures of Ser-P complexes prior ionization and after (**a**) Deaminatiom (**b**) Decarboxylation (HF/6-31G(d) level optimized). Distances in Å.



Ser-$PH^{\bullet+}$ + H_2O + e^- → Ser-$PH^{\bullet}$ + H_2O

Ser-$PH^{\bullet}$ + H_2O → [Ser---$PH^{\bullet}$---H_2O]$^{\neq}$

[Ser---$PH^{\bullet}$---H_2O]$^{\neq}$ → Ser-$P(-P)^{\bullet}$ + H_3PO_4 + H_2O

RC	**I**	**TS**	**PC**
(a)	**(b)**	**(c)**	**(d)**

Figure 5.12. B3LYP/6-31G(d) optimized structure of Ser-P and H_3O^+ (distances in Å). (**a**) Prior to electron uptake. (**b**) Phosphoranyl radical intermediate formed after electron uptake. (**c**) Transition state for dephosphorylation. (**d**) Product complex.

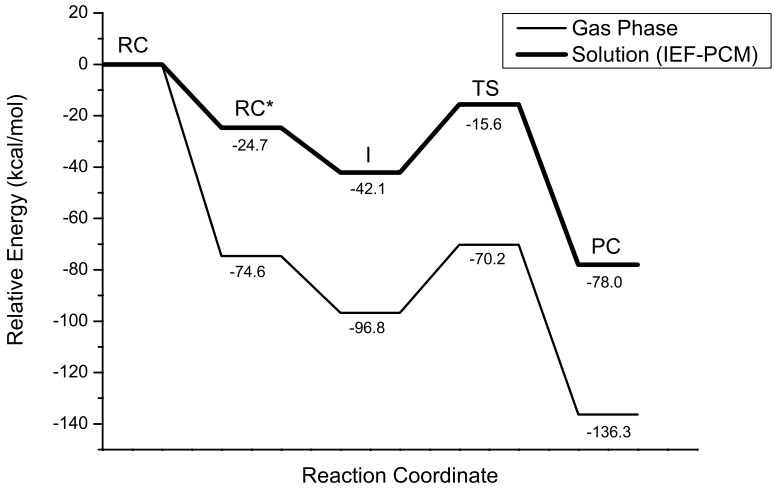


Figure 5.13. B3LYP/6-311+G(2df,p)//B3LYP/6-31G(d) profile for dephosphorylation of Ser-P. **RC**, reactant complex (Fig. 5.11a); **RC**$^{\bullet}$, reactant complex after electron uptake (single point); **I**, phosphoranyl intermediate (Fig. 5.11b), **TS**, transition state (Fig. 4.11c); **PC**, product complex (Fig. 5.11d).

The formation of the phosphoranyl intermediate is exothermic and barrierless, but the subsequent transition state leading to dephosphorylation has a barrier of 26.5 kcal/mol (Figure 5.13). This barrier is likely to be overestimated and is probably sensitive to interactions with the environment. It is furthermore evident from the calculations that the environment strongly influences to localize the unpaired spin on the phosphorous atom, thus determining whether Ser-P after electron capture will proceed towards Ser-P(–N)$^{\bullet}$ or Ser-P(–P)$^{\bullet}$. Formation of Ser-P(–P)$^{\bullet}$ (corresponding to direct strand break) does not seem to depend on stereoelectronic effects other than the formation of the trigonal bipyramidal phosphoranyl anion intermediate facilitated by the environment.

Since Ser-P is only a small model system, these reaction mechanisms are not directly applicable to the situation in irradiated DNA. However, the present work does confirm the hypothesized presence of a phosphoranyl intermediate leading to the cleavage of the phosphate–ester bond in Ser-P. It can be speculated that a similar reaction mechanism could be responsible for the cleavage of the phosphate–ester bond in DNA in order to give direct strand breaks. Clearly, however, in these systems, a large number of competing mechanisms are at play, involving the nucleobases acting as electron sinks, as primary ionization targets, or as targets of the radiolytic products of water.

5.6. Concluding summary

In this chapter, we have addressed the problems of the photophysical and photochemical properties of nucleobases and the subsequent formation of radical anions and cations in solution. These primary radicals react by pH dependent electron transfer reactions with the hydrated electrons, protons, and hydrogen atoms present in solution. As we have determined an internally consistent set of values for the chemical potentials of these species in aqueous solution, their energy levels are well-defined and known in the condensed phase. As a result, we can make computational radiation chemistry as well as computational electrochemistry.

Applying the first-principles electrochemical scheme, we have been able to describe qualitatively and simultaneously some relevant aspects of ET, PT, and PT–ET reactions in solution involving solvated protons, hydrated electrons, electrons in the SHE, and solvated hydrogen atoms. First principles Gibbs energies of redox processes in solution are provided and lie on the same scale as those derived from experimental measurements (standard electrode potentials and pK_a values).

We have connected the energetics of a cascade of reactions on the same energy scale: the formation of excited states, radical ions (including their secondary redox reactions and the chemical repair by thiols), OH radical adducts, as well as the ring-fragmentation and oxidation of 8-hydroxy-purine radical adducts. All these processes are feasible in the order we have integrated them, and they may lead to mutations if the damage is not properly repaired. The combination of DFT and WFT calculations enables us to link and organize a variety of mechanistic proposals based on a large number of specific radiolysis experiments, difficult otherwise to resolve when integrated in a complex set of chain reactions in the DNA double helix.

Our model of direct strand breakage of the phosphate–ester bond emphasizes the role of the environment in the localization of the radical character on the electron-withdrawing group or on the phosphate-ester bond. Our calculations suggest that the direct strand breakage can go by a trigonal bipyramidal radical intermediate.

6. Photochemistry of Skin-Sensitizing Psoralens

Photochemistry deals with the chemical effects of infrared, visible, and ultraviolet electromagnetic radiation, as distinguished from radiation chemistry, which is associated with ionizing radiation. *Photosensitization* is the process by which a photochemical or photophysical alteration occurs in one molecular entity as a result of initial absorption of radiation by another molecular entity referred to as a *photosensitizer*. In mechanistic photochemistry the term is limited to cases where the photosensitizer is not consumed in the reaction.[127] Dye-sensitized photoreactions are widely applied to material and biomedical sciences. In photochemotherapy, for instance, oral or topical photosensitizing agents are applied with subsequent exposure to light so as to trigger cytostatic or photodynamic action.

Cytostatic therapy makes use of drugs to inhibit cell proliferation by hindering DNA replication and transcription. Many of these substances interfere with the biosynthesis of deoxyribonucleotides and thus inhibit DNA replication in the S phase of the cell cycle.[128] Other compounds covalently bind to DNA in order to block its replication and transcription. Furocoumarins are intercalators that insert themselves into the double helix and photobind to DNA to yield interstrand crosslinks responsible for the blocking action.

Photodynamic therapy combines the selective accumulation of a photosensitizer in a target tissue, microorganism, or virus with the exposure to light of precise wavelength to trigger the formation of singlet oxygen from molecular oxygen. The conversion of triplet into singlet oxygen leads to cell death or viral inactivation, since the concentration of molecular oxygen in the tissue depletes, and singlet oxygen degrades vital molecular components of the target.

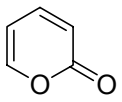
The design of efficient sensitizers for photochemotherapy should be based on the ultimate goal of sensitization, i.e., cytostatic or photodynamic action. Photosensitizing agents should ideally be single compounds with high selectivity for malignant over normal tissue and with no dark toxicity. For cytostatic action, the photobinding of furocoumarins to DNA requires that the photosensitizer shows a strong binding constant to specific DNA sites and has higher singlet excitation energies than the pyrimidine nucleobases. For photodynamic action, the photosensitizer should have higher energy than

singlet oxygen (> 0.98 eV), should preferably have high absorbance in the red region of visible light (since red light is more penetrating than blue light), and should readily form triplet excited states with long lifetimes (> 100 μs) to ensure quantitative reaction with molecular oxygen.[129] The photophysical properties of furocoumarins and some insights into their photodynamic action are discussed in Ref. **S-VII**.

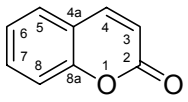
6.1. Mechanisms of photosensitization

Furocoumarins are tricyclic aromatic compounds of fused furan and bicyclic coumarin. In the classification according to pharmacological function, this family of compounds is named "furocoumarins" or "psoralens", even though it consists of substituted derivatives of coumarin, psoralen, and angelicin. In psoralen, the 2,3-furan bond is fused to the 6,7-bond of coumarin, whereas in angelicin, the 2,3-furan bond is fused to the 7,8-bond of coumarin.

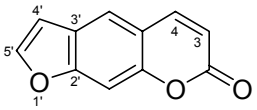
Furan (Fu)



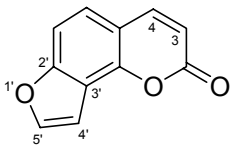
Pyrone (Pyr)



Coumarin (Cou)



Psoralen (Ps)



Angelicin (Ang)

Psoralens are versatile sensitizers prone to photodynamic and cytostatic actions. The fate of psoralens under the action of UV light can be outlined as follows:

- Formation of excited states

$$Ps(S_0) \xrightarrow{h\nu} Ps(S_1) \qquad \text{(Excitation)}$$

$$Ps(S_1) \longrightarrow Ps(T_1) \qquad \text{(Intersystem crossing)}$$

- Pathways of photosensitization

  Type I. Electron transfer mechanism: Photosensitizer triplet state reacts first with an electron donor D that is not molecular oxygen.

  $$Ps(T_1) + D \longrightarrow Ps^{\bullet -} + D^{\bullet +} \quad \text{(Reduction)}$$
  $$Ps^{\bullet -} + {}^3O_2 \longrightarrow Ps + O_2^{\bullet -} \quad \text{(Electron transfer, ET)}$$
  $$D^{\bullet +} + O_2^{\bullet -} \longrightarrow D(O_2) \quad \text{(Back ET)}$$

  Type II. Electron-exchange type mechanism: Photosensitizer triplet state reacts first with molecular oxygen (photodynamic action).

  $$Ps(T_1) + {}^3O_2({}^3\Sigma) \longrightarrow Ps(S_0) + {}^1O_2({}^1\Delta)$$

  Type III. Photobinding to, e.g., DNA: Intercalated psoralen in a singlet excited state cross-links to a DNA site (cytostatic action).

  $$Ps(S_1) + DNA(S_0) \longrightarrow Ps = DNA \quad (C_4\text{-cycloadduct})$$

- Direct electron transfer.

  $$Ps(S_0) + {}^3O_2 \xrightarrow{h\nu} Ps^{\bullet +} + O_2^{\bullet -}$$

- Autoionization.

  $$Ps(T_1) + Ps(S_0) \longrightarrow Ps^{\bullet -} + Ps^{\bullet +}$$
  $$Ps(T_1) + Ps(T_1) \longrightarrow Ps^{\bullet -} + Ps^{\bullet +}$$

Calculations at the PCM/B3LYP/6-31+G(d,p)//B3LYP/6-31G(d,p) level of theory combined with TD-DFT were performed on substituted derivatives of furocoumarins with methyl groups and groups supplying mesomeric effects (cf. Ref. **S-VII**). Psoralen and angelicin show similar photophysical properties, i.e., transition energies to the lowest singlet and triplet excited states, as well as transition probabilities. The differences in their photochemical behaviour to show cytostatic action should therefore be due to different patterns of intercalation in the double helix. Psoralen derivatives tend to form diadducts, whereas coumarin and angelicin derivatives tend to form monoadducts.

The photophysical effects of substitution (cf. Ref. **S-VII**) are a lowering of the excitation energies and, in most cases, an increase in the transition probabilities. Moreover, the gap between the first singlet and the first triplet reduces with increasing size of the compound. For bicyclic compounds the gap is ~2 eV, and narrows to ~1 eV for tricyclic and more substituted compounds.

Guanine is the nucleobase with the lowest ionization potential, and hence, the best electron donor in DNA. The reduction step of triplet furocoumarins with guanine in type-I sensitization is endothermic by ~10 kcal/mol. The subsequent reaction of the furocoumarin anions with triplet oxygen is exothermic by more than 35 kcal/mol on average. As a result, the presence of molecular oxygen in the vicinity could drive type-I reactions.

Along an alternative reaction pathway, triplet furocoumarins readily undergo electron exchange with molecular oxygen to yield singlet oxygen. The type-II reaction is exothermic by more than 40 kcal/mol on average.

6.2. Concluding summary

We have shown that psoralen and angelicin have similar photophysical properties; therefore, the observed differences in photobinding should be owing to their particular patterns of intercalation in the double helix and to the molecular mechanism of photobinding. Our calculations suggest that all furocoumarins taken into account herein have strong photodynamic action. The stronger the photodynamic action, the weaker the cytostatic action, since both mechanisms of photosensitization compete. Moreover, these furocoumarins should exhibit high levels of dark toxicity when used as cytostatic agents. For these reasons, the mechanistic details of photobinding are also necessary to design new furocoumarins with improved photophysical and cytostatic properties.

Acknowledgements

The temptation to read the acknowledgements usually comes before 'the burning desire' to read science. This epilogue aims at broadening the readership of the dissertation, a document attempting to be a comprehensive scientific summary and, at the same time, a Hollywood ending of an adventure dating back to last century. But the latter is the introducing commercial for a best-seller coming next…

It is a pleasure to thank those people who helped me with scientific aspects of this dissertation. I am indebted to my friend and supervisor Leif A. Eriksson for his excellent supervision and polishing, his endless support and patience, and all the freedom I have had to do research over these years. It has actually been fruitful our peregrination round several scientific institutions.

I have also enjoyed the company of present and former members of the group: Bo Durbeej—travelling companion and advisor—James G., Fahmi H., Maria L., Martin A., Christer E., Yanni, Xiaogi, Chuanbao, Ismael T., Peter, Anders B.… It has been stimulating to work in projects together with very skillful guys as Johan Hattne, Jan Lipfert, and Estíbaliz López.

One of the advantages of a peregrination is the people one meets on the road. Many thanks to the members of staff at the Department of Quantum Chemistry: O. Goscinki, E. Brändas, S. Lunell, Hans K., and Nesima S., and all students (Sasha, Mauritz, Marie, Johan, Peter, Björn…). From the very beginning, it has been fun to run this race together with my dear friends *don* Magnus Jansson and Per-Erik Larsson.

At BMC, Janos Hajdu and Inger Andersson facilitate the interaction among researchers and students of several fields. I thank both for this excellent scientific and social environment, and I should also thank Janos for his mentorship. My warmest thanks to David van der Spoel, Karin V., Anke T., Renco, Magnus B., Jenny, Al, Gunnar, Calle Caleman, Magnus B., Gösta H., Martin S., Diane L., Talal G., Gunilla C., Nic T., Alexandra P., Sara L., Michiel vL., Tex B.… for being so friendly. Christina R. and Susanne Ö. always have been very helpful in taking care of and sorting out all formalities.

I would like to thank Luciano O. Triguero for sharing his expertise in physics and computers. I am very grateful to Mario Piris for teaching me plenty of the physics applied in this dissertation. My special thanks go to Bernardo Llano for constantly filling my lacunae in mathematics with

definitions and theorems. This trinity has always rescued me from many difficult theoretical—and also very practical—problems.

I would like to thank all these people always ready for discussion, debate, good time, and help: Ana Elia G., Ana S., Arantxa F., Karin K., Joaquin, Juan Ramon L., Raquel M., Miguel Angel R., Helena, Inés M., Estitxu, Ignacio, Anabel Lam, Nelaine M., Betty Pita, the Vidal-Fajardo family, Freya M., Orestes T., and the gang from my corridor…

I am very lucky to have an extremely extended family that is very supportive in many ways: from Sevilla, Madrid, and Santander in Spain; from Goliad Texas and Miami in the US; from Guadalajara and Mexico City in Mexico; from Havana, Camagüey, and other Cuban cities… To all of them, and to my brothers, my deepest gratitude. All my respect for my mother, the spinal cord of our family.

Finally, my stay in Sweden can be summed up in a thought of Francis Bacon: "Travel, in the younger sort, is a part of education; in the elder, a part of experience. He that travelleth into a country before he hath some entrance into the language, goeth to school, and not to travel".

The Swedish Institute is gratefully acknowledged for financial support.

Uppsala, March 2004.

Summary in Swedish

Modern fysikalisk-kemisk beräkningsmetodik: En introduktion till biomolekylära strålningsskador och fototoxicitet.

Alla naturens livsformer utsätts ständigt för strålning av olika slag. Av stor relevans ur ett biokemiskt perspektiv är solens UV-strålning (våglängder mellan 290 och 380 nm), som på basis av fotobiologisk effekt delas upp i UV-B (290 – 320 nm) och UV-A (320 – 380 nm). Det solljus vi utsätts för utgörs till ca 0.3% av UV-A och till ca 8% av UV-B. Solstrålningen orsakar förändringar i arvsmassan, DNA, vilka i vissa fall kan leda till celldöd eller inducera mutationer i de celler som överlever. UV-B-strålningen är ansvarig för merparten av de direkta hudskador som solens strålning förorsakar. Även om UV-A strålningen orsakar färre direkta skador, verkar denna indirekt genom fotosensitisering. Detta innebär att vissa naturliga substanser i huden eller i kosmetiska produkter fungerar som "antenner" genom att fånga upp UV-A-strålning och överföra dess energi till biomolekyler, och framför allt till DNA. Dessa molekyler benämns sensitiserare.

Många kemikalier i maten, i luften, och som även produceras som naturliga metaboliter i cellerna, kan också skada olika biomolekyler, företrädelsevis DNA. De flesta av dessa substanser reagerar med vatten vilket genererar reaktiva syre-föreningar som i sin tur orsakar skadorna.

Kort sagt, vi lever i en mycket oxidativ omgivning vilken inkluderar UV-strålning och kemikalier, som på många sätt kan starta en kaskad av komplexa molekylära kedjereaktioner. Dessa reaktioner kan leda till mutationer i DNA om de molekylära skadorna inte upptäcks och repareras på tillbörligt sätt av enzymatiska reparations-system. Mutationer kan leda till okontrollerad celltillväxt och differentiering, vilket ligger till grund för exempelvis åldrande, cancer, psoriasis, m.m.

Vi känner således till de primära orsakerna (UV-A, UV-B, och kemikalier) till, och de slutliga effekterna (degenerativa sjukdomar och tillstånd) av, biomolekylära strålningsskador. Det arbete som presenteras i denna avhandling handlar om att försöka förklara och relatera till varandra, de olika reaktioner som ett antal välkända reagens orsakar på DNA. Extra vikt har fästs vid absorption av UV-strålning, vilket genererar exciterade tillstånd i DNAs baspar och som sedan omfördelas till att ge radikaljoner

(laddade molekyler med oparade elektroner). Radikaljoner reagerar gärna med vatten och dess fotodissociations-produkter. De sistnämnda inkluderar t.ex. hydrerade elektroner, protoner, väteradikaler, hydroxyl-radikaler, och superoxid-joner.

Radikaljoner ändrar elektronfördelningen inuti DNA-molekylen, vilket kan leda till att DNA-strängens bindningar mellan socker och fosfat bryts. I detta arbete presenteras en förklaringsmodell för hur denna reaktion sker på molekylär nivå.

Även hydroxyl-radikaler är mycket reaktiva, och angriper DNAs byggstenar där de genererar mutagena föreningar. Denna avhandling undersöker de molekylära mekanismer som leder till hydroxylradikal-addukter med DNA-baserna adenin och guanin, samt hur dessa addukter oxideras till 8-oxo-adenin och 8-oxo-guanin. De sistnämnda är skador på arvsmassan som ledder till den typ av mutationer som kallas transversioner.

Avhandlingen innefattar även en studie av fotosensitiserings-mekanismer inom kemoterapi. I fotokemoterapi appliceras fotosensitiserande ämnen på huden eller intas oralt, varefter de aktiveras med strålning för att initiera cytostatiska eller fotodynamiska reaktioner. En grupp av föreningar som infogar sig i DNAs dubbbelhelix och fotobinder till DNA genom att bilda korslänkar är psoralenerna. Korslänkarna håller ihop de två kedjorna i DNA, och hindrar därvid DNA-replikation och transkription. Dess cytostatiska effekt är således att de kan blockera celldelningen.

Om man vill använda psoralener som cytostatiska reagens, bör dessa molekyler uppvisa selektivitet för malign snarare än normal vävnad, och bör ej vara toxiska i mörker. De bör således binda starkt till DNA snarare än att reagera med andra föreningar som kan öka den oxidativa stressen. Denna studie visar att psoralener, då de väl absorberat UV-strålning, kan överföra sin energi till molekylärt syre i sin nära omgivning för att generera en högreaktiv förening kallad singlett-syre. Denna process är känd som en fotoallergisk reaktion, och svarar för psoralenernas mörker-toxicitet.

Samtliga resultat i denna avhandling har erhållits med metoder inom beräkningskemins område. Dessa tekniker gör det möjligt att beräkna och förutsäga molekylers egenskaper med hjälp av kvantkemiska teorier och datorer. Molekylära simuleringar är helt enkelt experiment som utförs i datorer. Det är således även ytterst viktigt att utveckla metoder för att kunna simulera vissa typer av experiment. En grundläggande ny metod inom fotoelektrokemi och jon-termokemi presenteras häri, vilken möjliggör studier av reaktioner i lösning, som de vi vet äger rum inom experimentell strålnings-kemi och fotokemi. Denna approximation medger inte bara en hanterbar teoretisk modell för reaktioner i vilka elektroner och protoner deltar som oberoende joner i en reaktion, utan ger även ett enkelt och (beräkningsmässigt) kostnadseffektivt vektyg för att understödja experimentella studier, för att förklara experimentella observationer, och, på

samma gång, för att ge insikt på molekylär nivå i problem av biomedicinsk och teknisk relevans.

Översättning av Leif A. Eriksson och Bo Durbeej.

References

1. *Oxford English Dictionary*, 2nd ed.; Oxford University Press: Oxford, 1989. *Oxford Reference Online*, <http://www.oed.com>.
2. Clapham, C. *The Concise Oxford Dictionary of Mathematics*; Oxford University Press: Oxford, 1996. *Oxford Reference Online*, <http://www.oxfordreference.com>.
3. *A Dictionary of Computing*; Oxford University Press: Oxford, 1996. *Oxford Reference Online*, <http://www.oxfordreference.com>.
4. Duhem, P. *The Aim and Structure of Physical Theory* (English translation by P. P. Wiener); Princeton University Press: Princeton, 1991.
5. *A Dictionary of Physics*; Isaacs, A. Ed.; Oxford University Press: Oxford, 2000. *Oxford Reference Online*, <http://www.oxfordreference.com>.
6. Zewail, A. H. *J. Phys. Chem. A* **2000**, *104*, 5660.
7. DiMauro, L. F. *Nature* **2002**, *419*, 789.
8. Drescher, M.; Hentschel, M.; Kienberger, R.; Uiberacker, M.; Yakovlev, V.; Scrinizi, A.; Westerwalbesloh, T.; Kleineberg, U.; Heinzmann, U.; Krausz, F. *Nature* **2002**, *419*, 803.
9. *Nanotechnology Research Directions: IWGN Workshop Report*, IWGN Division at Loyola Collage: Baltimore, 1999, <http://www.wtec.org>.
10. Ghoniem, N. M.; and Cho, K. *Computer Modeling in Engineering & Sciences* **2002**, *3*, 147.
11. Ghoniem, N. M.; Busso, E. P.; Kioussis, N.; Huang, H. *Philosophical Magazine* **2003**, *83*, 3475.
12. Ingram, G. D.; Cameron, I. T. In *Proceedings of the 9th APCChE Congress and CHEMECA 2002*; Gostomski, P. A.; Morison, K. R. Eds.; University of Canterbury: Christchurch, NZ, 2002, p. 254.
13. Crawshaw, J.; Windle A. H. *Fibre Diffraction Rev*. **2003**, *11*, 52.
14. Jensen, F. *Introduction to Computational Chemistry*; Wiley: Chichester, UK, 1999.
15. Leach, A. R. *Molecular Modelling: Principles and Applications*, 2nd ed.; Pearson Education: Harlow, UK, 2001.
16. Bieri, O.; Kiefhaber, T. *Biol. Chem*. **1999**, *380*, 923.
17. Chan, H. S.; Dill, K.A. *Physics Today* **1993**, 46, 24.
18. *WTEC Panel on Applications of Molecular and Material Modelling*; ITRI at Loyola Collage: Baltimore, 2002, <http://www.wtec.org>.

19. Vreven T.; Morokuma, K. *J. Comp. Chem*. **2000**, *21*, 1419.
20. Carloni, P.; Rothlisberger, U.; Parrinello, M. *Acc. Chem. Res*. **2002**, *35*, 455.
21. Aktah, D.; Passerone, D.; Parrinello, M. *J. Phys. Chem. A* **2004**, *108*, 848.
22. Werner, B. T. *Science* **1999**, *284*, 102.
23. Healey, R. In *The Stanford Encyclopedia of Philosophy (Fall 1999 Edition)*; Edward N.; Zalta, E. N. Ed., <http://plato.stanford.edu>.
24. Kelley, D. *The Foundations of Knowledge* (Unabridged); The Objectivist Center: New York, 1988.
25. Blackburn, S. *The Oxford Dictionary of Philosophy*; Oxford University Press: Oxford, 1996. *Oxford Reference Online*, <http://www.oed.com>.
26. *Encyclopædia Britannica*, *Encyclopædia Britannica Online*, 2004, <http://search.eb.com>.
27. Kohn, W.; Becke, A. D.; Parr, R. G. *J. Phys. Chem*. **1996**, *100*, 12974.
28. Head-Gordon, M. *J. Phys. Chem*. **1996**, *100*, 13213.
29. Szabo, A; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, revised 1st ed.; McGraw-Hill: New York, 1989.
30. Steinfeld, J. I.; Francisco, J. S.; Hase, W. L. *Chemical Kinetics and Dynamics*; Prentice-Hall: New Jersey, 1989.
31. Truhlar, D. G. In *The Reaction Path in Chemistry: Current Approaches and Perspectives*; Heidrich, D. Ed.; Kluwer Academic Publishers: The Netherlands, 1995, p. 229.
32. McQuarrie, D. A.; Simon, J. D. *Molecular Thermodynamics*; University Science Books: Sausalito, California, 1999.
33. Atkins, P.W., Friedman, R. S. *Molecular Quantum Mechanics*, 3rd ed.; Oxford University Press: Oxford, 1996.
34. Minkin, V. I. *Pure Appl. Chem*. **1999**, *71*, 1919.
35. Fulde, P. *Electron Correlations in Molecules and Solids*, 3rd ed.; Springer-Verlag: Berlin, 1995.
36. Raghavachari, K.; Anderson, J. B. *J. Phys. Chem*. **1996**, *100*, 12960.
37. Löwdin, P.-O. *Adv. Chem. Phys*. **1959**, *2*, 207.
38. Pettifor, D. G. *Bonding and Structure of Molecules and Solids*; Clarendon Press: Oxford, UK, 1995.
39. Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic Structure Theory*; John Wiley & Sons: Chichester, UK, 2000.
40. Kohn, W. *Rev. Mod. Phys*. **1999**, *71*, 1253.
41. Kohn, W.; Becke, A. D.; Parr, R. G. *J. Phys. Chem*. **1996**, *100*, 12974.
42. Chakravorty, S. J.; Corongiu, G.; Flores J. R.; Sonnad, V.; Clementi, E.; Carravetta, V.; Cacelli, I. In *Modern Techniques in Computational Chemistry: MOTECC-89*; Clementi, E. Ed.; ESCOM: Leiden, 1989.

43. Dupuis, M.; Mougenot, P.; Watts, J. D.; Hurst, G. J. B.; Villar, H. O. In *Modern Techniques in Computational Chemistry: MOTECC-89*; Clementi, E. Ed.; ESCOM: Leiden, 1989.
44. Becke, A. D. In *Modern Electronic Structure Theory*; Yarkony, D. R. Ed.; *Advanced Series in Physical Chemistry*, Vol. 2; World Scientific: Singapore, 1995.
45. Tomasi, J.; Persico, M. *Chem. Rev*. **1994**, *94*, 2027.
46. Rivail, J.-L.; Rinaldi, D. In *Computational Chemistry, Review of Current Trends*; Leszczynski, J. Ed.; World Scientific: New York, 1995, p 139.
47. Cramer, C. J.; Truhlar, D. G. *Chem. Rev*. 1999, 99, 2161.
48. Tapia, O.; Goscinski, O. *Mol. Phys*. 1975, 29, 1653.
49. Bransden, B. H.; Joachain, C. J. *Physics of Atoms and Molecules*, 2nd ed.; Prentice Hall: Harlow, UK, 2003.
50. Helgaker, T. In *European Summer School in Quantum Chemistry*; Roos, B. O.; Widmark, P.-O. Eds.; University of Lund: Lund, 1999.
51. Kauzmann, W. *Quantum Chemistry*; Academic Press: New York, 1957.
52. Casida, M. E. In *Recent Advances in Density Functional Methods*; Chong, D. P., Ed.; World Scientific: Singapore, 1995; p. 155.
53. Taylor, P. In *European Summer School in Quantum Chemistry*, Vol. 2; Roos, B. O.; Widmark, P.-O. Eds.; University of Lund: Lund, 1999; p. 401.
54. Rappoport, D. In *Winter School in Theoretical Chemistry: Lecture Notes*; University of Helsinki: Helsinki, 2003, <http://www.chem.helsinki.fi>.
55. Bauernschmitt, R.; Ahlrichs, R. *Chem. Phys. Lett*. **1996**, *256*, 454.
56. Tozer, D. J.; Handy, N. C. *Phys. Chem. Chem. Phys*. **2000**, *2*, 2117.
57. Furche, F. *J. Chem. Phys*. **2001**, *114*, 5982.
58. Furche, F.; Ahlrichs, R. *J. Chem. Phys*. **2002**, *117*, 7433.
59. Shao, Y.; Head-Gordon, M. *J. Chem. Phys*. **2003**, *118*, 4807.
60. Dreuw, A.; Weisman, J. L.; Head-Gordon, M. *J. Chem. Phys*. **2003**, *119*, 2943.
61. Marques, M. A. L.; Gross, E.K.U. *Annu. Rev. Phys. Chem*. **2004**, *55*, 427.
62. Stratmann, R. E.; Scuseria, G. E.; Frisch, M. J. *J. Chem. Phys*. **1998**, *109*, 8218.
63. Pople, J. A. *Rev. Mod. Phys*. **1999**, *71*, 1267.
64. Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; and Pople, J. A. *Ab Initio Molecular Orbital Theory*; John Wiley & Sons: New York, 1986.
65. Pople, J. A.; Binkley, J. S.; Seeger, R. *Int. J. Quantum Chem.: Quantum Chem. Symp*. **1976**, *10*, 1.
66. Pople, J. A.; Seeger, R.; Krishnan, R. *Int. J. Quantum Chem.: Quantum Chem. Symp*. **1977**, *11*, 149.

67. Bartlett, R. J.; Purvis, G. D. *Int. J. Quantum Chem*. **1978**, *14*, 561.
68. Taylor, P. In *European Summer School in Quantum Chemistry*, Vol. 2; Roos, B. O.; Widmark, P.-O. Eds.; University of Lund: Lund, 1999; p. 393.
69. Parsons, R. Pure Appl. Chem. 1974, 37, 500.
70. Trasatti, S.; Parsons, R. *Pure Appl. Chem*. **1983**, *55*, 1251.
71. Bockris, J. O'M.; Reddy, A. K. N. *Modern Electrochemistry*; Plenum Press: New York, 1970.
72. Reiss, H. *J. Electrochem. Soc*. **1988**, *135*, 247C.
73. Prigogine, I.; Defay, R. *Chemical Thermodynamics*; Longmans: London, 1954.
74. Everett, D. H. *An Introduction to the Study of Chemical Thermodynamics*; Longmans: London, 1959.
75. McQuarrie, D. A. *Statistical Mechanics*; Harper & Row: New York, 1976.
76. Reiss, H. *J. Phys. Chem*. **1985**, *89*, 3783.
77. Reiss, H.; Heller, A. *J. Phys. Chem*. **1985**, *89*, 4207.
78. Herring, C.; Nichols, M. H. *Rev. Mod. Phys*. **1949**, *21*, 185.
79. Trasatti, S. *Pure Appl. Chem*. **1986**, 58, 955.
80. von Sonntag, C. *The Chemical Basis of Radiation Biology*; Taylor & Francis: London, 1987.
81. Kantor, G. J. In *Encyclopedia of Molecular Biology and Molecular Medicine*, Vol. 2; Meyers, R. A. Ed.; VCH: Weinheim, 1996, p. 40.
82. von Sonntag, C.; Schuchmann, H-P. In *Encyclopedia of Molecular Biology and Molecular Medicine*, Vol. 3; Meyers, R. A. Ed.; VCH: Weinheim, 1996, p. 354.
83. Newcomb, T. G.; Loeb, L. A. In *DNA Damage and Repair*, Vol. 1: *DNA Repair in Prokaryotes and Lower Eukaryotes*; Nickoloff, J. A.; Hoekstra, M. F.; Eds.; Humana Press Inc.: Totowa, 1998, p. 65.
84. Coohill, T. P. In *CRC Handbook of Organic Photochemistry and Photobiology*; Horspool, W. M.; Song, P.-S. Eds.; CRC Press: Boca Raton, 1995, p. 1267.
85. Peak, M. J.; Peak, J. G. In *CRC Handbook of Organic Photochemistry and Photobiology*; Horspool, W. M.; Song, P.-S. Eds.; CRC Press: Boca Raton, 1995, p. 1318.
86. Bernas, A.; Guthier, M.; Grand, D.; Parlant, G. *Chem. Phys. Lett*. **1972**, *17*, 439.
87. Bernas, A.; Blais, J.; Guthier, M.; Grand, D. *Chem. Phys. Lett*. **1975**, *30*, 383.
88. Bernas, A.; Grand, D.; Amouyal, E. *J. Phys. Chem*. **1980**, *84*, 1259.
89. Itaya, K.; Kawai, M.; Toshima, S. *J. Am. Chem. Soc*. **1978**, *100*, 5996.
90. Bard, A. J.; Itaya, K.; Malpas, R. E.; Teherani, T. *J. Phys. Chem*. **1980**, *84*, 1262.
91. Schuster, G. B. *Acc. Chem. Res*. **2000**, *33*, 253.

92. Sprecher, C. A.; Johnson, W. C. *Biopolymers* **1977**, *16*, 2243.
93. Samuni, A.; Neta, P. *J. Phys. Chem*. **1973**, *77*, 1629.
94. Lassmann, G.; Eriksson, L. A.; Lendzian, F.; Lubitz, W. *J. Phys. Chem. A* **2000**, *104*, 9144.
95. Eriksson, L. A.; Himo, F. *Trends Phys. Chem*. **1997**, *6*, 153.
96. (a) Hoog, A. J. de; Buys, H. R.; Altona C.; Havinga, E. *Tetrahedron* **1969**, *25*, 3365. (b) Edward, J. T. *Chem. Ind*. **1955**, 1102. (c) Lemieux, R. U. In *Molecular Rearrangements*, P. de Mayo Ed.; Interscience: New York, 1964, p. 723.
97. Kirby, A. J. *The Anomeric Effect and Related Stereoelectronic Effects at Oxygen*; Spriger-Verlag: Berlin, 1983.
98. Tvaroška, I.; Bleha, T. *Adv. Carbohydr. Chem. and Biochem*. **1989**, *47*, 45.
99. Deslongchamps, P. *Stereoelectronic Effects in Organic Chemistry*; Pergamon: New York, 1983.
100. (a) Salzner, U.; Schleyer, P. v. R. *J. Am. Chem. Soc*. **1993**, *115*, 10231. (b) Salzner, U.; Schleyer, P. v. R. *J. Org. Chem*. **1994**, *59*, 2138. (c) Salzner, U. *J. Org. Chem*. **1995**, *60*, 986.
101. Llano, J.; Montero, L. A. *J. Comput. Chem*. **1996**, *11*, 1371.
102. Perrin, C. L.; Armstrong, K. B.; Fabian M. A. *J. Am. Chem. Soc*. **1994**, *116*, 715.
103. Krol, M. C.; Huige, C. J. M.; Altona, C. *J. Comp. Chem*. **1990**, *11*, 765.
104. David, S.; Eisenstein, O.; Hehre, W. J.; Salem, L.; Hoffmann, R. *J. Am. Chem. Soc*. **1973**, *95*, 3806.
105. Schleyer, P. v. R.; Kos, A. J. *Tetrahedron* **1983**, *39*, 1141.
106. Wolfe, S.; Whangbo, M.-H.; Mitchell, D. *J. Carbohydr. Res*. **1979**, *69*, 1.
107. (a) Montagnani, R.; Tomasi, J. *Int. J. Quantum Chem*. **1991**, *39*, 851. (b) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc*. **1991**, *113*, 8305. (c) Cramer, C. J.; Truhlar, D. G. *J. Comp. Chem*. **1992**, *13*, 1089. (d) Cramer, C. J. *J. Org. Chem*. **1992**, *57*, 7034. (e) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc*. **1993**, *115*, 5745.
108. Burrows, C. J.; Muller, J. G. *Chem. Rev*. **1998**, *98*, 1109.
109. Candeias, L. P.; Steenken, S. *Chem. Eur. J*. **2000**, *6*, 475.
110. Hagen, U. *Radiat. Environ. Biophys*. **1986**, *25*, 261.
111. Kasai, H.; Crain, P. F.; Kuchino, Y.; Nishimura, S.; Ootsuyama A.; Tanooka, H. *Carcinogenesis* **1986**, *7*, 1849.
112. Teoule, R. *Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med*. **1987**, *51*, 573.
113. Hatahet Z.; Wallace, S. S. In *DNA Damage and Repair*, Vol. 1: *DNA Repair in Prokaryotes and Lower Eukaryotes*; Nickoloff, J. A.; Hoekstra, M. F.; Eds.; Humana Press Inc.: Totowa, 1998, p. 229.
114. Cheng, K. C.; Cahill, D. S.; Kasai, H.; Nishimura, S.; Loeb, L. A. *J. Biol. Chem*. **1992**, *267*, 166.

115. Pavlov, Y. I.; Minnick, D. T.; Izuta, S.; Kunkel, T. A. *Biochemistry* **1994**, *33*, 4685.
116. Cullis, P. M.; Malone, M. E.; Merson-Davies, L. A. *J. Am. Chem. Soc*. **1996**, *118*, 2775.
117. Steenken, S. *Chem. Rev*. **1989**, *89*, 503.
118. Vieira, A. J. S. C.; Steenken, S. *J. Am. Chem. Soc*. **1990**, *112*, 6986.
119. Steenken, S.; Jovanovic, S. V.; Bietti, M.; Bernhard, K. *J. Am. Chem. Soc*. **2000**, *122*, 2373.
120. Pogozelski, W. K.; Tullius, T. D. *Chem. Rev*. **1998**, *98*, 1089.
121. Nelson, D. J.; Symons, M.C.R.; Wyatt, J. L. *J. Chem. Soc., Faraday Trans*. **1993**, *89*, 1955.
122. Giese, B.; Beyrich-Graf, X.; Erdmann, P.; Petretta, M.; Schwitter, U. *Chem. Biol*. **1995**, *2*, 367.
123. Becker, D.; Razskazovskii, Y.; Sevilla, M. D. *Radiat. Res*. **1996**, *146*, 361.
124. Becker, D.; Bryant-Friedrich, A.; Trzasko, C. A.; Sevilla, M. D. *Radiat. Res*. **2003**, *160*, 174.
125. Boon, P. J., Cullis, P. M., Symons, M. C. R., Wren, B. W. *J. Chem. Soc., Perkin Trans. 2* **1984**, 1393.
126. Sanderud, A.; Sagstuen, E. *J. Phys. Chem*. **1996**, *100*, 9545.
127. Inczédy, J.; Lengyel, T.; Ure, A. M. *Compendium of Analytical Nomenclature: Definitive Rules, 1997*, 3rd ed.; Blackwell Science: Oxford, 1998, <http://www.iupac.org/publications/analytical_compendium>.
128. Koolman, J.; Röhm, K-H. *Color Atlas of Biochemistry*; Thieme: Stuttgart, 1996.
129. Harriman, A. In *CRC Handbook of Organic Photochemistry and Photobiology*; Horspool, W. M.; Song, P.-S. Eds.; CRC Press: Boca Raton, 1995, p. 1374.

Acta Universitatis Upsaliensis

Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology*. (Prior to October, 1993, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science".)

Distribution:
Uppsala University Library
Box 510, SE-751 20 Uppsala, Sweden
www.uu.se, acta@ub.uu.se

ISSN 1104-232X
ISBN 91-554-5940-4