A Thesis Submitted for the Degree of PhD at the University of Warwick

Permanent WRAP URL:

http://wrap.warwick.ac.uk/87995

Copyright and reuse:

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it.

Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

Theoretical studies in dye-sensitized solar cells and photo-catalysis at metal oxide interfaces

Ip Chung Man

A thesis submitted in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy in Chemistry

Department of Chemistry, University of Warwick

July 2016

Contents

List of figures and tables

Acknowledgements

I would like to express my most sincere gratitude to Prof. Alessandro Troisi for all the guidance and support provided during my pursuit of this doctoral degree. My gain from my supervisor is beyond scientific knowledge; his work attitude, critical thinking and open-mindedness to knowledge have certainly inspired me. I believe I have become a more sophisticated person under his supervision and I am very thankful to him for all the learning opportunities.

I would also like to thank Dr. Natalia Martsinovich, Dr. Juan Arago, Dr. Myeong Lee and Rocco Fornari for insightful academic discussions. I would like to acknowledge Dr. Antonio Eleuteri for his contribution on the statistical modeling as part of this thesis. I would like to thank my parents, sister and friends who have been consistently supportive and encouraging throughout my PhD study.

Finally I would like to kindly acknowledge the financial support offered by EPSRC.

Declaration

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree. The work presented in this thesis was conducted by the author except:

The statistical analysis presented in section A4.2 and section A4.3, and Figure 4.1, 4.4 and 4.5, which were contributed by Dr. Antonio Eleuteri and Prof. Alessandro Troisi.

Parts of this thesis have been published or submitted for publication by the author:

1. C. M. Ip, A. Eleuteri and A. Troisi, Predicting with confidence the efficiency of new dyes in dye sensitized solar cells, *Phys. Chem. Chem. Phys*, **2014**, 16, 19106

2. C. M. Ip and A. Troisi, Does the donor-π-acceptor character of dyes improve the efficiency of dye-sensitized solar cells?, *J. Phys. Chem. Lett.*, **2016**, 7, 2989

3. C. M. Ip and A. Troisi, Computational study of competing reaction mechanisms of photo-catalytic reduction of CO_2 on anatase(101), *Phys. Chem. Chem. Phys*, **2016**, 18, 25010

Abstract

In this thesis we present theoretical and computational studies for the p-type and n-type dye-sensitized solar cells (DSSCs), and the photo-catalytic reduction of carbon dioxide to generate methane. For p-type DSSCs, we computed the hole injection and recombination rate at the semiconductor-dye (NiO-C343) interface based on theories applied previously for electron injection and recombination in n-type DSSCs, derived from Fermi golden rule and non-adiabatic charge transfer theory. Our analysis showed that the faster recombination in p-type in comparison to n-type was due to difference in Franck-Condon factor of the relevant transitions.

For n-type DSSCs, we devised a statistical model based on the electronic structural properties of dyes for predicting the efficiency of this device with confidence when a new dye was employed. The approach for constructing the model was QSAR-like and involved examining correlations between the efficiency of the device and a number of predictors that were properties of the dye. We also quantified the donor-π-acceptor (D-π-A) character of dyes and showed statistically that increasing the strength of this character was ineffective for improving the efficiency of n-type DSSCs.

For photo-catalytic reduction of carbon dioxide to produce methane, we studied with DFT calculations three competing reaction mechanisms on TiO_2 anatase(101) proposed in literature on the basis of experimentally observed reaction intermediates. By comparing the thermodynamics of mechanisms we showed that the formaldehyde pathway was the most favorable reaction mechanism. The computational methodology employed was useful for testing mechanistic hypotheses for reactions on the surface of solid catalysts.

Abbreviations

AC	Alternating Current
AIC	Akaike Information Criteria
AM	Air Mass
ANOVA	Analysis of Variance
B3LYP	Becke 3-parameter exchange and Lee-Yang-Parr correlation functional
BEP	Bells-Evans-Polanyi
BET	Brunauer-Emmett-Teller
C343	Coumarin 343
CB	Conduction Band
cDFT	Constrained Density Functional Theory
CHE	Computational Hydrogen Electrode
CI	Charge Injection
CI-NEB	Climbing Image Nudged Elastic Bands
D-π-A	Donor-π-Acceptor character (of dyes)
DFT	Density Functional Theory
DRIFTS	Diffuse Reflectance Infrared Fourier Transform Spectroscopy
DSSC	Dye-Sensitized Solar Cell
DZP	Double-ζ Polarized (basis set)
EDM	Excitation Dipole Moment
EIS	Electrochemical Impedance Spectroscopy
ESR/EPR	Electron Spin Resonance/Electron Paramagnetic Resonance
ET	Electron Transfer
FC	Franck-Condon
HER	Hydrogen Evolution Reaction
HOMO	Highest Occupied Molecular Orbital
HT	Hole Transfer
IEA	International Energy Agency
IR	Infra-Red
LR	Likelihood Ratio
L(S)DA	Local (Spin) Density Approximation
LUMO	Lowest Unoccupied Molecular Orbital
MLF	Maximization of the Likelihood Function
NDD	Normalized Dipole Density
OA	Orbital Asymmetry
OD	Orbital Density

OER	Oxygen Evolution Reaction
PBE	Pardew-Burke-Ernzerhof
PCE	Power Conversion Efficiency (of dye-sensitized solar cells)
PCET	Proton-Coupled Electron Transfer
PCM	Polarized Continuum Model
(P)DOS	(Partial) Density of States
PM-IRAS	Polarization Modulation Infrared Reflection Absorption Spectroscopy
PT	Proton Transfer
pzc	point of zero charge
QSAR	Quantitative Structure-Activity Relationship
QSPR	Quantitative Structure-Property Relationship
SES	Solvent Excluding Surface
SHE	Standard Hydrogen Electrode
(S)GGA	(Spin-polarized) Generalized Gradient Approximation
TAS	Transient Absorption Spectroscopy
TDDFT	Time-Dependent Density Functional Theory
TPD	Temperature-Programmed Desorption
UHF	Unrestricted Hartree Fock
UV	Ultra-Violet
VB	Valence Band
VOC	Volatile Organic Compounds
WoS	Web of Science

Chapter 1　Introduction

A main objective of this work is to employ theories and computational chemistry to aid the design of new materials for dye-sensitized solar cells (DSSCs) and photo-catalysis. Both technologies are currently developing to gain a stronger position in the sustainable energy market. In this chapter we will first provide a brief description of the current challenge in the energy industry, followed by the motivations for developing DSSCs and photo-catalysis, as well as their respective challenges for attaining high efficiencies. We will then briefly introduce the possible roles of theories and computational chemistry in the development of these technologies. Subsequently we will provide introductions of DSSCs and photo-catalysis, including their operational principles, some research focuses, common experimental techniques employed and examples of theoretical works. This chapter will end with an outline of this thesis.

1.1　Challenges in the energy industry

The new oil extraction[1] and recovery[2] techniques that have emerged in recent years have cast doubts on the existence of a global energy shortage.[3] The contemporary development of alternative energy sources to fossil fuels is therefore motivated more strongly by our concern over the climatic and environmental changes; the burning of fossil fuels is notorious for increasing the emission of greenhouse gases, leading to dramatic variation of climatic patterns relative to what the human race has experienced over the past decades.[4] Alternative sustainable energy sources, however, do not necessarily possess environmental advantages over fossil fuels. Nuclear energy, which provides around 11% of the global electricity production,[5] is far from being a popular

choice of energy source due primarily to the tragic catastrophes occurred in, for instances, Ukraine 1986[6] and Japan 2011.[7] The nuclear wastes and radiation released in these accidents had a far-reaching environmental impact, threatening the well-being of many living organisms by, for example, increasing the risk of cancer.[6]

An outstanding candidate for a sustainable energy source is solar energy. Such type of energy is perceived as clean, abundant and freely available, despite its dependence upon geography. The attractiveness of solar energy is well-reflected by the large sum of investments received, where in 2013 over half of the 214 billion US dollars invested in renewable energy worldwide were devoted to solar energy.[8] The international energy agency (IEA) have optimistically predicted that solar energy will provide about 16% of global electricity production by 2050, even though it is likely that solar energy provides only 2% by 2020.[9] Another striking advantage of solar energy is that the design of photovoltaic devices are highly flexible, leading to a wide range of applications: from pocket calculators using small-sized amorphous silicon solar cells to rooftop solar panels for powering a household using 'second generation' thin-film solar cells. This unique aspect that is often not possessed by other renewable energy offers large leeway for 'levelised cost' improvement, production scale, sites for installation and even aesthetic of the devices. It is expected that the price/performance ratio will keep reducing with the advancement of 'third generation' cells, and there are explicit signs showing that the cost of solar power is plunging; for example, the cost of power generated by residential photovoltaic systems has dropped by 21% in Japan in 2013.[8] Edward Lucas, a senior journalist of *the Economist*, suggested that solar energy will bound to be cheaper and become 'a dagger in the heart of the fossil-fuel industry'.[8]

A key disadvantage of solar energy is that the supply of energy is mostly intermittent, where no device thus far is known to function efficiently under dim lights or nocturnally; hence energy storage is of paramount importance. Solar energy storage is not only for 'battening down the hatches' for a household at night, but also induces large savings in power consumptions and financial returns.[8,10] The storage of solar energy can be in many forms, be it warm water, ice or in the battery of an electric car,[8] which should lead to a crucial reminder that solar energy is not restricted to generating electricity with photovoltaic cells. It can be converted to thermal energy to drive turbines,[11] or it can also be employed in visible light chemical synthesis with the aid of a photo-catalyst,[12] where in this case solar energy can be stored in high energy-density chemical compounds such as methane.[13] Methane has long been a part of renewable energy, being a major constituent of natural gas, a major product in the bio-fuel industry,[14] and also a preferred 'stepping stone' for a cleaner usage of coal for energy application.[15]

There are, therefore, two pivotal elements for solar energy to beat fossil fuel, especially oil: the first is the further reduction of price/performance ratio, reaching a competitive level to that of the oil technology; and the second is a cost-effective approach for efficiently storing solar energy. Both elements can be strongly influenced by the materials employed in the technology for harvesting solar energy. The understanding of the physical and chemical processes, such as the charge dynamics in solar cells, and the chemical reactions for producing methane, would be important for informing the systematic search and design of new materials.

1.2 Motivations and challenges in the development of DSSCs and photo-catalysis

At present the solar cell market has been dominated by silicon-based solar cells,[16] where the price per Watt has been dropping over the last four decades.[17,18] The efficiency of silicon solar cells will be, however, ultimately limited by the factors considered in the Shockley-Queisser model[19] and the Auger recombination.[20] There has been therefore a surge of alternative solar cells with the potential of attaining efficiency higher than that can be achieved by silicon-based solar cells. One of the low-cost alternatives is dye-sensitized solar cells (DSSCs).

The first DSSC that was feasible for practical applications was developed by O'Regan and Grätzel,[21] in which they demonstrated a breakthrough that involved using a charge-transfer dye adsorbing on semiconductor particles with high surface area for light absorption, allowing the device to harvest solar power much more efficiently than previous attempts. DSSCs can be assembled with low-cost materials with relatively simple and inexpensive fabrication, and be available in various colours. The rather desirable price/performance ratio of DSSCs has already allowed the device to be commercialized, and at present they are likely to find applications in building-integrated photovoltaic systems, flexible electronics, and perhaps other niche markets.[16] The considerably special operation principles and constructs (see section 1.4.1) of DSSCs have also served as a model system for other innovative development in solar energy technology, such as dye-sensitized photo-catalysis.[22]

The current power conversion efficiency (PCE) of DSSCs, however, is still typically lower than that of silicon solar cells by approximately 10-15%.[23] The

key challenge for increasing DSSCs' competitiveness is therefore the improvement of PCE, among some other stability issues.[24,25] The improvement of PCE is closely related to the efficient generation of photo-current, which is pertained to the optical and electronic properties of dyes. The understanding of how charges can be transferred efficiently between the dye and the electrode is therefore important for the improved design of both components. It is, however, challenging to understand solely with macroscopic experiments due primarily to two elements: (i) the elementary processes at multiple time-scales that can span from seconds to femto-seconds, and (ii) the complicated devices' physics and chemistry that originates from the interactions between various components in the device.

Photo-catalysis is perceived as a clean method of storing solar energy in the form of chemical energy, most notably as hydrogen fuel and methane, synthesized using water and carbon dioxide as initial reactants respectively. The combination of such methods with photo-electrochemical cells, such as fuel cells, which generate electricity with chemical energy, can potentially lead to sustainable photo-driven energy cycles that can complement the intermittency of solar cells and prepare energy feedstock for urgent energy shortage. As such the technology would be advantageous against conventional combustion of fossil fuels in terms of environmentally friendliness, a strong motivation for developing photo-catalysis instead of improving existing fossil fuels technology; fossil fuels, however, would most likely possess advantage in terms of cost, in which even advanced photo-catalysis may still require external help such as legislation to reduce such disadvantage.[26]

The extensive research effort devoted to photo-catalysis[27] was initiated by the reports in the 1970s on the observation of water splitting[28] and methane from

reducing CO_2.[29] Over roughly 40 years, however, the technology for fuel generation is still far from commercialization.[30,31] The key challenge in photo-catalysis is to identify a catalytic system that is capable of (i) absorbing solar light efficiently, (ii) providing electrons and holes at the surface for redox reactions efficiently, and (iii) catalyzing the reaction efficiently. These requirements are difficult to achieve simultaneously. The different requirements have also led to different measures for characterizing the performance of a catalytic system, which generally would be either catalyst-based or light-based.[13] Both types of measure have their drawbacks and combining them would still be insufficient to encompass all the important limiting factors for the performance of a photo-catalytic system. For example, if one considers using turnover frequency with Brunauer-Emmett-Teller (BET) surface area to characterize a photo-catalyst, some errors may be introduced in the determination of surface area since not all surface area is photo-active.[13,32] All in all, the lack of a single measure of efficiency for catalytic systems does not allow a universal ranking of photo-catalytic systems, which causes difficulty for identifying the best catalytic systems among a large number of candidates.

In any case the systematic improvement of requirement (iii) for a catalytic system should only be possible when there is sufficient knowledge of the reaction mechanism. The understanding of the reaction mechanism would help identify the rate-limiting step of the reaction, and further understanding of the physical origin of the kinetic barrier associated with this step should provide hints for changing and/or modifying the catalytic system to improve the reaction efficiency. Many proposed reaction mechanisms are educated guesses based on intermediates identified in experiments, in which the connections and roles of the intermediates are not given directly by the experimental techniques.[33,34] Incorrect

guesses can lead to wasteful analysis and proposed mechanisms should therefore seek agreement from further studies on, for example, the thermodynamics and kinetics of the reactions.

1.3   The roles of theories and computational chemistry

One of the possible roles of theories for DSSCs would be to complement the understanding of interfacial charge transfer processes from macroscopic experiments. Theories for electron transfer[35] are well-documented and have been applied to electrochemical systems,[36] and the modelling of organic-solid interfaces is now possible with a number of computational tools.[37,38] Models for charge transfer events are therefore both possible and highly desirable, especially when capable of revealing the determining factors that control these processes. Computational chemistry is also invaluable for rationalising the correlations between the chemical structures of materials and device performance.[39] The optical and electronic properties of dyes are now commonly explored with density functional theory (DFT) and its time-dependent version, and when in conjunction with quantum dynamics calculations, these methods have also contributed to the understanding of elementary processes (see section 1.4.3).

For photo-catalysis, computational chemistry can be useful to improve the certainty of a proposed reaction mechanism. Examples of using computational chemistry to understand reaction mechanisms are abundant, typically by allowing examination of the potential energy surfaces of the reactions,[40–42] although the same objective can perhaps be achieved by considering other methods such as the united reaction valley approach.[43] In many cases the chemical structure and the energetics of the intermediates and transition states are now computationally

accessible, offering understanding of the kinetics and thermodynamics of a reaction mechanism. A number of examples are given in section 1.5.3 to demonstrate how computational chemistry has been employed in this respect for photo-catalytic reactions.

1.4 An introduction to dye-sensitized solar cells

1.4.1 Background

The most common types of DSSCs are the n-type DSSCs, where the photo-electrode is constituted by n-type semiconductor nano-particles such as TiO_2, which are covered with molecular dyes and submerged in electrolyte solution. The ideal operation principles for a n-type device is (Fig.1.1), initially, the injection of a photo-excited charge from the lowest unoccupied molecular orbital (LUMO) of the dye to the conduction band (CB) of the semiconductor. Subsequently this injected charge should migrate through the external circuit and interact with the electrolyte at the counter-electrode. The electrolyte then donates an electron to the oxidized dye, completing a charge transfer cycle. The power conversion efficiency (PCE) of the cells is hampered when serious charge recombination occurs, which can be between the dye and the semiconductor, the electrolyte and the semiconductor and/or within the dye immediately after photo-excitation.
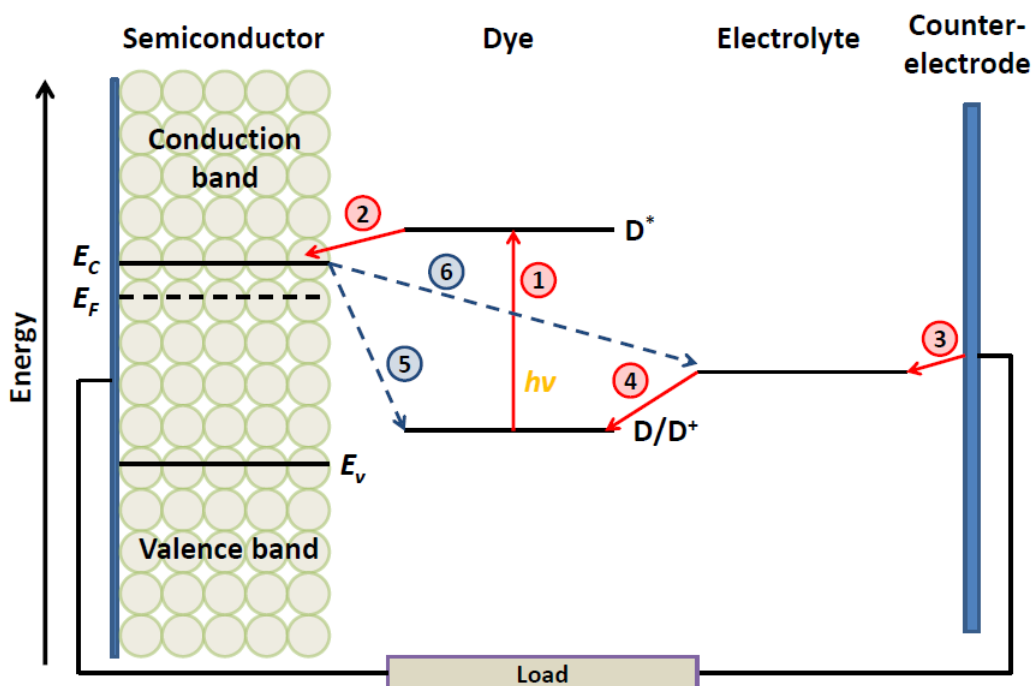
Figure 1.1 Illustration of the operation principles of n-type DSSCs. The ideal electron transfers (red arrows) start from 1.) photo-excitation of the dye, followed by 2.) charge injection, 3.) reduction of electrolyte and finally 4.) reduction of oxidized dye. The unwanted electron transfers (blue dashed arrows) include 5.) charge recombination between semiconductor and dye and 6.) between semiconductor and electrolyte. E_c is the conduction band edge, E_F is the quasi-Fermi level and E_v is the valence band edge.

A n-type DSSC with a Zn-based porphyrin dye (SM315) and cobalt (II/III) shuttle currently holds the record PCE of about 13% for liquid-electrolyte-based DSSCs,[44] which is still far from the theoretical limit of 32% calculated by Shockley and Queisser.[19] The dye in the device has been widely conceived as one of the most important components for maximizing the PCE of the device, and large amount of research effort[45] has been devoted to the discovery and/or the creation of the best dye. An ideal dye should demonstrate excellent electronic and optical properties, such as wide absorption spectrum, long excited state lifetime and appropriate energy level alignment relative to TiO_2. One of the earliest and renowned examples that exhibits these properties is the Ru(II)

complexes,[45] such as dye N3[46] and dye N719,[47] which were the model dyes 'to-beat' for two decades. Another popular class of dye is the organic dyes, which typically exhibit higher molar extinction coefficients (50,000 to 200,000 $M^{-1}cm^{-1}$)[48] than the Ru(II) complexes, and can be more environmentally-friendly by allowing metal-free designs.[49] Many organic dyes are designed in accordance with the donor-π-acceptor (D-π-A) scheme,[45,48,49] which allows flexible structural modification of the dyes for attaining more desirable dyes' properties. Such a design scheme, however, does not guarantee high PCE and we will discuss in Chapter 5. Apart from the D-π-A scheme, there is no widely adopted design rules for organic dyes, despite the large amount of dyes created. The synthesis of dyes can be both laborious and time-consuming, and quick estimation of resulting PCE using certain hypothetical new dye should facilitate the development of dyes, reducing wasteful attempts. A tool designed for this purpose is demonstrated in Chapter 4.

The n-type device utilizes only a single photo-active electrode, but in theory the counter-electrode (Fig.1.1) can also be photo-active by employing a p-type semiconductor, harvesting more solar light and so resulting in PCE[50] higher than an n-type device. A tandem DSSC with two photo-active electrodes has been shown to exhibit higher photo-voltage but lower PCE than a comparable n-type device, where one of the attributions to the difference observed in PCE is that the dyes at different electrodes were competing for photons.[51] Apart from harmonizing the spectral response of the two dyes, the PCE of a tandem cell can also be improved by more efficient photo-current generation, which will be limited by the weaker electrode when the photo-electrodes are connected in series. The current record PCE of a p-type device is about 10-fold[52] less than that of the n-type, and therefore substantially limits the ability of a tandem cell to

generate high photo-current.
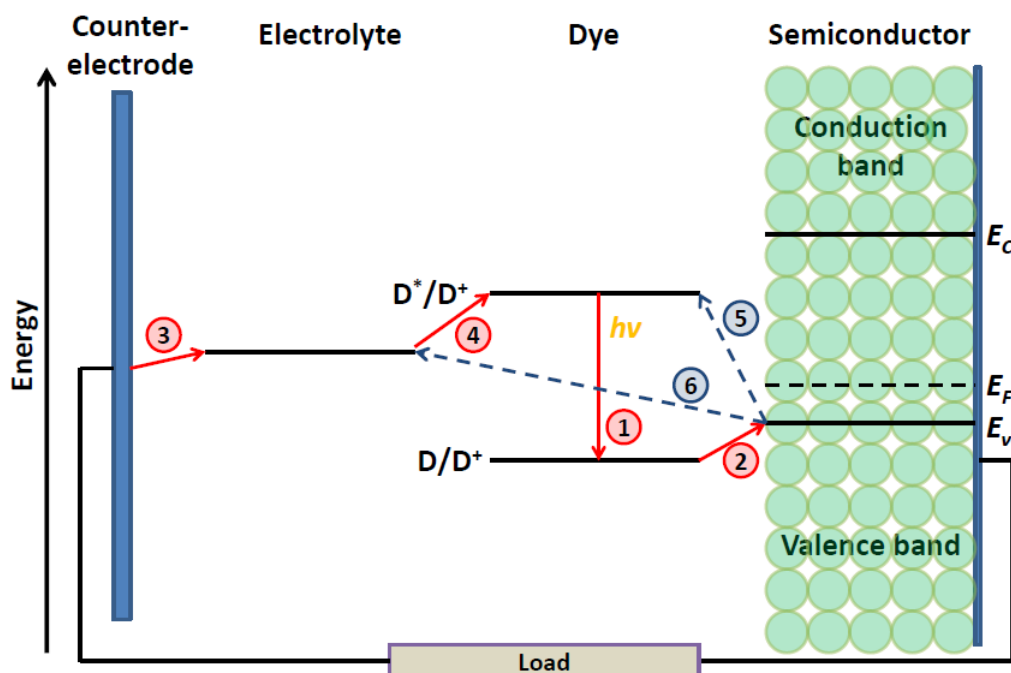
Figure 1.2 Illustration of the operation principles of p-type DSSCs. The ideal hole transfers (red arrows) start from 1.) creation of photo-excited hole, followed by 2.) hole injection, 3.) oxidation of electrolyte and finally 4.) oxidation of reduced dye. The unwanted electron transfers (blue dashed arrows) include 5.) hole recombination between semiconductor and dye and 6.) between semiconductor and electrolyte.

The operation principles of a p-type DSSC (Fig. 1.2) are similar to those introduced for the n-type (Fig. 1.1). Ideally, photo-excitation would lead to hole injection from the dye to the semiconductor, which is typically NiO, and the dye is oxidized by the redox mediator thereafter. The hole can recombine with the electron from the LUMO of the dye and/or from the electrolyte. Notice that since the flow of the current is reversed, the appropriate electronic level alignments for efficient charge transfer at the semiconductor-dye interface are different from those in the n-type DSSCs. In this case the HOMO of the dye should be lower in energy than the valence band (VB) edge of the semiconductor, while the CB edge

of the semiconductor should be higher in energy than the LUMO of the dye. The hole injection and recombination at the semiconductor-dye interface are modeled in Chapter 3 and we will discuss possible strategies to alleviate recombination in p-type devices.

1.4.2 Experimental techniques for PCE and charge dynamics measurements in DSSCs

In DSSCs, the PCE (η) is the 'figure of merit' and it is related to the short-circuit current density (J_{sc}), open-circuit voltage (V_{oc}), the fill factor (FF) and the power of incident radiation (P_{in}) by:

$$\eta = \frac{J_{sc} V_{oc} FF}{P_{in}}$$
(Eq. 1.1)

The J_{sc} is the maximum current density achievable by a device at zero voltage, and it is related to the electronic and optical properties of a dye. The V_{oc} is the maximum voltage at zero current, given by the energy difference between the quasi-Fermi level and the redox potential of the electrolyte. The FF is the ratio of maximum attainable power to the product of J_{sc} and V_{oc}, and P_{in} is the power of the incident light.

The measurement of η is typically performed under 1 Sun (1000 Wm^{-2}) irradiation at air mass (AM) of 1.5G, where AM, or the optical path length of sunlight through the atmosphere, of 1.0 is equivalent to the path length from the Zenith to sea level. A major source of error in estimating η is the determination of the active area, a quantity required when converting the measured short-circuit current in unit of A, to J_{sc} in unit of Am^{-2}.[53] When the best experimental practice for this conversion is not used, one should expect higher η of the device as a result.[53]

In many cases the timescales of charge transfer dynamics in DSSCs are in the range from nano- (ns) to femto-seconds (fs), and a suitable technique to detect these processes is transient absorption spectroscopy (TAS).[54–56] This technique monitors the variation in absorbance and/or transmission of the sample as a function of time by applying a 'probe' light after flashing UV or visible 'pump' light to excite the sample. Species can be identified by relating the absorbance to the molar extinction coefficients in accordance with the Beer-Lambert Law. The time constants can be obtained, for instance, by characterizing the transient absorbance curves with optimized Gaussian curves.[56] An important source of error would simply be that the timescales of the charge transfer processes do not fall within the instrument and/or method response function, such as when charge injection is in the range of less than 100 fs.[57] The 'probe' light should also have a fluence (number of incident photons per unit area) much lower than that of the 'pump' light to avoid influencing the population of the excited states when measuring the absorbance.[58]

Electrochemical impedance spectroscopy (EIS) is also adequate for characterizing charge transfer processes in DSSCs.[59,60] This method involves applying a small AC potential to the device, and recording the resultant current response due to this perturbation as a function of frequency. The frequency-dependent impedance can be expressed in Euler's form and visualized as a Nyquist plot, or with a Bode phase plot where the frequency can be read directly. Typically two or three semicircles would be observed in measurements of DSSCs under illumination, and the time constant of charge injection and recombination is characterized by intermediate frequency range of 10 to 100 Hz.[60–62]

1.4.3 Theoretical studies for charge dynamics in DSSCs

While experiments are capable of measuring fast electron transport time and short electron lifetime in femto-seconds in DSSCs,[45] various phenomena can occur concurrently at multiple timescales,[45,63] which can create difficulty for analyzing results. The inability of experiments to disentangle processes at different timescales implies difficulty to understand the effect of the condensed-phase environment on a particular process. Kinetic modeling has accompanied experimental measurements since the early developing stage of DSSCs for studying the electrical characteristics of the device.[45] Another important role of theoretical modeling is to establish correlations between the electronic structure of materials and the elementary processes in the device, which complements the understanding of elementary mechanisms based on experimental studies, and aids the design of materials by predicting materials' properties.[39] Although the complexity of DSSCs cannot be fully captured in current theoretical models, agreements are commonly seen between theoretical and experimental charge transfer times.[64,65] Current theoretical models are therefore deemed capable of providing reliable understanding of charge dynamics in DSSCs.[39]

Models for electron injection at the semiconductor-dye interface in n-type DSSCs are particularly abundant.[39,66–70] One of the common computational approaches for modeling injection was to employ a model Hamiltonian parameterized from first-principle electronic structure calculations and/or experimental references.[71] Thoss et. al.,[64,72,73] for instance, studied injection in a number of semiconductor-dye systems based on this approach and combined with multi-configuration time-dependent Hartree method[74,75] for acquiring a quantum dynamical description of the process. In their studies the injection

process was characterized as ultrafast and agreed with experimental measurements. They had also revealed the influence of the electron-vibration coupling on the nuclear degree of freedom at different frequency modes of the dye, demonstrating an example of the quantum Zeno effect.[67,76] May et.al.[77–79] also adopted the model Hamiltonian approach, using parameters obtained from experimental linear absorption spectra of some semiconductor-dye systems, and computed the injection rate with a Fermi-Golden rule type expression. They showed that electronic coupling can be altered by changing the bridge-anchoring group of the dye, and generalized the Fano effect with their models. Deriving an injection rate expression from a model Hamiltonian within the Green's function framework, Jones et. al.[80] and Martsinovich et. al.[65] had employed this rate expression for comparing injection rates of a set of dyes. Such approach will be further discussed in Chapter 2.

Another type of method that has been commonly used to model electron transfer processes in DSSCs is the non-adiabatic molecular dynamics (NAMD) method. NAMD can provide detailed descriptions of the system evolution but restricted by high computational cost for slow processes, such as those slower than ~10 ps.[39] Prezhdo et. al.,[63,67,81–84] for instance, simulated electron dynamics at the semiconductor-dye interface with a few different approaches, but in general based on NAMD with approximations such as the quantum-classical mean-field approach. In one of the studies[81] they showed that the non-adiabatic transfer pathway for photo-induced electron injection was always present after photo-excitation and dominated the short transfer times, whereas the adiabatic pathway was usually slow. Batista et. al.[85,86] employed *ab initio* DFT molecular dynamics, in combination with quantum dynamics calculations, for characterizing electron transfer at the semiconductor-dye interface and the

influence of thermal fluctuations on these processes. They showed ultrafast interfacial electron transfer was sensitive to the symmetry of the initial electronic state in the dyes, and the thermal nuclear fluctuations would elevate the rate of injection.

In contrast to electron injection, fewer attempts[87,88] have been made to model charge recombination in isolation. A possible reason is that the accuracy of the computed charge recombination rate is known to be highly sensitive to the relative energy levels alignment at the semiconductor-dye interface.[39] Standard first-principle computational methods such as DFT may not be capable of satisfying such requirement, and large uncertainty in the computed charge recombination rates could be introduced as such. Maggio et.al.[89] had, for instance, modeled charge recombination based on non-adiabatic charge transfer theory at semiconductor-dye[87] and semiconductor-electrolyte interfaces,[88] and reduced the errors of the energy levels alignment at the interface by introducing experimental parameters (also see Chapter 2). A phenomenological model[90] based on Marcus theory of electron transfer for describing charge recombination was developed in parallel with experimental measurements, which had accounted for both recombination from electronic states in the conduction band, as well as from surface states. With this model it had been shown that the Marcus 'inverted' region can be exploited by employing a redox shuttle with a strongly positive potential to slow recombination from the conduction band, but would at the same time accelerate recombination from the surface states.

The introduced theoretical and computational methods have all been developed consistently for charge dynamics in n-type DSSCs. On the other hand, there has been a lack of theoretical efforts towards understanding the charge dynamics at the semiconductor-dye interface for p-type devices, the weaker

component in a tandem cell. A possible reason is that less research effort has been devoted to p-type devices than for n-type devices. However, both experimental charge transfer time constants and well-defined theoretical framework for transfer rate constants were available, and it was possible to study theoretically the interfacial charge dynamics for p-type devices. Such type of study is given in Chapter 3.

1.5 An introduction to photo-catalysis for fuel generation

1.5.1 Background

Photo-catalysis presents a clean and environmentally-friendly opportunity for alleviating future and global energy demand in terms of chemical energy storage. The main objective for photo-catalysis in the context of fuel generation is to perform efficient photo-driven solar fuels synthesis. Two prominent examples are photo-catalytic water splitting and photo-catalytic reduction of CO_2, which can generate high-specific-energy hydrogen fuel (~129 - 152 MJ/kg)[91] and methane (~55 - 61 MJ/kg)[91] respectively.

The ideal operation principle for photo-catalysis (Fig. 1.3) is, as a first step, the creation of an electron-hole pair in the bulk due to photo-excitation of the catalyst. Following charge separation is the migration of electron and hole towards the surface of the catalyst, and subsequently the electron and hole undergo interfacial charge transfer with surface adsorbates. Holes and electrons, however, can also recombine at the surface and/or in the bulk, especially when these particles are trapped at surface defect sites, instead of reacting with surface species as an outcome. Recombination at the surface are generally serious,[27,92,93] hence in practice electron or hole scavengers, such as molecular oxygen and

simple amines and alcohols respectively,[94] are incorporated in the system to alleviate these unwanted processes. The interfacial redox reactions are also dependent on the relative electronic energy level alignment, which can be sensitive to a number of factors such as solvent environment and surface morphology.

Semiconductor Nano-particle
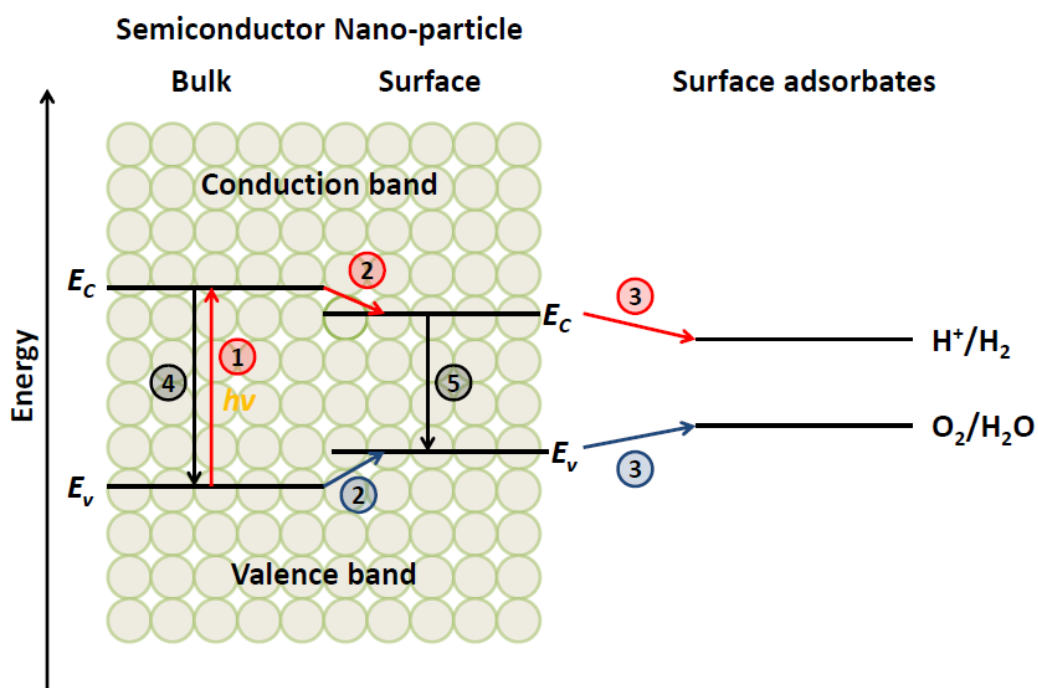
Figure 1.3 Illustration of the operational principles of photo-catalysis, using oxidation (H^+/H_2) and reduction (O_2/H_2O) of water as an example. The ideal electron (red arrows) and hole (blue arrows) transfers start from 1.) photo-excitation leads to charge separation, 2.) excited electrons and holes migrate to the surface, and 3.) undergo redox reactions with surface adsorbates. Charge recombination can occur 4.) in the bulk and 5.) on the surface of semiconductor nano-particles.

Photo-catalytic water splitting has been one of the main research focuses in photo-catalysis, due not only to the generation of hydrogen fuel but also to the frequent presence of water in other photo-catalytic reactions.[27,95] In a photo-electrochemical water splitting cell, the splitting of water to produce

molecular oxygen and hydrogen requires an energy supply of 1.23 eV. The overall reaction consists of two half-reactions, the hydrogen evolution reaction (HER):

$$2H^+ + 2e^- \rightarrow H_2 \qquad \text{(Eq. 1.2)}$$

and the oxygen evolution reaction (OER):

$$2H_2O + 4h^+ \rightarrow O_2 + 4H^+ \qquad \text{(Eq. 1.3)}$$
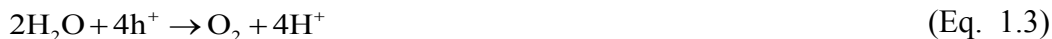
The HER is generally inefficient on a conventional TiO_2 catalyst, due to the considerably positive CB edge of TiO_2, and weak absorption of solar light in the visible light spectrum. Attempted solutions include anionic doping with N and S atoms,[96] attaching dyes,[22,97] or replacing with p-type semiconductors.[98] All these modifications are not ideal, with either limited improvements seen or other problems arise, such as dye regeneration in dye-sensitized systems[22] and sacrificing reactive photo-holes for OER when p-type semiconductors are employed.[98] The kinetics of the HER is also unfavorable on a bare semiconductor; hence co-catalysts such as Pt metal are frequently incorporated.[98,99]

A large over-potential is typically associated with OER on TiO_2, leading to the requirement of an energy supply of greater than 1.23 eV in practice to drive the reaction. The OER reaction mechanism is still under debate[67,94,100–102] but it is generally agreed that the surface-trapped photo-holes are essential and have sufficient oxidizing power to produce oxygen via four sequential hole transfer steps. Without a clearly elucidated reaction mechanism the physical origin of the OER over-potential has therefore remained implicit. A number of dopants have, however, already been tested for promoting OER but without significant improvements.[96,101] Co-catalysts such as Mn_3O_4, IrO_2 and RuO_2 often only lead to moderate improvement of water-splitting rate by a factor of 1.4 at most when

co-loaded with co-catalysts for HER.[103] The more recently designed Z-scheme hybrid configured catalytic systems may present an opportunity for simultaneously promoting HER and OER, where now the two half-reactions are performed upon two separate semiconductor nano-particles.[103–105] Yet, current efficient configurations for photo-catalytic water splitting[106] such as NaTaO$_3$:La(2%) under UV-spectrum (with hydrogen production rate of 19.8 mmolg^{-1}h^{-1}),[107] are far from comparable with the efficiency of hydrogen generation with steam reform of natural gas, the process that produces the majority of global hydrogen fuel.[108,109]

In photo-catalytic reduction of CO$_2$ in water, the transformation to methane is a reduction half-reaction that requires 8 protons and 8 electrons, or 8 hydrogen radicals, to complete:[13]

$$CO_2 + 8H^\bullet \rightarrow CH_4 + 2H_2O \qquad \text{(Eq. 1.4)}$$

The 8 holes generated at the VB edge of the semiconductor, e.g. TiO$_2$, should also react with surface species which would otherwise cause unwanted charge recombination[27,93] and hinder reaction efficiency. In theory the holes can be consumed in an OER (Eq. 1.3), where the protons generated can be reduced for acquiring hydrogen radicals. The total ideal reaction would be:

$$CO_2 + 2H_2O \rightarrow CH_4 + 2O_2 \qquad \text{(Eq. 1.5)}$$
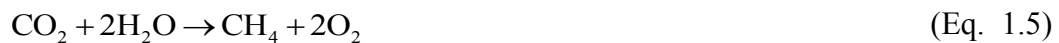
To study the photo-reduction of CO$_2$ in isolation, additional hole scavengers such as amines and alcohols are typically employed to remove surface holes.[13] The generation of oxygen (Eq. 1.3) is not common due to difficult four-hole chemistry.[93]

There is currently no photo-catalyst that allows this conversion to be efficient for industrial practice. The formation rate of CH$_4$, one of the common measures for the efficiency, is extremely low, typically less than tens of μmol g^{-1} h^{-1}.[13,110]

This is lower than the aforementioned hydrogen production rates from efficient photo-catalytic water splitting. There is a growing interest in developing new catalysts for improving efficiency, with attempts already made that are not limited to conventional concepts such as doping, adding metal co-catalysts and replacing with other transition metal oxides,[13,110] but also with more innovative organic approaches such as enzyme mimicking.[111] There is, however, no widely accepted reaction mechanism of photo-catalytic reduction of CO_2 to CH_4 and the development of catalytic systems are largely based on trial-and-error.

The main research focus in both water splitting and photo-reduction of CO_2 is evidently the development of catalytic systems with improved efficiency in terms of light absorption, interfacial charge transfer and catalyzing the reaction. Trial-and-error has been the main approach for catalyst development[112–114] but this approach has been ineffective, where catalytic systems with sufficient efficiency for commercialization have yet to be identified after almost 40 years of research. Systematic improvement of catalysts, on the other hand, should require at least (i) a widely adopted metric for comparing the performance of catalysts; (ii) knowledge of the reaction mechanism, which should reveal the rate-limiting step and provide possible hints for controlling product selectivity; and (iii) some design rules for catalysts, where only rather few of them[114–117] have been proposed thus far. These requirements are currently far from being completely fulfilled, and can be time-consuming to achieve. However, successful fulfillment of these requirements should not only accelerate the discovery and/or the invention of efficient catalytic systems for water splitting and photo-reduction of CO_2, but also serve as important examples for other photo-catalytic reactions.

1.5.2 Experimental techniques for identifying reaction intermediates in photo-catalysis

Electron paramagnetic resonance (EPR) spectroscopy and infra-red (IR) spectroscopy are two of the most common techniques to identify reaction intermediates in photo-catalysis, among a few others.[118] Both techniques have been employed in the experimental studies to identify intermediates in photo-catalytic reduction of CO_2.[119–122] EPR spectroscopy is designed for the detection of open-shell intermediates, which are likely to evolve upon interfacial charge transfer and hydrogen radical transfer on the surface. In the presence of an external magnetic field, unpaired electron of an intermediate would have spin quantum number (m_s) of either ½ or −½, each of them with a specific energy that can be computed in accordance with the Zeeman equation. The energy difference between the two spin states is the absorption energy that is typically recorded as first derivative in EPR spectrum at a fixed microwave frequency. Intermediates can be identified by analyzing the g-factors and/or the hyperfine structures, in which the difference of these factors between intermediates originates from the different couplings between the electron spins and local magnetic fields, and between electron spins and the nucleus spins respectively. EPR experiments are typically performed at 77K[120] for the purpose of avoiding low sensitivity due to high population of the excited electron spin states. The technique is incapable of detecting closed-shell species,[119] hence should be employed together with techniques such as gas chromatography to complement this weakness.

IR spectroscopy represents a number of techniques that reveal the structure of an IR-active intermediate upon absorption at specific frequencies that correspond to the motion of normal vibrational modes. These techniques vary from one another in terms of, for instance, experimental setup and/or data manipulation.

Diffuse reflectance infrared Fourier transform spectroscopy (DRIFTS),[123] for example, allows measurements with powder samples and hence records diffuse reflection patterns alongside absorbance for characterizing materials. Polarization modulation infrared reflection absorption spectroscopy (PM-IRAS),[118] on the other hand, isolates the contribution of the surface-bound species to the spectra by taking the difference between spectra obtained with *s*-polarized and *p*-polarized light, in which gas phase species are sensitive to both types of light but surface species are much more sensitive to the *p*- than *s*-polarized light.

Often the ultimate objective of identifying reaction intermediates is to aid the understanding of a molecular reaction mechanism, but the introduced methods themselves are not capable of informing the role of the identified intermediates in the reacting systems. Major sources of errors in proposed reaction mechanisms are therefore the inaccurate chemical intuition from the researcher, and/or the incorrect assignments of spectral features. Kinetic modeling,[124] isotope labeling experiments,[125] and computed potential energy surfaces,[40] are all complementary tools that have shown to add confidence to hypothetical mechanisms in the past. We will demonstrate in Chapter 6 how potential energy surfaces computed with DFT can aid the discrimination of competing mechanisms of photo-catalytic reactions.

1.5.3 Theoretical studies for reaction mechanisms in photo-catalysis

As introduced in section 1.2 and section 1.5.1, the research field of photo-catalysis is largely focusing on the discovery of a photo-catalytic system that can efficiently (i) absorb solar light, (ii) deliver electron and holes to the surface for redox reactions and (iii) catalyze the reaction. In terms of improving

criteria (i) and (ii), theoretical studies have taken chiefly the role of revealing the electronic structures, the optical properties and the relative electronic energy level alignment with surface adsorbates of photo-catalytic materials. The knowledge of these properties would be useful for informing the experimentalists about the appropriate tuning of the electronic structure of a given photo-catalyst, such as the appropriate energy range for electronic states to be introduced within the band gap by doping.[126]

Another important role of theoretical studies is to reveal the thermodynamics and kinetics of the surface molecular reaction, as well as the adsorption geometries of reaction intermediates, in order to complement the understanding of the reaction mechanism proposed based on experiments. The understanding of reaction mechanism allows identification of the rate-limiting step, and further analysis to understand the origin of the kinetic barrier of this step should provide valuable insights into modifying catalysts for the improvement of criterion (iii). The focus of this section will be placed on the role of theoretical studies to understand reaction mechanisms in photo-catalysis.

A relatively large amount of theoretical work has been devoted to study photo-catalytic water splitting,[67,126] and we will therefore use studies on the reaction mechanism of OER to demonstrate the role played by theoretical studies. Valdes et.al.,[127] for example, had employed the computational hydrogen electrode (CHE) method[128] and analyzed a few possible reaction mechanisms on rutile(110) with different surface terminations. They showed that the surface fully covered with oxygen was the relevant surface for reaction, and the rate-limiting step was a deprotonation step:

$$H_2O + S \rightarrow OH - S + H^+ + e^- \qquad \text{(Eq. 1.6)}$$

where S was a coordinatively unsaturated site on the surface. They had also pointed out that water oxidation could be light-driven on ideal TiO_2 rutile(110) surface without external bias. Li et.al.[101] had determined the microscopic mechanisms for OER on 3 different anatase surfaces. They showed that visible light provided sufficient energy to drive OER kinetically, that the first deprotonation was the rate-limiting step, and that the reaction was not sensitive to the surfaces investigated. They also discussed the physical origin of the over-potential by analyzing the electronic structures of rutile, anatase and co-doped anatase with (Mo, C) and (Nb, N). The origin of the over-potential was attributed to the instability of the surface-adsorbed OH state in the rate-limiting step, and introducing occupied states higher in energy than the VB states in TiO_2 by doping could enhance OER activity. Nguyen et. al.[129] and Hellman et. al.[130] had studied OER on α-Hematite(0001) with different surface terminations by adopting the CHE method. Both studies had agreed that one of the oxygen-terminated surfaces was the most stable surface termination under photo-illumination. Reaction mechanisms had been proposed for this surface accordingly. Hellman et.al.[130] suggested that the formation of O_2 could be considered as the rate-limiting step, while Nguyen et.al.[129] had suggested the formation of surface O from the oxidation of surface OH was the most energetically demanding step. The main objective of these studies was to identify the surface termination responsible for OER activity, which was important for establishing the reaction mechanism on α-Hematite(0001). Although further analysis of possible rate-limiting steps had not been given in these studies, their findings and proposed reaction mechanisms had offered suggestions for further experimental studies.

Theoretical studies had also contributed to the understanding of the

mechanism of elementary electron and proton transfer reactions in OER, such as sequential or concerted transfer mechanisms. Chen et. al.,[102] for instance, studied the chemical dynamics of the first proton and electron transfer step of the OER with DFT-based first principle molecular dynamics, and showed that the mechanism of this step was a proton transfer (PT) step followed by an electron transfer (ET) step. They explained the influence of pH on the rate of OER, in which PT was a rate-limiting step at low pH (pH < point of zero charge (pzc)) while at high pH the rate-limiting step was the ET. The barrier for ET at high pH was much lower than the barrier of PT at low pH and therefore the rate of OER was faster at high pH, where the surface was covered by hydroxide. Cheng et. al.[131] studied the elementary mechanisms of four assumed elementary steps of OER by computing the reaction energies of these steps. They showed that the mechanisms should be sequential ET and PT for all four steps, and the second step had the highest reaction energy. They also separated the protonic and electronic component of the thermodynamic over-potential and showed that this over-potential was attributed mainly to the scattered alignment of the electronic levels, not to the component from deprotonation. The devised representation of the electronic and protonic component for the reaction could also be useful to understand a catalyst's activity or inactivity.

Although a substantial amount of theoretical studies have been devoted to study OER, the reaction mechanism of OER currently remains debatable (see also section 1.5.1). Some details of the resulting reaction mechanism from different theoretical studies are often in disagreement. For instance, Liu et. al.[101] have not included adsorbed OOH in their identified reaction mechanism, while OOH is a common intermediate in other theoretical studies.[127,131,132] The discrepancy can be caused by the different catalysts' surfaces, computational

methods and experimental evidences employed in these studies. On the other hand, an important agreement between different theoretical studies[101,127,132] is that, the identified rate-limiting steps have appeared to be similar, which is the first deprotonation step, similar to (Eq. 1.6). Hence, the analysis of the origin of the over-potential associated with this step by Liu et. al., and the explanation of the effect of pH on the elementary mechanism of this step by Chen et. al., are seemingly important understanding for new design of photo-catalyst to achieve higher OER activity.

The purpose of understanding the OER mechanism is ultimately to facilitate water-splitting. Hitherto theoretical studies have been able to provide some understanding on a single process in isolation, but the reaction system of water splitting involves various processes occurring simultaneously, including the HER, and charge recombination on the surface of the catalyst.[67] Current theoretical models are incapable of examining the mutual influence of various important concurrent processes involved in water splitting, and the development of multi-scale models[133,134] to provide such detailed picture remains a challenge. Another challenge will be the identification of suitable materials. There is currently a large number of catalytic materials available,[106] and more are expected to be synthesized and studied. It is therefore important to develop a method for identifying materials that may match the requirement of an efficient photo-catalyst. In this respect, it is perhaps useful to continue the development of computational models designed previously for screening catalysts, such as the volcano plots.[67,98,116]

While there is an accumulating amount of theoretical studies on photo-catalytic water-splitting, there are comparatively much fewer studies on the photo-catalytic reduction of CO_2. These theoretical studies either focused on

the reduction of CO_2 to form products other than CH_4, or considered a subset of mechanistic hypotheses proposed for the conversion to CH_4. Indrakanti et.al.,[135] for instance, studied the adsorption of CO_2 on small clusters using both post-Hartree-Fock and DFT methods, and they showed that transferring an excited electron from a stoichiometric TiO_2 surface to CO_2 was energetically unfavourable, but the charge transfer might be more favourable when oxygen vacancies were present. He et. al.[136] studied 2-electron reduction of CO_2 on anatase(101), using a periodic slab model with the GGA+U scheme, and they identified competitive pathways to form HCOOH and CO. In the same study a simple model was devised to screen a large number of dopants for lowering the reaction barriers. Ji and Luo[137] studied a proposed reaction pathway on the anatase(101) surface, and subsequently proposed a new pathway.

Photo-catalytic reduction of CO_2 to gain CH_4 (Eq. 1.4) is a complicated reaction that many intermediates can possibly form in the course of the reaction, which has led to several proposed reaction mechanisms.[13] For TiO_2, three reaction mechanisms proposed based on experimentally identified intermediates have been summarized in the literature,[13] but only one of the proposed reaction mechanisms has been studied theoretically.[137] Without comparing all three proposed reaction mechanisms, it is not possible to rule out mechanistic hypotheses with supporting evidence given on a common basis, such as thermodynamics, and identify the most likely mechanism and/or the reaction conditions that favor a particular mechanism. In chapter 6, we have therefore considered all three proposed reaction mechanisms on defect-free TiO_2 anatase(101) and compared their thermodynamics, with the objective to identify the most favorable mechanism. The method employed in the study is general for testing mechanistic hypotheses for reactions occur on the surface of solid

catalysts.

1.6 Thesis Outline

In Chapter 2 we introduce the theories and methods employed in this thesis in a general context. In Chapter 3 we demonstrate how we modeled the charge injection and recombination in p-type DSSCs with non-adiabatic charge transfer theories and suggest possible modification for PCE improvement. In Chapter 4 we present a statistical model for predicting the PCE of DSSC based on quantum-chemically computable properties of dyes. In Chapter 5 we present our work on attempting to answer the question of whether increasing the strength of D-π-A character would improve PCE of n-type DSSC, based upon statistical grounds. In Chapter 6, we show the thermodynamic landscapes of three proposed reaction mechanisms of the conversion of CO_2 to CH_4 on TiO_2 anatase(101), and discuss the most favorable pathway for this reaction. We conclude this thesis in Chapter 7.

Chapter 2   Theory and methodology

The theories and methodologies to be introduced in this chapter include some standard computational theories in quantum chemical calculations, the charge transfer theories for modeling charge dynamics in DSSCs (Chapter 3), and the general methodologies for statistical modeling with regression (Chapter 4).

2.1 Standard computational methods

2.1.1 Density functional theory

The majority of electronic structure calculations performed in this thesis, including the geometry optimization of dyes (Chapter 4 and 5) and relaxation of periodic systems (Chapter 3 and 6), were based on Density Functional Theory (DFT).[138–140] DFT is a variational ground-state theory in which the computation of the total energy functional ($E[\rho]$) using the Kohn-Sham equations[141] is given by:

$$E[\rho] = T[\rho] + V_{ext}[\rho] + V_H[\rho] + E_{xc}[\rho] \qquad \text{(Eq. 2.1)}$$

where $T[\rho]$ is the kinetic functional, $V_{ext}[\rho]$ is the external potential functional, $V_H[\rho]$ is the Hartree energy functional and $E_{xc}[\rho]$ is the exchange-correlation functional. Computed results may depend on the choice of $E_{xc}[\rho]$, which is the sum of the difference in kinetic energy and electron-electron interaction between real system and the fictitious system with non-interacting Fermions under the Kohn-Sham equations. DFT with functional-type such as Local Density approximation (LDA) and Generalized Gradient Approximation (GGA), possesses the crucial advantage of producing some predictions with acceptable accuracy but low computational cost (scale to $\sim N^{2\text{-}3}$, where N is the relative measure of system size) in comparison to alternatives.[142] Hence, often used in

combination with Bloch wave-functions, DFT is widely adopted in the simulations of large periodic systems.[143]

A challenge for DFT is the modeling of the electronic energy level alignment at the semiconductor (inorganic) –dye (organic) interface. Such kind of interface is common in photo-voltaic devices such as DSSCs, since organic materials can often provide tunable and desirable optical properties with rather poor charge carrier mobility, while it is *vice versa* for inorganic materials. The combination of the two types of material can therefore complement the weaknesses of one another.[37] The efficient charge separation at the interface, however, is still very much depending on the electronic energy level alignment. Standard DFT/GGA is incapable of describing simultaneously the electronic properties of both inorganic and organic parts with sufficient accuracy.[37] The method is notorious for over-delocalizing wave-functions in organic materials,[37,140] and generally wrongly predicts the band gap of a semiconductor (typically underestimation is ~30% to 100% for solids),[144,145] as well as the HOMO-LUMO gap of the organic material (typical error for non-hybrid functional is ~0.73 eV).[146] These weaknesses of DFT can be problematic for the computation of charge transfer rates at these types of interfaces, especially when the rate has an exponential dependence on the energy gap between the initial and final states of the reaction. There are a number of alternatives which are capable of providing a comparatively accurate alignment, such as the *GW* approach[37,147,148] and time-dependent density functional theory (TDDFT),[149,150] but these methods typically require much higher computational cost. It is also possible to compute the interfacial charge transfer rates with sufficient accuracy based on DFT/GGA alignment; for instance, if the errors introduced by DFT in the band gap of the semiconductor and the HOMO-LUMO gap of the organic material cancel out, an

acceptable alignment may be acquired on a relative scale;[65] and/or when experimental parameters are introduced to the DFT/GGA alignment to correct for the error.[87]

For strongly-correlated materials, such as NiO, standard DFT cannot describe the band structure with satisfaction and special treatment of the strong on-site Coulomb interactions in total energy calculations is required. For instance,

$$E_{DFT} = E_{DFT}[\{\varepsilon_i\}] + \frac{(U-J)}{2} \sum_{l,j,\sigma} \rho_{lj}^{\sigma} \rho_{jl}^{\sigma} \qquad \text{(Eq. 2.2)}$$

where $\{\varepsilon_i\}$ is the set of Kohn-Sham eigenvalues, U and J are the spherically averaged matrix elements of the screened Coulomb electron-electron interactions, and ρ_{lj}^{σ} is the density matrix of d-electrons with spin σ. The development of Eq. 2.2 leads to the DFT+U[151–153] computational scheme, a widely applied method for treating strongly-correlated materials, among other alternatives[154] such as dynamical mean field theory[155] and many-body perturbation theory based on Green's function.[148] A limitation of DFT+U is that the parameter U is normally chosen rather arbitrarily by, for instance, fitting properties of interest against experimental results, instead of based on physical understanding.[154] Other limitations of DFT include systematic underestimation of reaction barriers[140] and inability of describing charge-localized states, in which the latter may be dealt with by using constrained DFT (cDFT).[156]

2.1.2 Other standard methods

Time-dependent density functional theory (TDDFT)[150] was employed for providing excitation energies and oscillator strength for simulating absorption spectrum of dyes in Chapter 4. This method has been commonly applied in the field of DSSC for understanding dyes' optical properties,[39,157] as well as the

relative energy level alignment between the semiconductor and the excited dye.[149] TDDFT is the time-dependent analogue of DFT, providing the energy of the system by solving the time-dependent Kohn-Sham equation.

The climbing-image nudged elastic band (CI-NEB)[158] was employed for the search of transition states of reactions in Chapter 6. The NEB[159–161] is a chain-of-state elastic-band type method for the acquisition of a minimum energy path, while the CI is implemented for the purpose of improving the accuracy of the location of the transition state. The NEB optimizes the intermediate images along a path by moving images in accordance with the total force acting on images, which is the sum of the spring force along the local tangent ($\mathbf{F}_i^s\big|_{\parallel}$) and the true force perpendicular to the local tangent ($\nabla E(\mathbf{R}_i)\big|_{\perp}$):

$$\mathbf{F}_i = \mathbf{F}_i^s\big|_{\parallel} - \nabla E(\mathbf{R}_i)\big|_{\perp} \qquad \text{(Eq. 2.3)}$$

where $\mathbf{R}_i$ denotes the coordinates of image i. The CI then identifies the image with the highest energy i_{max}, and the total force of this image, instead of computing with Eq. 2.3, is now given by the full force of the potential with inverted component along the band:

$$\mathbf{F}_{i_{max}} = -\nabla E(\mathbf{R}_{i_{max}}) + 2\nabla E(\mathbf{R}_{i_{max}})\big|_{\parallel} \qquad \text{(Eq. 2.4)}$$

2.2 Physical modeling of charge transfer at the semi-conductor/dye interface in DSSC

2.2.1 Theory of charge injection

The computation of hole injection rate is based on a theory of charge injection derived based on the Fermi-Golden rule formalism,[162] within the framework of Green's function. The theory stems from the News-Anderson model

Hamiltonian,[163,164] and was developed and applied for identifying dyes[65,80] and their possible anchoring groups[165] that promoted fast charge injection. Using charge injection in n-type DSSC as an example, we illustrate key points of the theory of charge injection employed to study hole injection in Chapter 3. A detailed account of the theory is given in ref. [166].

The charge injection process in n-type DSSC at the semiconductor-dye interface is the transfer of electron from a localized state $|s\rangle$ in the dye to the manifold of states $\{|l\rangle\}$ in the semiconductor (Fig 1.1), which ideally would be the states in the conduction band. The time-dependent probability of finding a system at state $|s\rangle$ ($P_s(t)$) can be given by:

$$P_s(t) = |G_{ss}(t)|^2 \tag{Eq. 2.5}$$

where $G_{ss}(t)$ is the forward time propagator of state $|s\rangle$, and can be written in terms of Fourier transform of retarded Green's function matrix elements. The (causal) Green's operator is a common way of expressing the time evolution of a system.[167] Since the retarded Green's function operator can be related to a system Hamiltonian by:

$$G(E) = \frac{1}{E - H + i\varepsilon} \tag{Eq. 2.6}$$

where ε is a positive infinitesimal quantity, $G_{ss}(E)$ can therefore be evaluated for the Newns-Anderson model Hamiltonian[163,164] by writing:

$$G_{ss}(E) = \frac{1}{E - E_s - \sum_{ss}(E)} \tag{Eq. 2.7}$$

where E_s is the energy of state $|s\rangle$ and $\sum_{ss}(E)$ is the self-energy of state $|s\rangle$, which consists of a real part and an imaginary part:

$$\sum_{ss}(E) = \Delta_{ss}(E) - \frac{i}{2}\hbar\Gamma_{ss}(E) \tag{Eq. 2.8}$$

The imaginary part of the self-energy is related to the charge injection rate (Γ_{ss}) in a golden-rule type relation:

$$\hbar\Gamma_{ss}(E) = 2\pi\sum_{l}V_{sl}V_{sl}^{*}\delta(E-E_{l})$$ (Eq. 2.9)

where V_{sl} is the coupling between states $|s\rangle$ and $|l\rangle$, and $\delta(E-E_l)$ is the Dirac delta function, where the sum of this function over all states $|l\rangle$ is the density of states of the semiconductor.

A crucial property for efficient injection is the appropriate relative energy alignment of the state $|s\rangle$ in the dye to the state $|l\rangle$ in the semiconductor (Fig. 1.1). Errors can therefore be introduced by the inaccuracies in the computation of the energies of both states $|s\rangle$ and $|l\rangle$. For instance, ground-state DFT/GGA is known to perform weakly in the computation of energies of virtual orbitals,[65] which can result in incorrect relative energy alignment as introduced in section 2.1.1. The coupling between the dye and the semiconductor is also important for determining the injection efficiency.[165] Many dyes have been designed with LUMOs localized on the anchoring group of the dye, in order to attain proximity between the LUMOs and surface electronic states for maximizing the coupling. This coupling, however, is also subject to other variables such as adsorption geometries, which are typically determined with computational studies.

Eq. 2.9 can be conveniently applied for screening dyes with the same anchoring group for fast injection, when combined with a practical computational partitioning scheme.[65] In this case Eq. 2.9 is rewritten as:

$$\hbar\Gamma_{ss}(E) = \sum_{m,n}\Gamma_{mn}(E)c_{m}c_{n}^{*}$$ (Eq. 2.10)

where c_{m} are the molecular orbital coefficients, and:

$$\Gamma_{mn}(E) = \frac{2\pi}{\hbar}\sum_{k,k'}V_{mk}V_{nk'}^{*}\rho_{kk'}(E)$$ (Eq. 2.11)

where V_{mk} is the coupling between the anchoring group and the substrate, when the initial state $|s\rangle$ and final state $|l\rangle$ are expressed as linear combination of localized basis functions $\{\chi_m\}$ and $\{\varphi_k\}$ respectively, and $\rho_{kk'}(E)$ is the energy-dependent local density of states of the substrate. Generally the most computationally expensive calculations will be those involving the substrate, but in this case $\Gamma_{mn}(E)$ can be reused for different dyes with the same anchoring groups, where the coupling V_{mk} can be reasonably assumed to be the same. The screening can therefore be performed without repeating the calculations involving the substrate for different dyes, and requires only geometry optimization of dyes to attain c_m for Eq. 2.10.

2.2.2 Theory of charge recombination

The computation of hole recombination in Chapter 3 is based on the theory of charge recombination derived from the Fermi Golden rule formalism and Marcus theory for electron transfer, developed and applied to the electron recombination in n-type DSSCs by Maggio and Troisi.[87,88,168] In a n-type device the recombination process at the semiconductor-dye interface is an electron in the manifold of the semiconductor, $\{|l\rangle\}$, transfers to a localized state of the oxidized dye, $|i\rangle$ (Fig 1.1). A distinction between charge recombination to charge injection is the typical time frames of the two processes, where the former is slow (s to ns) and the latter is fast (fs). Such discrepancy in the duration of the processes suggests different importance of the nuclear motions; essentially the equilibrium nuclear configurations for the initial and final states in injection would be unchanged, while the relaxation of the nucleus is necessary to be part of the rate expression of charge recombination.

Charge recombination in n-type DSSCs is generally deemed as a non-adiabatic

charge transfer process, which has a small Landau-Zener transition probability at the crossing region of (at least) two electronic potential energy surfaces, or the coupling between the donor state and the acceptor localized state is weak. The rate expression of charge recombination in a n-type device, which follows the energy gap law,[169] can be written as:

$$k_n = \int_{cbm} \Gamma_{ii}(E) f(E - E_F) F(E, \Delta G_n, \lambda_n) dE \qquad \text{(Eq. 2.12)}$$

where cbm is the conduction band minimum of the semiconductor, $\Gamma_{ii}(E)$ is the transfer rate in the absence of nuclear relaxation, similar to Eq. 2.9 but recast as:

$$\Gamma_{ii}(E) = \frac{2\pi}{\hbar} \sum_l V_{il} V_{il}^* \delta(E - E_l) \qquad \text{(Eq. 2.13)}$$

since the recombination is between state $|i\rangle$ and $\{|l\rangle\}$. $f(E - E_F)$ is the Fermi-Dirac distribution, which provides the probability of an electron occupying a state with energy E in the manifold:

$$f(E - E_F) = \frac{1}{1 + e^{(E - E_F)(k_B T)^{-1}}} \qquad \text{(Eq. 2.14)}$$

where E_F is the quasi-Fermi level at non-equilibrium, such as when the device is under illumination, k_B is the Boltzmann constant and T is the temperature. $F(E, \Delta G_n, \lambda_n)$ is the thermally-average Franck-Condon term that can be expressed with classical nuclear modes as:

$$F(E, \Delta G_n, \lambda_n) = \frac{1}{\sqrt{4\pi\lambda_n k_B T}} \exp\left[\frac{(-E + \Delta G_n + \lambda_n)^2}{4\lambda_n k_B T}\right] \qquad \text{(Eq. 2.15)}$$

where E is the energy of the electron in the conduction band for recombination, ΔG_n is the free energy change of the oxidation of dye and λ_n is the reorganization energy of the same process. ΔG_n and λ_n will be further discussed in the following section. When the vibrations are treated quantum mechanically the Franck-Condon term can be written as:

$$F(E, \Delta G_n, \lambda_n, \omega^{eff}) = \frac{1}{\sqrt{4\pi\lambda_n k_B T}} \sum_v e^{-S^{eff} \frac{(S^{eff})^v}{v!}} \exp[\frac{(-E + \Delta G_n + \lambda_n + vh\omega^{eff})^2}{4\lambda_n k_B T}] \text{(Eq. 2.16)}$$

where S^{eff} is the effective Huang-Rhys factor and $h\omega^{eff}$ is the energy of the effective quantum mode. Both Eq. 2.15 and Eq. 2.16 are able to reproduce experimental recombination times at reasonable values of the cbm.[87] Further details on the theoretical background of the Franck-Condon term in the semi-classical limit and bridge-mediated electronic coupling are available elsewhere.[89]

The derivation of Eq. 2.12 is based on the assumption that the initial and final multi-electronic states of the process can be represented by Slater determinants differing from each other only in the occupation of one orbital. The coupling between the two states is therefore approximated by the coupling between a conduction band orbital of the semiconductor and the HOMO of the dye.[87] The theory is also not appropriate for treating situations where the trap state is in the immediate vicinity of the dye,[87] unless the theory is combined with another model such as a tight-binding model.[168] The rate expression (Eq. 2.12) is conveniently related to a number of parameters that are practically tunable by modifying the structure of dyes, such as V_{il} (Eq. 2.13), ΔG_n and λ_n (Eq. 2.15), which has led to some new design concepts for dyes to minimize charge recombination in recent years, including using the symmetry of orbitals and tailoring a chemical bridge.[170,171]

2.2.3 Free energy change and reorganization energy

We first clarify the symbols for denoting these two parameters in various contexts. In n-type DSSCs, these two parameters correspond to the single-electron transfer to the oxidized dye (*D*) at the semiconductor-dye

interface:

$$D^+ + e^- \rightarrow D \qquad\qquad (Eq. 2.17)$$

and will be denoted as ΔG_n and λ_n respectively. In p-type DSSCs, these two parameters correspond to the single-hole transfer to the reduced dye at the semiconductor-dye interface:

$$D^- + h^+ \rightarrow D \qquad\qquad (Eq. 2.18)$$
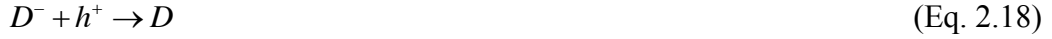
and will be denoted as ΔG_p and λ_p respectively. For discussion in a general context, where the distinction between p-type and n-type devices is not necessary, the free energy variation and reorganization energy are denoted as ΔG and λ respectively, unless specified otherwise.

In the general rate expression of charge recombination, where ΔG_n and λ_n in Eq. 2.12, Eq. 2.15 and Eq. 2.16 are replaced by ΔG and λ, the ΔG is given by the free energy difference between the equilibrium neutral dye and the equilibrium singly-charged dye. The λ is given by the sum of the internal vibrational λ_i of the dye, and the external solvational λ_s. In a two-sphere model of reactants the internal λ_i can be computed as:[172]

$$\lambda_i = \frac{1}{2} \sum_j k_j (Q_j^r - Q_j^p)^2 \qquad\qquad (Eq. 2.19)$$

where Q_j^r and Q_j^p are the jth normal mode coordinates Q for reactant and product at equilibrium respectively, and k is the reduced force constant. In the presence of non-classical vibrations λ_i can be computed as:[169,173]

$$\lambda_i = \sum_j \hbar \omega_j S_j \qquad\qquad (Eq. 2.20)$$

where ω_j is the vibrational frequency of jth normal mode, and S is the Huang-Rhys factor, a measure of the strength of the electron-phonon coupling.

The λ_s is the reorganization energy due to solvent rearrangement upon electron transfer. The solvent environment is also likely to influence ΔG, since the energy

of the charged dye is sensitive to polar solvent medium. The influence of polar medium on slow charge transfer has in fact been demonstrated in various studies. Vaisser et.al.,[174] for instance, showed that λ_s constituted at least 80% of the total λ for Ruthenium dyes in their calculations. They achieved a much improved matching with experimental references when the total λ instead of the λ_i alone was used in their computation of diffusion coefficients of hole transfer between dyes in DSSCs. Fletcher[175] showed that polar solvents could catalyze electron transfer reactions in the exergonic region, or the 'inverted' region, where the driving force of the reaction was strongly negative, as well as in the endergonic region, or the 'superverted' region, where the driving force of the reaction was strongly positive.

For a two-sphere model of reactants for electron transfer based on Marcus theory[35,36,172] λ_s can be computed as:

$$\lambda_s = (\Delta e)^2 (\frac{1}{2a} + \frac{1}{2b} - \frac{1}{R})(\frac{1}{\varepsilon_o} - \frac{1}{\varepsilon_s}) \qquad \text{(Eq. 2.21)}$$

where ε_o and ε_s are the optical and static dielectric constant respectively, and the term $(1/\varepsilon_o - 1/\varepsilon_s)$ is known as the Pekar factor which represents the longitudinal nuclear polarization,[35,176] a and b are the ionic radii of the reactants and R is the separation of the reactants. Eq. 2.21 can be further developed to provide analytical expression suitable for charge recombination in n-type DSSCs.[90,177] The effect of solvent dynamics on electron transfer can also be incorporated in this term, where in Zusman theory[178–181] λ_s is a dynamical term computed in the classical limit as:

$$\lambda_s = \frac{1}{8\pi}(\frac{1}{\varepsilon_m} - \frac{1}{\varepsilon_s})\int \Delta D^2(r)dr \qquad \text{(Eq. 2.22)}$$

where ε_m is the dielectric constant of the solvent at intermediate frequencies, ε_s is the dielectric constant of the solvent, and $\Delta D(r)$ is the difference of inductions in

the first and second dynamics states.

In this thesis a computational scheme[89,182] for total λ is followed, instead of computing λ_s and λ_i separately. λ_s is static and λ is given by:

$$\lambda = [0.5(E_{q_1}^r - E_{q_1}^p) + 0.5(E_{q_2}^p - E_{q_2}^r)]$$ (Eq. 2.23)

where $E_{q_1}^r$ and $E_{q_1}^p$ are the energies of the reactant and product in solution at reaction coordinate q_1, in which the product is in ground-state equilibrium; and $E_{q_2}^r$ and $E_{q_2}^p$ are the energies of the reactant and product in solution at reaction coordinate q_2, in which the reactant is at ground-state equilibrium. This method of computing the λ has been tested using Polarized Continuum Model (PCM) for mimicking the solvent environment.[89] Eq. 2.23 can also be used for computing λ_i by acquiring the energies in vacuum instead of solution, and hence obtaining λ_s by simply taking the difference between λ and λ_i. When evaluating charge recombination rate based on Eq. 2.16 only λ_s is necessary. Stabilizing effects due to H-bonds between solvent molecules and solutes, however, are not taken into account with this computational scheme since only a continuum model, a 'reaction field' controlled by a dielectric constant without explicit molecules, is employed to mimic solvent environment. In addition, these two terms are typically approximated from total energy differences in this thesis, unless specified otherwise.

The accurate evaluation of λ for an adsorbing dye is deemed difficult, since the solvent effect on the substrate and the exact exposure of dye to the solvent typically cannot be determined. While it is difficult to estimate the impact of solvent on the substrate, the partial exposure of an adsorbing dye suggests the total λ computed in accordance with Eq. 2.23 will likely be an overestimate of the total λ in reality. In previous evaluation of λ_n for charge recombination,[87] an

approximation[183] was introduced: only half of the adsorbing dye was exposed to the solvent, and numerically the solvent part of the λ_n was also halved. Such approximation was unnecessary when λ_n was acting as a predictor of the PCE for a set of dyes in Chapter 4, since all dyes included were assumed to subject to the same condition.

2.3 Statistical modeling for material screening

Statistical methods were employed to predict PCE (Chapter 4) and understand correlations between properties (Chapter 5), and this type of study was rare in the field of DSSCs. The objective of this section is to provide a brief introduction of quantitative structure-activity relationships (QSAR), a dominant type of application of statistics in science, as well as some specific statistical techniques used in the validation of the regression model in Chapter 4.

2.3.1 An introduction to QSAR

QSAR is a quantitative method to correlate hypothetically relevant physical or chemical properties of a chemical or material, known as the predictors, to the response value of interest such as the biological activity of a drug, resulting in a predictive model for the response. Such an approach is particularly important in combinatorial drug design, especially when the possibilities for substituent are numerous.[184,185] An early example of QSAR was the prediction of biological activity based on its linear relationship with the octanol-water partition coefficient, log P, which was a representative quantity for the hydrophobic character of a drug.[186] Another example was the synthesis of Norfloxacin,[187] designed for targeting Gram-positive and Gram-negative bacteria.[188] On the basis

of QSAR the antibacterial activities were shown to be correlated parabolically with steric parameters, which assisted in identifying substituents that were effective for decreasing bacterial activities.

In a typical QSAR study, the first step is the selection of predictors. There are in general no strict rules for selection, where the number and type of predictors to be included in the model is a choice of the researcher. The type of predictors can be highly diverse, from experimentally-derived to computable,[189,190] from classical to quantum-mechanical,[189] from 2-dimensional (topological) to 3-dimensional (conformational),[190] and depends upon the targeting activity or property and the regression model. In many cases, a small set of predictors that offers sufficient accuracy is desirable, implying relatively short time and little input for reliable prediction. Increasing the number of predictors does not imply increase in accuracy, due to the occurrence of over-fitting, meaning that the model is overly complex and therefore fitting random errors or noise rather than revealing true correlations. Over-fitting can typically be avoided by following some rules of thumb, such as 1 fitting parameter for 10 data points, or the ratio of the number of molecules of interest to the number of predictors in QSAR should be greater or equal to 5.[191] The methods to reduce predictors can be roughly classified into either the wrapper method or the filter method.[191] The wrapper method involves using an optimization algorithm for selecting predictors for an objective function; whereas the filter method does not rely on algorithms but removes predictors from an initially large set by considering, for instance, the correlation between predictors. Other methods such as the mutual information-based method[192] and the χ^2 method[193] are perhaps less commonly applied.

For a given set of predictors, the next step is the examination of

structure-activity relationship (SAR), which can be based on computing various correlation coefficients such as Person's r coefficient or Spearman's ρ rank coefficient. A model can be constructed if correlations can be established, and the model can take different mathematical forms depending on the SARs. In a simple multiple linear regression model for linear SAR the response (y) is given by:

$$y = n_0 + \sum_i n_i x_i \qquad\qquad \text{(Eq. 2.24)}$$

where n_i is the fitting coefficient for i^{th} predictor x_i, and n_0 is the intercept. One can consider other types of model such as polynomial,[194,195] neural-network[196] or kernel[197] for non-linear SARs. Generalized linear models can be considered when the response distribution is regarded as arbitrary, but still within the exponential family of distributions, instead of simply normal.[198]

Similar applications of QSAR, or QSPR, are also well-known in materials discovery,[195] although perhaps the same level of usage as in drug discovery was not reached. For examples, various scaling relations between adsorption energies[113] and neural network models[195,199] were developed for computational screening and design of catalysts; QSPR was also common for predicting the melting points of ionic liquids and the glass transition temperatures for polymers.[195]

Thanks to the increasing quality and quantity of data on dyes in DSSCs in recent years, it is now also possible to conduct statistical studies to understand their structure-property relationships, and generate tools for predicting the performance of new dyes. A group of such models were built for coumarin and phenothiazine dyes, with the objectives to reveal structure-photovoltaic performance relationships and designing new molecular structures with optimal properties for these families of dyes, using descriptors relating to properties such

as vibrational normal modes and topology of dyes, among other predictors.[200–203] We have developed a QSPR-like model for DSSCs' dyes and will be introduced in chapter 4.

2.3.2 Some statistical techniques for validating regression models

In this section we introduce a number of statistical techniques employed for validating the statistical model developed in Chapter 4. Consider a generalized linear model, which allows incorporation of non-normal response distribution and with the linear predictor defined by:

$$\eta_i = g[E(y_i)] \tag{Eq. 2.25}$$

where $E(y_i)$ is the expected value of the response y_i and g is the link function. g can take any mathematical form in the exponential family,[198] such as the Gaussian distribution, which is also known as the identity link. One can test if the data can fit well with a particular distribution by comparing their probability distributions, which is typically represented with a quantile-quantile (Q-Q) plot.

To assess the precision of estimates, such as predictions and/or regression coefficients, one can consider calibrating the model with iterative bootstrap re-sampling.[204,205] Such a procedure is particularly useful when no standard procedure is available for estimates or they are approximated based on asymptotic theory.[198,206] In general, the bootstrap re-sampling starts with selecting a random sample of size n, where n is the size of an original sample for a fitted model. The random sample is the original sample with some observations omitted or duplicated, and will be fitted with the model developed initially with the original sample. By repeatedly re-sampling and fitting for, for example, 100 times, where a bootstrap estimate is obtained in each iteration, a bootstrap standard deviation of the sampling distribution of the estimate, or the precision of

the estimate, can be acquired.[198] The calibration can be assessed by a bootstrap shrinkage estimator, where the ideal results should be the shrinkage estimator taking the value of 1. The reliability of estimating the shrinkage can be improved by incorporating shrinkage in the original estimation process, where different shrinking factors can be assigned to difference parameters.[207]

In the event of a lack of complexity, the model can encompass more fitting parameters, but such action would also increase the risk of over-fitting. A possible solution is to consider statistical penalization of the model fitting criteria for complexity. Consider estimating parameters for a model following the common maximization of the (log of) likelihood function (MLF), such as solving the first derivative of joint probability of observations with respect to a parameter, or equivalently, maximizing the likelihood ratio (LR) χ^2 statistic of the model with respect to the "null" model, a model without predictors. The likelihood ratio is the ratio of the likelihood with the hypothesized parameters to the likelihood of the data at the parameters given by MLF.[207]

A method for penalizing the LR achieved by a given model for its complexity, for obtaining a more unbiased assessment of the model's worth, is the employment of the Akaike's information criterion (AIC), which in 'adjusted χ^2' can be written as:[207]

$$\text{AIC} = \text{LR } \chi^2 - 2p \qquad\qquad \text{(Eq. 2.26)}$$

where p is the number of parameters in the model. It is evident from Eq. 2.26 that, a model with large p will reduce the effective LR, so the optimal model will result from a trade-off between maximizing LR and reducing p. It turns out the above criterion can still be biased when the sample size is small, so one can consider a corrected AIC that takes into account the sample size n:

$$\text{AIC}_C = \text{LR } \chi^2 - 2p\left(1 + \frac{p+1}{n-p-1}\right)$$

(Eq. 2.27)

From the formula we can see that a small sample size incurs a greater penalty than a large sample size as the sample size tends to infinity $\text{AIC}_C \underset{n\to\infty}{\to} \text{AIC}$.

When some parameters appear to have low predicting power, such as with low *F*-statistics in analysis of variance (ANOVA), it is tempting to remove these parameters to simplify the model. This, however, can introduce bias and produce optimistic estimates of the parameters' covariance matrix. An alternative is to simplify the full model by approximating it.[207] For instance, a new set of estimated responses can be produced as predicted by the full model, and the new response is then regressed against a chosen subset of predictors. Such approach is termed 'simplification by approximation'.

Chapter 3 Computation of hole transfer rates in p-type dye-sensitized solar cells: why is charge recombination faster than in n-type devices?

3.1 Introduction

The development of p-type dye-sensitized solar cells (p-type DSSCs) was motivated by the recognition of the high theoretical efficiency that could be achieved by tandem cells with two photo-active electrodes,[50] as introduced in Chapter 1. In such types of device the generation of photo-current is limited by the weaker electrode, which is the photo-cathode that operates in accordance with the principles of p-type DSSCs (Fig. 1.2). Understanding the origin of the weak photocurrent generation in p-type DSSCs was therefore deemed important for improving the performance of tandem DSSCs.

The primary charge generation process in p-type DSSCs is promoted by a photo-excited dye that injects a hole into the semiconductor. Hole injection is the transfer of a hole from a localized orbital of the dye, typically the HOMO, to a one-electron state of the VB of the semiconductor. The analogous process in n-type DSSCs is the electron injection from the LUMO of the dye to the CB of the semiconductor. Hole recombination, on the other hand, is the transfer of a hole from a one-electron state of the valence band of the semiconductor to an orbital of the dye, typically the LUMO, or it could be more intuitive to visualize as an electron transfer from the LUMO of the dye to an empty orbital in the VB. This process is an analogue to charge recombination in n-type DSSCs, where electrons back-transfer from the CB of the semiconductor to the HOMO of the dye. Charge recombination processes contribute to the decrease of the solar cell

efficiency and should be minimized.

While hole injection in p-type DSSCs appeared to be as efficient as electron injection in n-type DSSCs, hole recombination was considered as one of the main causes for the low photo-current observed in p-type DSSCs.[208] A number of experimental investigations on the hole transfer kinetics in p-type DSSCs demonstrated that the hole recombination was much faster than desirable. TAS and kinetic analysis of Coumarin 343 (C343) sensitized NiO showed that, while charge injection had an ultrafast component and was similar to common n-type DSSCs,[209,210] charge recombination occurred at the timescale of tens of pico-seconds,[56,211] and it was faster than the recombination times observed in C343 sensitized TiO$_2$. Electrochemical impedance spectroscopic studies also addressed the charge recombination problem in NiO p-type DSSCs.[60,212] Attempts to suppress recombination included adjusting molecular dipole alignment on NiO surface,[212] increasing the tunneling distance by extending oligo-thiophene linker in donor-π-acceptor dyes,[51] and substituting conventional I_3^-/I^- redox shuttle with alternatives such as the Co(II)/(III) pair.[52] While some success on retarding recombination had been achieved, such as slowing down the process from ns to μs,[51,213] the extent of improvement was insufficient for producing p-type DSSCs with PCE comparable to that of n-type DSSCs. The origin of the fast recombination had in fact remained unclear, and intuitive suggestions on retarding recombination might negatively influence other processes in the device, such as hole injection,[213] hence resulting in small improvement of PCE even with slower recombination. In this chapter we attempt to explain why recombination is faster in p-type DSSCs than in n-type DSSCs. Based on the injection and recombination theories introduced in section 2.2, we evaluate the hole injection and recombination rates at the NiO-C343 interface, a

typical semiconductor-dye interface in p-type DSSCs, and compare results with analogous interface in n-type DSSCs.

3.2 Theoretical background and computational methods

3.2.1 Theory of hole injection

The method for computing the rate of hole injection in this chapter is essentially identical to that for computing electron injection in n-type DSSCs (see section 2.2.1) except for the orbitals involved.[65] The rate ($\Gamma_{ss}(E)$) of a hole transfers from a one-electron state $|s\rangle$, such as the HOMO of the dye, to a manifold of states $\{|l\rangle\}$, such as an orbital in the VB of the semiconductor, can be expressed as Eq. 2.9. In this case, in Eq. 2.9, E is the energy of the HOMO of the dye, E_l is an eigenvalue of the manifold of eigenstates in the semiconductor, such as the valence band (VB) maximum; V_{sl} is the coupling between state $|s\rangle$ and state $|l\rangle$ and $\delta(E - E_l)$ is the Dirac delta function, which in this work is approximated by a normalized Gaussian function with 0.1 eV broadening. The states $|s\rangle$ and $|l\rangle$ are conveniently represented as linear combinations of basis functions $\{\chi_m\}$ and $\{\varphi_k\}$ localized on the dye and the semiconductor respectively:

$$|s\rangle = \sum_m c_{ms} \chi_m \qquad \text{(Eq. 3.1)}$$

$$|l\rangle = \sum_k a_{kl} \varphi_k \qquad \text{(Eq. 3.2)}$$

where $\{c_{ms}\}$ and $\{a_{kl}\}$ are the orbital coefficients derived from the calculations of isolated dye and semiconductor respectively. The coupling V_{sl} can be expressed in terms of the coupling V_{mk} between localized orbitals χ_m and φ_k:

$$V_{sl} = \sum_m \sum_k c_{ms} a_{kl} \langle \chi_m |V| \varphi_k \rangle \qquad \text{(Eq. 3.3)}$$

and the rate can be expressed as:

$$\Gamma_{ss}(E) = \sum_{m,n} \Gamma_{mn}(E) c_{ms} c_{ns}^* \qquad \text{(Eq. 3.4)}$$

where

$$\Gamma_{mn}(E) = \frac{2\pi}{\hbar} \sum_l \sum_{k,k'} V_{mk} V_{nk'}^* a_{kl} a_{k'l}^* \delta(E - E_l) \qquad \text{(Eq. 3.5)}$$

The equation above can be re-written by defining the local density of states $\rho_{kk'}(E)$ as:

$$\rho_{kk'}(E) = \sum_l a_{kl} a_{k'l}^* \delta(E - E_l) \qquad \text{(Eq. 3.6)}$$

which gives Eq 2.11. For non-orthogonal basis sets, $\Gamma_{mn}(E)$ should be modified to become:

$$\Gamma_{mn}(E) = \frac{2\pi}{\hbar} \sum_{k,k'} (ES_{mk} - V_{mk})(ES_{nk'} - V_{nk'})^* \rho_{kk'}(E) \qquad \text{(Eq. 3.7)}$$

In practice, to compute the charge injection rate one needs (i) the energies, E_l, of the considered NiO Kohn-Sham eigenstates, $|l\rangle$, relative to the HOMO energy E; (ii) the coefficients of the basis functions of NiO, a_{lk}, and of C343, c_{ms}; (iii) the coupling term V_{mk}; and (iv) the overlap matrix element S_{mk}.

3.2.2 Theory of hole recombination

The approach to calculate hole recombination in p-type DSSCs is similar to that for calculating charge recombination in n-type DSSCs.[87] The hole recombination rate ($k_{h,rec}$) is given by:

$$k_{h,rec} = \int \Gamma_{ii}(E)(1 - f(E - E_F)) F(E, \Delta G_p, \lambda_p) dE \qquad \text{(Eq. 3.8)}$$

The first term of the integrand, $\Gamma_{ii}(E)$, is given by Eq. 2.13, where the coupling V_{il} is between one-electron states $\{|l\rangle\}$ in the semiconductor and a one-electron state $|i\rangle$ in the dye, which corresponds to the LUMO of the dye. $\Gamma_{ii}(E)$ can therefore be expressed in terms of the orbital coefficients of the LUMO, $\{c_{mi}\}$:

$$\Gamma_{ii}(E) = \sum_{m,n} \Gamma_{mn}(E) c_{mi} c_{ni}^* \qquad \text{(Eq. 3.9)}$$

In Eq. 3.8, $f(E - E_F)$ is the Fermi-Dirac distribution and E_F is the Fermi level, or the quasi-Fermi level in non-equilibrium situations. The term $1 - f(E - E_F)$ is therefore the probability of a hole occupying a state at a given energy. $F(E, \Delta G_p, \lambda_p)$ is the thermally averaged Franck-Condon integral between initial and final vibrational states:

$$F(E, \Delta G_p, \lambda_p) = \frac{1}{\sqrt{4\pi\lambda_p k_B T}} \exp[-\frac{(E - \Delta G_p + \lambda_p)^2}{4\lambda_p k_B T}] \qquad \text{(Eq. 3.10)}$$

which is similar to Eq. 2.15 and differs most notably with the use of different ΔG and λ in these expressions, as discussed in section 2.2.3. In practice, to compute the charge recombination rate given by Eq. 3.8 one needs (i) the function $\Gamma_{ii}(E)$, which is calculated with the same procedure used in hole injection, except that the MO coefficients in Eq. 3.9 are those of the LUMO instead of the HOMO of the dye; (ii) the quasi-Fermi level, E_F; (iii) λ_p for Eq. 2.18; and (iv) ΔG_p for Eq. 2.18.

3.2.3 Electronic structure calculation of the semiconductor and the interface

The electronic structure of NiO is known to be dominated by strong on-site Coulombic repulsion between d-electrons,[214–216] which cannot be described satisfactorily by standard DFT methods. Bredow and Gerson,[217] for instance, compared the performance of unrestricted Hartree-Fock (UHF) and DFT with several different functionals on predicting the bulk properties of NiO. They showed that UHF overestimated the band gap, whereas DFT underestimated the band gap, except when the computation was performed with B3LYP functional.

DFT with B3LYP, however, described incorrectly the electronic structure near the VB maximum. This part of the valence band should be dominated by the O $2p$ orbitals,[218] but B3LYP predicted a larger weight of Ni $3d$ orbitals in this region. In this case we considered modeling NiO with DFT/SGGA + U scheme as introduced in section 2.1.1., which has the advantage of low computational cost when compared to other methods (see section 2.1.1) for modeling strongly correlated materials. In the rest of this section we will provide the computational details and methods for modeling the NiO slab and NiO-C343 interface that were employed in our evaluations of the hole transfer rates.

All DFT and DFT+U calculations were performed using the SIESTA code[219] with Pardew-Burke-Ernzerhof (PBE)[220] exchange-correlation functional. The atomic core potentials were approximated by the Troullier-Martins norm-conserving pseudo-potentials[221] and the basis set was double-ζ polarized (DZP) for all atoms. All the +U calculations were spin-polarized, in which the total spin of NiO was fixed as zero throughout these calculations. The initial spin density was set with an anti-ferromagnetic order. A rather large effective U (U_{eff} = $U - J$) of 8.25 eV (J = 0.95 eV) was employed with the intent to attain a computed band gap close to the experimental band gap, as the energy level alignment was an important property in this study. This U_{eff} value was close to that suggested by Anisimov et. al (U_{eff} = 7.1 eV, J = 0.95 eV).[222]

The surface considered was NiO(100), which is known to be non-polar and rather stable.[223,224] To model the adsorption of the dye C343 on NiO, we first tested two adsorption modes of a much smaller molecule (acetic acid) that shares the same anchoring group (carboxylic acid anchoring group) with C343 (Fig. 3.1 (c)). The anchoring group is the part of the dye that facilitates adsorption of dye on the surface. Generally the surface energy and the band gap of NiO(100) were

considerably invariable with slab thickness,[225] and we had adopted a 3-layer slab, constructed with surface unit cell extended by 1 unit, with a total (2 × 2) extended surface unit cells. The extension of unit cell along the y-axis was necessary for describing the anti-ferromagnetism of NiO.[153] The adsorption modes tested were non-dissociative molecular (Fig. 3.1(a)) and dissociative bridging (Fig. 3.1(b)). In the computations of dissociative adsorptions, it was assumed that the carboxylic hydrogen bound to the surface oxygen closest to the adsorption site after dissociation. The counterpoise corrections for eliminating basis set superposition error[226] were included in the calculations of the adsorption energy. The most stable adsorption mode was the molecular adsorption mode, with absolute adsorption energy of 0.86 eV, which was about 0.2 eV lower in adsorption energy than dissociative bridging. This value was close to the adsorption energy of carboxylic acid anchoring on anatase (101), which was about 0.9 eV.[165]
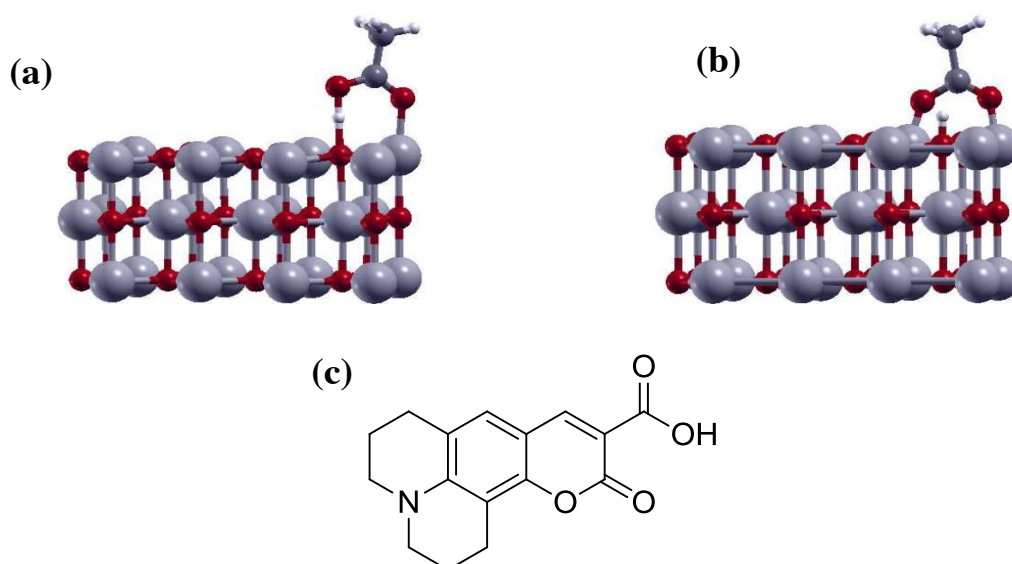

Figure 3.1 Acetic acid adsorbs on 3-layer NiO (100) in (a) non-dissociative molecular mode and (b) bidentate dissociative mode. (c): Chemical structure of C343.

The acetic acid-slab model was then used to evaluate the required matrix elements, V_{mk} and S_{mk}, the eigenvalues E_l, and the coefficients of the basis functions of the NiO slab, a_{lk}, for computing the charge transfer rates (Eq. 3.7 and Eq. 3.8). The model was initially optimized with DFT calculation, and subsequently with the U_{eff} correction implemented in a single-point calculation of the optimized structure at the Gamma point. We noticed that convergence in DFT+U relaxation of slabs was difficult to achieve, where self-consistent field (SCF) calculations generally required long computational time and large number of cycles to reach convergence.

3.2.4 Computational methods and details for the dye component

C343 dye (Fig. 3.1(c)) was one of the simplest and considerably efficient organic dyes that had been studied experimentally for both p-type[56] and n-type DSSCs.[227] The MO coefficients needed for evaluating $\Gamma_{ss}(E)$ (Eq. 3.4) and $\Gamma_{ii}(E)$ (Eq. 3.9) were obtained from structure optimization using SIESTA[219] at the GGA/PBE[220] level with DZP basis set. The starting lattice vectors were 20 Å × 20 Å × 20 Å, which should provide a simulation box that is large enough to avoid interactions between individual C343 molecules under periodic boundary conditions.

The computational strategy for the total reorganization energy λ_p of C343 was discussed in section 2.2.3, based on Eq. 2.23 and the approximation that only half of the adsorbing dye was exposed to the solvent. In the computation of λ_p, the required optimization and single-point calculations were performed with B3LYP/6-31++G**, and the solvent environment (acetonitrile) was simulated with the Polarized Continuum Model (PCM).[228] These calculations were performed with Gaussian03.[229]

The computed charge recombination rates could be very sensitive to the energy levels of the dye component relative to the energy levels of the semiconductor. Results that were more comparable to the experimental literature may be obtained using data deriving from the experiments, especially for the dye's free energy change for recombination, $\Delta G_{p,}$. We would therefore discuss the most convenient method to evaluate this quantity in the result section, after comparing the computed and experimental interfacial relative energy levels.

3.3   Results

3.3.1   Hole Injection

In this study hole injection was a hole transferred from the oxidation potential of C343, E_s, to one of the VB states of NiO, where the maximum of these states was denoted as E_v, and the energy difference between E_v and E_s represented the driving force of the hole injection. The experimental and computed E_v and E_s are both reported in Table 3.1, where the redox potentials provided were converted from their values against the standard hydrogen electrode (SHE) to absolute electrode potentials with an additive constant of −4.6 eV.[45] For the computed E_v reported we were unable to check the energy of the zero-energy level for the DFT+U slab calculation due to technical difficulties of retrieving electrostatic potential data,[230] but the computed E_v was not significantly different to another previous report of E_v (−5.11 eV).[231] However, the computed E_s based on the HOMO energy given by B3LYP/DFT calculation, was 5.7 eV, implying that the energy difference between computed E_v and E_s did not match that evaluated from experimental E_v and E_s. In this respect, we employed the experimental energy levels by shifting consistently the band energies from DFT calculations to render

the calculated VB maximum coincided with the experimental VB maximum, and used the experimental oxidation potential in place of the computed result (Table 3.1).

Table 3.1 Numerical values of interfacial energy levels and parameters that entered the computation of the rates. The numbers underlined are values employed for computing the injection rates, and the numbers in bold are values employed for computing the recombination rates. (d): property of the dye, (s): property of the semiconductor. All values are in eV.

Energy levels	Symbols	Numerical values	
		Experimental	Computed
Reduction potential (d)	ΔG_p	**−3.4**[a]	−2.9
Quasi-Fermi level (s)	E_F	**−4.95**[b]	
Valence band maximum (s)	E_v	<u>−5.0</u>,[c] −4.8[d]	−5.4[e]
Oxidation potential (d)	E_s	<u>−5.8</u>[a]	−5.7
Other parameter			
Reorganization energy (d)	λ		**0.7**

a: ref. [232]; b: see main text below for derivation; c: ref. [233]; d: ref. [234]; e: uncertainty is associated with this value as we were not able to check the energy of the zero-energy level for DFT+U calculations.

The computed injection time τ_{inj}, which is given by the reciprocal of the injection rate (Eq. 2.9), is reported in Fig. 3.2 (a) as a function of E_s, in order to provide an estimate of the uncertainty in the computed injection time associated with the uncertainty in the relative energy levels. The injection time of $E = E_s$ is 16 fs, a value that is rather insensitive to the uncertainty in E_s: the injection rate is changed by a factor of about 3 at most for changes of E_s up to 1 eV. Fig. 3.2(b) illustrates the density of states (DOS) and the projected density of states (PDOS) on O and Ni atoms, and shows that the dependence of the hole injection rate on the energy (Fig. 3.2(a)) follows the PDOS of the O atoms, which dominates the

region close to the VB maximum. This implies that the DFT+U describes the band structure correctly (see section 3.2.3) near the valence band maximum of NiO, and is necessary for the computation of the hole transfer rate at the interface involves NiO.

Experimentally, the injection time constant for C343 had been determined quantitatively with a kinetic analysis based on TAS measurements. Our computed injection time was faster than this experimentally measured injection time for C343 (~210 fs;[56] the experimental set up of this measurement, however, was unable to record injection time faster than ~150 fs). Other TAS measurements also show that the injection time was generally fast (~180 fs, measured with a Ruthenium(II) complex;[235] and <200 fs with a triphenylamine-based chromophore).[236] In addition, the standard deviations in measured injection times could also be large for TAS (up to $\pm$ 50 fs).[210] It was therefore likely that there were injection times faster than the injection time constant reported from TAS experiment for the C343-NiO interface.[56]

An important observation was that the computed hole injection time was close to the computed electron injection time in n-type DSSCs (~11 fs),[65] essentially because, in both cases, the couplings and the DOSs that entered the expression of the injection time were not too dissimilar. This observation, which agreed with experimental observation, suggested that the semiconductor-dye interaction, as measured by $\Gamma(E)$, was not significantly different for the two classes of DSSCs.
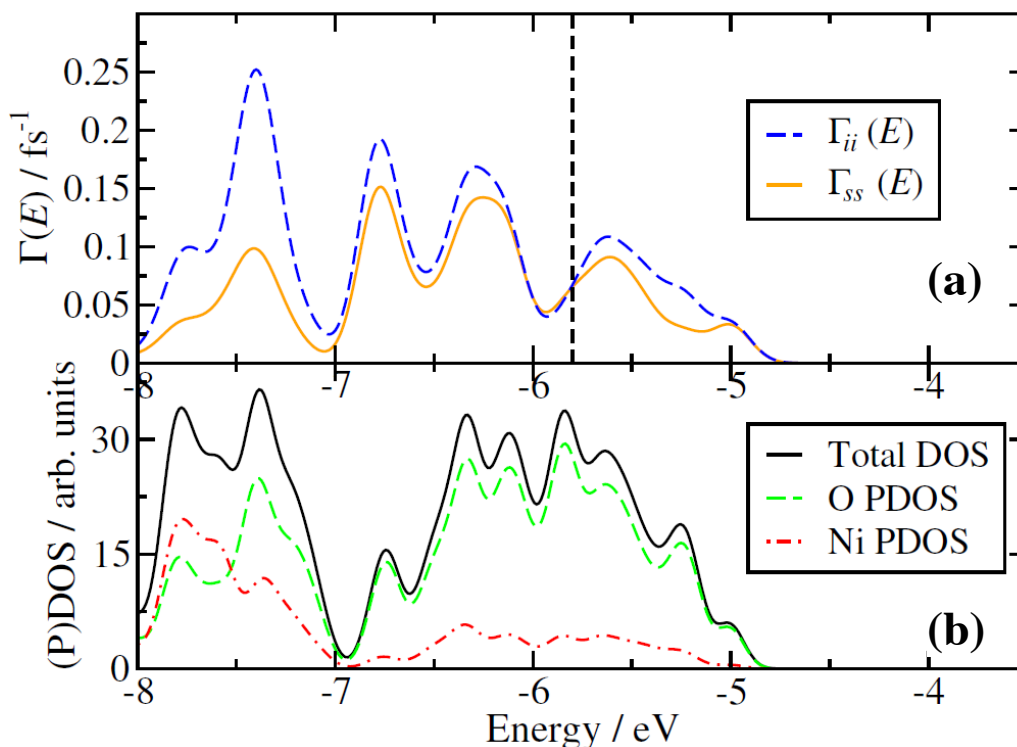
Figure 3.2 (a) Computed hole injection rates (orange line) against injection energies in eV. The black-dotted line indicates experimental E_s. (b) Plots of total density of states (black line), partial density of Ni atoms (red dotted line) and O atoms (green dotted line) of 3-layer NiO slab.

3.3.2 Hole recombination

The computation of hole recombination rates is known to be highly sensitive to the accuracy of relative energy levels at the semiconductor-dye interface.[39] This sensitivity is evident from Eq. 3.8 and Eq. 3.10, where the relative energy difference enters the rate expression as a quadratic quantity in the exponential term, and hence the impact of any associated uncertainty on the rate is likely to be magnified significantly. For this reason, the experimental energy gap was employed by shifting consistently the band energies from DFT calculations, same as the procedure used in the computation of injection rate. The computed reduction potential in acetonitrile (PCM) at B3LYP/6-31++G** with Gaussian 03[229] yielded a result that deviated substantially from the experimental value. The

experimental value was underestimated by 0.5 eV without zero-point energy and entropic correction; these effects were minor for computing redox potentials,[237] and no improvement was observed when the basis set was increased to aug-cc-pVTZ.[237] The sensitivity of the computed rates to the relative energy levels would be examined by evaluating the rates as a function of a range of ΔG_p.

The computed $\Gamma_{ii}(E)$ function in the rate expression (Eq. 3.8) followed the DOS and PDOS of the O atoms in NiO (Fig. 3.2), in line with Eq. 2.13. Another crucial parameter affecting the hole recombination rates is the quasi-Fermi level E_F of the semiconductor (Table 3.1), controlling the probability distribution of the holes under illumination. E_F is related to the V_{oc} of DSSCs as $V_{oc} = (E_F - E_{F,redox})/e$, where $E_{F,redox}$ is the Fermi level of the electrolyte and e is the elementary charge. Evidently, E_F depends on light intensity and our hole recombination rate would be computed under 1 sun illumination, where V_{oc} is close to 0.1 V in these devices.[212] The effect of the illumination was therefore introduced in our model by setting the quasi-Fermi level of the semiconductor (−4.95 eV) to an energy of 0.1 eV below the redox potential of I_3^-/I^- redox shuttle (−4.85 eV).

Fig. 3.3 shows the relationship between the computed recombination time against a range of ΔG_p (Eq. 3.8) and for three different values of λ_p, in order to illustrate the sensitivity of the computed charge recombination time on these parameters. Using the experimental ΔG_p of C343 and the computed reorganization energy (0.7 eV), the computed recombination time was ~500 ps, which was in reasonable agreement with the experimentally determined time of ~20 ps,[7] and other reported recombination times of a few tens of ps.[8] This level of agreement suggests that the main physical phenomenon of charge recombination was captured correctly by the model.

The high sensitivity of recombination time to the relative energy levels is highlighted in Fig. 3.3, where one order-of-magnitude difference in recombination time could be produced by shifting ΔG_p by only 0.1 eV with respect to a constant E_v. Similarly, the results could also be strongly affected by λ_p, as illustrated by the curves showing recombination times computed with λ_p increased or decreased by 0.2 eV with respect to the computed value for C343. The results in Fig. 3.3 suggest how both parameters could be tuned via chemical synthesis to produce dyes with longer recombination times.

Similar to what had been observed for n-type DSSCs,[87] the strong dependence of the charge recombination rate on parameters that could not be computed with sufficient accuracy suggested that a prediction of charge recombination rate entirely from first principles (i.e. without experimental input) was currently unfeasible. On the other hand, once experimental reference data were introduced, it was possible to predict how the charge recombination rate would change with respect to the chemical modifications of the dye, where these modifications would alter the oxidation free energy change, reorganization energy and the orbital shape.
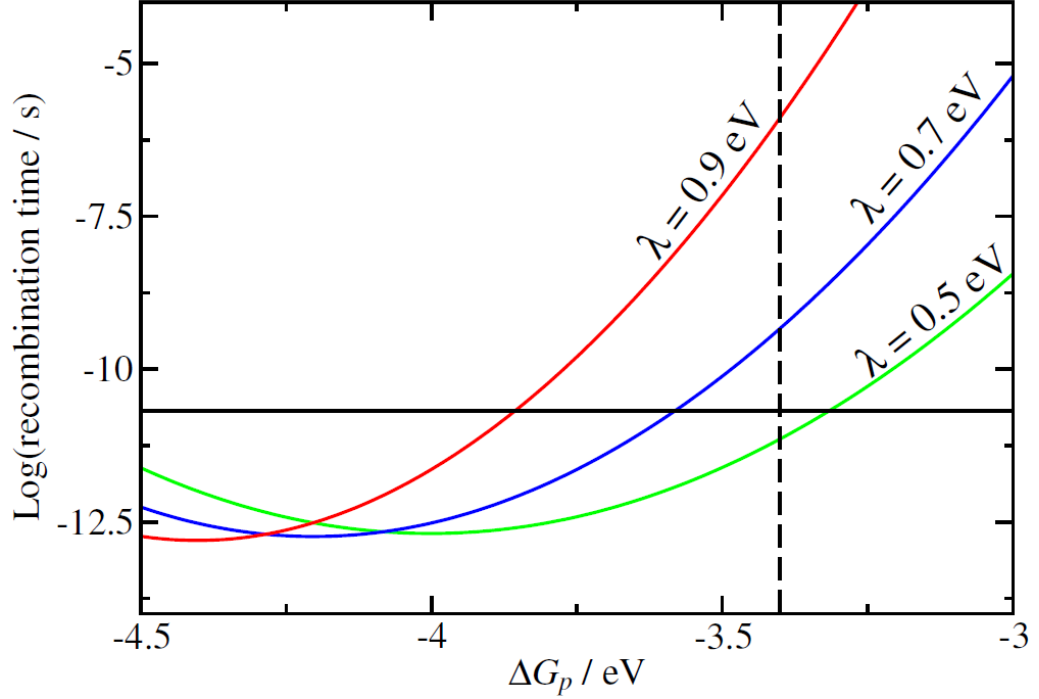
Figure 3.3 Log of computed recombination time (s) against ΔG_p (eV). The black solid line represents the experimental recombination time and the black dotted line represents the ΔG_p of C343 in acetonitrile.

In general, charge recombination time in p-type DSSCs are substantially (~3 order-of-magnitudes)[208] faster than the charge recombination time in n-type. To identify the source of this difference, we compared the terms in the charge recombination rate expression for p-type and n-type devices. The first term in the integrand, $\Gamma(E)$, was similar between the two types as discussed earlier in section 3.3.1. In the Fermi-Dirac distribution, the separation between the quasi-Fermi levels and the band edges are different in the two types under 1 sun illumination, but we had verified that under a 'like-for-like' comparison (by setting the separation equal to 0.1 eV in both cases), the effect on recombination time was negligible (<< 1 order of magnitude). The remaining term was the Franck Condon (FC) term, in which the p-type and n-type differed from each other by the driving force for recombination:

$$FC_{p-type} \sim \frac{1}{\sqrt{4\pi\lambda_p k_B T}} \exp[-\frac{(E_v - \Delta G_p + \lambda_p)^2}{4\lambda_p k_B T}]$$

(Eq. 3.11)

$$FC_{n-type} \sim \frac{1}{\sqrt{4\pi\lambda_n k_B T}} \exp[-\frac{(-E_c + \Delta G_n + \lambda_n)^2}{4\lambda_n k_B T}]$$

(Eq. 3.12)

where in Eq. 3.12, E_c is the conduction band edge of the semiconductor.

Eq. 3.11 and Eq. 3.12 assumed for simplicity that the carrier recombined by transferring from the dye to an orbital at the relevant band edge of the semiconductor. Given the results thus far, and at a given temperature, the observed difference in recombination time between the two types of cell could only be due to the difference in reorganization energy and/or the driving force of the recombination, i.e. $E_v - \Delta G_p$ and $-E_c + \Delta G_n$. In section A3.1 we have demonstrated that the reorganization energies were similar for dyes with similar size, and therefore the different kinetics observed in p-type and n-type was unlikely due to different reorganization energies.

It was therefore expected that the origin of the difference in recombination times between n-type and p-type devices lies in their different driving forces. To verify this hypothesis we considered 30 dyes for which the experimental oxidation potential and reduction potential are known, and combined them with the experimental E_c (−4.0 eV for TiO$_2$)[238] and E_v (−5.0 eV for NiO)[233] respectively to obtain averages and standard deviations for the driving forces in n-type and p-type recombination. The detailed data are reported in section A3.2. Figure 3.4 illustrates the range of FC values that were pertinent to the two types of devices for different values of reorganization energy. The few order-of-magnitudes slower recombination rate in n-type seemed to be entirely due to the different typical driving force in the two types of devices, leading to a

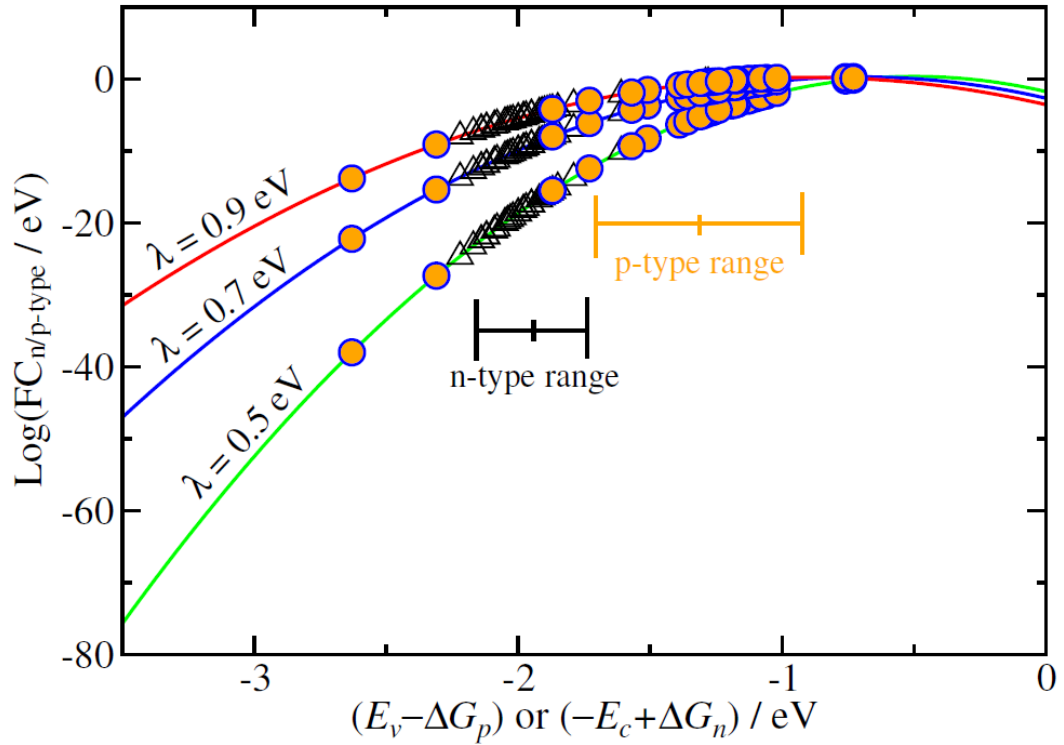few order-of-magnitudes smaller FC factors for n-type devices.

Figure 3.4 Log of FC terms against the driving force of recombination in eV, where $(E_v - \Delta G_p)$ is the driving force for p-type and $(-E_c + \Delta G_n)$ is the driving force for n-type. Three λ values have been considered (0.5 eV (green), 0.7 eV (blue) and 0.9 eV (red)). The n-type and p-type ranges represent the means and standard deviations of the experimental oxidation potentials (black) and reduction potentials (orange) of 30 dyes designed for p-type DSSCs.

On the basis of our results, there are two general strategies for lengthening the recombination time at the semiconductor-dye interface in p-type DSSCs, either increasing the driving force or decreasing the reorganization energy. The first strategy could be achieved by either lowering the VB maximum of the semiconductor and/or shifting the reduction potential of the dye up the energy scale. Practically the former could be realized by substituting NiO with another semiconductor, but the search for a cost-effective alternative could be difficult. The reduction potential of a dye could be adjusted by introducing chemical

structural modifications, such as systematically changing the structure of the conjugation bridge of the linker in donor-π-acceptor dyes.[239] The adjustment should at the same time (i) retain the reduction potential lower in energy than the conduction band edge, (ii) attain a reasonable optical gap for light absorption, (iii) avoid excessive energy loss due to neutralization of the dye by the redox shuttle. The reorganization energy of a dye could perhaps be decreased by increasing the size of the dye (Fig. A3.2), such as by adding simple inactive alkyl chain, remembering however that most changes that affected the reorganization energy would also alter the redox levels.

Apart from the strategies based on our results, a number of strategies proposed to improve the efficiency of n-type devices by maximizing the coupling for charge injection and minimizing it for charge recombination[170,171] could also be adapted for p-type devices.[171] In addition, this study had not considered the impact of defects in the semiconductor, which are known to reduce the driving force and accelerate charge recombination.

3.4   Conclusions

We evaluated hole injection and recombination times at the semiconductor(NiO)-dye(C343) interface in p-type DSSCs with non-adiabatic charge transfer theories, and explained the similarities and differences in interfacial charge transfers between p-type and n-type DSSCs. The methodology was adapted from a similar one developed for the study of charge injection and recombination in n-type devices. Reasonable charge injection times were obtained from electronic structure calculations of the dye, the semiconductor and a model system to evaluate the coupling between the two. The computed charge

injection time was not critically affected by the uncertainty in the alignment of the energy levels. The calculation of the charge recombination rate was instead very strongly dependent on the alignment of the redox and semiconductor energy levels and we have therefore performed the evaluation of the rate by combining computational data and experimental data pertaining to the energy level alignment.

The computed rates were broadly in agreement with the experimentally available data. The methodology was therefore used to explain the reason for the faster charge recombinations observed in p-type devices. Having verified that the dye-electrode coupling was similar for both devices, and also considering a larger number of experimental data on n-type and p-type devices, we concluded that the difference in the rates were derived from different Franck-Condon factors in the rate expressions, which ultimately depended on the different driving forces for the charge recombination processes. As the charge recombination took place in the Marcus inverted region, we suggested that, to slow down the charge recombination in p-type devices, the simplest strategy was to modify either the dye or the semiconductor to increase the charge recombination driving force.

3.7 Appendix

A3.1 The effects of sizes and functional groups of dyes on reorganization energy

Correlation between reorganization energy and the size of a dye was observed for n-type devices (Fig. A3.2; see also Fig. 4.2(b)), and we have also tested whether the same correlation could be observed for p-type devices. The reorganization energies for hole and electron recombination in p-type and n-type devices were computed for four dyes (Fig. A3.1(a) and Fig. A3.2). From Fig. A3.2 we expected that this correlation would hold for p-type devices (blue data points). The spread of data illustrated quantitatively that the size was an important factor, although other molecule-specific properties may also be important. The differences observed between reorganization energies for hole and electron recombination of the same dye was from 0.07 eV to 0.16 eV (Table A3.1). These differences were chiefly originated from the internal part of the total reorganization energies which account for approximately 18-38 % of the total reorganization energy, based on the assumption that only half of a dye is exposed to the solvent.

We have also examined the effect of different functional groups on reorganization energies. Reorganization energies of a number of hypothetical derivatives of C343 (Fig A3.1(b)) were computed and it had been observed that the variations in reorganization energies were small (0.02 eV to 0.07 eV). Change of functional groups peripheral to the main chromophore seemed to have a smaller effect than the sizes of dyes. We therefore concluded that the different kinetics observed in p-type and n-type was unlikely due to different reorganization energies.
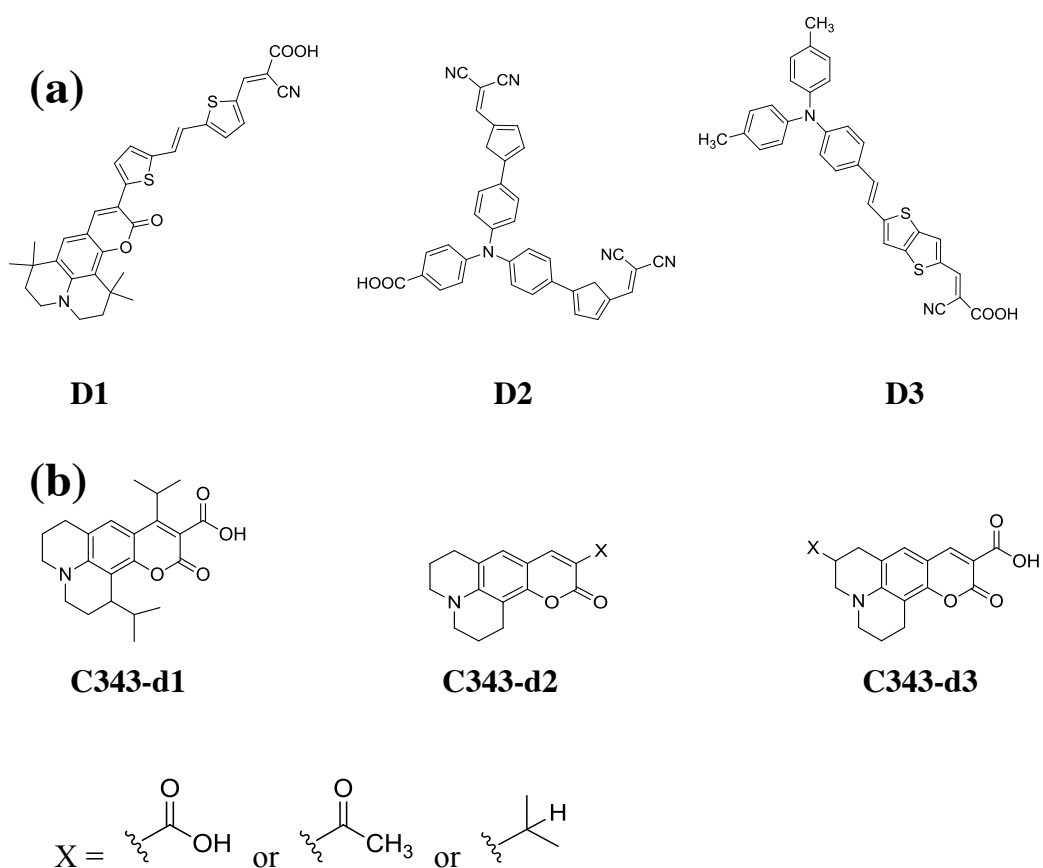
Figure A3.1 (a) Chemical structures of dyes (**D1** to **D3**) for testing the correlation between reorganization energies and the size of a dye in Fig. A3.2; and (b) Chemical structures of a range of hypothetical C343 derivatives (**C343-d1** to **C343-d3**) for examining the effect of functional groups on reorganization energies.

Table A3.1 Internal and solvent λ for hole and electron recombination of dyes C343, **D1**, **D2** and **D3**. The total reorganization energies are computed with the solvent part being halved.

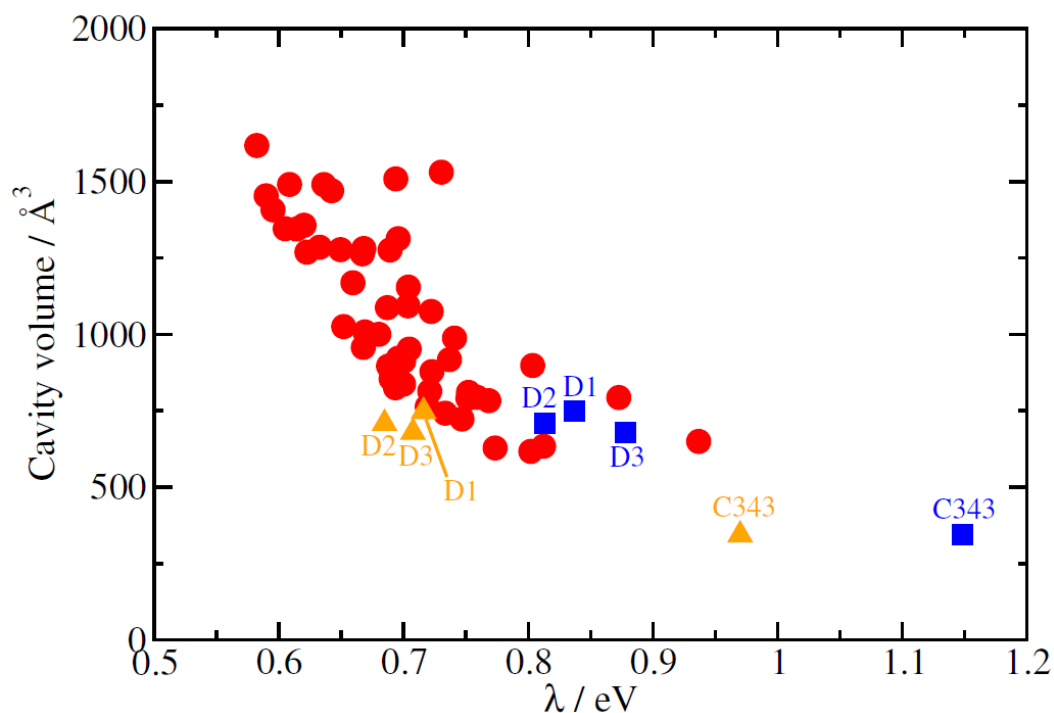| Dye	Hole recombination			Electron recombination		
	$\lambda_{i,p}$ / eV	$\lambda_{s,p}$ / eV	λ_p / eV	$\lambda_{i,n}$ / eV	$\lambda_{s,n}$ / eV	λ_n / eV
C343	0.26	0.89	0.71	0.12	0.85	0.55
D1	0.14	0.70	0.49	0.12	0.60	0.42
D2	0.11	0.70	0.46	0.07	0.61	0.38
D3	0.21	0.67	0.55	0.07	0.64	0.39

Figure A3.2 Correlation between λ (eV) and cavity volume (Å^3). Red circles were λ computed for n-type DSSCs with the dataset employed in Chapter 4 (see section 4.2). Blue squares and orange triangles were λ computed for hole and electron recombination respectively with the dyes in Fig. A3.1(a). The cavity volume is the volume of the cavity generated in PCM, where the computational details are given in section 4.3.1.

A3.2 Oxidation and reduction potentials data in Fig. 3.4

Table A3.2 shows the redox potentials data used in Fig. 3.4 in section 3.5, which were acquired from 30 dyes designed for p-type DSSCs. The data were converted from V (vs NHE) to eV (vs absolute electrode potential) with the additive constant of -4.6.

Table A3.2 Redox potentials data of 30 dyes designed for p-type DSSCs used in Fig. 3.4 in section 3.5.

Dye	ΔG_p / eV	ΔG_n / eV	Ref.
TPPC	-3.70	-5.61	[233]
Erythrosin B	-3.49	-5.79	[233]
Triphenylamine dye P1	-3.73	-5.98	[240]

Triphenylamine dye P2	−3.90	−5.90	236
Triphenylamine dye P3	−4.24	−6.00	236
Triphenylamine dye P4	−3.87	−5.92	241
Triphenylamine dye P7	−3.94	−5.88	236
Triphenylamine dye O2	−3.82	−6.00	242
Triphenylamine dye O6	−3.98	−6.04	242
Triphenylamine dye O7	−3.76	−5.95	242
Ruthenium polypyridine 1	−3.88	−6.22	243
Ruthenium polypyridine 2	−3.43	−6.08	243
Ruthenium polypyridine 3	−3.74	−6.17	243
Ruthenium polypyridine 4	−3.61	−6.12	243
Ruthenium NDI Dyad O25	−3.27	−5.29	235
Ruthenium NDI Dyad O26	−4.27	−5.28	235
PMI 18	−3.83	−6.03	244
PMI 19	−3.81	−6.05	244
PMI NDI	−3.94	−6.02	244
PMI PhNDI	−3.83	−5.98	244
PMI PHC60	−3.87	−6.12	244
Arylamine dye 1	−3.64	−5.89	245
Arylamine dye 2	−3.71	−5.87	245
Arylamine dye 3	−3.83	−5.91	245
Arylamine dye 4	−3.74	−6.09	245
Arylamine dye 5	−3.92	−5.97	245
Arylamine dye 6	−3.69	−5.85	245
IrPhen	−2.37	−6.14	213
IrDPQCN2	−3.13	−6.14	213
IrBpystyryl	−2.69	−5.95	213
MEAN / SD	$−3.69 \pm 0.39$	$−5.94 \pm 0.22$	

Chapter 4   Predicting with confidence the PCE of new dyes in DSSC

4.1 Introduction

In this chapter an attempt has been made to construct a predictive model with confidence levels for the PCE (η) of a n-type DSSC with a new dye, based on the known performance of existing dyes. In the development of n-type DSSCs an important fraction of the research effort was devoted to the synthesis and testing of new dyes as introduced in Chapter 1. The very few design rules[49,246–248] emerged over the past years have guided the exploration of a large set of dyes that, when tested under standardized conditions and fabrication methods, should inform the development of new dyes. As such, this provided the motivation to develop a tool for predicting the resulting PCE by 'considering the history of dyes in n-type DSSCs'.

   In the field of DSSCs it was very common to build models for systems under investigation starting from physical principles, and this type of predictive modeling, such as the model introduced in Chapter 3 and many others,[67,85,249] has been part of the development since the early days. The understanding of the physical principles that governed the PCE of the device was however very challenging, due primarily to the intervening interactions between different components in the device and largely remained unclear. QSAR, one of the main approaches currently used to rationalize large medicinal chemistry data sets (see Chapter 2), could be effective for identifying correlations or lack of correlations between properties, and contribute to the understanding of the underlying physical principles for a given problem. In any case, it can be used to narrow

down the exploration of materials, when synthesis and testing constitute the slower step.

4.2 Dataset and general computational strategy

To build a structure–property relation for dyes in n-type DSSCs, a sufficiently large database of dyes tested under similar conditions, such as with the same electrolyte and similar fabrication methods, was needed. In any convincing statistical analysis the data cannot be handpicked, and it was also desirable that they were derived from a relatively uniform set; in this case, for example, a set of dyes with related chemical characteristics such as a shared anchoring group would be desirable. To address both issues we considered dyes listed in Table 1 of the review by Mishra et al.,[49] all being synthetic neutral organic dyes tested in similar devices. To avoid the risk of involuntary bias, new dyes reported after the publication of Mishra et. al.[49] review were not included. Dyes were also excluded from the review that did not have the common carboxylic anchoring group.

We aimed to find some correlations between the properties of the dyes that can be accessed very easily via routine quantum chemistry calculations and PCE. We could then compute these properties for a new dye and predicted the probability that its PCE in a DSSC was larger than a given threshold. The calculations should be relatively inexpensive so that all the calculations could be performed semi-automatically for all dyes considered. Importantly, such a procedure was useful only if new potential dyes could be screened rapidly after the statistical regression. It has been tested that, for dyes that required more than 760 basis functions for the electronic structure calculation, our devised automatic procedure would be inefficient and a manual optimization would be required.

These dyes with over 760 basis functions were also excluded in our dataset employed. The systematic and random errors in these computed properties would be fully accounted for in a statistical analysis.

By considering all these criteria a dataset of 52 dyes was established. The dyes employed should be representative of the chemical structures used for organic dyes and the PCE were very broadly distributed, with an average of 5.61% and a standard deviation of 1.95%. This suggested that, until the data have been collected, there was no evident bias toward reporting only high performing dyes, where our dataset contained the same number dyes of high (> 8%) and low (< 3%) PCE.

4.3 Predictors

As demonstrated by a broad range of QSAR, there was no best or conclusive way to select predictors for statistical analysis (see Chapter 2), and it was expected that the selections of predictors could be improved in the future. In this attempt, computable predictors that were sufficiently independent from one another, easy to evaluate, and expected to influence the PCE of the device from physical considerations were included.[250] Importantly, it was not possible to increase the number of predictors for a given data set without risking an over-fitting of the data. Therefore, the common rule-of-thumb of not having more than 1 fitting parameter for every 10 data points was followed. A list of the predictors considered in this analysis with motivation and a description of the computational methods is given in the following sections. All quantum chemical calculations required for the predictors were performed with Gaussian03.[229]

4.3.1 Free energy of oxidation and reorganization energy of dye

These properties of dyes entered into the theory of charge recombination (see Chapter 2) and were therefore deemed relevant to the PCE. The free energy of oxidation (ΔG_n) and reorganization energy (λ_n) were approximated from the total energy differences computed using the B3LYP functional[251] and the 3-21G* basis set, where all structures were optimized at this level. The Polarized Continuum Model (PCM)[228] was included to mimic solvent effects in DSSCs, using the solvent parameters appropriate for acetonitrile (dielectric constant = 36.64). The model was built with a solvent excluding surface (SES), where the overlapping index between two interlocking spheres was 0.8 and the minimum radius was 0.5 Å. Construction of the SES was based on the GePol method[252] and the set of atomic radii was defined according to the UAKS model.[253] The computational approach to calculate ΔG_n and λ_n was identical to that adopted by Maggio et. al.,[87] except the approximation that only part of the molecule was exposed to the solvent was not taken into account (see section 2.2.3). It should be noted that the component of the reorganization energy associated with the iodide/triiodide redox shuttle in a number of possible chain of reactions[254] was neglected as it was expected to be common to all dyes.

4.3.2 Spectral overlap

It was expected that high PCE was associated with strong absorption of solar radiation. The absorption spectrum for each dye was computed with the inclusion of solvent, and the overlap ($\tilde{S}_k$) between the computed absorption spectrum for dye k ($\varepsilon_k(E)$), and the solar spectrum ($P(E)$):

$$\tilde{S}_k = \int P(E)\varepsilon_k(E)dE \tag{Eq. 4.1}$$

was evaluated. To have convenient data, $\tilde{S}_k$ was normalized to the value (3.27) of one of the dyes ($S_k = \tilde{S}_k / S_0$; dye 50 in table 1 in Mishra et. al.[49] review).

The excitation energies and their corresponding oscillator strength for the computation of absorption spectrum were obtained by single-point TDDFT calculations with the 6-31G* basis set, B3LYP functional[251] and solvent effects mimicked based on PCM.[228] The input geometries were the optimized geometries of neutral dyes from B3LYP/3-21G* calculations. The number of excited states included in the calculation was set to 11 as the lowest energy of the 11th excited state is typically closer to 4 eV with the lowest energy among all dyes being 3.5 eV. The calculation of the dye absorption above 3.5 eV was however not very important for DSSCs since the optical band gap of TiO_2 is just above 3 eV.[255] The solar spectrum employed was the AM 1.5 direct normal plus circumsolar spectrum taken from ASTM G173-03.[256,257] The simulated absorption spectrum of dye k, $\varepsilon_k(E)$, was computed by:

$$\varepsilon_k(E) = \left(2\pi\sigma^2\right)^{-1/2} \sum_i f_i \exp(-(E_i - E)^2 / 2\sigma^2) \qquad \text{(Eq. 4.2)}$$

where E_i and f_i were the excitation energy and the dimensionless oscillator strength for i[th] transition respectively, σ was a broadening parameter of 0.2 eV. The integral for computing $\tilde{S}_k$ was evaluated numerically in the range between 0 and 10.3 eV.

4.3.3 Orbital Asymmetry

A good fraction of dyes, often referred to as donor-π-acceptor (D-π-A) dyes, were synthesized to have a large orbital density of the LUMO and a small orbital density of the HOMO on the anchoring group, such as a carboxylic acid group, so that charge injection was favored and charge recombination was prevented. To

represent the strength of this orbital distribution of dyes, a quantity known as orbital asymmetry (*OA*) was defined as the log of the ratio of the orbital density of LUMO (OD_{LUMO}) to the orbital density of HOMO (OD_{HOMO}) on the anchoring group.[170] The *OD* of the molecular orbitals was computed by:

$$OD_{MO} = \sum_{i(anchor),j} c_{MO,i} c_{MO,j} S_{ij} \qquad \text{(Eq. 4.3)}$$

where S_{ij} was the overlap between basis set function i and j, $c_{MO,i}$ were the molecular orbital coefficients, *MO* was either LUMO or HOMO of the dye k. i ranged over all basis functions on the anchoring group of the molecule and j ranged over all basis functions of the molecule. The calculations on the neutral dyes were used to obtain the molecular orbital coefficients.

4.3.4 Normal dipole density

The dipoles of the ground-state dyes were thought to affect the conduction band level of the semiconductor if the dyes were in a similar orientation with respect to the surface, and hence affecting the V_{oc} of the cell. Such effect of the dipole on V_{oc} has been demonstrated with a few acid derivatives adsorbed on TiO2 surface.[258,259] We assumed that the orientation of the dye was guided by the carboxylic anchoring group oriented on the surface as in a calculation of a benzoic acid on TiO$_2$. The following components were therefore evaluated: 1.) the component of the dipole of the dye k perpendicular to the surface $\mu_{k,z}$ and 2.) the area of the same dye on the TiO$_2$ surface A_k. The normalized dipole density (*NDD*) for dye k was given by:

$$NDD_k = \frac{\mu_{k,z}}{A_k} \qquad \text{(Eq. 4.4)}$$

In the computation of *NDD*, all dyes had the same (carboxylic) anchoring

group and were assumed to adopt the same orientation on the anatase(101) surface as a model compound benzoic acid,[65] where the adsorption mode was mono-dentate non-dissociative molecular (Fig. 4.1(a)). The optimized geometry of each dye was rotated into the orientation of the benzoic acid on anatase(101), so that the plane of the C-COO group of the reoriented molecule was the same as in the benzoic acid and the C-COO bond was pointing in the same direction as the reference (Fig. 4.1(b)). The anatase(101) plane in the reference structure was perpendicular to the z Cartesian axis so that the molecular electrical dipole moment of the rotated molecule in the z direction was used to establish any correlation between surface dipole and PCE. The area occupied by the molecule in the xy plane was estimated by projecting the atomic coordinates on the plane, assuming that each atom spanned a circle with radius of 1.5 Å.
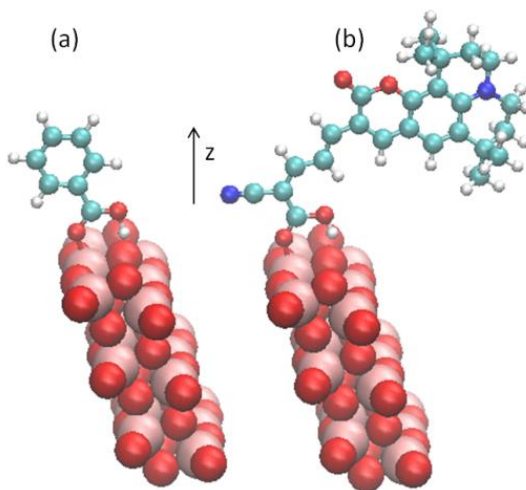
Figure 4.1 (a) Structure of the optimized model anchoring molecule on anatase(101) and (b) assumed structure of a dye[260] on the surface for computing the surface coverage and electrical dipole moment perpendicular to the surface.

4.3.5 Other predictors

Since the rule-of-thumb of 1 predictor for 10 data point to avoid over-fitting was

followed, a number of other initially considered predictors, which were either correlated strongly with the chosen predictors, or lack of a strong hypothetical reasoning to correlate with PCE, were therefore excluded in the model. The alternative predictors that were correlated with chosen predictors included 1.) The HOMO-LUMO gap of a dye, which correlated strongly with the absorption spectrum (first excitation energy), and acquired from single-point TDDFT calculations with B3LYP/6-31G* in PCM; 2.) the volume of a dye, or the cavity volume of a dye in PCM calculations, which correlated strongly with λ_n and acquired from B3LYP/3-21G* optimization of neutral dyes with PCM based on the GePol method; 3.) the HOMO energy of a dye, which correlated strongly with ΔG_n and acquired from B3LYP/3-21G* optimization of neutral dyes with PCM. Fig. 4.2 shows the correlations between these alternative predictors and the chosen predictors. For indicative purpose, the correlation coefficients are provided in Table 4.1 for the 5 predictors included in the initial model. The correlations were much weaker than those illustrated in Fig. 4.2, with the largest correlation found between λ_n and S.

Other considered predictors that were not computed include: (i) Electronic properties of the excited states, which may affect the charge injection rate, but considering charge injection was the fastest process in the device it rarely determined the PCE;[45] (ii) single-triplet splitting,[261] which may be important for metal-containing dyes, but these dyes were not included in our dataset; (iii) simple topological predictors, such as the number of aromatic rings in a dye,[189] in which their physical relevance to PCE were not clearly known.
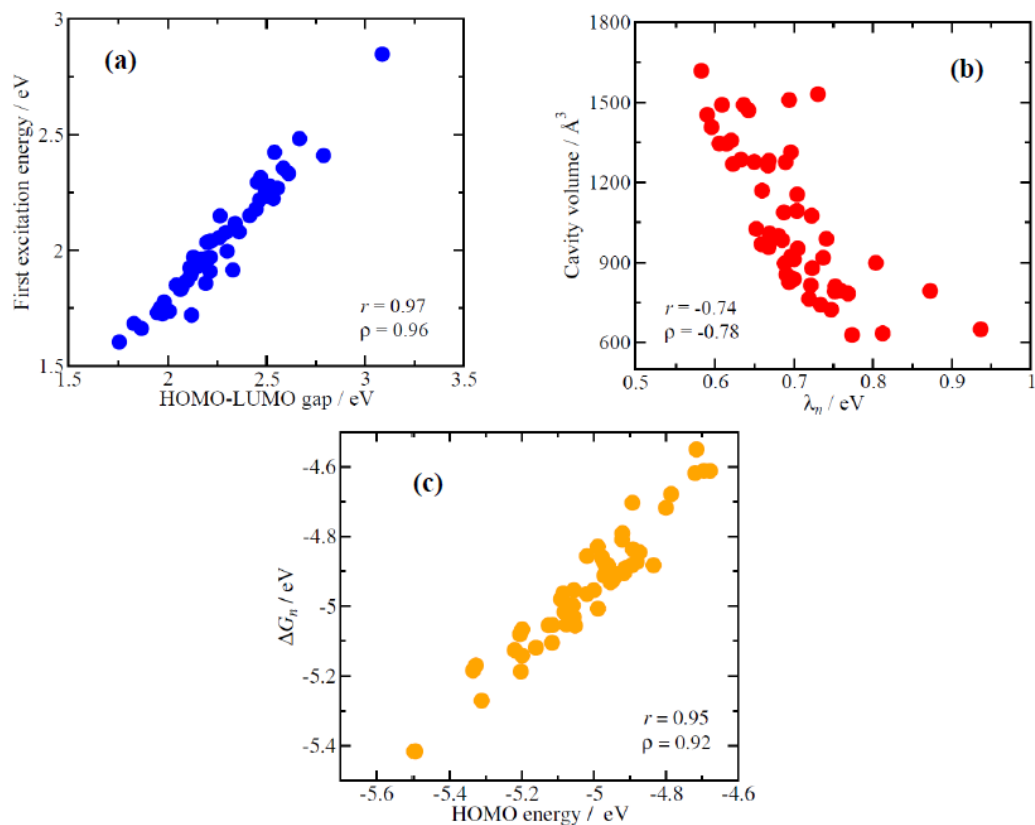
Figure 4.2 Correlations between (a) first excitation energy (eV) and HOMO-LUMO gap (eV) of a dye (b) cavity volume ($Å^3$) and λ_n (eV) of a dye and (c) ΔG_n (eV) and HOMO energy (eV) of a dye. r is the Pearson's correlation coefficient and ρ is the Spearman's correlation coefficient.

Table 4.1 Pearson's r and Spearman's ρ correlation coefficients between the five predictors used in the main statistical analysis. The Pearson's r (Spearman's ρ) coefficients are in higher (lower) triangle of the square table.

	λ_n / eV	ΔG_n / eV	S	NDD / D/$Å^2$	OA
λ_n / eV	-	−0.20	−0.60	0.25	0.05
ΔG_n / eV	−0.10	-	0.36	−0.04	0.03
S	−0.50	0.25	-	−0.17	−0.36
NDD / D/$Å^2$	0.36	0.04	−0.24	-	0.22
OA	−0.23	0.16	−0.01	0.13	-

4.4 Correlations between selected predictors and PCE

Since the correlation between PCE and J_{sc} was stronger than that between PCE and V_{oc} (Fig. A4.1), we have attempted to establish a relationship between the five computed properties and the expected PCE (η_{exp}) in the form of a function $\eta_{exp}(\Delta G_n, \lambda_n, S, NDD, OA)$. A diagram of the measured PCE against the computed parameters (Fig. 4.3) immediately gave some useful indication. It appeared that there was an important correlation between λ_n and η, with high η associated with small λ_n as suggested by phenomenological models. A correlation was also evident between the computed ΔG_n and η, i.e. it seemed that high η were found in a range of ΔG_n as expected from microscopic theories, which also suggested that ΔG_n would affect η nonlinearly.[168] Surprisingly, no correlation was evident in the plots of measured η against S, NDD and OA. More quantitatively, Fig. 4.3(f) shows the Spearman ρ^2 statistics[207] for each predictor and suggested a higher degree of correlation, potentially non-linear and non-monotonic, between η and the predictors ΔG_n and λ_n.
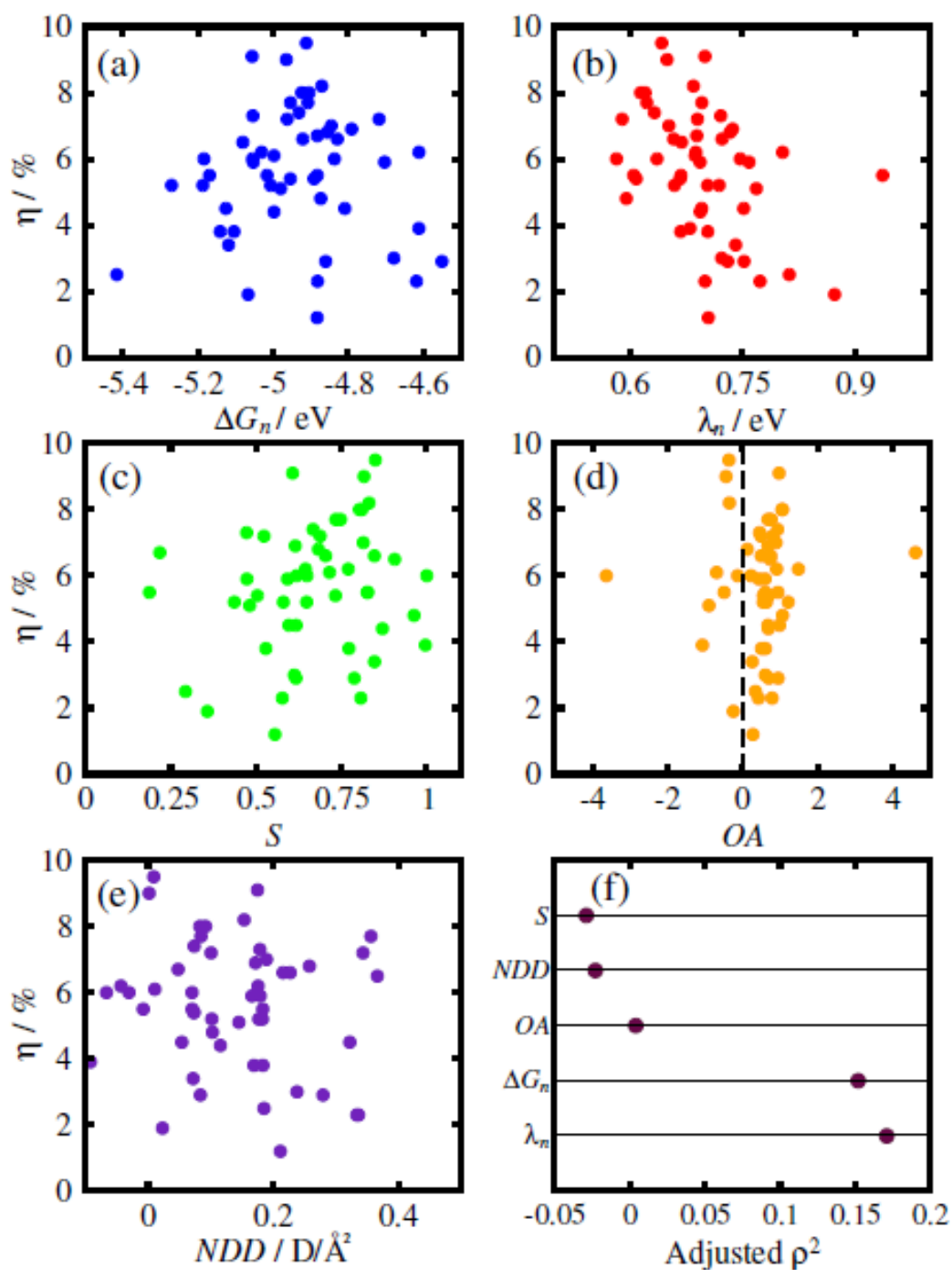
Figure 4.3 (a)-(e) Correlations between PCE of the dyes in our dataset against five computable parameters ΔG_n, λ_n, S, OA, NDD. (f) Strength of marginal relationship between predictors and PCE using the Spearman ρ^2 statistics.

4.5 Construction of predictive model for PCE

A model was first constructed between η_{exp} and the predictors with an intuitive approach, and subsequently a more rigorous statistical procedure was considered. Both models would be introduced in the following, with detailed account of the procedure and computed statistics provided in the appendix (sections **A4.2** and **A4.3**).

4.5.1 An intuitive model

On the basis of the visual inspection of Fig. 4.3(c-e), the roles of *S*, *NDD* and *OA* were ignored in this intuitive approach, and a model (Fig. 4.4(a)) was constructed with the simplest 5-parameter non-linear function of ΔG_n and λ_n:

$$\eta_{\text{exp}} = a + b\Delta G_n + c\Delta G_n^2 + d\lambda_n + e\lambda_n^2 \tag{Eq. 4.5}$$

where the parameters *a*, *b*, *c*, *d*, *e* can be uniquely determined to minimize the squared difference between η_{exp} and measured η. The residuals, or the difference between predicted and actual values, were normally distributed with standard deviation $\sigma_n = 1.67$ % (Fig. 4.4(b)), and therefore it was possible to predict the probability that, for a dye with a computed $(\Delta G_n, \lambda_n)$ pair, the η was higher than a given threshold η':

$$P(\eta > \eta') = \left(2\pi\sigma_\eta^2\right)^{-1/2} \int_{\eta'}^{\infty} \exp\left(-\left(\eta - \eta_{\text{exp}}\right)^2 \big/ 2\sigma_\eta^2\right) d\eta \tag{Eq. 4.6}$$
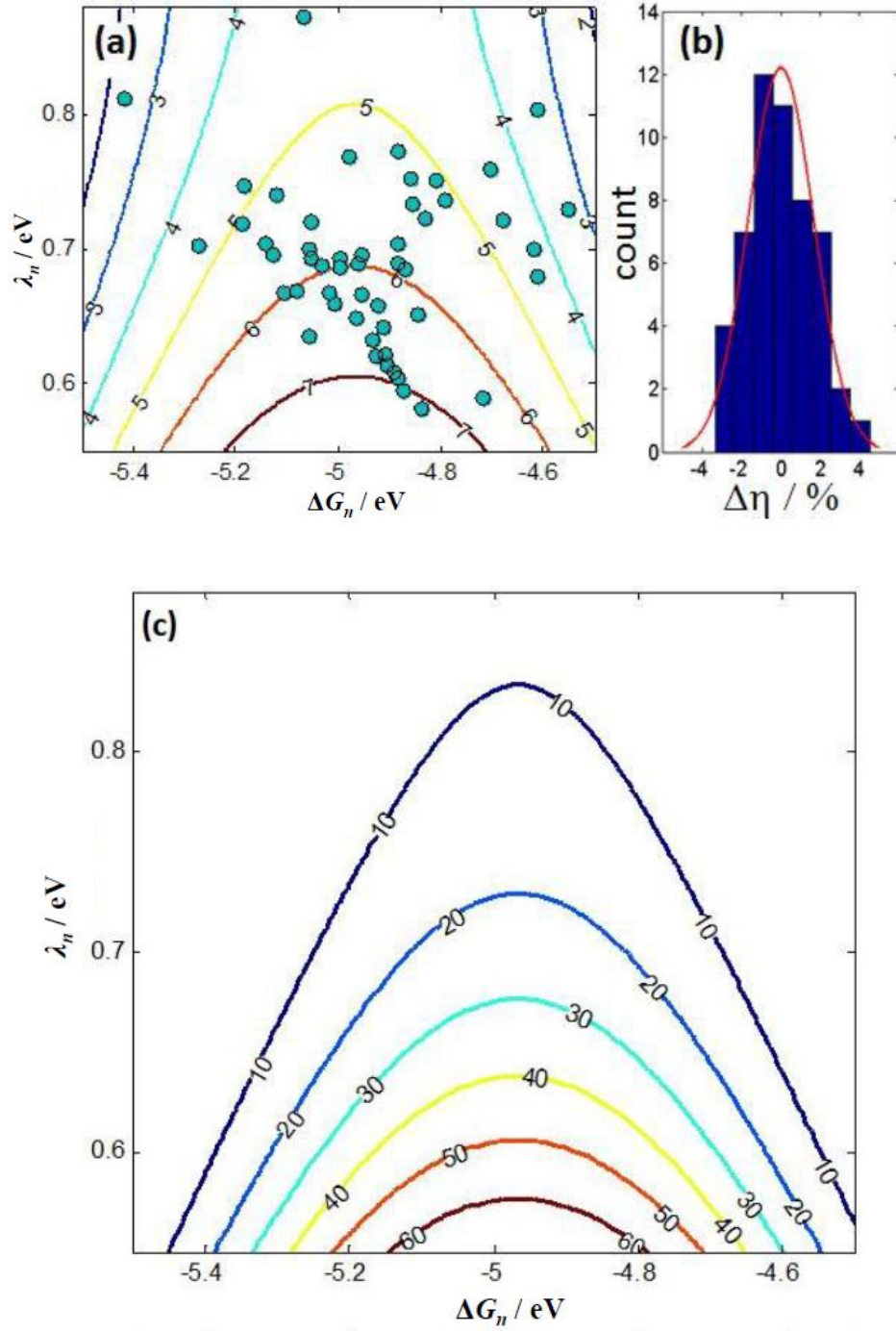
Figure 4.4 (a) Predicted η (%) from the fitting in Eq. 4.6 with an indication of the data points included in the fitting procedure. (b) Distribution of the difference between "predicted" and actual η values. (c) Map of the probability (%) that η exceeds 7% as a function of the computed parameters ΔG_n, λ_n following the polynomial fit in Eq. 4.5.

Fig. 4.4(c) shows the probability that the η was higher than 7% as a function of computed ΔG_n and λ_n. Interestingly, there were "accessible" regions of the map with probability higher than 60% and lower than 20%, implying that the map was a good tool for planning the synthesis of new dyes considering it took only a few minutes to set up the calculation of ΔG_n and λ_n, and perhaps a few hours for their execution on a standard desktop computer.

4.5.2 A rigorous model

A more rigorous procedure was based on the construction of a generalized linear model where the η_{exp} was initially expressed as

$$\eta_{\text{exp}} = \beta_0 + g_{\Delta G_n}(\Delta G_n; \beta_1) + g_{\lambda_n}(\lambda_n; \beta_2) + \beta_3 S + \beta_4 NDD + \beta_5 OA \qquad \text{(Eq. 4.7)}$$

where $g_{\Delta Gn}(\Delta G_n; \beta_1)$ was the link function for ΔG_n with parameter β_1 to be estimated, $g_{\lambda n}(\lambda_n; \beta_2)$ was the link function of λ_n with parameter β_2 to be estimated, and β_3, β_4 and β_5 were the parameters to be estimated for S, NDD and OA respectively. Eq. 4.7 was linear in S, NDD, and OA and contained linear and non-linear components in ΔG_n and λ_n, although the overall function would still be linear in all the parameters. In particular, $g_{\Delta Gn}(\Delta G_n; \beta_1)$ and $g_{\lambda n}(\lambda_n; \beta_2)$ expanded ΔG_n and λ_n into restricted cubic splines with parameter vectors β_1, β_2 respectively.[207] The spline expansions were defined uniquely from the data for ΔG_n and λ_n; this was a well-established methodology to include non-linear terms in regression procedures where the analytical form of the non-linearity could not be derived from a physical basis.

Eq. 4.7 contained too many fitting parameters with respect to the 52 data points available, which would violate the empirical rule of thumb of allowing one parameter for, at most, 10 data points. The initial fitting was therefore

performed using a statistical penalization procedure,[207] based on using the corrected AIC (Eq. 2.26) to penalize the likelihood ratio (LR) of the model. The analysis of variance (ANOVA) (Table A4.3) of the fitting confirmed that there was no evidence of correlation between the predictors *S*, *NDD*, *OA* and η. A reduced model was therefore built from the total model (Eq. 4.7) by "simplification by approximation",[207] as introduce in section 2.3.2, which produced the fitting as

$$\eta_{\exp} = \beta_0 + g_{\Delta G_n}(\Delta G_n; \beta_1) + g_{\lambda_n}(\lambda_n; \beta_2) \qquad\qquad \text{(Eq. 4.8)}$$

The standard deviation of the residuals for this more advanced model was 1.71% and, as before, it was possible to predict the probability that η was higher than a given threshold for any values of computed ΔG_n and λ_n. Fig. 4.5 shows a map with the probability of η higher than 7% with this more accurate model. The differences between the intuitive and the rigorous procedures were not large but the rigorous procedure guaranteed that the effect of potentially more complex nonlinearities was not neglected. In addition, the functional form in Eq. 4.8 produced more conservative estimates outside the region where data points were present, while the polynomial fit of Eq. 4.5 produced unphysical estimates in these regions. The calibration graph of the model in (Eq. 4.8), obtained by bootstrap re-sampling (Fig. A4.4), shows that this model is much more appropriate than the model based on (Eq. 4.5) (*c.f.* Fig. A4.3).[207]

The proposed map could be used to either direct the synthesis of new dyes where the maximum η was predicted, or prepare dyes in the region of the map where there were few or no data points, to learn more about the system in these conditions. The advantage of this statistical approach was that the confidence intervals of the prediction included both the existence of effects and parameters

that were not included in the predictors, as well as the inaccuracies of both the computational chemistry and the experimental measurements: all inaccuracies and missing effects would simply decrease the level of confidence of the prediction. Considering that new families of DSSC were being used, for instance, with different electrolytes, it was believed that the construction of a similar map, with perhaps a larger set of data and predictors, should constitute a priority in the rational exploration of the chemical space.
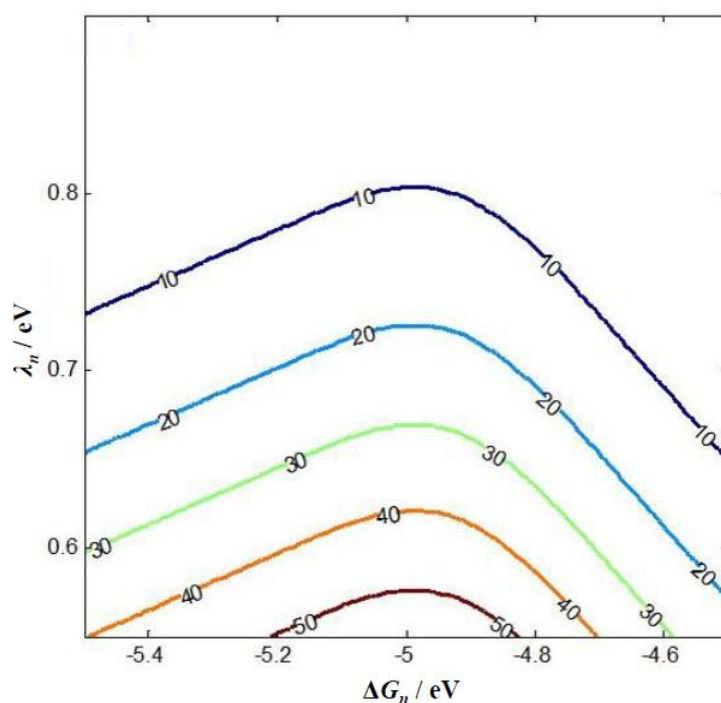
Figure 4.5 Map of the probability (%) that η exceeds 7% as a function of the computed parameters ΔG_n, λ_n based on Eq. 4.8.

It is also important to stress the difference between our approach, where correlations were searched for between computables and a target experimental property, and the alternative computational tools for material discovery that generated a large set of "theoretical" materials and directly computed the property of interest, such as the band gap or other electronic properties. The latter

approach was particularly suitable when the underlying physics was relatively well-understood and the direct computation of the property of interest was possible. For n-type DSSCs it was not possible to compute the η from first principles and a closer alliance between theory and experiment was therefore necessary.

Finally, such analysis in larger and unbiased data sets offered the best opportunity to validate some hypotheses put forward to describe the physics of DSSC. After considering the results, it was not too surprising that the overlap with the solar radiation did not correlate with the PCE, possibly due to cells with small absorptance were not even reported and beyond a threshold of absorptance the PCE did not change. On the other hand, it was quite surprising to see that there was no effect in having HOMO and LUMO localized in different regions of the dye, considering the enormous effort put into the preparation of large families of D-π-A dyes. The efficacy of D-π-A character on PCE of dyes was further investigated in Chapter 5.

4.6 Conclusions

A general method was proposed to predict the PCE of n-type DSSCs with new dyes from easily computable quantities, including, for the first time, the degree of confidence of such predictions. Carboxylated organic dyes studied with iodide/tri-iodide electrolyte were considered, but the method could be applied to a different family of DSSCs and the accuracy of its prediction could be improved over time by expanding the set of data and/or the set of predictors.

4.7 Appendix

A4.1 Correlations between PCE and J_{sc}, and PCE and V_{oc}

Fig. A4.1 shows the correlations between PCE and J_{sc}, and PCE and V_{oc} of the 52 dyes in the dataset. The Pearson's r indicated that the PCE was more strongly influenced by the J_{sc}, suggesting that it was more suitable to use predictors that were relevant to the dye's properties, rather than other components that influenced the V_{oc} of the device, such as the electrolyte's properties.
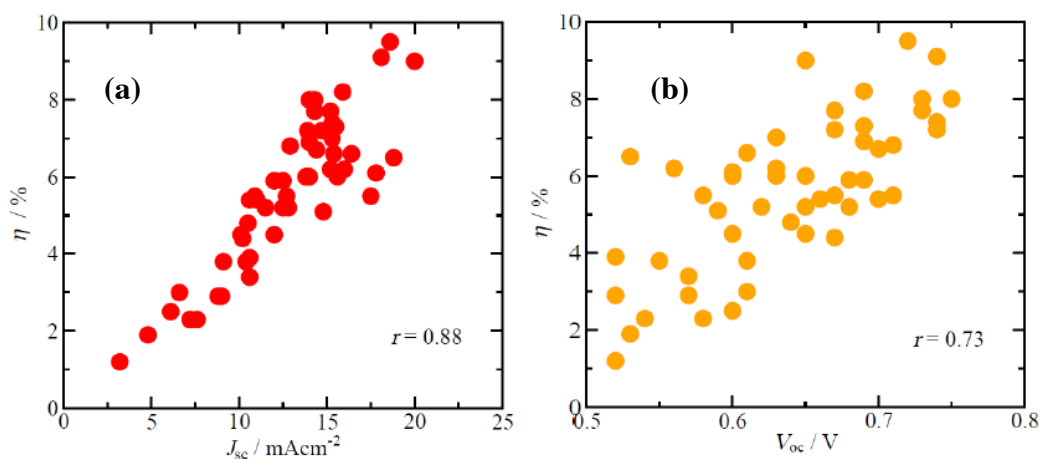
Figure A4.1 (a) Correlation between PCE and J_{sc} of the 52 dyes in the dataset. (b) Correlation between PCE and V_{oc} of the 52 dyes in the dataset.

A4.2 Model construction with an intuitive approach

The results of fitting the data with the expression $\eta_{exp} = a + b\Delta G + c\Delta G^2 + d\lambda + e\lambda^2$ were given in Table A4.1, including the 95% confidence interval in the fitting parameters. The interval was reported for completeness but it was not meaningful considering the nature of the data, i.e. it was known that additional effects beyond ΔG_n and λ_n contribute to the coupling. A better way to evaluate the quality of the fitting was to consider the distribution of the difference between actual PCE and PCE computed by the fitting above (Fig. 4.4(b), the standard

deviation was 1.67 %).

Table A4.1. Results of polynomial fitting for the intuitive model.

Parameter	Fitted value	95% confidence interval	Units
A	-290.1	-585.6 : 5.32	%
B	-125.9	-243.0 : -8.77	% eV^{-1}
C	-12.67	-24.54 : -0.804	% eV^{-1}
D	-37.22	-128.28 : 53.8	% eV^{-2}
E	19.32	-42.64 : 81.38	% eV^{-2}

A4.3 Model construction with a rigorous approach

A quantile plot of the response variable has been preliminarily constructed with respect to the Gaussian distribution (Figure A4.2), showing that its distribution could be reasonably approximated by a Gaussian distribution, as implied in the simplified model.
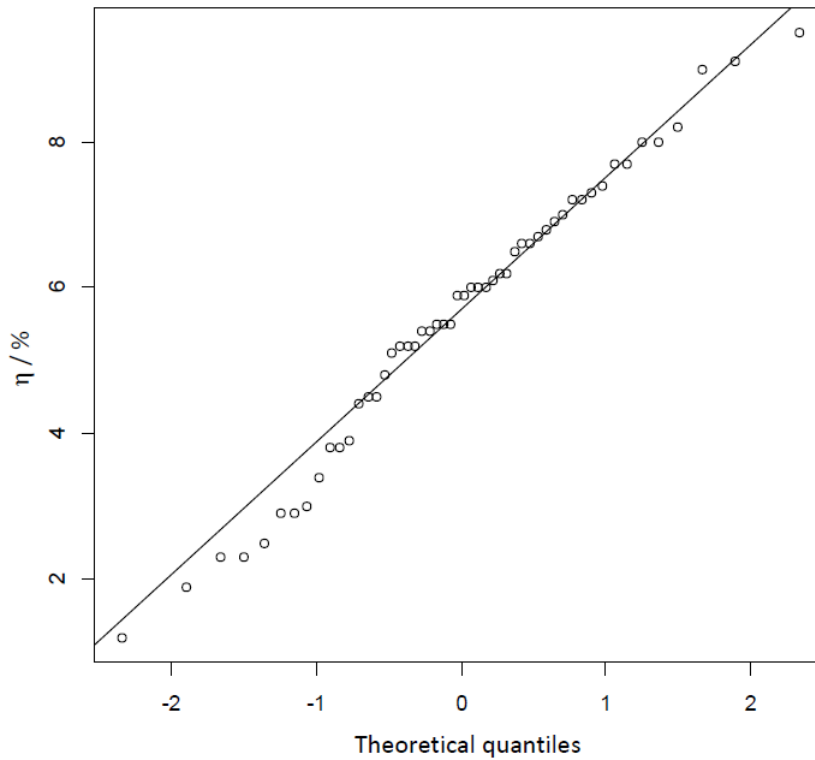
Figure A4.2. Q-Q plot of the variable with respect to the Gaussian distribution.

Based on Fig. A4.2, the approximation of the distribution of η_{exp} is satisfied by a Gaussian distribution. In a generalized linear model the vector of the predictors and possible non-linear functions of them was indicated by $\mathbf{X}$, where in our case, for example, $\mathbf{X} = (\Delta G, \lambda, S, DD, OA, g_1(\Delta G), g_2(\lambda), \ldots)$. The conditional expected value of the PCE on a particular set of predictors was expressed formally as

$$E[\eta \,|\, \mathbf{X}] = f(\mathbf{X}; \beta) \qquad\qquad\qquad \text{(Eq. A4.1)}$$

where β was a parameter vector to be estimated and f was a nonlinear additive function of the predictors, which was $f = g_1(X_1; \beta_1) + g_2(X_2; \beta_2) + \ldots$. However, each of the g_k was a linear function of β_k, so overall f was a linear function of β. A simple linear regression model, which was when g was the identity function and $\mathbf{X}$ was the vector of the five linear predictors, gave rather poor results, as illustrated by the calibration graph obtained with bootstrap resampling[207] (Figure A4.3). The bootstrapped shrinkage estimate is 0.75,[207] which denotes rather poor validation performance (i.e. about 25% lack of fitting).

The poor fit might be explained with effective lack of complexity (as the above model is linear in the data). We would therefore consider a more complex model; given the information provided by the Spearman ρ^2 statistic (Fig. 4.3(f)) and the limited sample size, nonlinearities for ΔG_n and λ_n were introduced with penalisation also considered. In the intuitive model higher powers of ΔG_n and λ_n were introduced in the fitting function but this was far from ideal when there was no physical reason for such an expansion; furthermore, polynomials have some undesirable properties, most notably non-locality, where the fit in one region could be greatly affected by data in another; polynomials also tend to infinity considerably rapidly when predicting outside of the range of the data used for the fit.
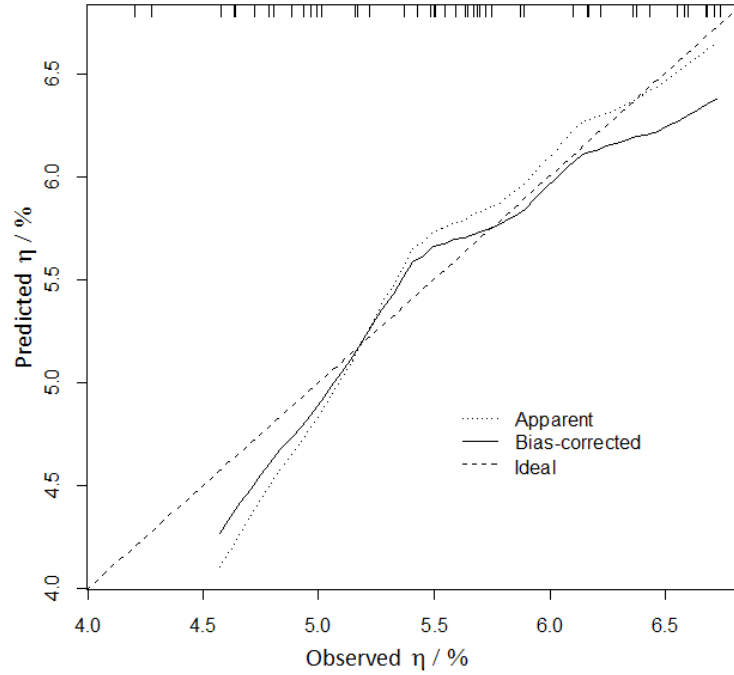
Figure A4.3. Predicted η against observed η for a simple linear regression. One hundred bootstrapped predictions of the model were averaged. Apparent and bias-corrected estimates were reported. The mean absolute error for the bias-corrected estimates is 0.149. A perfect fit would lie on the dashed line. The tags on the upper part of the graph denote the coordinates of the predicted points (the graphs are obtained by smoothing).

A more general approach was to construct restricted, or natural, cubic regression splines, where if using third order restricted cubic spline as an example the nonlinear function took the form:

$$g(X;\boldsymbol{\beta}) = \beta_0 + \beta_{\text{lin}} X_1 + \beta_{\text{nonlin}} X_2 \qquad \text{(Eq. A4.2)}$$

where $X_1 = X$,

$$X_2 = (X - t_1)_+^3 - \frac{t_3 - t_1}{t_3 - t_2}(X - t_2)_+^3 + \frac{t_2 - t_1}{t_3 - t_2}(X - t_3)_+^3 \qquad \text{(Eq. A4.3)}$$

and $(z)_+$ was equal to z if $z > 0$ and zero otherwise. Note $\boldsymbol{\beta} = (\beta_0, \beta_{\text{lin}}, \beta_{\text{nonlin}})$ was a parameter vector. Higher order splines could be defined but were not used here. Note that constant and linear terms were included in the expansion, and that for $X \geq t_3$ the function was linear. The parameters t_1, t_2 and t_3 were known as the knots of the spline and were determined uniquely for the data set, being

located respectively at the 0.1, 0.5 and 0.9 quantile of the data for the predictor X. For the predictor ΔG_n, the function $g_{\Delta G_n}(\Delta G_n)$ contained nodes $t_{1,\Delta Gn}$, $t_{2,\Delta Gn}$ and $t_{3,\Delta Gn}$ at -5.1685, -4.9277 and -4.6784 respectively; for the predictor λ_n the function $g_{\lambda_n}(\lambda_n)$ has nodes $t_{1,\lambda n}$, $t_{2,\lambda n}$ and $t_{3,\lambda n}$ (at 0.6086, 0.6936 and 0.7734 respectively).

A general model containing all five predictors and the non-linearity in ΔG_n and λ_n would have the following form:

$$E[\eta \,|\, \Delta G_n, \lambda_n, S, DD, OA]$$
$$= \beta_0 + g_{\Delta G}(\Delta G_n; \boldsymbol{\beta}_1) + g_\lambda(\lambda_n; \boldsymbol{\beta}_2) + \beta_3 S + \beta_4 DD + \beta_5 OA \qquad \text{(Eq. A4.4)}$$

The number of parameters in the above model was 8, where two for each of the spline expansions; note that the intercept β_0 was common to both and was made explicit, so it was not double-counted to avoid identifiability issues. A widely accepted heuristic was to consider $1/20^{\text{th}}$ to $1/10^{\text{th}}$ of the sample size as the upper bound for the number of parameters in the model. Violation of this limit could produce biased results and over-fitting. However, it was possible to go beyond this limit by penalizing model fitting criteria for complexity. The likelihood ratio (LR) χ^2 statistic of the model was maximized with respect to the "null" model, which was the model without any predictors, and only β_0 was trivially "fit" to the average of the response. The correct AIC (Eq. 2.26) was then employed to penalize the LR.

The ANOVA table of the penalized model (Table A4.2) shows that there is some predictive power in ΔG_n and λ_n, since they have the highest F statistics. Notice that the relation with λ_n seemed to be mostly linear. Although it would be tempting to simplify the model and remove the remaining variables, this would also bias the results and produced optimistic estimates of the parameters'

covariance matrix, since we had "cheated" by looking at a more complex hypothesis, and had already expended the degrees of freedom.

Table A4.2 ANOVA table of full model.

Factor	Degrees of freedom	*F* test statistics	*p*-value
ΔG_n	2	1.39	0.26
nonlinear	1	2.57	0.12
λ_n	2	3.43	0.04
nonlinear	1	0.37	0.55
NDD	1	0.07	0.79
S	1	0.44	0.51
OA	1	0.24	0.62

By following the approach of "simplification by approximation", as introduced in section 2.3.2, a final model was produced in which confidence limits and statistical tests were unbiased, and included the effects of model selection. The simplified model, with a new response value *Z*, could be written as

$$E[Z \mid \Delta G_n, \lambda_n] = \beta_0 + g_{\Delta G_n}(\Delta G_n; \boldsymbol{\beta}_1) + g_{\lambda_n}(\lambda_n; \boldsymbol{\beta}_2) \qquad \text{(Eq. A4.5)}$$

The calibration graph in Fig. A4.4 shows the good performance of the reduced model. The bootstrapped shrinkage estimate was 0.99 and denoted good validation performance, with about 1% lack of fitting. The ANOVA table of the reduced model and its final parameters were given in Tables A4.3 and Table A4.4.
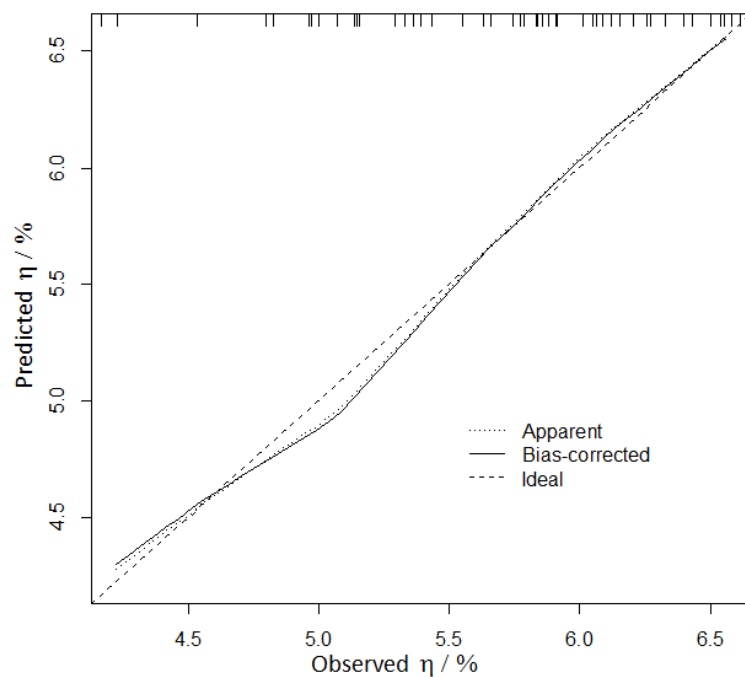
Figure A4.4. Predicted η against observed estimated PCE Z. One hundred bootstrapped predictions of the model were averaged. Apparent and bias-corrected estimates were reported. The mean absolute error for the bias-corrected estimates was 0.044. A perfect fit would lie on the dashed line. The tags on the upper part of the graph denoted the coordinates of the predicted points (the graphs are obtained by smoothing.)

Table A4.3. ANOVA table of reduced model.

Factor	Degrees of freedom	*F* test statistic	*p*-value
ΔG_n	2	42.48	<0.0001
nonlinear	1	48.86	<0.0001
λ_n	2	141.04	<0.0001
nonlinear	1	4.32	<0.0001

Table A4.4. Estimated coefficients of general linear regression model fit (reduced model).

Parameter	Fitted value	95% confidence interval	Units
β_0	20.26	15.54:24.99	%
$\beta_{1,\text{lin}}$	1.531	0.576:2.486	$\% \text{ eV}^{-1}$
$\beta_{1,\text{nonlin}}$	-3.699	-4.860:-2.537	$\% \text{ eV}^{-3}$
$\beta_{2,\text{lin}}$	-9.613	-12.373:-6.853	$\% \text{ eV}^{-1}$
$\beta_{2,\text{nonlin}}$	-0.036	-2.556:2.483	$\% \text{ eV}^{-3}$

In conclusion, this more rigorous procedure for the regression of η against a set of predictors shows that the ones with the greatest predictive capability were ΔG_n, which had a strong nonlinear component, and λ_n, which was mostly linear.

Chapter 5 Does the donor-π-acceptor character of dyes improve the PCE of DSSC?

5.1 Introduction

In this chapter we discuss whether the D-π-A design scheme for organic dyes in DSSCs is effective for elevating the PCE by observing simple statistical correlations. The development of organic dyes has in fact been very much dominated by this single design rule.[45,48,49] It has been postulated that a good dye should contain an electron acceptor (A) portion (where the LUMO is localized) close to the surface of the semiconductor and an electron donor (D) portion (where the HOMO is localized) close to the solution, and the two should be connected by a π-conjugated bridge, in which this overall design was then known as the D-π-A design scheme (Fig. 5.1).[45,49,262] In theory such a design scheme should be effective for facilitating charge injection from the dye to the semiconductor (the dye's LUMO interacted strongly with the conduction band states of the semiconductor), and for increasing the efficiency of charge neutralization by the electrolyte with respect to charge recombination (the HOMO should accept more readily an electron from the electrolyte than from the semiconductor).

It was well-known that D-π-A scheme provided not only the advantage of creating long-lived charge separated state, but also great flexibility to modify the chemical structure of dyes. This in turn led to myriad of possibilities to idealize dye's performance, such as adjusting the amount of π conjugation to refine a dye's ability to absorb solar radiation,[49] and inserting alkyl side chain to avoid dye aggregation.[263] As such, there has been a large volume of work devoted to

the widely adopted D-π-A scheme for metal-free organic dyes,[45,49,264] but the PCE based on these dyes has seen little improvement in recent years, in which the pinnacle achieved has remained at around 9%[45,49] over some time.
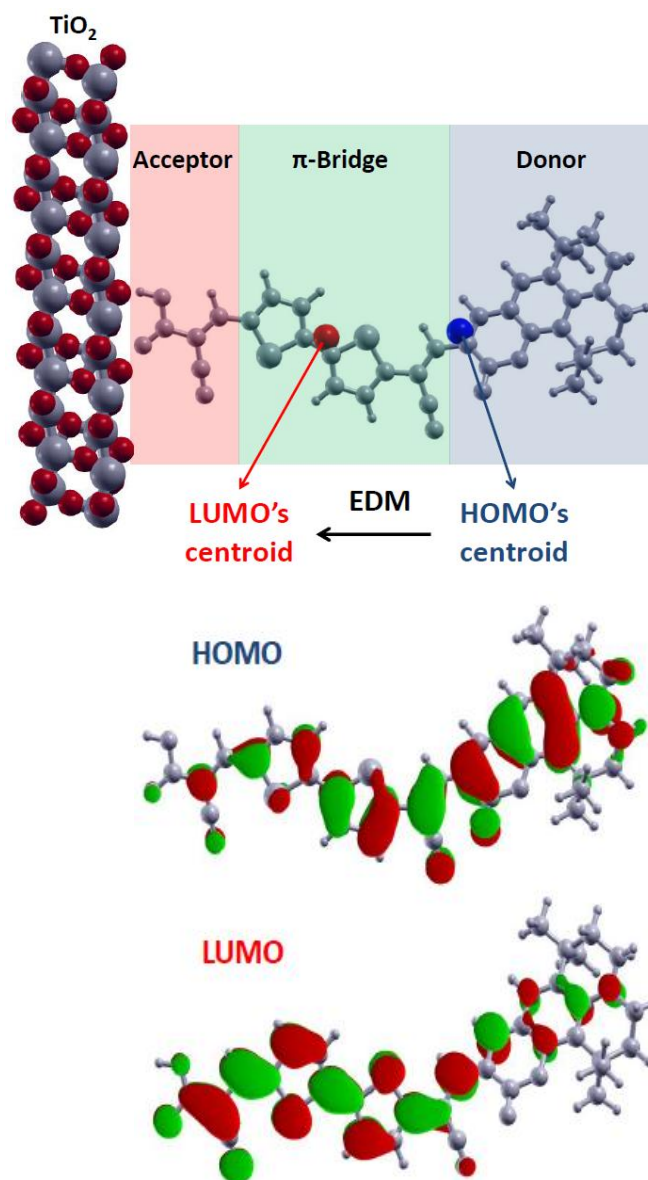
Figure 5.1 (top) Illustration of D-π-A character of a dye[265] adsorbed onto a semiconductor such as TiO$_2$. The acceptor part is close to the semiconductor to facilitate charge injection, whereas the donor part is far from the semiconductor to avoid charge recombination. The *EDM* of a dye is the separation between the LUMO's centroid and the HOMO's centroid. (bottom) Illustration of the HOMO and LUMO of the dye.

It has also been suggested that D-π-A character would supposedly lead to a dye with high reorganization energy, which would negatively impact its performance.[45] All in all this led to the question of whether D-π-A character could truly pave the way to DSSC with higher PCE. We therefore assessed the influence of D-π-A character on PCE with simple statistical analysis. We devised a quantitative measure for D-π-A character, computed this measure for 116 metal-free organic dyes, and examined the correlation between the D-π-A character and PCE of these dyes.

5.2 Constructing a new dataset

The critical prerequisite for any statistical analysis was the availability of an unbiased and sufficiently large set of experimental data, which has not been available for DSSC. A strategy to develop such a set was therefore pivotal for our analysis. The following properties would be essential for the data set: (i) it should represent a good account of historical data but also account for more recent progresses in the field; (ii) it should refer to data collected under similar experimental conditions with sufficient chemical similarity between the dyes; (iii) the data set, which would be unavoidably a subset of all those available, should be collected without being handpicked by the analyst, to exclude any type of bias and, in particular, *confirmation bias*[266] in the dataset.

In order to satisfy these criteria, our data were collected in accordance with the following strategy. We initially considered, as done in Chapter 4, the neutral metal-free organic dyes reported in Mishra et.al.'s review in 2009[49] to sample work precedent to the publication of this review. To sample more recent work but maintain uniformity in the data set we collected data on the PCE of new dyes

subject to the following constraints: (i) dyes contained only a single carboxylic anchoring group; (ii) the measurements of PCE were performed under AM1.5G illumination; (iii) the redox shuttle was iodide/triiodide redox shuttle in acetonitrile. The first criterion was implemented for avoiding strong deviation in PCE due to adsorption geometry and charge injection rates of dyes with different anchoring groups;[165] the second and third were to ensure the data were originated from measurements under similar conditions. It has been noticed that the best current cells were based on cobalt(II/III) redox shuttle,[44] but for the benefit of obtaining a larger statistical sample the iodide/triiodide redox shuttle was considered instead.

Next, a strategy was required to collect data that have been reported after the publication of Mishra et.al.'s review.[49] A pseudo-random procedure devised which guaranteed the absence of bias, but was also fully reproducible by other researchers with the aim of collecting the chemical structure and the PCE of ~60 dyes appeared in the literature over the past 6 years. All data were collected from the search results provided by Web of Science (WoS), using the keywords 'dye-sensitized' and 'solar' and 'organic' and 'synthesi*' (the asterisk represents any number of characters in WoS) by year, starting from year 2010 to year 2015. For each year the search results were sorted according to 'relevance' and were refined by 'articles only'. Only dyes reported in journals with an impact factor of over 4 were considered, to limit the number of hits and ensure in a rough way that the data were of sufficient quality. The first 10 dyes for each year were then included in the search results. Many studies, however, reported more than one dye and it was generally difficult to achieve exactly 10 dyes for each year without omitting dyes from some studies. Hence, all dyes reported within a study were included, except when a dye did not satisfy the criteria stated previously.

The number of dyes for each year was therefore equal or greater than 10 dyes. The search results and the papers used have been recorded to ensure reproducibility and given in appendix A5.1. From this preliminary data set of 133 dyes we excluded dyes which required more than 860 basis functions in the quantum chemical calculation (approximately those with more than 155 atoms) to make sure all calculations could be performed automatically and could converge without human intervention. The final data set contained 116 dyes.

5.3 Quantification of donor-π-acceptor character

It should be noted that, D-π-A character was not a well-defined property of dyes. Many dyes have been classified as D-π-A dyes based primarily on chemical intuition. A number of qualitative and quantitative indices[267–269] have been devised for D-π-A character, but they have not yet been applied consistently to examine the efficacy of D-π-A character on improving PCE. We have considered three possible descriptors for the D-π-A character including (i) the excitation dipole moment (*EDM*) (Fig. 5.1), which was similar to some descriptors introduced previously in the literatures;[267,269] (ii) the "orbital asymmetry" (*OA*) parameter, defined in Chapter 4 as the ratio of the weight of LUMO and HOMO at the anchoring group of the dye; and (iii) the dipole difference between the ground-state and first excited-state dipole moments of the dye ($\Delta\mu$). The *EDM* was computed as the difference between the centroid of the weight of HOMO ($\mathbf{C}_{HOMO}$) and the centroid of the weight of LUMO ($\mathbf{C}_{LUMO}$):

$$EDM = \pm \left| \mathbf{C}_{HOMO} - \mathbf{C}_{LUMO} \right| \qquad \text{(Eq. 5.1)}$$

$$\mathbf{C}_{MO} = \sum_i a_{iMO} a_{iMO} \mathbf{r}_i + \sum_i \sum_{j>i} a_{iMO} a_{jMO} S_{ij} \mathbf{r}_i \qquad \text{(Eq. 5.2)}$$

where a_{iMO} was the atomic orbital coefficient of either HOMO or LUMO, S_{ij} was the overlap matrix elements of basis functions i and j, and $\mathbf{r}_i$ was the 3-dimensional Cartesian coordinates where the atomic orbital i was centred. The sign of *EDM* was determined by the dot product of the vector between the centroids ($\mathbf{C}_{HOMO} - \mathbf{C}_{LUMO}$) and the vector of the Carbon-Carbon bond at the anchoring group, which indicated the direction of charge excitation.

The orbital coefficients, atom coordinates and overlap matrix required for computing *EDM* were acquired from single-point calculations of geometrically optimized dyes in acetonitrile. Geometry optimizations were performed with a relatively small basis set (3-21G*), and the energies of the optimized structures were subsequently re-evaluated with 6-31G* basis set. The orbital coefficients and overlap matrix for computing *OA* (Eq. 4.3) were acquired from the geometry optimization with 3-21G*. For $\Delta\mu$, the ground state dipole moments of dyes were acquired from the 6-31G* single-point calculations of the optimized geometries, and the first excited state dipole moments were obtained from TDDFT/6-31G* calculations. All calculations were performed with DFT/B3LYP hybrid functional, with the acetonitrile environment mimicked with polarized continuum model (PCM).[228] All calculations were performed with the Gaussian 03 package.[229]

5.4 Correlations

Fig. 5.2 shows the correlations between different descriptors, the correlation between *OA* and PCE (η), and the correlation between $\Delta\mu$ and PCE. In Fig. 5.2(a), it could be seen that there was good correlation between *OA* and *EDM* when some outliers (hollow data points) were omitted. However, whether these outliers were included or not we observed no correlation between *OA* and η (Fig. 5.2(c)).

In Fig. 5.2(b) we observed excellent correlation of $r = -0.98$ between *EDM* and $\Delta\mu$ for the subset of dyes included in Mishra's review.[49] This excellent correlation suggested that *EDM* and $\Delta\mu$ were equivalent for describing the D-π-A strength, and for saving computational expense we would only compute *EDM* for the full dataset considered (116 dyes). Fig. 5.2(d) shows the correlation between $\Delta\mu$ and η of a subset of dyes as employed in Fig. 5.2(b), and the correlation was negligible, which was qualitatively consistent with Fig. 5.2(c).
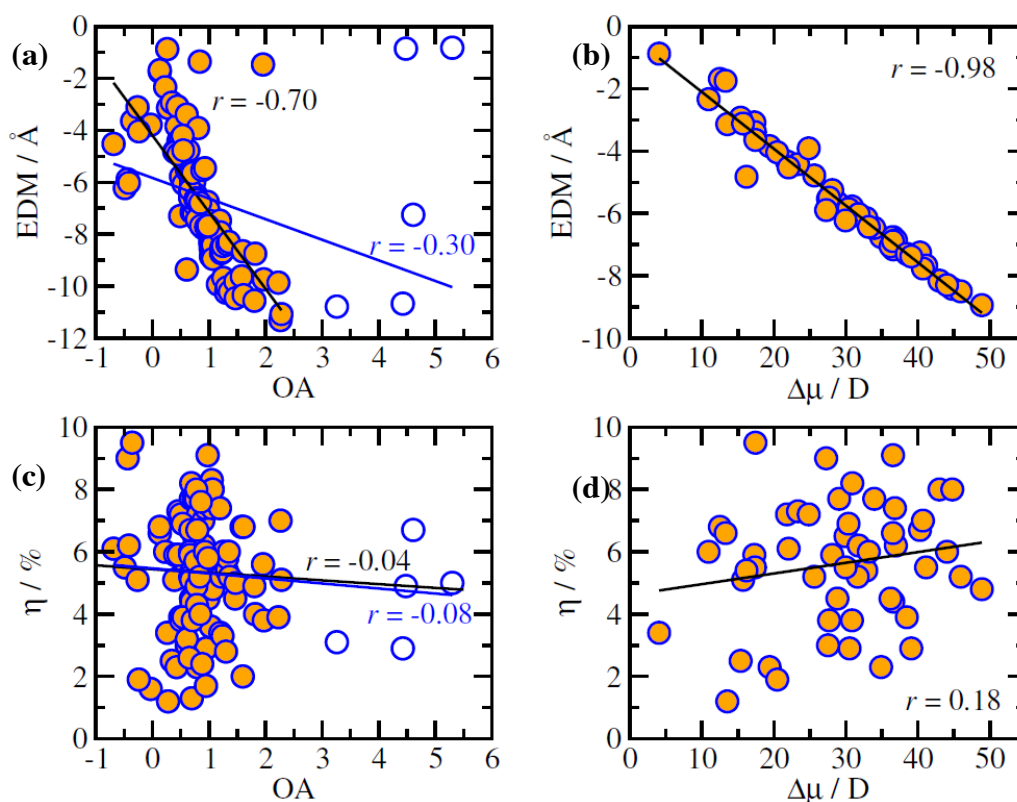
Figure 5.2 (a) Correlation between *EDM* and *OA* of 116 (blue line) and 111 (black lines) dyes. The black line is fitted with the date set without *OA* greater than 3 (hollow data points). (b) Correlation between *EDM* and $\Delta\mu$ of 52 dyes; (c) Correlation between η and *OA* of 116 (blue line) and 111 (black lines) dyes, follows the same fitting method as described in (a); and (d) Correlation between η and $\Delta\mu$ of 52 dyes.
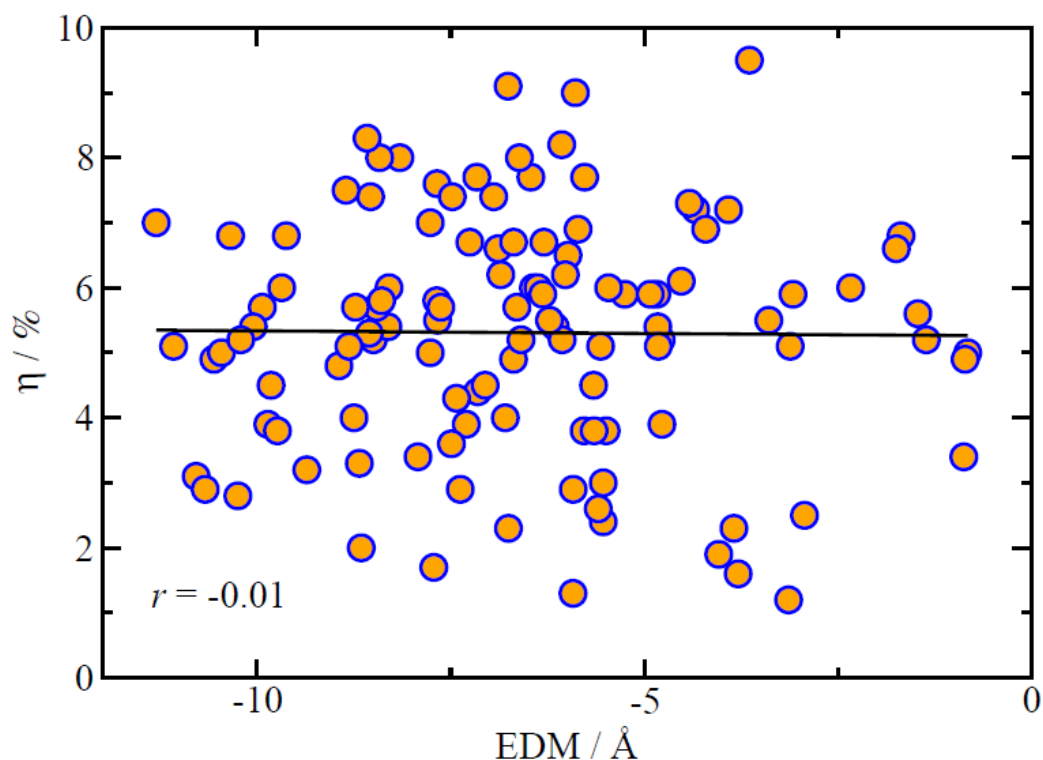
Figure 5.3 Correlation between η (%) and *EDM* of 116 dyes. Mean of *EDM* ($\mu(EDM)$) = −6.61. Standard Deviation (S.D.) of *EDM* = 2.43. $\mu(\eta)$ = 5.31 %. S.D. of η = 1.80 %. Pearson's correlation coefficient (r) = −0.01.

Fig. 5.3 shows the correlation between PCE and *EDM* with 116 data points. A more negative *EDM* represented a stronger D-π-A character. It could be seen that there was no linear correlation between the two properties, which was supported by an extremely low Pearson's correlation coefficient (r = −0.01) from simple linear regression analysis. By visual inspection it was unlikely that the fitting can be improved with higher order polynomials. The results did not directly imply that the D-π-A design did not work but that the combination of all other effects was so dominant that there was no measurable benefit in introducing a D-π-A character. Suppose, for example, that D-π-A character was only helpful (i.e. it correlated positively with the efficiency) if the dye molecule was correctly oriented on the surface and with the D-π-A director perpendicular to the surface;

we had not included orientation effect in the analysis. Our results would therefore imply that the D-π-A character could be helpful or detrimental for the efficiency depending on the orientation and that helpful and detrimental orientations were equally likely. A completely opposite interpretation was that the D-π-A character had no influence. In the first case one should work toward better orientation of the dyes rather than strengthening the D-π-A character and in the latter one should consider a different design rule.

To illustrate that some correlations could be found and could provide new physical insights we revisited the correlation between η and λ_n, first explored statistically in Chapter 4 (Fig. 4.3). The computational procedure of λ_n was the same as described in Chapter 4. As shown in Fig. 5.4, dyes with low λ_n were more often associated to large η. Such an observation agreed with predictions originated from physical principles, where in the evaluation of charge recombination rate decreasing λ_n would retard charge recombination,[87] and hence should increase η. The correlation shown in Fig. 5.4 illustrated how statistical analysis could be useful to validate any given physical model of DSSCs.

Due to the correlation observed between η and λ_n, it should be meaningful to examine also the correlation between *EDM* and λ_n, in order to study whether D-π-A character could influence η through associated λ_n. Fig. 5.5 shows the correlation between *EDM* and λ_n, where large absolute value of *EDM* (strong D-π-A character) would generally associate with low λ_n. Such an observation clearly disagreed with the previous suggestion[45] that a stronger D-π-A character was associated with higher λ_n (another example of how a statistical analysis could support or disprove physical hypotheses).
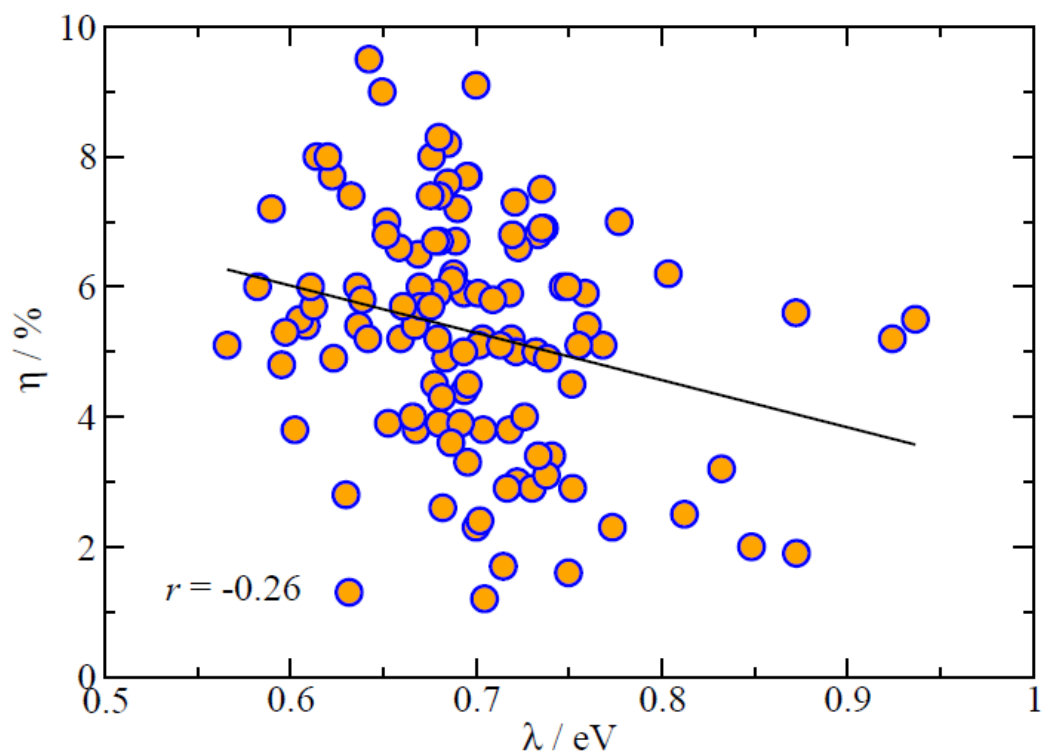
Figure 5.4 Correlation between η (%) and λ_n (eV) of 116 dyes. $\mu(\lambda_n) = 0.70$ eV. S.D. of $\lambda_n = 0.06$ eV. $r = -0.26$.
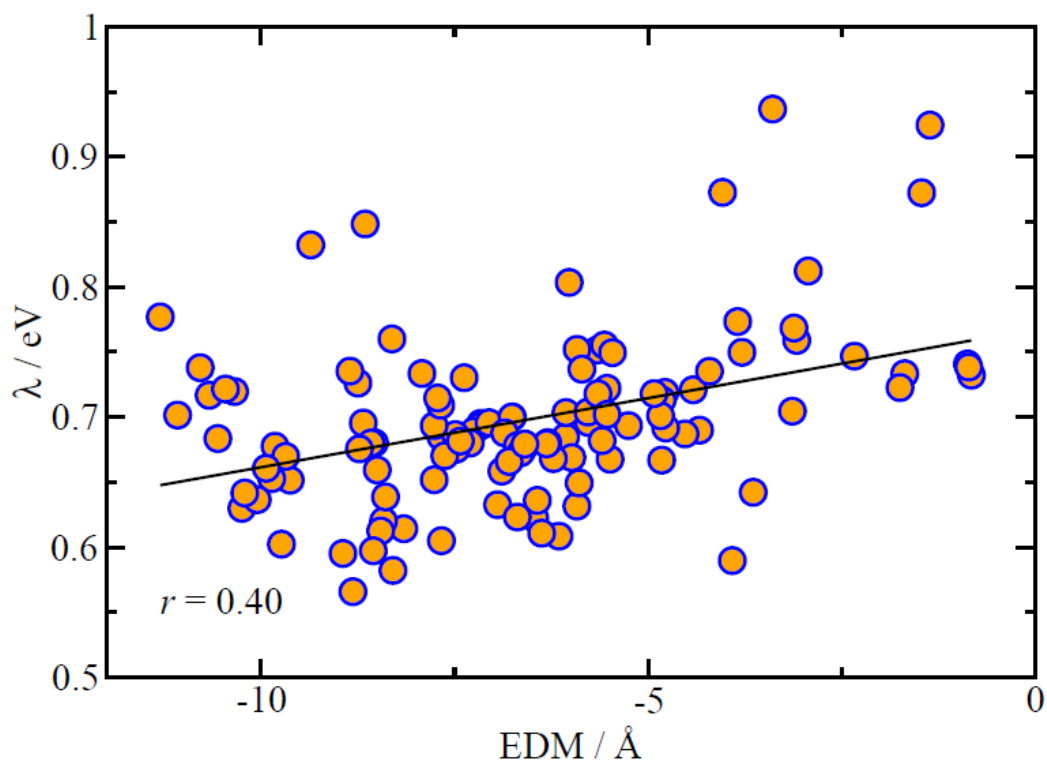
Figure 5.5 Correlation between λ_n (eV) and *EDM* of 116 dyes. $r = 0.40$.

In view of the observed correlations between λ_n and η, and λ_n and *EDM*, one could argue that an increased D-π-A character should lower λ_n and therefore increased η. If there were no other benefits from the D-π-A character apart from reducing λ_n, we would expect the correlation coefficient between *EDM* and η to be still a significant -0.22, which was estimated based on the correlation coefficients between λ_n and η ($r = -0.26$), and between λ_n and *EDM* ($r = 0.4$) (see appendix A5.2 for method of estimation). Having observed a correlation coefficient of $r = -0.01$ between *EDM* and η one could even speculate that the D-π-A character was detrimental to efficiency (the standard error on the correlation coefficient with this sample size was approximately 0.09, as discussed in the Appendix A5.2).

5.5 Conclusions

In conclusion, we have considered a large set of dyes and found no correlation between their D-π-A character and the experimental PCE. Equivalently, this meant that other effects were much more important and washed out any potentially beneficial effects of the D-π-A character. We have shown that such type of statistical analysis could be used to validate structure-property hypotheses derived from basic physical principles. For examples we have seen that, as predicted, there was an improved PCE for dyes with reduced λ_n, while it was not true that the increased D-π-A character increased the λ_n. We could speculate that the erroneous emphasis on the D-π-A character has been due to a number of factors. Probably there has been an element of confirmation bias (i.e. good D-π-A dyes were used to support the theory but counter examples were ignored). The lack of quantification of D-π-A character and the predominance in literature of
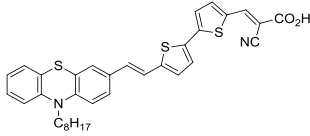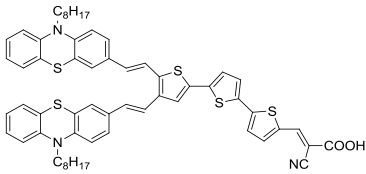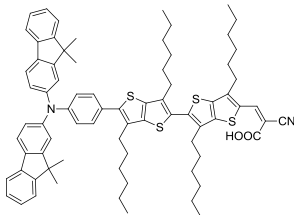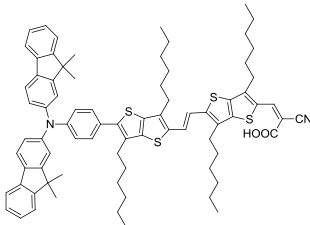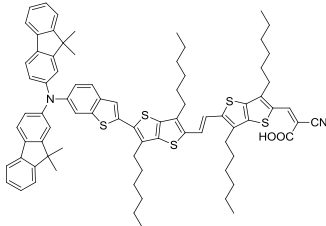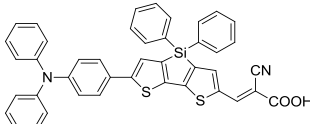
papers comparing the performances of just few dyes have certainly contributed to this misconception. More importantly, we thought, that the popularity of this design rule has caused a large majority of new dyes to be synthesized with the D-π-A character built in. High performing dyes have therefore been discovered with higher probability within this class of compounds, i.e. confirming the erroneous hypothesis by a bias in the sampling. A retrospective look at a uniform set of data was probably the best way to test the validity of any given design rule.
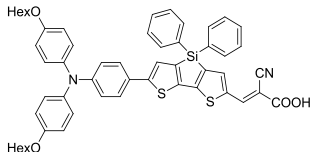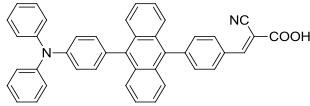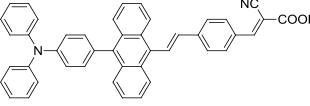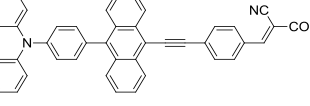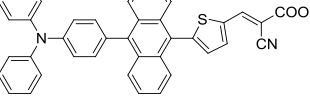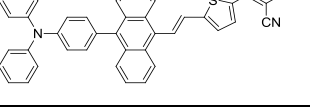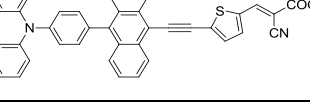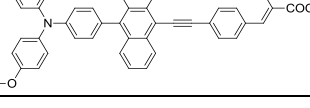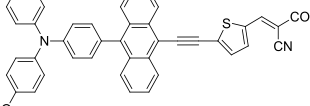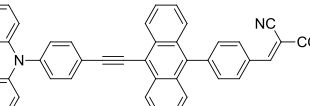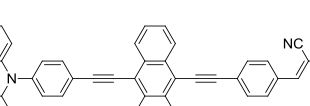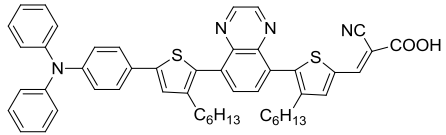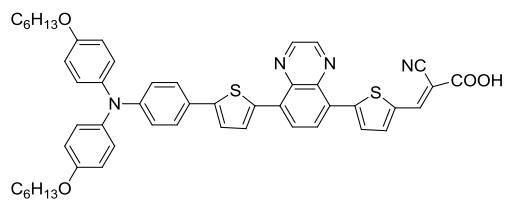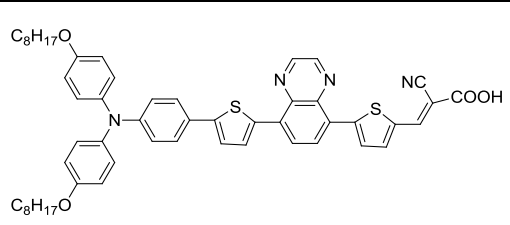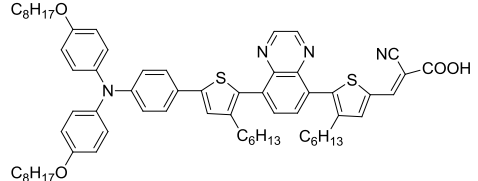
5.7 Appendix
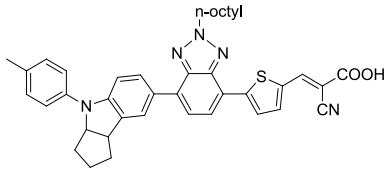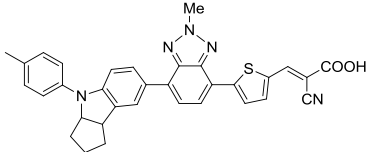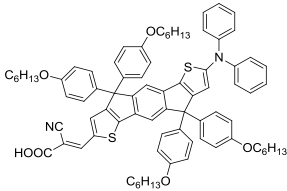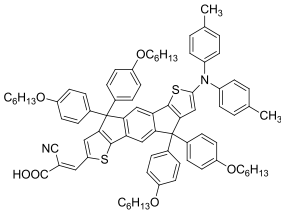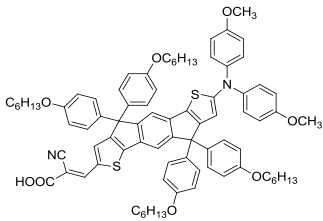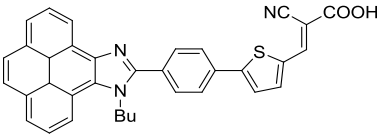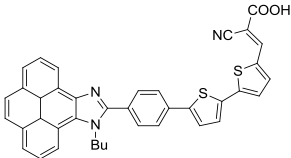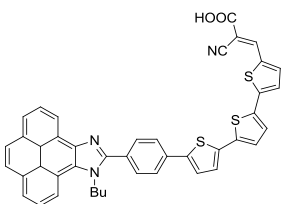
A5.1 New dataset
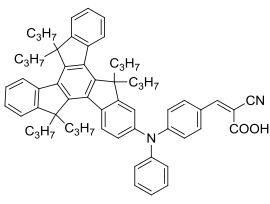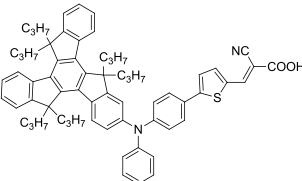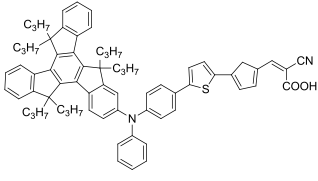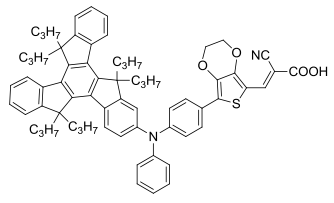
As discussed a new dataset of dyes was developed in accordance with the procedure described in section 5.2. Table A5.1 provided the chemical structures, corresponding PCE (η (%)), year of the report and references of dyes in the dataset (excluding dyes taken from Mishra's review[49]).
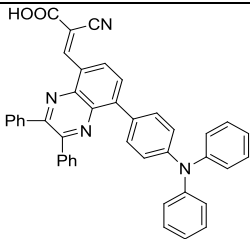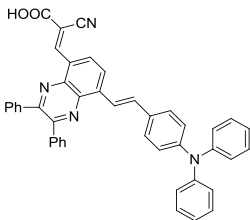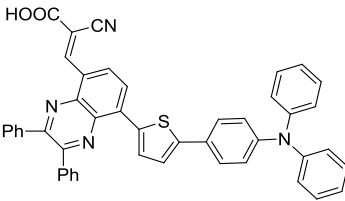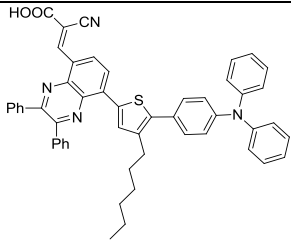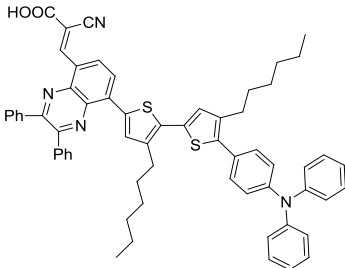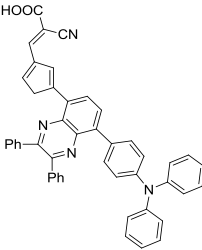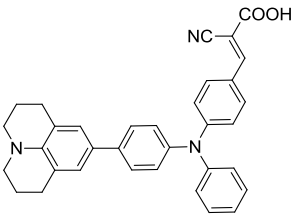
Table A5.1 A new dataset with chemical structures and PCE of dyes.
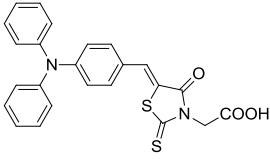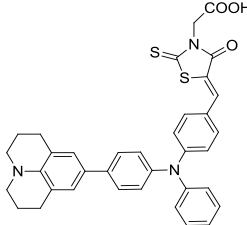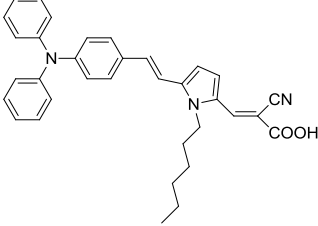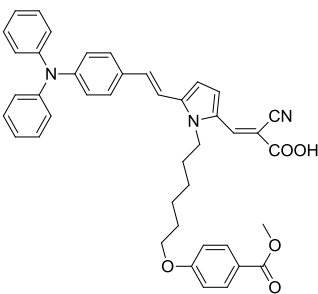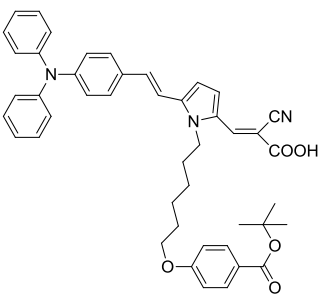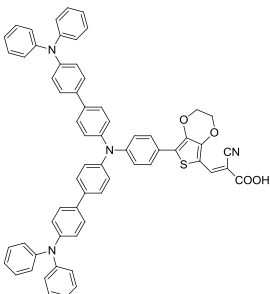
Label	Dye	η / %	Year	Ref.
53		5.4	2010	[270]
54		2.8	2010	[270]
55		5.1	2010	[271]
56		8.0	2010	[271]
57		9.1	2010	[271]
58		6.7	2010	[272]
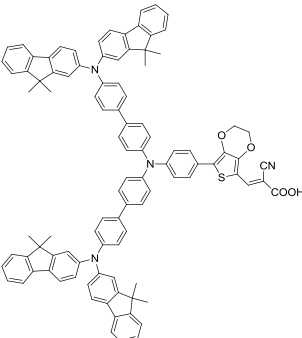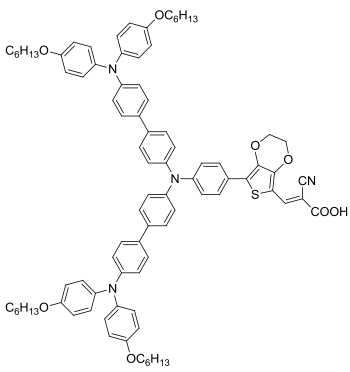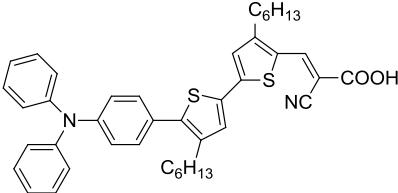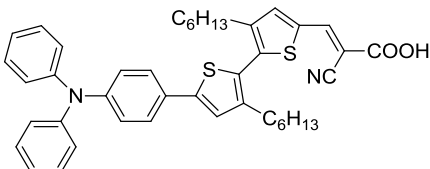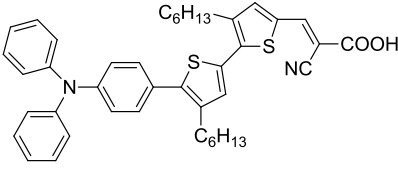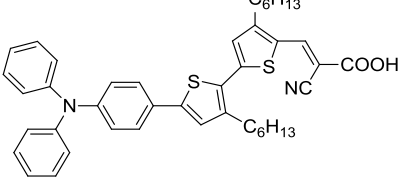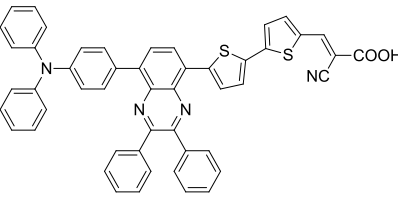
59		7.6	2010	[272]
60		3.1	2010	[273]
61		5.1	2010	[273]
62		6.8	2010	[273]
63		2.9	2010	[273]
64		4.9	2010	[273]
65		4.5	2010	[273]
66		7.0	2010	[273]
67		6.8	2010	[273]
68		3.9	2010	[273]
69		5.7	2010	[273]
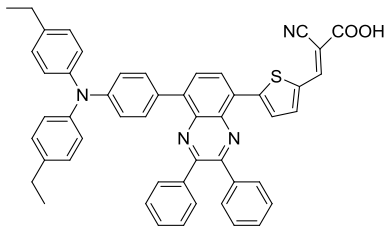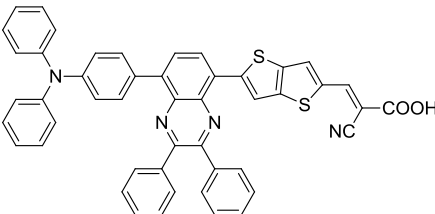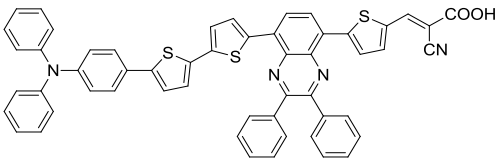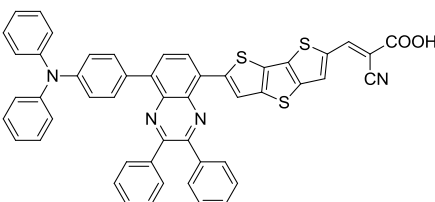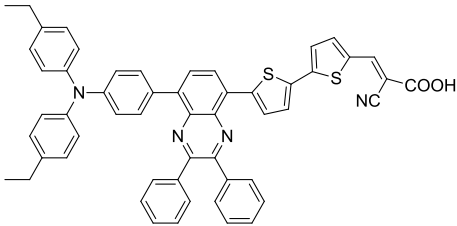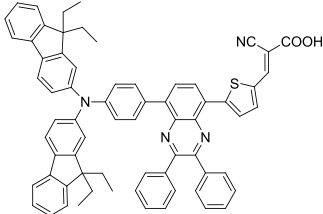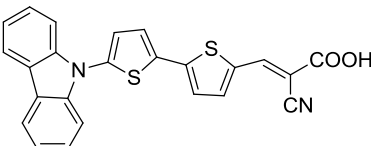
70		8.0	2011	[274]
71		6.7	2011	[274]
72		6.0	2011	[275]
73		6.7	2011	[275]
74		6.3	2011	[275]
75		3.4	2011	[276]
76		3.8	2011	[276]
77		5.7	2011	[276]
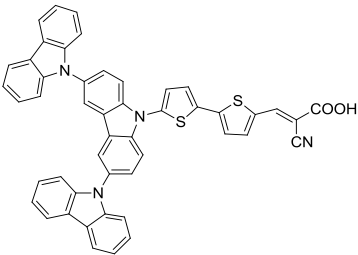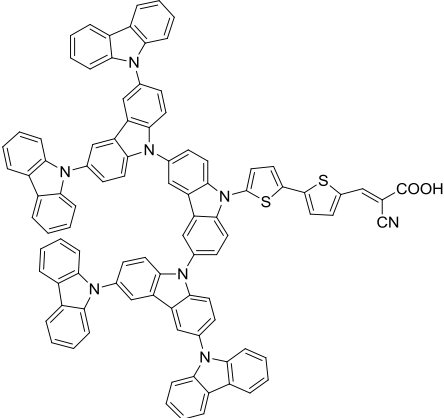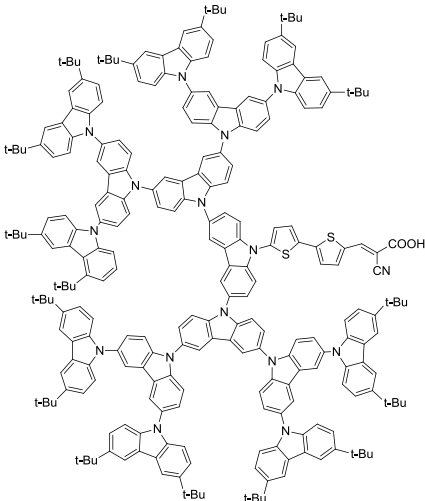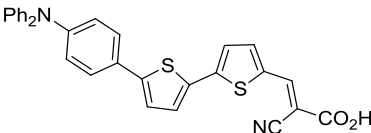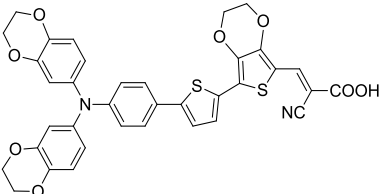
78		3.9	2011	[277]
79		4.9	2011	[277]
80		5.3	2011	[277]
81		6.0	2011	[277]
82		3.3	2012	[278]
83		7.4	2012	[278]
84		8.3	2012	[278]
85		3.5	2012	[278]
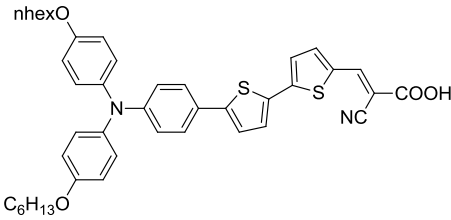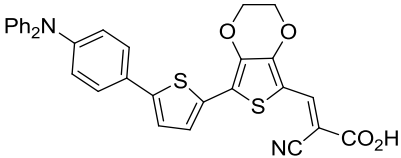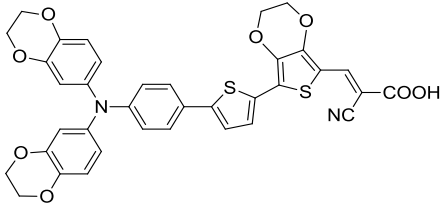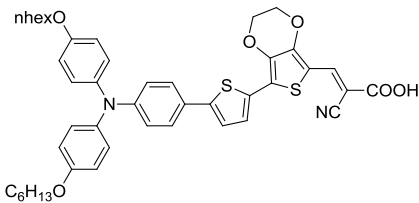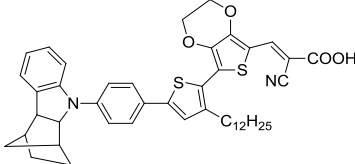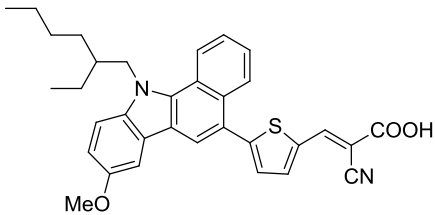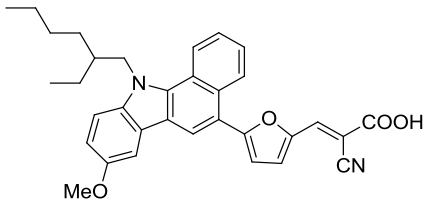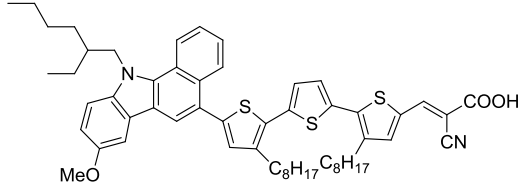
86		2.4	2012	[279]
87		2.6	2012	[279]
88		4.0	2012	[279]
89		3.6	2012	[279]
90		5.0	2012	[279]
91		7.4	2012	[279]
92		2.0	2013	[280]

93		1.6	2013	[280]
94		3.2	2013	[280]
95		5.1	2013	[281]
96		5.9	2013	[281]
97		5.9	2013	[281]
98		5.1	2013	[282]

99		4.4	2013	[282]
100		3.3	2013	[282]
101		5.0	2013	[283]
102		4.0	2013	[283]
103		5.8	2013	[283]
104		5.9	2013	[283]
105		6.0	2014	[284]

106		5.7	2014	[284]
107		5.7	2014	[284]
108		5.7	2014	[284]
109		5.4	2014	[284]
110		5.2	2014	[284]
111		5.8	2014	[284]
112		5.1,	2014	[285]

113		3.8	2014	[285]
114		2.6	2014	[285]
115		1.8	2014	[285]
116		5.0	2015	[286]
117		1.7	2015	[286]

118		5.6	2015	[286]
119		4.9	2015	[286]
120		4.3	2015	[286]
121		5.2	2015	[286]
122		5.2	2015	[286]
123		6.0	2015	[287]
124		6.9	2015	[287]
125		7.5	2015	[287]

It should be noted that:

1. Dyes reported on *chemsuschem* and *chemistry: an asian journal* were not included since we did not have straightforward access to literatures from these journals;

2. For studies included comparison with the addition of additives, such as CDCA, we used the data obtained without the additives.

A5.2 Estimation of Correlation coefficients of mutually dependent variables

If there was a linear correlation r_{xy} between variables x and y (due to normally distributed random fluctuations of variable y) and a linear correlation r_{yz} between the variable y and z (due to normally distributed random fluctuation of variable z not correlated to the fluctuations of variable y) the linear correlation between x and z was given by:

$$r_{xz} = \frac{r_{xy} \cdot r_{yz}}{\sqrt{r_{xy}^2 + r_{yz}^2 - r_{xy}^2 \cdot r_{yz}^2}}$$
(Eq. A5.1)

Furthermore, it has been shown that the standard error on the computed correlation was approximately given by[288]

$$\sigma_r = \frac{(1 - r^2)}{\sqrt{n-1}}$$
(Eq. A5.2)

where n is the number of data points. The two equations above are used in section 5.4. The first one could be derived simply noticing that, if $y = x + R_x$, where R_x was a random variable normally distributed around zero with standard deviation σ_x, the correlation between r_{xy} could be expressed as:

$$r_{xy} = \frac{1}{\sqrt{1 + \sigma_x^2}}$$
(Eq. A5.3)

If $z = y + R_y$, where R_y was a random variable normally distributed around

zero with standard deviation σ_y. If R_x and R_y were uncorrelated the relation between z and y would be $z = x + R_x + R_y$ whose correlation would be

$$r_{xy} = \frac{1}{\sqrt{1 + \sigma_x^2 + \sigma_y^2}}$$

(Eq. A5.4)

Combining Eq. A5.3 and Eq. A5.4 would yield Eq. A5.1. Note that the result was insensitive to scaling the variable x or y by any arbitrary constant and it was therefore valid for the general case $y = a(x + R_x)$ and $z = b(y + R_y)$.

Chapter 6 Computational study of competing reaction mechanisms of photo-catalytic reduction of CO_2 on TiO_2 anatase(101)

6.1 Introduction

In this chapter we study three proposed reaction mechanisms for photocatalytic reduction of CO_2 to gain CH_4 on a defect-free TiO_2 anatase(101) surface with DFT calculations. There was no widely accepted reaction pathway for the molecular transformation of CO_2 to CH_4 on TiO_2, or Eq. 1.4., and each proposed reaction pathway had some weaknesses. The effects of reaction conditions on the reaction mechanisms were also not well-understood. The main objective of this study is to identifying the most favorable pathway among the three proposed reaction pathways. We would compare the thermodynamic landscape of the reaction mechanisms, suggest methods for verifying the identified pathway, and note possible mechanistic hypotheses.

6.2 Proposed reaction mechanisms

In this section we report briefly the key intermediates in the mechanisms as described in the literature[13] with the evidence in support and the possible doubts that have been cast. The pathways, named in accordance with an intermediate along the pathway, are known as the carbene pathway, the formaldehyde pathway and the glyoxal pathway. These pathways mainly differentiate from one another after the presumable activation of carbon dioxide (CO_2) via one-electron reduction. In the carbene pathway,[120] the subsequent steps are the formation of carbon monoxide (CO), carbon residue (C), and a series of step-wise abstractions

of 4 H radicals by the carbon to gain methane (CH_4), in which carbene (CH_2) is expected in the last series of steps. In the formaldehyde pathway,[29] the main intermediates are formic acid (HCOOH), formaldehyde (HCHO) and methanol (CH_3OH), forming in this order along the pathway. The glyoxal pathway is the most complex pathway among the three proposed mechanisms, postulated based on a series of Electron paramagnetic resonance (EPR) experiments and DFT calculations.[119] After the activation of CO_2, it is expected that the HCO radical would form and dimerize to give glyoxal (HOCCOH). A series of electron and proton transfers thereafter lead to the formation of trans-ethane-1,2-semidione, glycol-aldehyde, vinoxyl radical and acetalaldehyde. The next step is a hole transfer to acetalaldehyde, to produce unstable acetyl radicals. The C-C bond in acetyl radical is then cleaved to give CO and CH_3, of which the latter can consume another H radical to produce CH_4.

While the proposed mechanisms appeared to be reasonable based on chemical intuition and supported by some experimental evidence, there were details in each pathway that have not been elucidated or verified. For the carbene pathway, some important reaction intermediates (C, CH_3, CO, H) have been confirmed with EPR studies, and experimental data were able to fit with the kinetic model for this mechanism.[289] This mechanism, however, did not explain the absence of HCOOH, a competitive product to CO that has been reported earlier in the proposed formaldehyde pathway.[29] Although there were reports of C residual detected on surfaces,[121] in theory the formation of a carbon atom was expected to be a thermodynamically unfavorable process, and the reaction to form C from CO has not been verified.

For the formaldehyde pathway, the closed-shell products were commonly observed and supported by the electrochemical reduction potential data (Eq. 6.1

to Eq. 6.4, vs NHE, pH = 7),[290] where reduction potentials of CO_2 to form a number of products (HCOOH, H_2CO, CH_3OH and CH_4) via multiple electron and proton transfer were generally less negative than the conduction band potentials of TiO_2 ($E = -0.50V$):

$$CO_2 + 2H^+ + 2e^- \rightarrow HCOOH \qquad E = -0.61V \qquad \text{(Eq. 6.1)}$$

$$CO_2 + 4H^+ + 4e^- \rightarrow H_2CO + H_2O \qquad E = -0.48V \qquad \text{(Eq. 6.2)}$$

$$CO_2 + 6H^+ + 6e^- \rightarrow CH_3OH + H_2O \qquad E = -0.38V \qquad \text{(Eq. 6.3)}$$

$$CO_2 + 8H^+ + 8e^- \rightarrow CH_4 + 2H_2O \qquad E = -0.24V \qquad \text{(Eq. 6.4)}$$

There was, however, no experimental evidence for multiple-electron transfer and it was more likely that the reaction proceeded via single-electron transfer, but the expected radical intermediates from these reductions were not observed.[13,291] Experimental data were also unable to fit the kinetic model of this pathway.[289]

The glyoxal pathway encompassed some possibilities that were not considered in the other two proposed pathways, including the formation of C_2 compounds as intermediates, and having oxidation with the photo-generated hole as an elementary step. In addition, the proposed mechanism took into account the formation of commonly observed products such as CO, H_2CO and CH_3OH, even though these intermediates were not included in the conversion pathway to CH_4. For H_2CO and CH_3OH, both species were expected to undergo oxidation with surface holes to generate HCO and CH_2OH radicals, taking mainly the role of sacrificial hole scavengers, and would not undergo one-electron reduction.[119,291] Therefore this proposed mechanism suggested the formaldehyde pathway would only occur with 2-electron 2-proton transfer reactions. The main weakness of this pathway was that glyoxal and glycolaldehyde have not been reported.[119] In the study with EPR that led to the postulation of this pathway, the transformations of the radicals to molecules were not demonstrated, and the identification of the

vinoxyl radical was rather tentative.[119] Table 6.1 summarized the key points of the proposed mechanisms.

Table 6.1. Summary of the detected species in the proposed mechanisms, the supporting evidence and the uncertain details for the proposed mechanism.

Pathway	Detected species in proposed mechanisms	Supporting evidence	Uncertain details
Carbene	C, CO, H, CH_3OH, CH_4, CH_3[120]	EPR experiments; kinetic model fitting	Formation of C atom; No explanation on the absence/role of HCOOH
Formaldehyde	HCOOH, H_2CO, CH_3OH, CH_4[29]	Electrochemical reduction potentials	Lack of experimental support e.g. EPR; No explanation on the absence/role of CO
Glyoxal	CH_3, CH_4, H [290] $HCOCH_3$[292]	EPR experiments; DFT calculations	Glyoxal and glycolaldehyde are deduced from EPR,[119,293] but are not reported

Although there was no accepted mechanism, various views regarding some important aspects of the reaction, such as rate-limiting steps, have been advocated. It was generally believed that the activation of CO_2 was a difficult reaction to realize and likely to be the rate-limiting step, suggested also by the electro-chemical potential of CO_2/CO_2^- ($E = -1.90V$). Adsorption of CO_2 on the surface seemed to be able to make the reaction viable, but there were very few reports of such anion on TiO_2.[294] Some studies have therefore suggested that the

reduction of CO_2 may occur only under specific circumstances, such as when oxygen vacancies were introduced.[122,135] Another possible rate-limiting step for the reaction was the adsorption of reactants, where a pseudo-first order relationship has been observed between photo-reduction of CO_2 and initial CO_2 concentration.[295]

6.3   Computational method and details

All calculations were performed with GGA/PBE functional with ultra-soft pseudo-potentials with Quantum Espresso,[296] unless specified otherwise. The total energy of the intermediates were computed with the anatase(101) surface, which was modeled as a two tri-layer slab with (2 × 2) surface unit cells. The two-layer slab was a rather thin slab but has been shown to be sufficient for modeling adsorption of molecules on the surface,[297] and we employed this slab also with the intention to save computational expense, in view of the rather large number of computations required in this study. For the most important results, where the thermodynamic landscapes of different pathways were being compared, we computed also the energies of the intermediates with a 5-layer slab for checking the main conclusions (see section A6.5). The Monkhorst-Pack k-point grid used was 2 × 2 × 1, and the kinetic energy cutoffs for wave-functions and charge density were 35Ry and 280Ry respectively.[297] For intermediates involving radical species, spin-polarized calculations were performed. The energies of the transition states were computed using the climbing-image nudged elastic band (CI-NEB) method.[158] We have first performed NEB based on an initial guess of the path, and subsequently improved the transition state energy of the NEB minimum energy path by using the climbing-image technique.

Our reaction (Eq. 1.4) involved many reactants in sequence. To compare different reaction paths it was important to include the energy of the unreacted species in the intermediates' energy. To illustrate the procedure we considered the simple reaction:

$$CO_2 + 2H^{\bullet} \rightarrow CO + H_2O \qquad \text{(Eq. 6.5)}$$

We first computed the total energy of CO_2 adsorbed on the surface (denoted as (E_1)), which was the sum of the total energy of the surface adsorbate (CO_2), the surface slab employed and the associated adsorption energy. Similarly and separately, we computed also the total energy of a H atom adsorbed on the surface (E_H), which was the sum of the total energy of the H atom, the surface slab and the associated adsorption energy. The total energy of the reactant in Eq. 6.5, which was one adsorbed CO_2 and two adsorbed H atoms far away from CO_2 (E_{int1}), was simply:

$$E_{int1} = E_1 + 2E_H \qquad \text{(Eq. 6.6)}$$

In the next step, we assumed a H atom has diffused and was adsorbed close to CO_2, and we computed the energy of this configuration (E_2). The diffusion process was not studied explicitly. In addition, the total energy of the surface (denoted as MS) that has initially accommodated the diffused-away hydrogen (E_{MS}) was also computed. The total energy of this intermediate, which was one adsorbed CO_2 with one adsorbed H atom close to it and the other far away from it (E_{int2}), was:

$$E_{int2} = E_2 + E_H + E_{MS} \qquad \text{(Eq. 6.7)}$$

Thereafter the total energy of COOH on the surface (E_3), as well as the transition state of the hydrogen transfer to produce COOH, were both computed and aligned with other intermediates by adding appropriate energies of E_H and E_{MS}. The same procedure was repeated until CO and H_2O were formed.

The elementary steps in the proposed pathways considered could have different mechanisms, such as sequential Electron transfer (ET) and Proton transfer (PT), concerted coupled proton-electron transfer (CPET) and hydrogen atom transfer, which may lead to the computation of an unmanageable number of possible intermediates for this study. We therefore modeled the photo-reduction process based on the hypothesis that the conversion from CO_2 to CH_4 proceeded via a series of hydrogen atom transfer reactions. In support of this hypothesis there was an expectation that hydrogen atom transfer reactions very often exhibit low-energy barriers (less than 1.0 eV).[298,299] On the more practical side, this hypothesis would also allow a consistent computational setup to be applied for the three reaction mechanisms considered. We noted that the Lowdin charge for an adsorbed H atom was typically ~0.6 in our calculations, and partial reduction of the TiO_2 slab was observed when there was no co-adsorbate. We have also observed that co-adsorbates were partially reduced when in the presence of a nearby adsorbed H atom. We noticed that the transfer of an adsorbed H atom had been considered as sequential ET and PT in some studies,[136,137] but in this study we would refer to this process as H atom transfer, as we preferred to define a pure electron transfer reaction when two potential energy surfaces with different localization of the electron were involved.

Once reactants and product were connected via the computed intermediates under the hypothesis of sequential H atom transfer it was possible to assess the potential impact of alternative sub-mechanisms connecting the same intermediates. If two reaction paths had relatively low barriers and very different energy of the intermediates, one could conclude that the lower energy intermediate path was more likely even if new lower energy paths connecting the intermediates could be found. If, on the contrary, the energy of the intermediates

was similar and the reaction path was determined by the barrier heights, it would be important to study the details of the rate determining steps. The approach presented here, depending on the results, would provide either the most likely mechanisms among those proposed or an indication of the elementary steps to be studied in greater detail.

The hypothesis of sequential H atom transfer required some minor modification of the carbene pathway as described in the literature.[13] We have considered alternative intermediates in this study to connect CO_2 to CO (connected via COOH) and CO to C atom (connected via HCO and HCOH) in the carbene pathway, to be compatible with our scheme. These alternative intermediates could be reasonably formed when an adsorbed H atom was close to the reactant on the surface (see section A6.2 for adsorption geometries), and/or have also been observed in other related reactions, which would be discussed in the results section. In the figures where our computed reaction pathways were defined (Fig. 6.1, Fig. 6.4(Top) and Fig. 6.5(Top)) we would use different representations to distinguish intermediates in proposed reaction pathways and intermediates suggested in this study.

Due to the rather large number of calculations required in this study, we have opted for the PBE functional for its relatively low computational cost. The adsorption energy of H atom on anatase(101), an important quantity to the relative energies in our energy profiles, may be dependent on the choice of functional. We have noticed that the PBE approach was able to produce the adsorption energies of H atom at the surface bridging O site of anatase(101) similar to that with the PBE+U approach (the difference was in between the range of 0.04 eV to 0.28 eV).[300,301] A limitation of the PBE approach was that the method was not able to describe localized excess electron, but pure, or

non-adiabatic, electron transfer reactions were not considered in this study. In additions, there were reports of both localization and delocalization of the excess electron in TiO_2 anatase,[302] and the localized and delocalization arrangements were found to be close in energy.[303] Hence localized excess electrons of the slab were deemed unimportant in this study. An accurate description of the localized states at the expense of higher computational cost, such as using DFT+U, would be needed when considering non-adiabatic electron transfer reactions.

6.3 Reaction profiles

6.3.1 Carbene pathway

Fig. 6.1 shows how the carbene pathway was defined in our study, where the adsorption geometries of all the intermediates were reported in section A6.2. The pathway began with the formation of CO (**8**) from CO_2 (**1**) via the formation of COOH (**4**), a typical intermediate in the water-gas shift reaction on a solid surface.[41] CO then detached from the surface and reacted with a H atom (**11**) to spawn HCO (**12**), and a subsequent H atom transfer produced HCOH (**14**). HCO was a common species detected in experiments[119] and HCOH was a species expected in carbene chemistry[304] and the photo-catalytic formation of carbohydrates.[305] HCOH was then rearranged to produce C atom and water (**16**), and this C atom would sequentially abstract 4 H atoms on the surface to generate CH_4 (**26**).
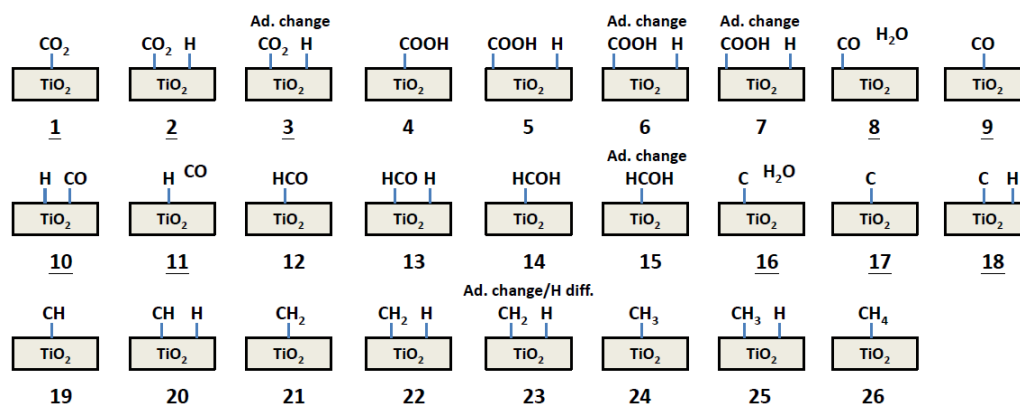
Figure 6.1 Illustration of the intermediates involved in the carbene pathway in this study. 'Ad. change' meant the intermediate has undergone adsorption mode and/or site change. 'H diff.' meant diffusion of H atom between sites close to the adsorption of the intermediate. The underlined intermediates include species in the proposed reaction pathway,[13,120] while other intermediates are suggested in this study.

Fig. 6.2 shows the energy profile of the carbene pathway. The solid black lines denote the energies of the intermediates, the blue solid curves denote the kinetic barriers, the blue dashed lines represent the processes where hydrogen diffused and adsorbed close to the intermediate, or a water molecule diffused away from the intermediate, for which the barrier was not computed. The zero energy in Fig. 6.2 corresponds to the linear adsorption of CO_2 on anatase(101) **(1)**.

We first noted that the rate-limiting step was the formation of a C atom from HCOH **(15 to 16)**, which was both thermodynamically (total energy difference (ΔE_{tot}) of 2.04 eV) and kinetically (kinetic barrier height (E_{act}) of 2.80 eV) unfavorable. Other possible paths to form C atoms, such as $CO + H_2O + e^-$,[306] were not examined (see section 6.3). Once the C atom was formed, the subsequent H atom transfers went steeply downhill. It should be noted that the ease of these H atom transfers were related to the adsorption geometries of the carbon-based radicals. For instances, if the carbon atom of CH_2 was bonded to both a surface unsaturated Ti atom and a bridging O atom, the transfer of a H

atom would have to be facilitated by breaking the C-O bond **(23)** (see section A6.2 for adsorption geometry); we were also unable to identify a reasonable transition state geometry for H atom transfer by CH_3, when CH_3 was strongly adsorbing on a surface unsaturated Ti atom via the carbon atom.
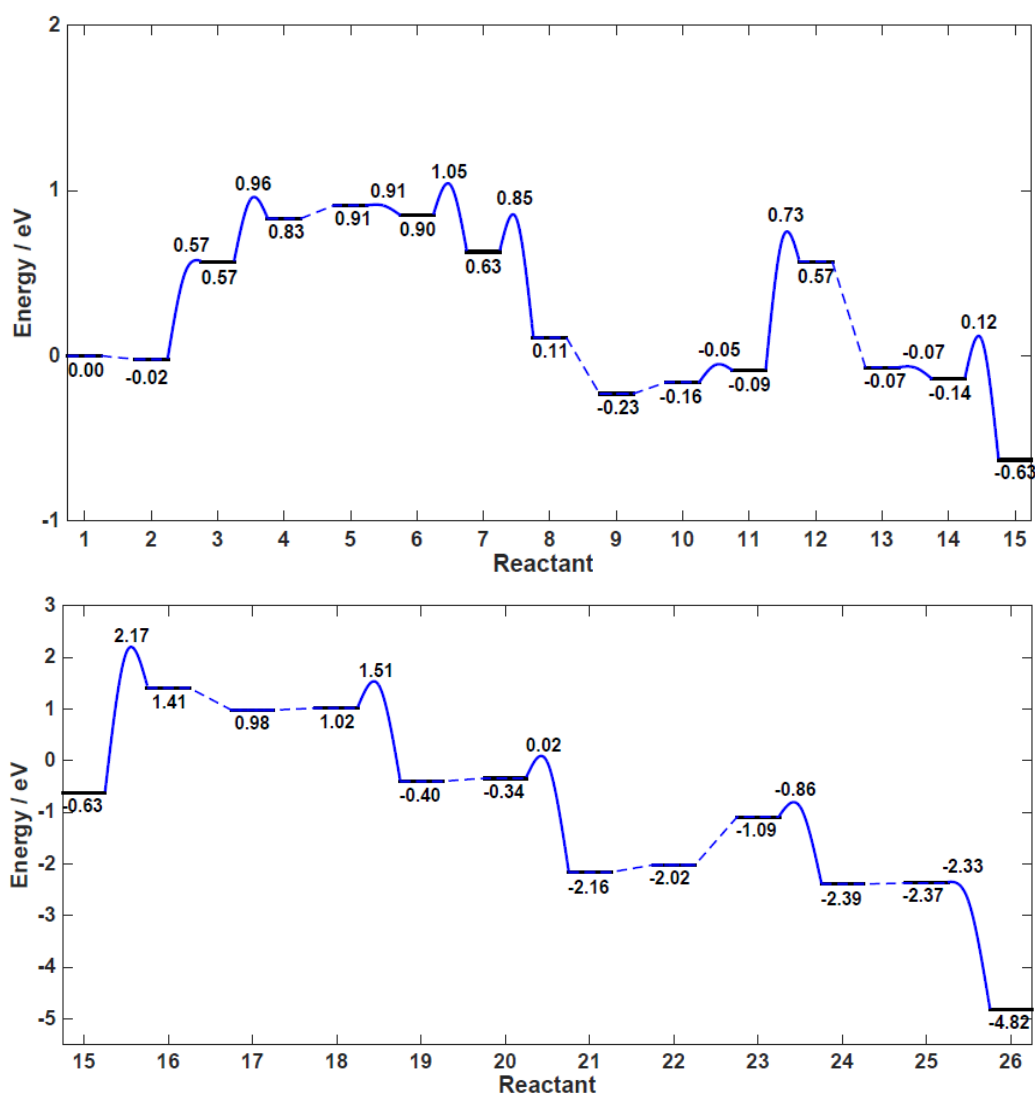
Figure 6.2 Energy profile of the Carbene pathway. The energy profile in the top panel was magnified along the energy scale.

6.4.2 Approximating kinetic barriers with total energy difference

In the course of building the energy profile for the carbene pathway the computation of kinetic barriers was the most time-consuming component, which

essentially limited the number of reaction paths that could be explored. There has been endeavor to establish simpler and quicker methods to determine rate-limiting steps without computing explicitly the kinetic barriers. For instance, some heterogeneous catalytic reactions[298,307] have been found to obey the Bells-Evans-Polanyi (BEP) relation,[308,309] a linear relationship established between reaction barrier and reaction enthalpy change, and hence the barriers of reactions of the same families that obeyed BEP relation could be computed easily. Norskov et al. have developed the Computational hydrogen electrode (CHE) method[127,128,310] that allowed estimation of the over-potential and the rate-limiting steps based on the assumed linear relationship between kinetic barrier and free energy difference, which appeared sometimes to be an acceptable estimate.[67]

To investigate whether the energetics of the intermediates alone was sufficient for identifying the rate determining step and assessing the plausibility of different reaction mechanisms, we reported in Fig. 6.3 the activation energy E_{act} against the ΔE_{tot} for all H atom transfer (both forward and backward) reactions. This was the main type of reaction that we would consider when constructing the energy profiles for the formaldehyde and glyoxal pathways. The black lines indicated the minimum acceptable activation energies (since $E_{act} \geq \min(0, \Delta E_{tot})$). It could be seen that high activation energies (1.0 eV to 2.8 eV) were always associated with very positive ΔE_{tot}, and the difference, $E_{act} - \Delta E_{tot}$, was much smaller than the typical range of activation energies. BEP-type relationships for H atom transfer reactions have in fact been observed in the past.[298,299]
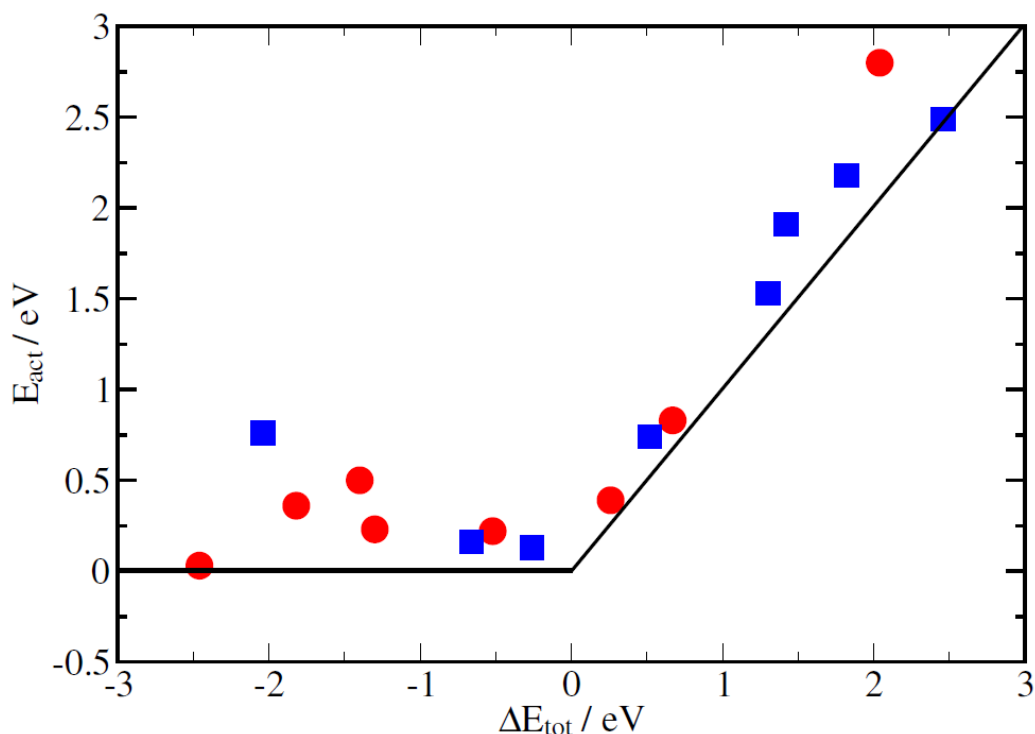
Figure 6.3 Correlation between ΔE_{tot} and E_{act} computed based on energies of the intermediates and transition states of H atom transfer reactions presented in Fig. 6.2. The red circles were E_{act} of the forward reactions (8) and the blue squares were E_{act} of the backward reactions (8).

This suggested that ΔE_{tot} often contained enough information to identify the slower steps in a reaction mechanism. More quantitatively, we could consider the following criteria to determine whether it was necessary to compute the kinetic barriers in the other two proposed pathways: (i) if the energy profile contained only the energies of the intermediates, and the resulting thermodynamic landscape was completely 'downhill' ($\Delta E_{tot} < 0$) or with 'uphill' steps having small ΔE_{tot}, e.g. $\Delta E_{tot} < 0.5$ eV, it would be necessary to compute the kinetic barriers in order to determine the rate-limiting step, as well as the potential maximum to assess the plausibility of the proposed reaction mechanism; (ii) if, however, the greatest ΔE_{tot} was $> \sim 1$eV, the energetics of the intermediates were likely to be sufficient to determine the rate-limiting step. We would then compute

the barriers for this likely rate-limiting step identified in each pathway.

6.4.3 Formaldehyde pathway

Fig. 6.4 shows the intermediates in the formaldehyde path in this study and the energy profile of this pathway. The path started with the formation of HCOOH **(5)** from linearly adsorbed CO_2 **(1)** via the formation of HCOO **(4)**. Subsequent step-wise abstractions of H atoms yield OCH_2OH **(8)**, H_2CO **(9)**, H_2COH **(12)**, CH_3OH **(14)**, CH_3 **(16)** and CH_4 **(20)**, in this order. Noted that, unlike other intermediates included in this pathway (**Table 6.1**), OCH_2OH **(8)** and H_2COH **(12)** have not been detected experimentally.[13] Noted also that, for the formation of HCOO **(4)** and OCH_2OH **(8)**, we have placed an extra H atom on the surface (**(3)** and **(7)**) before H atom transfer. This was due to the observation that, while the energies between intermediates with and without a co-adsorbed H atom nearby were typically very similar, such as intermediates **(1)** and **(2)** in Fig. 6.4, HCOO **(4)** and OCH_2OH **(8)** were much lower in energy in the presence of an adsorbed H atom nearby than when adsorbed H atom was absent. The adsorption of two H atoms on the surface, such as intermediates **(3)** and **(7)**, has been used as a model for the initial state of two-electron one-proton transfer reaction in DFT+U study elsewhere,[136,137] but we would consider this as a H atom transfer, since only a partial reduction of reactant was observed, as discussed previously in section 6.3.

The rate-limiting step was likely to be the formation of CH_3 (**15 to 16**). The ΔE_{tot} of this step was 1.29 eV, and the E_{act} was 1.49 eV. The second most unfavorable step was the H atom transfer to H_2CO (**11 to 12; ΔE_{tot} = 0.69 eV**). Considering the criteria given previously, it has been determined that the computation of kinetic barriers was not necessary for this pathway, except for the

rate-limiting step. It should be noted that the energies of the most thermodynamically stable adsorption mode for HCOOH[311] and H_2CO[312] were not used in Fig. 6.4. In this case we have employed adsorption modes that could facilitate H atom transfer without requiring changes in adsorption geometry or site that could yield OCH_2OH (**8**) and H_2COH (**12**).
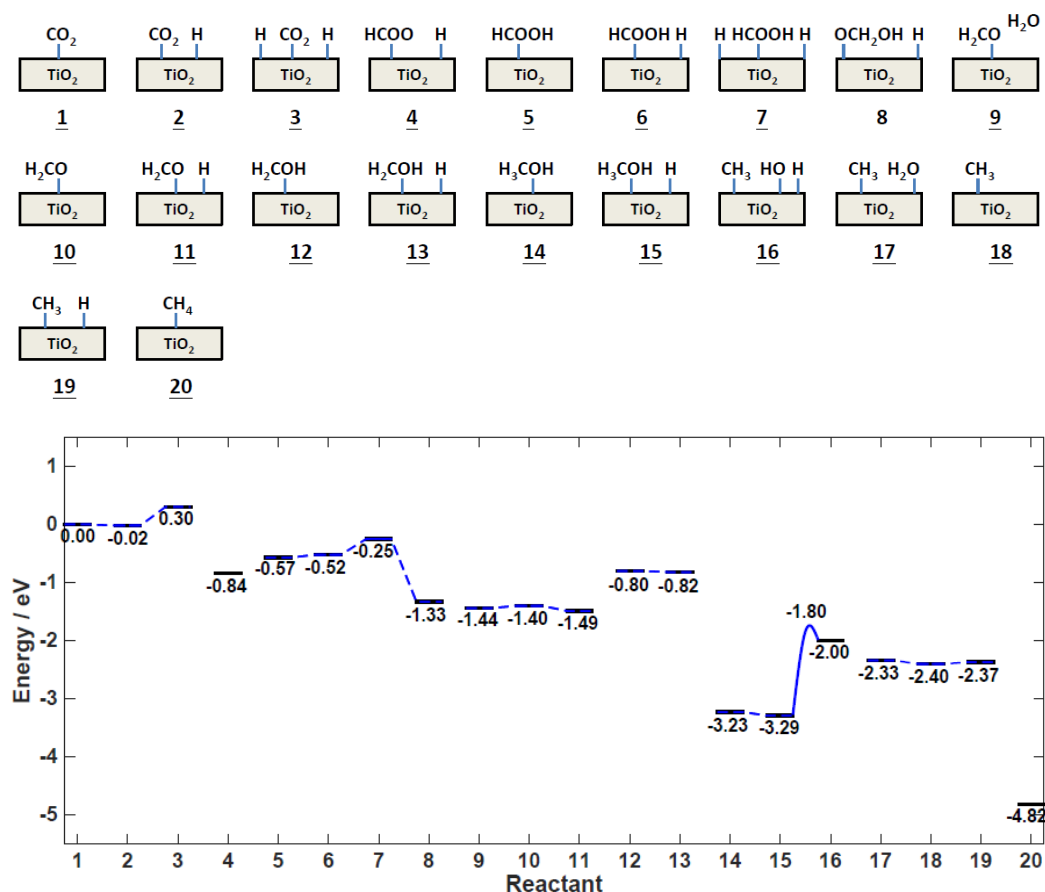
Figure 6.4 (Top) Illustrations of the intermediates involved in the formaldehyde pathway in this study. The underlined intermediates were species in the proposed reaction pathway,[13,29] while other intermediates were suggested in this study. (Bottom) Energy profile of the formaldehyde pathway.

It has been suggested in the literature[291] that H_2CO (**9**) and CH_3OH (**14**) were acting mainly as hole scavengers. From Fig. 6.4 it could be seen that it was more thermodynamically favorable for H_2CO (**9**) to form the oxidation product OCH_2OH (**8**) rather than being reduced to gain H_2COH (**12**), which was in

agreement with literature.[291] Similar preference to form oxidation product H$_2$COH (12) by CH$_3$OH was however not observed from Fig. 6.4. Such observation was in line with some experimental results. For instance, it has been demonstrated that the hole-scavenging power of CH$_3$OH was much weaker than other species such as H$_2$CO;[291] a temperature-programmed desorption (TPD) experiment[313] has suggested that methoxy, rather than molecularly adsorbed CH$_3$OH, was the effective hole-scavenging species in photo-oxidation of CH$_3$OH on TiO$_2$. Although in Fig. 6.4 the driving force for the forward reaction (15 to 16) was more favorable than the backward (14 to 13), the former was still highly thermodynamically unfavorable, which suggested CH$_3$OH (14) was more likely a product or by-product rather than an intermediate, a feature that has also been considered in the proposed carbene pathway.[13] The relative stability of CH$_3$OH would also suggest that this species was likely to be involved in other side reactions, such as the indirect photo-oxidation, which was an oxidation mechanism of CH$_3$OH that was equally supported in comparison to the direct oxidation,[94] and/or the molecular CH$_3$OH reacted with co-adsorb oxygen to generate methoxy.[313] As such CH$_3$OH may take mainly the role of a hole scavenger.

6.4.4 Glyoxal pathway

We first noted that in the glyoxal pathway CO was generated but did not act as a reaction intermediate. The fate of CO was however, not clear, where in theory it could also re-enter into the reaction cycle and act as an intermediate to produce CH$_4$. This postulation was supported by the general observation of CO being a typical minor or trace product, unless a co-catalyst or propan-2-ol was employed.[13] Therefore, in accordance with our postulation, when CO was formed

in the reaction of:

$$2CO_2 + 10H^\bullet \rightarrow CH_4 + 3H_2O + CO \qquad (Eq. \ 6.8)$$

It reacted with H atoms and produced CH_4:

$$CO + 6H^\bullet \rightarrow CH_4 + H_2O \qquad (Eq. \ 6.9)$$

and the total reaction was the doubling of Eq. 1.4. For the purpose of correctly aligning the energies of the intermediates and constructing an energy profile that was comparable with those of the other two pathways, we considered a total reaction which was the sum of (half) (Eq. 6.8) and (Eq. 6.9). The resulting ΔE_{tot} between the initial reactant, CO_2, and the final product, CH_4, would therefore be the same as computed for the other two proposed pathways shown previously (-4.82 eV). The mechanism of converting CO to CH_4 was, however, not studied, since they were not provided in the proposed pathway. In this case we would add a 'virtual step' of direct conversion of ½ CO to ½ CH_4 after the formation of the first ½ CH_4 in the profile.

Fig. 6.5 shows how the glyoxal path was defined in this study and the energy profile of this pathway. The red dashed line connects the initial (½ CO + 3 H) and final state (½ CH_4 + ½ H_2O) of the 'virtual step', in which the intermediates in between were not studied. The reaction began with the formation of HCO (**7**) from linearly adsorbed CO_2 (**1**) via the formation of formic acid (HCOOH) (**5**). Glyoxal (OHCCHO; (**9**)) was then produced from the dimerization of HCO (**7**). The subsequent step-wise H atom transfer produced trans-ethane-1,2,-semidione (**11**), glycolaldehyde (**13**), vinoxyl radical (**15**) and acetaldehyde (**18**). The acetaldehyde (**18**) was then transformed to give acetyl radical and a H atom (**19**). Thereafter the acetyl radical (**19**) was cleaved to form CO and methyl radical (**20**), which further abstracted a H atom to generate CH_4 (**23**). The reaction was complete when CO released (**20**) and re-entered the reaction cycle (**24**) and

generated CH_4 **(Final)**.
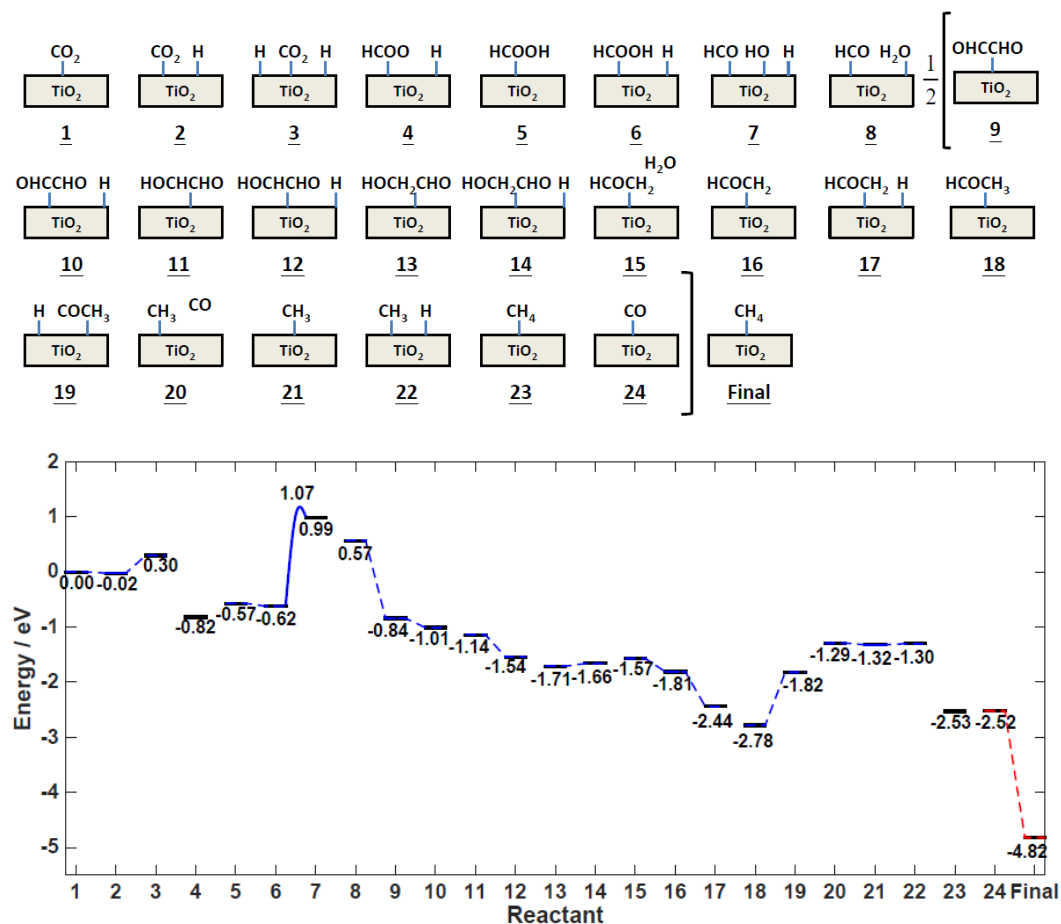
Figure 6.5 (Top) Illustrations of the intermediates involved in the glyoxal pathway in this study. The '½' brackets indicated the energies of which intermediates were halved. The underlined intermediates were species in the proposed reaction pathway,[13,119] while other intermediates were suggested in this study. (Bottom) Energy profile of the glyoxal pathway.

From Fig. 6.5 it could be seen that the rate limiting step was likely to be the formation of the considerably unstable HCO (**6 to 7**). The ΔE_{tot} of this step was 1.61 eV, and the E_{act} was 1.69 eV. The second most unfavorable step was the formation of CH_3 from acetaldehyde (**18 to 19;** 0.96 eV). Similar to the formaldehyde pathway, by considering the criteria given previously, we have omitted the computation of kinetic barriers for this pathway, except for the rate-limiting step. We also noted that acetaldehyde (**18**) would be the preferred

product over ½ CH$_4$ **(23)** and ½ CO **(20)**, but if ½ CO **(20)** was allowed to be further reduced and the product of this mechanism would be CH$_4$ **(Final)** only. From Fig. 6.5 acetaldehyde **(18)** was relatively stable on the surface, but it was a rather rare product.[13,292] This suggested acetaldehyde **(18)** was also likely to be involved in other side reactions, hindering the conversion to CH$_4$ **(Final)**. The mechanism in Fig. 6.5 would also be consistent with the difficulty of detecting glyoxal **(9)** and glycolaldehyde **(13)**,[119] as there was currently, to the best of our knowledge, no report of these species.

6.5 Discussion

Fig. 6.6 shows the simplified version of the three proposed pathways, where only important intermediates from H atom transfer reactions are included. Fig. 6.6 was constructed based on the computation of intermediates with a two-layer slab; a comparison with this figure based on a five-layer slab saw our qualitative conclusions remaining unchanged, and this comparison was given in section A6.5.

From Fig. 6.6 it was evident that the formaldehyde pathway was the most thermodynamically favorable pathway, where the energies in general go 'downhill' from one closed-shell product to another **(0th → 2nd → 4th → 6th → 8th H transfer, blue)**, whereas high-energy intermediates were involved in the other two pathways (HCO **(1.5th H transfer, green)** in the glyoxal pathway and COOH/C **(1st/4th H transfer, red)** in the carbene pathway). An observation was that H$_2$CO **(4th H transfer, blue)** was a favorable 'stepping stone' for all proposed pathways; it was seemingly more favorable for CO in the carbene pathway **(2nd H transfer, red)** and the HCO in the glyoxal pathway **(1.5th H**

transfer, green) to form H_2CO rather than a C atom **(4th H transfer, red)** and glyoxal **(3rd H transfer, green)** respectively. The formation of H_2CO from CO has also been suggested in other DFT studies of CO_2 methanation on Cu surface[310] and the anatase(101) surface.[137]
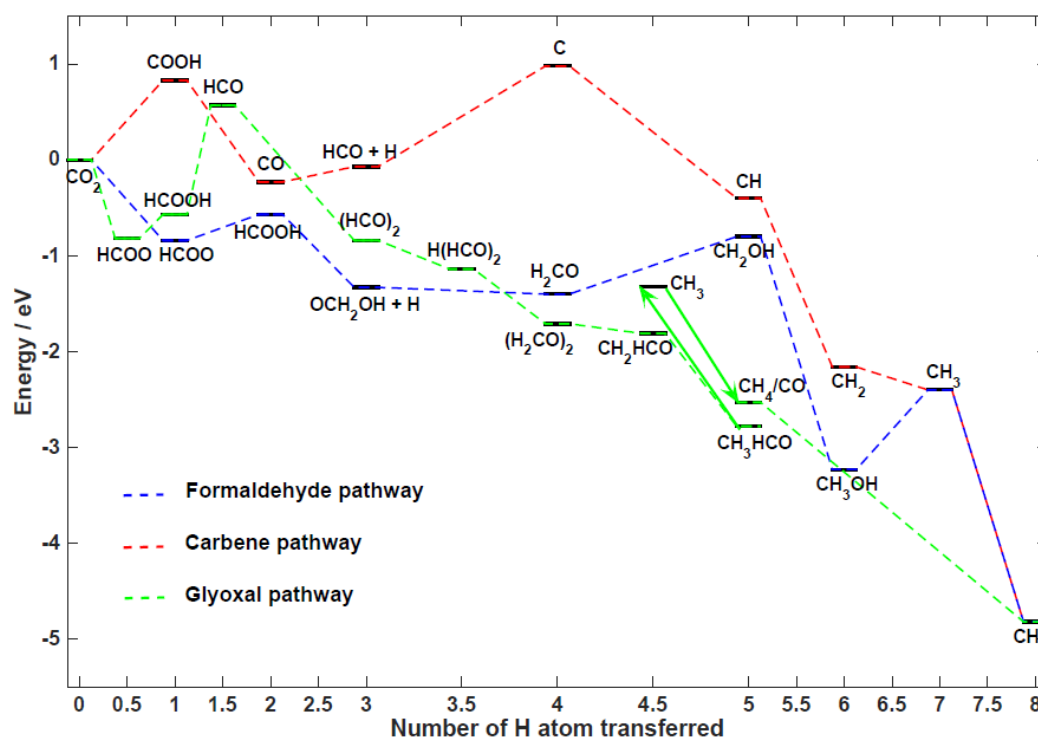
Figure 6.6 Comparison of the intermediates' energy for the carbene (red), formaldehyde (blue) and glyoxal (green) pathways. The green arrow from the 5th to the 4.5th H atom transferred indicated CH_3HCO lost a H atom to form CH_3, and the green arrow from the 4.5th to the 5th H atom transferred indicated CH_3 regained a H atom to form CH_4. Note that the region of 3rd to 5th H atom transferred was magnified. Intermediates with '+ H' meant a H atom co-adsorbs on the surface with the species.

We have also tested the dependence of our identification of the most favourable pathway on our chosen functional. We have computed the reaction energies between CO_2 and the highest-energy intermediates in each pathway with PBE + U (U = 4.0 eV)[137] and a five-layer slab, and compared them. The reaction energies in the carbene, glyoxal and formaldehyde pathways were 2.12 eV, 1.91

eV, and 1.21 eV respectively, and formaldehyde pathway should remain as the most favorable pathway.

We also noted that the uncertainty in our comparison of reaction mechanisms due to the introduction of some alternative intermediates for the carbene pathway (**1st and 3rd H transfer, red**) was deemed unimportant. The highest-intermediate in the carbene pathway was the C atom (**4th H transfer, red**), which was much higher in energy than the intermediates in the formaldehyde pathways. This observation would not be influenced by considering alternative intermediates in the carbene pathway.

Our observations based on Fig. 6.6 should be verifiable with currently available experimental techniques. For instance, according to the results in Fig. 6.6, OCH_2OH (**3rd H transfer, blue**) should be easier to detect than other radical species (**5th and 7th H transfer, blue**) in the formaldehyde pathway, and the identification of this species together with the following closed-shell products would be important supporting evidence for this pathway. OCH_2OH has been identified in the reaction between atomic hydrogen and formic acid in Kr matrix with IR spectroscopy,[314] giving rise to an IR signal, such as an intense band at around 3600 cm^{-1} due to O-H stretching. The assignment of the IR spectra might be simplified by using other versions of IR spectroscopic technique, e.g. polarization modulation infrared reflection adsorption spectroscopy (PM-IRAS).[118] The conversion of CO and HCO to H_2CO would require kinetic modeling and/or isotope labeling techniques for verification.

From Fig. 6.6, we would also expect the main bottleneck for the formaldehyde pathway to be the unfavorable formation of CH_3 (**6th → 7th H transfer, blue**). A new catalytic system should therefore improve the ease with which the C-O bond of CH_3OH cleaved. Alternatively, without changing the catalyst, the cleavage of

this C-O bond may be facilitated by adding hydrogen iodide to react with CH_3OH in order to generate CH_3I. TPD experiment showed that the adsorption of a CH_3I layer on $TiO_2(110)$ would produce CH_4.[315] CH_3I was also known to dissociate to produce CH_3 on other surfaces.[316,317]

We have noticed, on the basis of Fig. 6.6, two possible alternative sub-pathways that would avoid this unfavorable step. The first possibility was that H_2CO produced CH_2 instead of CH_3OH; but such reaction, to the best of our knowledge, was not known from literature. The second possibility was that the formaldehyde pathway proceeded via concerted 2-H atom transfers, i.e. the pathway incorporated only intermediates at 0th, 2nd, 4th, 6th, 8th H atom transfer **(blue)** in Fig. 6.6, without the formation of specific intermediates in between. This alternative pathway was similar to the mechanism suggested in another theoretical study[137] but there was a lack of experimental evidence,[13] and it was not clear from Fig. 6.6 what experiment could be performed to provide evidence. On the basis of literature, a third possibility was the formation of methoxy radical (OCH_3) from H_2CO, where OCH_3 has also been reported previously,[290] and was part of the mechanism of methanation of CO_2 on Cu surface determined from DFT calculations.[310] OCH_3 was not included in Fig. 6.6, but it was expected that this alternative path would also have the difficulty to cleave the C-O bond to generate CH_3, and the OCH_3 may prefer hole scavenging.[313]

A few effects of the reaction environment were not encompassed in our computation but might affect our prediction. The solvent molecules could stabilize strongly some of the species through, for example, H-bonding (the key intermediates were neutral so there were no major differences expected in the polarization energy). However, it could be noted that, considering the typical H-bond energy in water ($\sim < 0.4$ eV),[318] it was not possible to alter significantly

the relative energies of the landscape depicted in Fig. 6.6. The mechanism might also be interfered by other possible species formed in side reactions, such as carbonates, hydrogen molecules and hydroxyl radicals.[13] Surface defects might also alter the mechanism and/or the energetics of the mechanism prominently. It has been suggested that in the presence of oxygen vacancy CO was more easily formed,[122,319] and the pathway would be more similar to the carbene pathway.[137]

The results presented alongside the tests performed to rule out important computation errors allowed the identification of a reaction path among three that was clearly more favourable. As noted above the energy difference between intermediates was sufficiently large and the barriers sufficiently low that the detailed investigation of alternative mechanisms for the elementary reaction steps would not change the conclusion for the given energy landscape. It should be noted that the study of the direct non-adiabatic electron transfer from the defect required a very different type of study from the one we presented. Possibly the study should focus on a single reduction step and the other mechanisms it could follow, such as ET followed by PT, PT followed by ET, concerted proton-electron transfer, and multi-electron transfer as proposed in previous theoretical studies[136,137,290] but deemed unlikely[13,136,320] and without strong experimental support. Each elementary step would require adjustment to standard DFT, such as constrained DFT[156,321] for electron localization, and/or a periodic charged slab with background compensating charge that the energy may depend on the width of the vacuum layer in the simulation box.[322] Alternative model for the charged slab could perhaps be the adsorption of H atom on anatase(101),[136] but the treatment of localization of the excess electron would require relatively costly computational method such as DFT+U, and the results would be dependent on the U value employed.[136,323] Thus, a possible strategy for

investigating photo-catalytic reaction mechanisms was to consider initially a broad exploration of global reaction mechanisms, as we did here, followed by an in-depth study of some elementary steps if necessary.

6.6   Conclusions

In conclusion, we have investigated theoretically three proposed reaction mechanisms for the photo-catalytic reduction of CO_2 to gain CH_4 (Eq. 1.4) on defect-free TiO_2 anatase(101) with first-principles DFT calculations, and we determined that the formaldehyde pathway was the most likely on the basis of a greater thermodynamic stability of the intermediates. Formaldehyde was a thermodynamically preferred intermediate to form in the hydrogenation of CO and the hydrogenation of HCO, over C atom in the carbene pathway and glyoxal in the glyoxal pathway respectively. Our computational approach appeared to be useful for both developing sensible mechanistic hypothesis and designing experiments to validate them.

6.7 Appendix

A6.1 Dependence of reaction energy and adsorption energy on slab thickness

Fig. A6.1 shows the dependence of reaction energy (eV) on slab thickness (2 to 6 layers), using

$$CO_2 + 2H \rightarrow HCOOH \qquad \text{(Eq. A6.1)}$$

as an example. The reaction energy was converged at five-layer, where the reaction energy difference between five-layer and six-layer was 0.05 eV.

Table A6.1 shows the adsorption energies of CO_2, HCOOH and H atom on the slabs with different thicknesses. The adsorption energy computed with PBE changes negligibly for CO_2 and HCOOH (less than 0.1 eV), but for H atom the difference between two and five layers was 0.3 eV, and considered converged at four-layer. All adsorption energies of H atom computed with PBE were close to literature values, which were in between -2.01 eV[301] to -2.31 eV,[300] depending on the surface coverage and computational details apart from the functional employed.

The adsorption energy for CO_2, HCOOH and H atom were also computed with PBE+U,[324] with $U = 4.0$ eV.[137] The difference between PBE with and without U for CO_2 and HCOOH were considerably small (0.07 eV and 0.12 eV respectively). For H atom, this difference (0.2 eV) was similar to those observed in the literatures,[301] but the value for +U computation was ~0.3 eV larger than some literature values (~ -2.30 eV).[300,301]
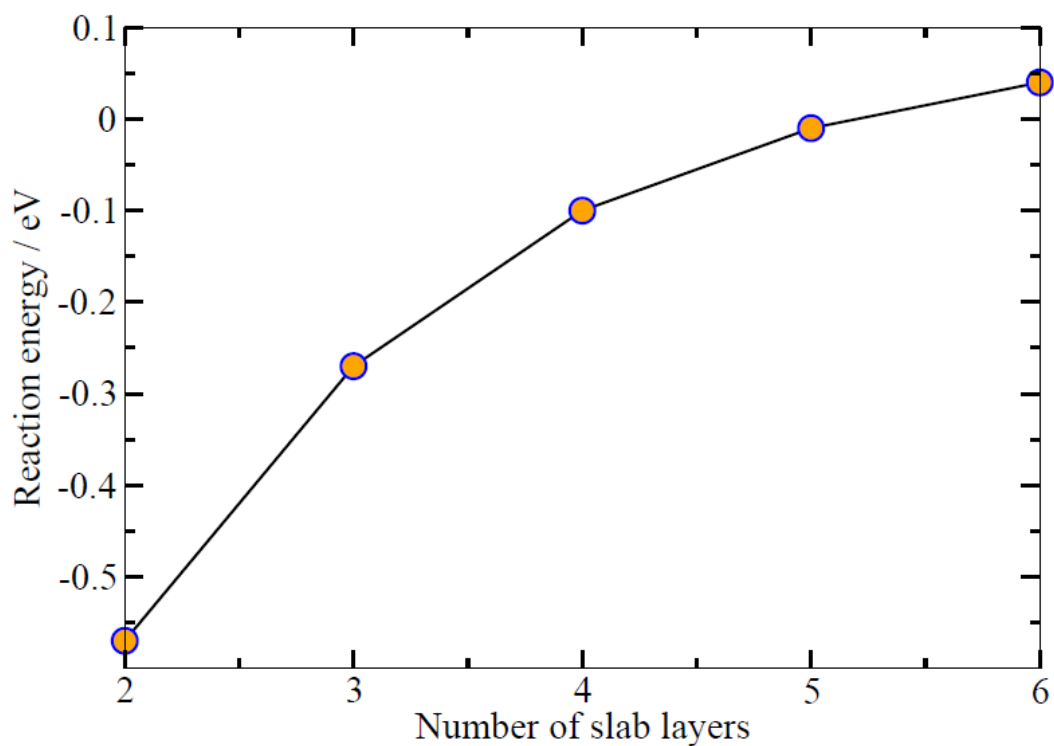
Figure A6.1 Dependence of reaction energy (eV) on slab thickness.

Table A6.1 Adsorption energies (eV) of CO_2, HCOOH and H on anatase(101) slab with 2 to 6 layers, computed with PBE. Numbers in bracket are computed with in PBE+U.

Number of layers	E_{Ad} of CO_2 / eV	E_{Ad} of HCOOH / eV	E_{Ad} of H / eV
2	−0.14	−0.41	−2.10
3	−0.14	−0.45	−2.27
4	−0.14	−0.44	−2.35
5	−0.14 (−0.21)	−0.46 (−0.58)	−2.40 (−2.61)
6	−0.15	−0.48	−2.43

A6.2 Adsorption geometries of intermediates and transition states in the carbene pathway

Fig. A6.2 shows the adsorption geometries of the intermediates and transition states in our computed carbene pathway (Fig. 6.1). 'TS' means transition states.
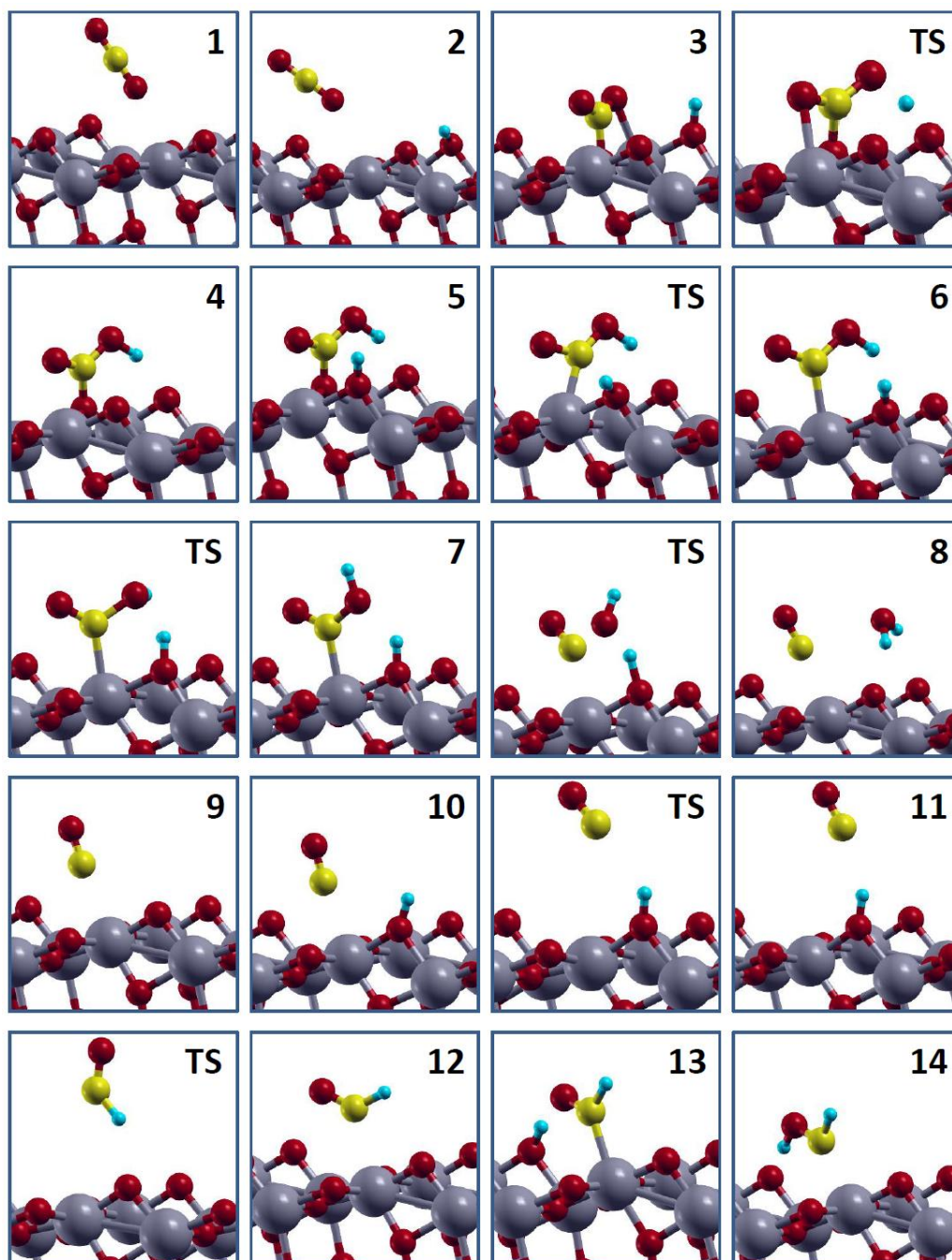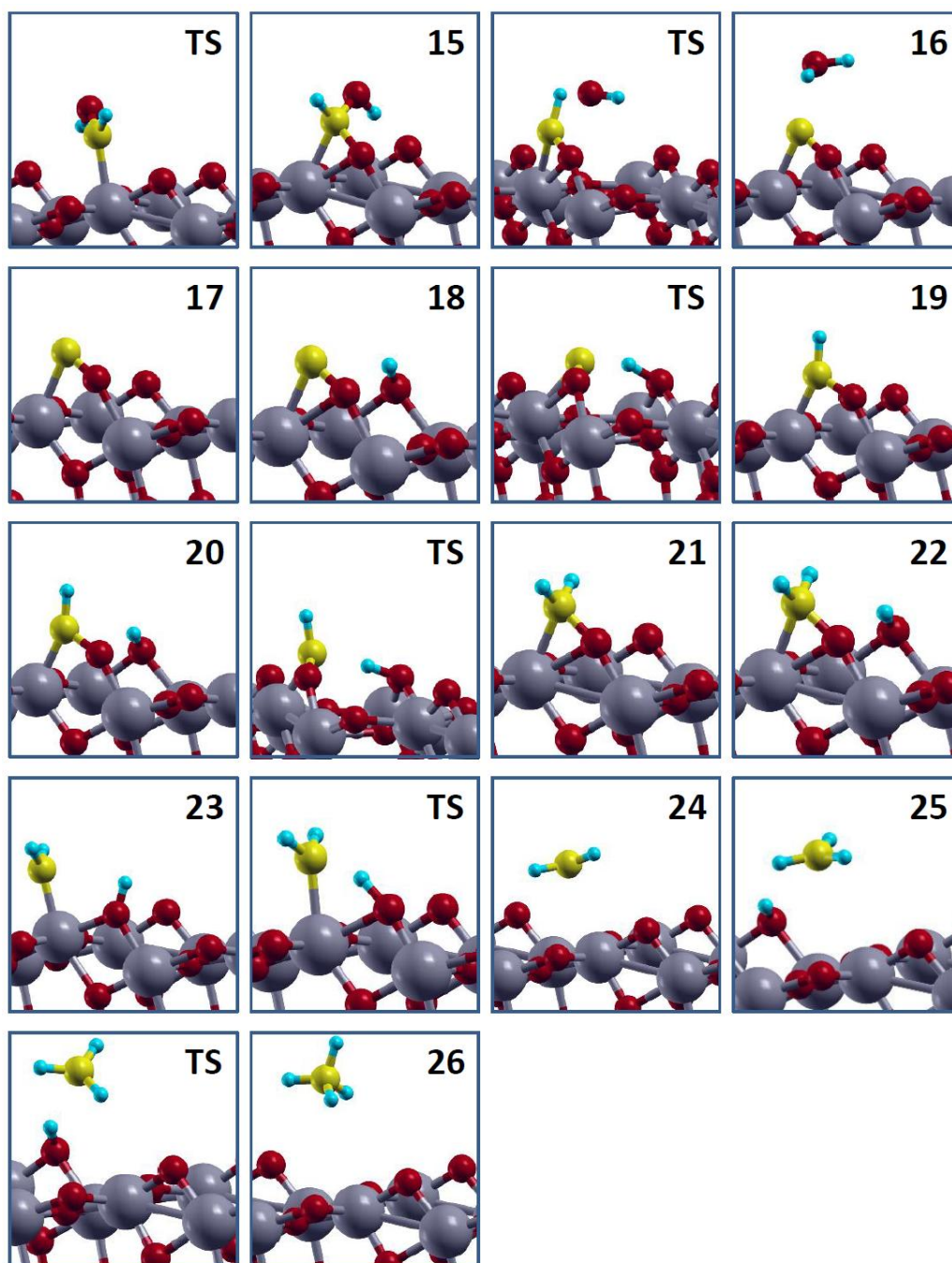
Figure A6.2 Adsorption geometries of the reaction intermediates and transition states in our computed carbene pathway.

A6.3 Adsorption geometries of intermediates in the formaldehyde pathway

Fig. A6.3 shows the adsorption geometries of the intermediates in the carbene pathway (Fig. 6.4(Top)). 'TS' means transition states.
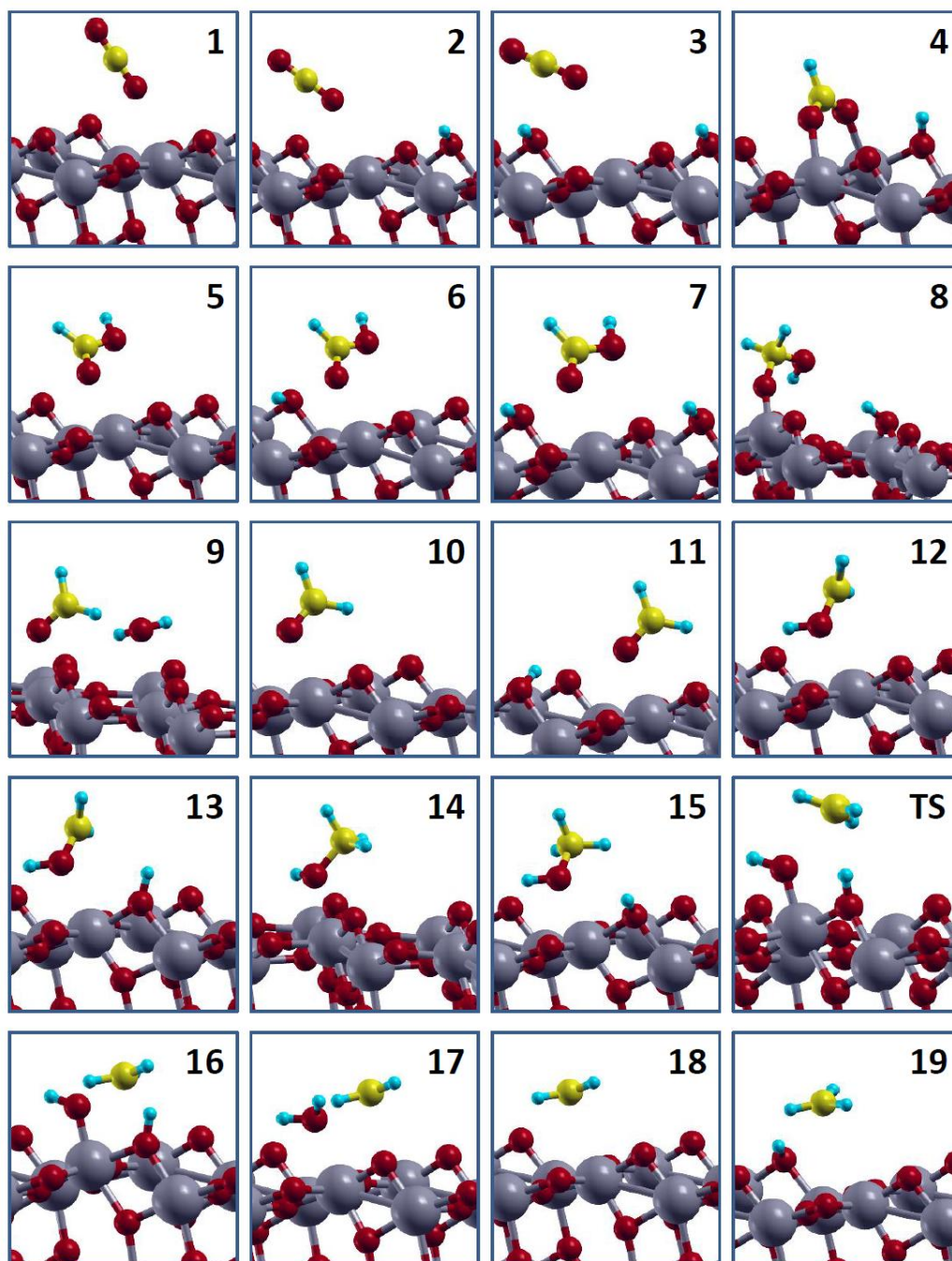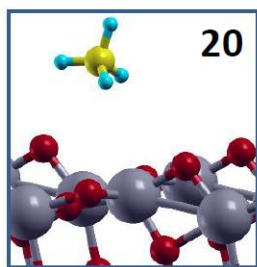
Figure A6.3 Adsorption geometries of the reaction intermediates in our computed formaldehyde pathway.

A6.4 Adsorption geometries of intermediates in the glyoxal pathway

Fig. A6.4 shows the adsorption geometries of the intermediates in the carbene pathway (Fig. 6.5(Top)). 'TS' means transition states.
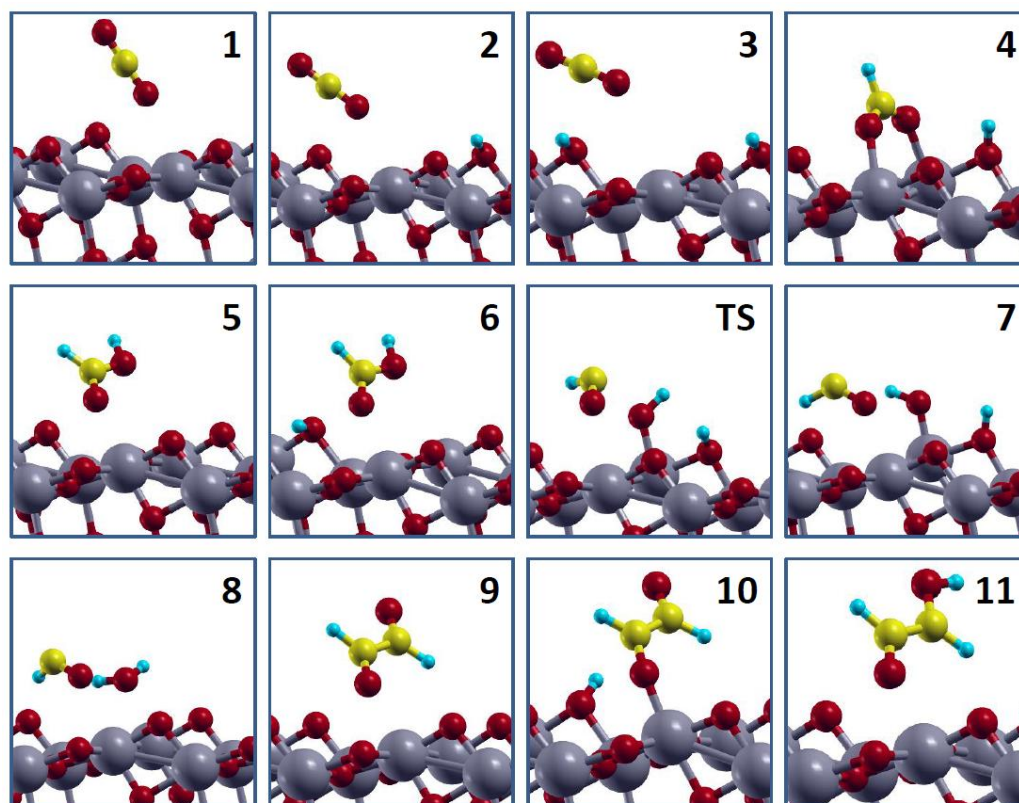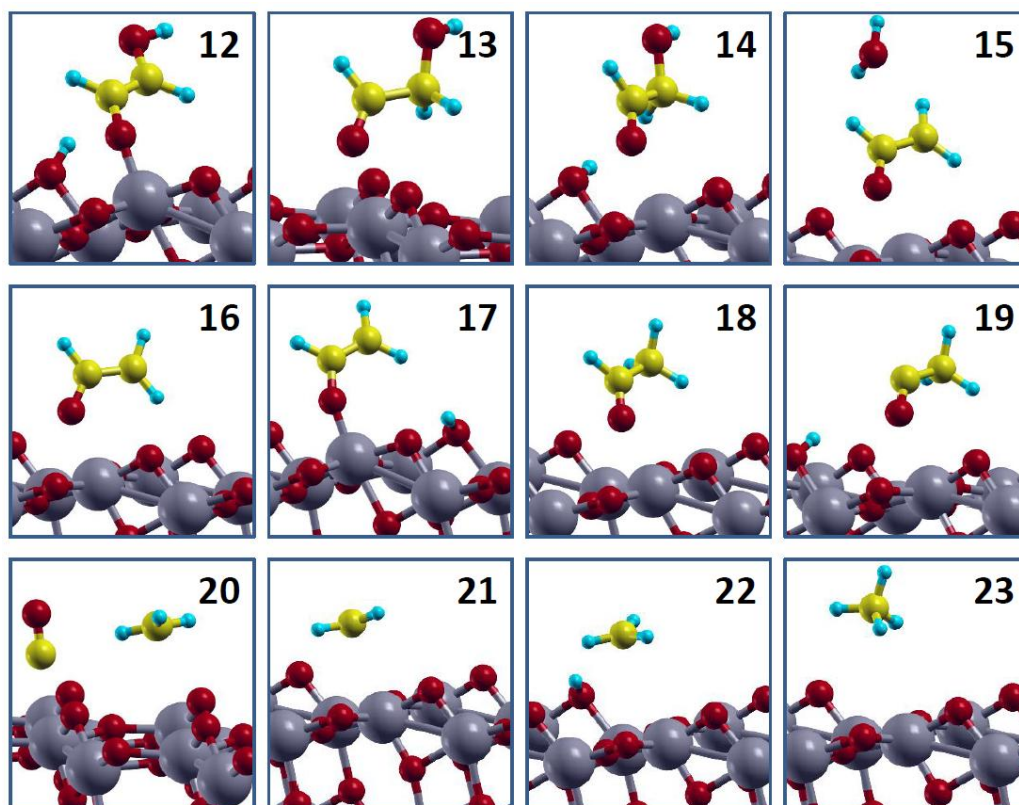
Figure A6.4 Adsorption geometries of the reaction intermediates in our computed glyoxal pathway.

A6.5 Effect of slab thickness on the identification of the most favorable mechanism

Our main objective in this study was to identify the most favorable mechanism by comparing three proposed mechanisms, as illustrated in Fig. 6.6 in section 6.5. As shown in Fig. A6.1 the reaction energy changed with the slab thickness, it was therefore deemed necessary to check if the formaldehyde pathway remained as the most favourable mechanism when thicker slab was employed.

Fig. A6.5 shows Fig. 6.6 in section 6.5 constructed with 5-layer-slab calculations. A comparison between the Fig. A6.5 and Fig. 6.6 in section 6.5 shows that the highest-energy intermediate in each pathway remained unchanged, and we would expect the formaldehyde pathway remained as the most favorable

pathway. In both cases, H_2CO (**4th H transfer, blue**) was a favorable 'stepping stone' for CO in the carbene pathway (**2nd H transfer, red**) and the HCO in the glyoxal pathway (**1.5th H transfer, green**).
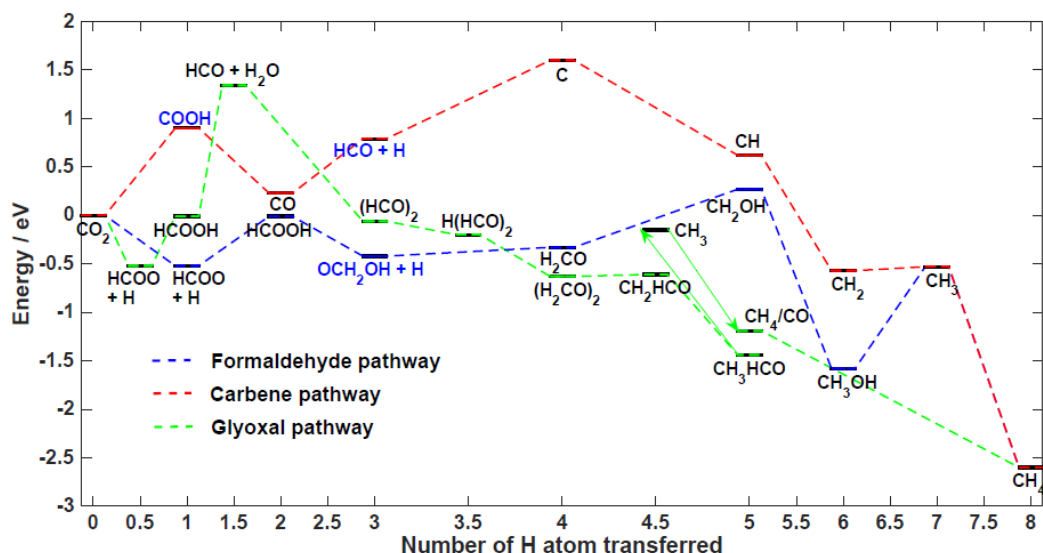
Figure A6.5 Modified Fig. 6.6 in section 6.5, constructed with computations of intermediates based on a 5-layer slab. The intermediates proposed in experiments were labeled in black, while intermediates labeled in blue were suggested in this study.

On the other hand, some differences in reaction energies were observed between the two profiles, which were mostly originated from the difference in the adsorption energy of H atom. For instance, the reaction energy of CO_2 to CO (**2nd H transfer, red**) had ~ 0.5 eV difference. This was similar to the example of HCOOH given in section A6.1, where the difference in the adsorption energy of H atom between two-layer and five-layer slab was 0.3 eV for one H atom (0.6 eV in this case due to two H atoms involved). The reaction energy difference in the two profiles for CO_2 to CH_4 was ~2.2 eV, which was also close to the accumulated difference of 2.4 eV (due to the requirement of 8 H atoms). As discussed, our adsorption energies of H atom with different slab thickness were close to the range of literature values observed (~-2 eV to ~-2.30 eV). Accurate

reaction energies were therefore difficult to be determined, due to the uncertainty contained within the reported values of adsorption energy of H atom, which was subject to the computational method and details, as well as surface coverage.

Chapter 7   Conclusion

In this work we attempted to aid the design of new materials in DSSCs and photo-catalytic reduction of CO_2, by specifically studying elementary processes (Chapter 3), providing predictive tools (Chapter 4) and testing hypotheses (Chapter 5 and 6). All attempts were based primarily on electronic structure calculations, but in conjunction with various strategies that were philosophically distinctive, such as statistical and physical modeling. One reason for the employing different strategies was that the technological statuses of the explored solar energy technologies were different, and a single strategy was not suitable.

For p-type DSSCs, the device is currently under-developed, with low PCE that is far from other commercialized photo-voltaic systems, and would also hinder the development of tandem DSSCs. The hole recombination at the semiconductor-dye interface of this type of device is deemed one of the main causes for the low PCE, and the reason for such process being much faster than the analogous recombination in relatively efficient n-type DSSCs is not known. Non-adiabatic charge transfer theories were used to evaluate the hole injection and recombination rates at the NiO-C343 interface, with the prime intention to understand the fast interfacial recombination in p-type DSSCs in chapter 3. It has been shown that the difference in the Franck-Condon factor for the two types of recombination was the main reason for the difference in the reaction rates. On the basis of this analysis it was therefore suggested that increasing the reduction potential and/or reducing reorganization energy of dyes should alleviate recombination. An area of this work that can be improved would be the modeling of the electronic structure of the surface. Our computation of the transfer rates rely on the introduction of experimental parameters, and it would be desirable if

a more accurate electronic structure can be obtained and reduce or even eliminate this dependence. A possibility would be to use higher level of theory such as the *GW* approach.[147,148]

For n-type DSSCs, the understanding of charge dynamics in the device is better in comparison to the p-type device. Simple physical models are in place and applied, and further unknown physical principles are likely to be unraveled with advanced models, which may be dependent on the availability of advanced computational power. On the other hand, many dyes have been attempted but the research progress in terms of PCE is slow. A method that allows quick estimation of the PCE of a hypothetical dye before time-consuming synthesis to avoid wasteful attempts would be desirable. We have adopted a QSAR-like statistical modeling approach to create a tool for this purpose. As such, given the reorganization energy and free energy change of dye oxidation of a hypothetical dye, it is possible to predict the probability of this dye achieving PCE over a certain threshold. This work can be extended further by considering other families of dyes in terms of anchoring groups and/or electrolyte, or possibly improved by considering a larger dataset or a different set of descriptors. In addition, a common design philosophy for dyes, the D-π-A scheme, has not been formally questioned. We have tested the efficacy of such design scheme by observing the correlation between some quantified measures for D-π-A strength and PCE with a large sample of 116 dyes, but no correlation was found. Such statistical study has demonstrated how statistics can be used to test postulation based on physical understanding.

For photo-catalytic reduction of CO_2 to form methane, the design of new catalyst is largely based on trial-and-error, and the molecular reaction mechanism, and therefore the rate-limiting step, is not clearly known. Since competing

proposed reaction mechanisms exist, it is important to first identify the most favorable pathway before any further attempts of in-depth understanding on the origin of the rate-limiting step. With this objective we have therefore compared the thermodynamic landscapes of three proposed reaction mechanisms, constructed with the DFT total energy of relevant reaction intermediates. The most favorable pathway was identified as the formaldehyde pathway. This work can be improved by incorporating solvent effects[101,325] and the impact of defects such as oxygen vacancies on the reaction mechanism.

In terms of technological aspects, only around 5% of PCE improvement was seen in the last 25 years for n-type DSSCs and therefore it seems reasonable to consider devoting more efforts into tandem cells for further PCE advancement. As discussed, improving p-type devices would be crucial for the realization of efficient tandem cells. An area that has not been explored theoretically was the hole recombination at the electrolyte-semiconductor interface, which was another important limiting factor for high PCE.[208,326] It appears that both NiO[326] and the dye-electrolyte interaction[327] can catalyze this process. The recombination theory introduced in this thesis can be adopted to study also the recombination at the electrolyte-semiconductor interface.[88]

Research progress for DSSCs must be accelerated in order to compete with the fast-growing perovskite solar cells, which at the moment suffer from stability issues[328] but can be low-cost and much more efficient than DSSCs.[329,330] To accelerate the research progress, it is perhaps important to develop first-principle tools that allow prediction of the PCE of DSSCs with hypothetical materials within short time. Apart from the example given in Chapter 4, the PCE may be predicted by reproducing the experimentally measured parameters that enter into Eq. 1.1, such as the short-circuit current density,[331] but both models are not the

most accurate.

The research progress of photo-catalysis is also considerably slow. Perhaps the most important aspect would be to establish a common characterization of the performance of catalytic systems, measured under similar well-defined conditions *c.f.* DSSCs. This is highly challenging since the performance of a catalyst cannot be characterized by contemplating only a single aspect, e.g. photo-response of the catalyst or the catalyzing power of the catalyst, and the effects of various interactions in the reaction system are largely unknown. This in turn can be an opportunity for theory since computations are not affected by the reaction environment. Possibly, a theoretical ranking of catalysts can be produced based on a selected predictor for catalysts' performance, and comparing this ranking with existing experimental ranking based on various efficiency measures[13,106] may reveal important factors that cause matching or mismatching. Recently a descriptor model for metal oxide catalysts has been developed, which was based on radial distribution function as a descriptor.[332] Alternative descriptor can be, for instance, adsorption energy.[113]

While DFT has demonstrated its ability to provide potential energy surfaces for complementing the understanding of photo-catalytic reaction mechanism, this approach is time-consuming, especially when the reaction is complex and transition state search is required. This approach would also be cumbersome if one considers studying mechanism on different surfaces or with different solvents. A simpler method to test mechanistic hypotheses would be desirable when a large amount of them exist, and ideally this method would avoid expensive transition state searches. Establishing linear free-energy relationships, such as the BEP relations, to allow estimation of activation energy can perhaps be the first step towards a simpler method that requires only the energy of the

intermediates. The number of calculations may be further reduced by employing scaling relationships for adsorption energy.[113]

In conclusion, this thesis serves as an example of using mixed strategies on the basis of electronic structure calculations for aiding the design of new materials in solar energy technologies, demonstrating that theories and computational chemistry can be flexibly applied in the development of these technologies. Researchers should contemplate different strategies instead of deciding intuitively before performing experiments; as demonstrated in this thesis, even in a subject dominated by physical modeling, statistical studies can be useful to, for example, test hypotheses and make predictions. We believe statistical studies similar to those presented in chapter 4 and 5 can be particularly helpful to shed light on future research direction when sufficient amount of suitable data is available. The different theories and methods presented in this work, such as the charge transfer theories and statistical methodologies, are all transferrable, and we will be excited to see these methods being adapted for future studies.

Bibliography

1   US Energy Information Administration, http://www.eia.gov/todayinenergy
        /detail.cfm?id=25372, (accessed July 2016).

2   S. Dale, presented in part at the Society of Business Economists Annual
        Conference, London, October, 2015.

3   M. K. Hubbert, *Science.*, 1949, **109**, 103.

4   National Aeronautics and Space Administration, http://climate.nasa.gov/
        effects/., (accessed July 2016)

5   World Nuclear Association, http://www.world-nuclear.org/information
        -library/current-and-future-generation/nuclear-power-in-the-world-today.a
        spx., (accessed July 2016).

6   A. V Yablokov, V. B. Nesterenko and A. V Nesterenko, *Chernobyl:
        Consequences of the Catastrophe for People and the Environment.*, New
        York Academy of Sciences, New York, 2009.

7   International Atomic Energy Agency, https://www.iaea.org/newscenter/
        news/five-years-after-fukushima-making-nuclear-power-safer, (accessed
        July 2016).

8   E. Lucas, *The Economist.*, 2015, Jan 17, 1.

9   C. Philibert, *Technology roadmap*, International energy agency, Paris,
        2014.

10   R. Kempener and E. Borden, *Battery Storage for Renewables : Market
        Status and Technology Outlook*, International renewable energy agency,
        Bonn, 2015.

11   B. Norton, *Harnessing solar heat*, Springer, Dordrecht, 2014.

12   H. Kisch, *Angew. Chem. Int. Ed.*, 2013, **52**, 812.

13   S. N. Habisreutinger, L. Schmidt-Mende and J. K. Stolarczyk, *Angew.
        Chem. Int. Ed.*, 2013, **52**, 7372.

14   R. Chandra, H. Takeuchi and T. Hasegawa, *Renew. Sust. Energ. Rev.*,
        2012, **16**, 1462.

15   C. O. Karacan, F. A. Ruiz, M. Cote and S. Phipps, *Int. J. Coal Geol.*, 2011,
        86, 121.

16   A. Polman, M. Knight, E. C. Garnett, B. Ehrler and W. C. Sinke, *Science*,
        2016, **352**, 307.

17   J. Kalkman, A. Merhaba, S. Bose and H. Bradley, *Emerging technologies
        in Solar PV : identifying and cultivating potential winners Content*, Arthur
        D. Little, Dubai, 2015.

18   J. Salvatore, *World Energy Perspective - Cost of Energy Technologies*, ed. G. Turner, H. Boyle, E. Nekhaev, A. Clerici and S. Ulreich, World Energy Council, London, 2013.

19   W. Shockley and H. J. Queisser, *J. Appl. Phys.*, 1961, **32**, 510–519.

20   C. Battaglia, A. Cuevas and S. De Wolf, *Energy Environ. Sci.*, 2016, **9**, 1552.

21   B. O'Regan and M. Gratzel, *Nature*, 1991, **353**, 737.

22   W. J. Youngblood, S. H. A. Lee, K. Maeda and T. E. Mallouk, *Acc. Chem. Res.*, 2009, **42**, 1966.

23   National center for photovoltaics, http://www.nrel.gov/ncpv/, (accessed July 2016).

24   H. S. Lee, S. H. Bae, Y. Jo, K. J. Kim, Y. Jun and C. H. Han, *Electrochim. Acta*, 2010, **55**, 7159.

25   K. Hara, Z. S. Wang, Y. Cui, A. Furube and N. Koumura, *Energy Environ. Sci.*, 2009, **2**, 1109.

26   E. J. Maginn, *J. Phys. Chem. Lett.*, 2010, **1**, 3478.

27   A. L. Linsebigler, G. Lu and J. T. Yates, *Chem. Rev.*, 1995, **95**, 735.

28   A. Fujishima and K. Honda, *Nature*, 1972, **238**, 37.

29   T. Inoue, A. Fujishima, S. Konishi and K. Honda, *Nature*, 1979, **277**, 637.

30   S. Protti, A. Albini and N. Serpone, *Phys. Chem. Chem. Phys.*, 2014, **16**, 19790.

31   A. Harriman, *Phil. Trans. R. Soc. A*, 2013, **371**, 20110415.

32   T. Maschmeyer and M. Che, *Angew. Chem. Int. Ed.*, 2010, **49**, 1536.

33   R. Nakamura and Y. Nakato, *J. Am. Chem. Soc.*, 2004, **126**, 1290.

34   F. Filippone, G. Mattioli and A. A. Bonapasta, *J. Am. Chem. Soc.*, 2006, **128**, 13772.

35   R. A. Marcus, *J. Chem. Phys.*, 1956, **24**, 966.

36   R. A. Marcus and N. Sutin, *Biochim. Biophys. Acta*, 1985, **811**, 265.

37   C. Draxl, D. Nabok and K. Hannewald, *Acc. Chem. Res.*, 2014, **47**, 3225.

38   W. Liu, A. Tkatchenko and M. Scheffler, *Acc. Chem. Res.*, 2014, **47**, 3369.

39   N. Martsinovich and A. Troisi, *Energy Environ. Sci.*, 2011, **4**, 4473.

40   K. N. Houk and P. H. Cheong, *Nature*, 2008, **455**, 309.

41   D. Cheng, F. R. Negreiros, E. Apra and A. Fortunelli, *ChemSusChem*, 2013, **6**, 944.

42   G. J. Cheng, X. Zhang, L. W. Chung, L. Xu and Y. D. Wu, *J. Am. Chem. Soc.*, 2015, **137**, 1706.

43   E. Kraka and D. Cremer, *Acc. Chem. Res.*, 2010, **43**, 591.

44   S. Mathew, A. Yella, P. Gao, R. Humphry-Baker, B. F. E. Curchod, N. Ashari-Astani, I. Tavernelli, U. Rothlisberger, M. K. Nazeeruddin and M. Grätzel, *Nat. chem.*, 2014, **6**, 242.

45   A. Hagfeldt, G. Boschloo, L. Sun, L. Kloo and H. Pettersson, *Chem. Rev.*, 2010, **110**, 6595.

46   M. K. Nazeeruddin, A. Kay, I. Rodicio, R. Humphry-Baker, E. Muller, P. Liska, N. Vlachopoulos and M. Gratzel, *J. Am. Chem. Soc.*, 1993, **115**, 6382.

47   M. K. Nazeeruddin, S. M. Zakeeruddin, R. Humphry-Baker, M. Jirousek, P. Liska, N. Vlachopoulos, V. Shklover, C. H. Fischer and M. Grätzel, *Inorg. Chem.*, 1999, **38**, 6298.

48   B. E. Hardin, H. J. Snaith and M. D. McGehee, *Nat. Photonics*, 2012, **6**, 162.

49   A. Mishra, M. K. R. Fischer and P. Bäuerle, *Angew. Chem. Int. Ed.*, 2009, **48**, 2474.

50   C. H. Henry, *J. Appl. Phys.*, 1980, **51**, 4494.

51   A. Nattestad, A. J. Mozer, M. K. R. Fischer, Y. B. Cheng, A. Mishra, P. Bäuerle and U. Bach, *Nat. mater.*, 2010, **9**, 31.

52   S. Powar, T. Daeneke, M. T. Ma, D. Fu, N. W. Duffy, G. Götz, M. Weidelener, A. Mishra, P. Bäuerle, L. Spiccia and U. Bach, *Angew. Chem. Int. Ed.*, 2013, **52**, 602.

53   H. J. Snaith, *Nat. Photon.*, 2012, **6**, 337.

54   A. N. M. Green, R. E. Chandler, S. A. Haque, J. Nelson and J. R. Durrant, *J. Phys. Chem. B*, 2005, **109**, 142.

55   C. Bauer, G. Boschloo, E. Mukhtar and A. Hagfeldt, *J. Phys. Chem. B*, 2002, **106**, 12693.

56   A. Morandeira, G. Boschloo, A. Hagfeldt and L. Hammarström, *J. Phys. Chem. B*, 2005, **109**, 19403.

57   R. Huber, J. E. Moser, M. Grätzel and W. Josef, *Chem. Phys.*, 2002, **285**, 39.

58   C. A. Schalley, *Analytical methods in Supramolecular chemistry*, Wiley-VCH, Weinheim, 2012.

59   J. Halme, P. Vahermaa, K. Miettunen and P. Lund, *Adv. Mater.*, 2010, **22**, E210.

60   Z. Huang, G. Natu, Z. Ji, P. Hasin and Y. Wu, *J. Phys. Chem. C*, 2011, **115**, 25109.

61   R. Kern, R. Sastrawan, J. Ferber, R. Stangl and J. Luther, *Electrochim. Acta*, 2002, **47**, 4213.

62   Q. Wang, J. E. Moser and M. Grätzel, *J. Phys. Chem. B*, 2005, **109**, 14945.

63   O. V. Prezhdo, W. R. Duncan and V. V. Prezhdo, *Acc. Chem. Res.*, 2008, **41**, 339.

64   I. Kondov, M. Thoss and H. Wang, *J. Phys. Chem. A*, 2006, **110**, 1364.

65   N. Martsinovich and A. Troisi, *J. Phys. Chem. C*, 2011, **115**, 11781.

66   T. Le Bahers, T. Pauporte, P. P. Laine, F. Labat, C. Adamo and I. Ciofini, *J. Phys. Chem. Lett.*, 2013, **4**, 1044.

67   A. V Akimov, A. J. Neukirch and O. V Prezhdo, *Chem. Rev.*, 2013, **113**, 4496.

68   G. Fratesi, C. Motta, M. I. Trioni, G. P. Brivio and D. Sanchez-Portal, *J. Phys. Chem. C*, 2014, **118**, 8775.

69   N. A. Anderson and T. Lian, *Annu. Rev. Phys. Chem.*, 2005, **56**, 491.

70   D. F. Watson and G. J. Meyer, *Annu. Rev. Phys. Chem.*, 2005, **56**, 119.

71   W. R. Duncan and O. V. Prezhdo, *Annu. Rev. Phys. Chem.*, 2007, **58**, 143.

72   I. Kondov, M. Cizek, C. Benesch, H. Wang and M. Thoss, *J. Phys. Chem. C*, 2007, **111**, 11970.

73   J. Li, M. Nilsing, I. Kondov, H. Wang, P. Persson, S. Lunell and M. Thoss, *J. Phys. Chem. C*, 2008, **112**, 12326.

74   H. D. Meyer and G. A. Worth, *Theo. Chem. Acc.*, 2003, **109**, 251.

75   H. Wang and M. Thoss, *J. Chem. Phys.*, 2003, **119**, 1289.

76   O. V. Prezhdo and P. J. Rossky, *Phys. Rev. Lett.*, 1998, **81**, 10.

77   L. Wang, R. Ernstorfer, F. Willig and V. May, *J. Phys. Chem. B*, 2005, **109**, 9589.

78   L. Wang, F. Willig and V. May, *J. Chem. Phys.*, 2006, **124**, 014712.

79   L. Wang, F. Willig and V. May, *J. Chem. Phys.*, 2007, **126**, 134110.

80   D. R. Jones and A. Troisi, *Phys. Chem. Chem. Phys.*, 2010, **12**, 4625.

81   W. Stier and O. V. Prezhdo, *J. Phys. Chem. B*, 2002, **106**, 8047.

82   W. R. Duncan, W. M. Stier and O. V. Prezhdo, *J. Am. Chem. Soc.*, 2005, **127**, 7941.

83   W. R. Duncan, C. F. Craig and O. V. Prezhdo, *J. Am. Chem. Soc.*, 2007, **129**, 8528.

84   S. A. Fischer, W. R. Duncan and O. V. Prezhdo, *J. Am. Chem. Soc.*, 2009, **131**, 15483.

85   L. G. C. Rego and V. S. Batista, *J. Am. Chem. Soc.*, 2003, **125**, 7989.

86   S. G. Abuabara, L. G. C. Rego and V. S. Batista, *J. Am. Chem. Soc.*, 2005, **127**, 18234.

87   E. Maggio, N. Martsinovich and A. Troisi, *J. Phys. Chem.C*, 2012, **116**, 7638.

88   E. Maggio, N. Martsinovich and A. Troisi, *J. Chem. Phys.*, 2012, **137**, 22A508.

89   E. Maggio, PhD thesis, University of warwick, 2013.

90   J. W. Ondersma and T. W. Hamann, *J. Am. Chem. Soc.*, 2011, **133**, 8264.

91   M. De Falco and A. Basile, *Enriched Methane: The First Step Towards the Hydrogen Economy*, Springer, Switzerland, 2016.

92   T. Hisatomi, K. Maeda, K. Takanabe, J. Kubota and K. Domen, *J. Phys. Chem. C*, 2009, **113**, 21458.

93   J. Tang, J. R. Durrant and D. R. Klug, *J. Am. Chem. Soc.*, 2008, **130**, 13885.

94   M. A. Henderson, *Surf. Sci. Rep.*, 2011, **66**, 185.

95   M. R. Hoffmann, S. T. Martin, W. Choi and D. W. Bahnemannt, *Chem. Rev.*, 1995, **95**, 69.

96   R. Asahi, T. Morikawa, H. Irie and T. Ohwaki, *Chem. Rev.*, 2014, **114**, 9824.

97   M. Ni, M. K. H. Leung, D. Y. C. Leung and K. Sumathy, *Renew. Sust. Energ. Rev.*, 2007, **11**, 401.

98   M. G. Walter, E. L. Warren, J. R. McKone, S. W. Boettcher, Q. Mi, E. A. Santori and N. S. Lewis, *Chem. Rev.*, 2010, **110**, 6446.

99   J. Yang, D. Wang, H. Han and C. Li, *Acc. Chem. Res.*, 2013, **46**, 1900.

100   A. Fujishima, X. Zhang and D. A. Tryk, *Surf. Sci. Rep.*, 2008, **63**, 515.

101   Y. F. Li, Z. P. Liu, L. Liu and W. Gao, *J. Am. Chem. Soc.*, 2010, **132**, 13008.

102   J. Chen, Y. F. Li, P. Sit and A. Selloni, *J. Am Chem. Soc.*, 2013, **135**, 18774.

103   T. Hisatomi, J. Kubota and K. Domen, *Chem. Soc. Rev.*, 2014, **43**, 7520.

104   K. Maeda and K. Domen, *J. Phys. Chem. Lett.*, 2010, **1**, 2655.

105   P. V. Kamat and J. Bisquert, *J. Phys. Chem. C*, 2013, **117**, 14873.

106   A. Kudo and Y. Miseki, *Chem. Soc. Rev.*, 2009, **38**, 253.

107   H. Kato, K. Asakura and A. Kudo, *J. Am. Chem. Soc.*, 2003, **125**, 3082.

108   G. W. Crabtree, M. S. Dresselhaus and M. V. Buchanan, *Phys. Today*, 2004, **57**, 39.

109   M. Ball and M. Weeda, *Int. J. Hydrog. Energ.*, 2015, **40**, 7903.

110   K. Li, X. An, K. H. Park, M. Khraisheh and J. Tang, *Catal. Today*, 2014, **224**, 3.

111 J. A. Macia-Agullo, A. Corma and H. Garcia, *Chem. Eur. J.*, 2015, **21**, 10940.

112 F. Fresno, R. Portela, S. Suárez and J. M. Coronado, *J. Mater. Chem. A*, 2014, **2**, 2863.

113 M. M. Montemore and J. W. Medlin, *Catal. Sci. Technol.*, 2014, **4**, 3748.

114 Z. P. Liu and P. Hu, *J. Chem. Phys.*, 2001, **115**, 4977.

115 Z. P. Liu and P. Hu, *J. Am. Chem. Soc.*, 2003, **125**, 1958.

116 J. K. Nørskov, T. Bligaard, J. Rossmeisl and C. H. Christensen, *Nat. chem.*, 2009, **1**, 37.

117 S. Bai, J. Jiang, Q. Zhang and Y. Xiong, *Chem. Soc. Rev.*, 2015, **44**, 2893.

118 A. J. Foster and R. F. Lobo, *Chem. Soc. Rev*, 2010, **39**, 4783.

119 I. A. Shkrob, T. W. Marin, H. He and P. Zapol, *J. Phys. Chem. C*, 2012, **116**, 9450.

120 M. Anpo, H. Yamashita, Y. Ichihashi and S. Ehara, *J. Electroanal. chem.*, 1995, **396**, 21.

121 C. C. Yang, Y. H. Yu, B. van der Linden, J. C. S. Wu and G. Mul, *J. Am. Chem. Soc.*, 2010, **132**, 8398.

122 L. Liu, H. Zhao, J. M. Andino and Y. Li, *ACS Catal.*, 2012, **2**, 1817.

123 C. P. S. Hsu, in *Handbook of Instrumental Techniques for Analytical Chemistry*, ed. F. Settle, Prentice Hall, New Jersey, 1997, ch. 15, 247-283

124 S. Kozuch and S. Shaik, *Acc. Chem. Res.*, 2011, **44**, 101.

125 M. E. Blake, K. L. Bartlett and M. Jones, *J. Am Chem. Soc.*, 2003, **125**, 6485.

126 M. D. Bhatt and J. S. Lee, *J. Mater. Chem. A*, 2015, **3**, 10632.

127 Á. Valdés, Z. W. Qu, G. J. Kroes, J. Rossmeisl and J. K. Nørskov, *J. Phys. Chem. C*, 2008, **112**, 9872.

128 J. K. Nørskov, J. Rossmeisl, A. Logadottir, L. Lindqvist, J. R. Kitchin, T. Bligaard and H. Jónsson, *J. Phys. Chem. B*, 2004, **108**, 17886.

129 M. T. Nguyen, N. Seriani, S. Piccinin and R. Gebauer, *J. Chem. Phys.*, 2014, **140**, 064703.

130 A. Hellman and R. G. S. Pala, *J. Phys. Chem. C*, 2011, **115**, 12901.

131 J. Cheng, X. Liu, J. a Kattirtzi, J. VandeVondele and M. Sprik, *Angew. Chem. Int. Ed.*, 2014, **53**, 12046.

132 X. Shen, Y. Small, J. Wang, P. Allen, M. Fernandez-Serra, M. Hybertsen and J. Muckerman, *J. Phys. Chem. C*, 2010, **114**, 13695.

133 A. Warshel, *Angew. Chem. Int. Ed.*, 2014, **53**, 10020.

134 M. Karplus, *Angew. Chem. Int. Ed.*, 2014, **53**, 9992.

135 V. P. Indrakanti, H. H. Schobert and J. D. Kubicki, *Energy Fuels*, 2009, **23**, 5247.

136 H. He, P. Zapol and L. A. Curtiss, *Energy Environ. Sci.*, 2012, **5**, 6196.

137 Y. Ji and Y. Luo, *ACS Catal.*, 2016, **6**, 2018.

138 P. Hohenberg and W. Kohn, *Phys. Rev.*, 1964, **136**, B864.

139 P. Geerlings, F. De Proft and W. Langenaeker, *Chem. Rev.*, 2003, **103**, 1793.

140 A. J. Cohen, P. Mori-Sánchez and W. Yang, *Chem. Rev.*, 2012, **112**, 289.

141 W. Kohn and L. J. Sham, *Phys. Rev.*, 1965, **140**, A1133.

142 W. Kohn, *Rev. Mod. Phys.*, 1999, **71**, 1253.

143 P. J. Hasnip, K. Refson, M. I. J. Probert, J. R. Yates, S. J. Clark and C. J. Pickard, *Phil. Trans. R. Soc. A*, 2014, **372**, 20130270.

144 C. S. Wang and W. E. Pickett, *Phys. Rev. Lett.*, 1983, **51**, 597.

145 M. K. Y. Chan and G. Ceder, *Phys. Rev. Lett.*, 2010, **105**, 196403.

146 G. Zhang and C. B. Musgrave, *J. Phys. Chem. A*, 2007, **111**, 1554.

147 L. Hedin, *Phys. Rev.*, 1965, **139**, A796.

148 G. Onida, L. Reining and A. Rubio, *Rev. Mod. Phys.*, 2002, **74**, 601.

149 F. De Angelis, S. Fantacci and A. Selloni, *Nanotechnology*, 2008, **19**, 424002.

150 E. Runge and E. K. U. Gross, *Phys. Rev. Lett.*, 1984, **52**, 997.

151 V. I. Anisimov, I. V Solovyev and M. A. Korotin, *Phys. Rev. B*, 1993, **48**, 16929.

152 S. L. Dudarev, G. A. Botton, S. Y. Savrasov, C. J. Humphreys and A. P. Sutton, *Phys. Rev. B*, 1998, **57**, 1505.

153 A. Rohrbach, J. Hafner and G. Kresse, *Phys. Rev. B*, 2004, **69**, 075413.

154 H. Jiang, *Int. J. Quantum Chem.*, 2015, **115**, 722.

155 N. Lin, C. A. Marianetti, A. J. Millis and D. R. Reichman, *Phys. Rev. Lett.*, 2011, **106**, 096402.

156 B. Kaduk, T. Kowalczyk and T. Van Voorhis, *Chem. Rev.*, 2012, **112**, 321.

157 P. N. Day, K. A. Nguyen and R. Pachter, *J. Phys. Chem. B*, 2005, **109**, 1803.

158 G. Henkelman, B. P. Uberuaga and H. Jónsson, *J. Chem. Phys.*, 2000, **113**, 9901.

159 G. Mills and H. Jónsson, *Phys. Rev. Lett.*, 1994, **72**, 1124.

160 G. Mills, H. Jónsson and G. K. Schenter, *Surf. Sci.*, 1995, **324**, 305.

161 G. Henkelman and H. Jónsson, *J. Chem. Phys.*, 2000, **113**, 9978–9985.

162 P. Dirac, *P. Roy. Soc. Lond. A Mat.*, 1927, **114**, 243.

163   D. M. Newns, *Phys. Rev.*, 1969, **178**, 1123.

164   S. Davydov and S. Troshin, *Phys. Solid States*, 2007, **49**, 1583.

165   F. Ambrosio, N. Martsinovich and A. Troisi, *J. Phys. Chem. Lett.*, 2012, **3**, 1531.

166   F. Ambrosio, PhD thesis, University of Warwick, 2014.

167   A. Nitzan, *Chemical dynamics in condensed phases*, Oxford University Press, 2006.

168   E. Maggio and A. Troisi, *J. Phys. Chem. C*, 2013, **117**, 24196.

169   M. Bixon, J. Jortner, J. Cortes, H. Heitele and M. E. Michel-Beyerle, *J. Phys. Chem.*, 1994, **98**, 7289.

170   E. Maggio, N. Martsinovich and A. Troisi, *Angew. Chem. Int. Ed.*, 2013, **52**, 973.

171   E. Maggio, G. C. Solomon and A. Troisi, *ACS Nano.*, 2014, **8**, 409.

172   R. A. Marcus, *Angew. Chem. Int. Ed.*, 1993, **32**, 1111.

173   N. E. Gruhn, D. A. Da Silva Filho, T. G. Bill, M. Malagoli, V. Coropceanu, A. Kahn and J. L. Bredas, *J. Am Chem. Soc.*, 2002, **124**, 7918.

174   V. Vaissier, P. Barnes, J. Kirkpatrick and J. Nelson, *Phys. Chem. Chem. Phys.*, 2013, **15**, 4804.

175   S. Fletcher, *J Solid State Electrochem*, 2008, **12**, 765.

176   D. V. Matyushov, *J. Chem. Phys.*, 2004, **120**, 7532.

177   D. Kuciauskas, M. S. Freund, H. B. Gray, J. R. Winkler and N. S. Lewis, *J. Phys. Chem. B*, 2000, **105**, 392.

178   L. D. Zusman, *Chem. Phys*, 1980, **49**, 295.

179   L. D. Zusman, *Chem. Phys.*, 1988, **119**, 51.

180   J. Casado-Pascual, M. Morillo, I. Goychuk and P. Hänngi, *J. Chem. Phys.*, 2003, **118**, 291.

181   Q. Shi, L. Chen, G. Nan, R. Xu and Y. Yan, *J. Chem. Phys.*, 2009, **130**, 164518.

182   X. Y. Li, J. Tong and F. C. He, *Chem. Phys.*, 2000, **260**, 283.

183   C. Venkataraman, A. V. Soudackov and S. Hammes-Schiffer, *J. Phys. Chem. C*, 2008, **112**, 12386.

184   R. P. Verma and C. Hansch, *Chem. Rev.*, 2009, **109**, 213.

185   H. Gao, J. A. Katzenellenbogen, R. Garg and C. Hansch, *Chem. Rev.*, 1999, **99**, 723.

186   C. Hansch, *Acc. Chem. Res.*, 1969, **2**, 232.

187   US Pat., 4146719 A, 1979.

188   H. Koga, A. Itoh, S. Murayama, S. Suzue and T. Irikura, *J. Med. Chem.*, 1980, **23**, 1358.

189   M. Karelson, V. S. Lobanov and A. R. Katritzky, *Chem. Rev.*, 1996, **96**, 1027.

190   A. R. Katritzky, M. Kuanar, S. Slavov, C. D. Hall, M. Karelson, I. Kahn and D. A. Dobchev, *Chem. Rev.*, 2010, **110**, 5714.

191   M. Shahlaei, *Chem. Rev.*, 2013, **113**, 8093.

192   L. A. Tarca, B. P. A. Grandjean and F. Larachi, *Ind. Eng. Chem. Res.*, 2005, **44**, 1073.

193   Y. Liu, *J. Chem. Inf. Comput. Sci.*, 2004, **44**, 1823.

194   J. M. Luco and F. H. Ferretti, *J. Chem. Inf. Comput. Sci.*, 1997, **37**, 392.

195   T. Le, V. C. Epa, F. R. Burden and D. A. Winkler, *Chem. Rev.*, 2012, **112**, 2889.

196   F. R. Burden and D. A. Winkler, *J. Med. Chem.*, 1999, **42**, 3183.

197   X. J. Yao, A. Panaye, J. P. Doucet, R. S. Zhang, H. F. Chen, M. C. Liu, Z. D. Hu and B. T. Fan, *J. Chem. Inf. Comput. Sci.*, 2004, **44**, 1257.

198   D. C. Montgomery, E. A. Beck and G. G. Vining, *Introduction to linear regression analysis*, Wiley, 5th edn., 2012.

199   T. Hattori and S. Kito, *Catal. Today*, 1995, **23**, 347.

200   V. Venkatraman, P. O. Åstrand and B. K. Alsberg, *J. Comput. Chem.*, 2014, **35**, 214.

201   V. Venkatraman and B. K. Alsberg, *Dye. Pigment.*, 2015, **114**, 69.

202   V. Venkatraman, S. Abburu and B. K. Alsberg, *Phys. Chem. Chem. Phys.*, 2015, **17**, 27672.

203   V. Venkatraman, M. Foscato, V. R. Jensen and B. K. Alsberg, *J. Mater. Chem. A*, 2015, **3**, 9851.

204   B. Efron, *Ann. Stat.*, 1979, **7**, 1.

205   B. Efron, *J. Amer. Stat. Assoc.*, 1987, **82**, 171.

206   B. Y. P. J. Bickel and D. A. Freedman, *Ann. Stat.*, 1981, **9**, 1196.

207   F. E. Harrel Jr., *Regression Modeling strategies: With applications to Linear Models, Logistic Regression, and Survival Analysis*, Springer, New York, 2006.

208   F. Odobel, L. Le Pleux, Y. Pellegrin and E. Blart, *Acc. Chem. Res.*, 2010, **43**, 1063.

209   H. N. Ghosh, J. B. Asbury and T. Lian, *J. Phys. Chem. B*, 2000, **2**, 6482.

210   J. M. Rehm, G. L. Mclendon, J. Moser and M. Gratzel, *J. Phys. Chem.*, 1996, **100**, 9577.

211 A. Morandeira, G. Boschloo, A. Hagfeldt and L. Hammarstro, *J. Phys. Chem. C*, 2008, **112**, 9530.

212 I. Hod, Z. Tachan, M. Shalom and A. Zaban, *Phys. Chem. Chem. Phys.*, 2013, **15**, 6339.

213 M. Gennari, F. Le, L. Zhang, Y. Pellegrin, E. Blart, A. M. Brown, A. Deronzier, M. Collomb, M. Boujtita, D. Jacquemin, L. Hammarstro and F. Odobel, *J. Phys. Chem. Lett.*, 2014, **5**, 2254.

214 G. A. Sawatzky and J. W. Allen, *Phys. Rev. Lett.*, 1984, **53**, 2339.

215 A. Fujimori and F. Minami, *Phys. Rev. B*, 1984, **30**, 957.

216 G. Lee, *Phys. Rev. B*, 1991, **43**, 674–682.

217 T. Bredow and A. Gerson, *Phys. Rev. B*, 2000, **61**, 5194.

218 S. Hufner, *Photoelectron spectroscopy*, Springer, Heidelburg, 1996.

219 M. J. Soler, E. Artacho, J. D. Gale, A. Garcia, J. Junquera, P. Ordejon and D. Sanchez-Portal, *J. Phys. Condens. Matter*, 2002, **14**, 2745.

220 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865.

221 N. Troullier and J. L. Martins, *Phys. Rev. B*, 1993, **43**, 1993.

222 V. I. Anisimov, J. Zaanen and O. K. Andersen, *Phys. Rev. B*, 1991, **44**, 943.

223 F. P. Netzer and M. Prutton, *J. Phys. C Solid State Phys.*, 1975, **8**, 2401.

224 M. R. Welton-Cook and M. Prutton, *J. Phys. C Solid State Phys.*, 1980, **13**, 3993.

225 F. Cinquini, L. Giordano, G. Pacchioni, A. Ferrari, C. Pisani and C. Roetti, *Phys. Rev. B*, 2006, **74**, 165403.

226 F. B. van Duijneveldt, J. G. C. M. van Duijneveldt-van de Rijdt and J. H. van Lenthe, *Chem. Rev.*, 1994, **94**, 1873.

227 K. Hara, T. Sato, R. Katoh, A. Furube, Y. Ohga, A. Shinpo, S. Suga, K. Sayama, H. Sugihara and H. Arakawa, *J. Phys. Chem. B*, 2003, **107**, 597.

228 J. Tomasi, B. Mennucci and R. Cammi, *Chem. Rev.*, 2005, **105**, 2999.

229 M. J. Frisch et. al., Gaussian 03 (Revision D.02), Gaussian Inc., Wallingford, CT, 2004.

230 J. Junquera, M. Zimmer, P. Ordejon and P. Ghosez, *Phys. Rev. B*, 2003, **67**, 155327.

231 M. C. Toroker, D. K. Kanan, N. Alidoust, L. Y. Isseroff, P. Liao and E. A. Carter, *Phys. Chem. Chem. Phys.*, 2011, **13**, 16644.

232 K. Murakoshia, S. Yanagidaa, M. Capelb and E. W. Castner, in *Nanostructured Materials: Clusters, Composites and thin films; ACS Symposium Series 679; American Chemical Society: Washington, DC*, 1997, p. 221.

233 J. He, H. Lindstro, A. Hagfeldt and S. Lindquist, *J. Phys. Chem. B*, 1999, **103**, 8940.

234 J. Szuber, *J. Electron Spect. Rel. Phen.*, 1984, **34**, 337.

235 Z. Ji and Y. Wu, *J. Phys. Chem. C*, 2013, **117**, 18315.

236 P. Qin, J. Wiberg, E. A. Gibson, M. Linder, L. Li, T. Brinck, A. Hagfeldt, B. Albinsson and L. Sun, *J. Phys. Chem. C*, 2010, **114**, 4738.

237 M. H. Baik and R. A. Friesner, *J. Phys. Chem. A*, 2002, **106**, 7407.

238 F. Fabregat-Santiago, G. Garcia-Belmonte, I. Mora-Seró and J. Bisquert, *Phys. Chem. Chem. Phys.*, 2011, **13**, 9083.

239 D. P. Hagberg, T. Marinado, K. M. Karlsson, K. Nonomura, P. Qin, G. Boschloo, T. Brinck, A. Hagfeldt and L. Sun, *J. Org. Chem.*, 2007, **72**, 9550.

240 P. Qin, H. Zhu, T. Edvinsson, G. Boschloo, A. Hagfeldt and L. Sun, *J. Am. Chem. Soc.*, 2008, **130**, 8570.

241 P. Qin, M. Linder, T. Brinck, G. Boschloo, A. Hagfeldt and L. Sun, *Adv. Mater.*, 2009, **21**, 2993.

242 Z. Ji, G. Natu, Z. Huang and Y. Wu, *Energy Environ. Sci.*, 2011, **4**, 2818.

243 Y. Pellegrin, L. Le Pleux, E. Blart, A. Renaud, B. Chavillon, N. Szuwarski, M. Boujtita, L. Cario, S. Jobic, D. Jacquemin and F. Odobel, *J. Photochem. Photobiol. A*, 2011, **219**, 235.

244 L. Le Pleux, A. L. Smeigh, E. Gibson, Y. Pellegrin, E. Blart, G. Boschloo, A. Hagfeldt, L. Hammarström and F. Odobel, *Energy Environ. Sci.*, 2011, **4**, 2075.

245 Y. S. Yen, W. T. Chen, C. Y. Hsu, H. H. Chou, J. T. Lin and M. C. P. Yeh, *Org. Lett.*, 2011, **13**, 4930.

246 Y. Wu and W. Zhu, *Chem. Soc. Rev.*, 2013, **42**, 2039.

247 J. N. Clifford, E. Martínez-Ferrero, A. Viterisi and E. Palomares, *Chem. Soc. Rev.*, 2011, **40**, 1635.

248 W. Zeng, Y. Cao, Y. Bai, Y. Wang, Y. Shi, M. Zhang, F. Wang, C. Pan and P. Wang, *Chem. Mater.*, 2010, **22**, 1915.

249 M. Pastore, S. Fantacci and F. De Angelis, *J. Phys. Chem. C*, 2013, **117**, 3685.

250 J. Bisquert and R. A. Marcus, *Top. Curr. Chem.*, 2014, **352**, 325.

251 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648.

252 J. L. Pascual-Ahuir and E. Silla, *J. Comput. Chem.*, 1990, **11**, 1047.

253 E. Silla, I. Tunon and J. L. Pascual-Ahuir, *J. Comput. Chem.*, 1991, **12**, 1077.

254 G. Boschloo and A. Hagfeldt, *Acc. Chem. Res.*, 2009, **42**, 1819.

255 H. Tang, K. Prasad, R. Sanjinès, P. E. Schmid and F. Lévy, *J. Appl. Phys.*, 1994, **75**, 2042.

256 C. A. Gueymard, *Sol. Energy*, 2001, **71**, 325.

257 C. A. Gueymard, *Sol. Energy*, 2004, **76**, 423.

258 S. Rühle, M. Greenshtein, S. G. Chen, A. Merson, H. Pizem, C. S. Sukenik, D. Cahen and A. Zaban, *J. Phys. Chem. B*, 2005, **109**, 18907.

259 C. P. Cho, C. C. Chu, W. T. Chen, T. C. Huang and Y. T. Tao, *J. Mater. Chem.*, 2012, **22**, 2915.

260 K. Hara, K. Sayama, H. Arakawa, Y. Ohga, A. Shinpo and S. Suga, *Chem. Commun.*, 2001, 569.

261 T. A. Niehaus, T. Hofbeck and H. Yersin, *RSC Adv.*, 2015, **5**, 63318.

262 Y. S. Yen, H. H. Chou, Y. C. Chen, C. Y. Hsu and J. T. Lin, *J. Mater. Chem.*, 2012, **22**, 8734.

263 N. Koumura, Z. S. Wang, S. Mori, M. Miyashita, E. Suzuki and K. Hara, *J. Am. Chem. Soc.*, 2006, **128**, 14256.

264 R. K. Kanaparthi, J. Kandhadi and L. Giribabu, *Tetrahedron*, 2012, **68**, 8383.

265 Z. S. Wang, Y. Cui, K. Hara, Y. Dan-oh, C. Kasada and A. Shinpo, *Adv. Mater.*, 2007, **19**, 1138.

266 R. S. Nickerson, *Rev. Gen. Psychol.*, 1998, **2**, 175.

267 T. Le Bahers, C. Adamo and I. Ciofini, *J. Chem. Theory. Comput.*, 2011, **7**, 2498.

268 T. Etienne, X. Assfeld and A. Monari, *J. Chem. Theory Comput.*, 2014, **10**, 3896.

269 T. Etienne, X. Assfeld and A. Monari, *J. Chem. Theory. Comput.*, 2014, **10**, 3906.

270 Z. Xie, A. Midya, K. P. Loh, S. Adams, D. J. Blackwood, J. Wang, X. Zhang and Z. Chen, *Prog. Photovolt. Res. Appl.*, 2010, **18**, 573.

271 H. Choi, I. Raabe, D. Kim, F. Teocoli, C. Kim, K. Song, J. H. Yum, J. Ko, M. K. Nazeeruddin and M. Grätzel, *Chem. Eur. J.*, 2010, **16**, 1193.

272 L. Y. Lin, C. H. Tsai, K. T. Wong, T. W. Huang, L. Hsieh, S. H. Liu, H. W. Lin, C. C. Wu, S. H. Chou, S. H. Chen and A. I. Tsai, *J. Org. Chem.*, 2010, **75**, 4778.

273 C. Teng, X. Yang, C. Yang, S. Li, M. Cheng, A. Hagfeldt and L. Sun, *J. Phys. Chem. C*, 2010, **114**, 9101.

274 Y. Cui, Y. Wu, X. Lu, X. Zhang, G. Zhou, F. B. Miapeh, W. Zhu and Z. S. Wang, *Chem. Mater.*, 2011, **23**, 4394.

275   J. H. Chen, C. H. Tsai, S. A. Wang, Y. Y. Lin, T. W. Huang, S. F. Chiu, C. C. Wu and K. T. Wong, *J. Org. Chem.*, 2011, **76**, 8977.

276   D. Kumar, K. R. J. Thomas, C. P. Lee and K. C. Ho, *Org. Lett.*, 2011, **13**, 2622.

277   M. Liang, M. Lu, Q. L. Wang, W. Y. Chen, H. Y. Han, Z. Sun and S. Xue, *J. Power Sources*, 2011, **196**, 1657.

278   X. F. Lu, Q. Y. Feng, T. Lan, G. Zhou and Z. S. Wang, *Chem. Mater.*, 2012, **24**, 3179.

279   S. R. Li, C. P. Lee, H. T. Kuo, K. C. Ho and S. S. Sun, *Chem. Eur. J.*, 2012, **18**, 12085.

280   G. Wu, F. Kong, J. Li, X. Fang, Y. Li, S. Dai, Q. Chen and X. Zhang, *J. Power Sources*, 2013, **243**, 131.

281   H. Li, Y. Hou, Y. Yang, R. Tang, J. Chen, H. Wang, H. Han, T. Peng, Q. Li and Z. Li, *Appl. Mater. Interfaces*, 2013, **5**, 12469.

282   Y. Hua, B. Jin, H. Wang, X. Zhu, W. Wu, M. S. Cheung, Z. Lin, W. Y. Wong and W. K. Wong, *J. Power Sources*, 2013, **237**, 195.

283   T. Duan, K. Fan, C. Zhong, X. Chen, T. Peng and J. Qin, *J. Power Sources*, 2013, **234**, 23.

284   S. R. Li, C. P. Lee, P. F. Yang, C. W. Liao, M. M. Lee, W. L. Su, C. T. Li, H. W. Lin, K. C. Ho and S. S. Sun, *Chem. Eur. J.*, 2014, **20**, 10052.

285   P. Thongkasee, A. Thangthong, N. Janthasing, T. Sudyoadsuk, S. Namuangruk, T. Keawin, S. Jungsuttiwong and V. Promarak, *Appl. Mater. Interfaces*, 2014, **6**, 8212.

286   K. Matsumura, S. Yoshizaki, M. M. Maitani, Y. Wada, Y. Ogomi, S. Hayase, T. Kaiho, S. Fuse, H. Tanaka and T. Takahashi, *Chem. Eur. J*, 2015, **21**, 9742.

287   X. Qian, Y. Z. Zhu, W. Y. Chang, J. Song, B. Pan, L. Lu, H. H. Gao and J. yu Zheng, *Appl. Mater. Interfaces*, 2015, **7**, 9015.

288   M. Kendall and A. Stuart, *The advanced theory of statistics, Vol.1: Distribution theory*, Griffin, London, 1977.

289   K. Koci, L. Obalova and O. Solcova, *Chem. Process Eng.*, 2010, **31**, 395.

290   N. M. Dimitrijevic, B. K. Vijayan, O. G. Poluektov, T. Rajh, K. A. Gray, H. He and P. Zapol, *J. Am. Chem. Soc.*, 2011, **133**, 3964.

291   N. M. Dimitrijevic, I. A. Shkrob, D. J. Gosztola and T. Rajh, *J. Phys. Chem. C*, 2012, **116**, 878.

292   G. R. Dey and K. Pushpa, *Res. Chem. Intermed.*, 2007, **33**, 631.

293   I. A. Shkrob, N. M. Dimitrijevic, T. W. Marin, H. He and P. Zapol, *J. Phys. Chem. C*, 2012, **116**, 9461.

294   J. Rasko and F. Solymosi, *J. Phys. Chem.*, 1994, **98**, 7147.

295   C. C. Lo, C. H. Hung, C. S. Yuan and J. F. Wu, *Sol. Energ. Mat. Sol.*, 2007, **91**, 1765.

296   P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari and R. M. Wentzcovitch, *J. Phys. Condens. Matter*, 2009, **21**, 395502.

297   N. Martsinovich, D. R. Jones and A. Troisi, *J. Phys. Chem. C*, 2010, **114**, 22659.

298   A. Michaelides, Z. P. Liu, C. J. Zhang, A. Alavi, D. A. King and P. Hu, *J. Am. Chem. Soc.*, 2003, **125**, 3704.

299   M. Saeys, M. F. Reyniers, G. B. Marin, V. Van Speybroeck and M. Waroquier, *ChemPhysChem*, 2006, **7**, 188.

300   M. M. Islam, M. Calatayud and G. Pacchioni, *J. Phys. Chem. C*, 2011, **115**, 6809.

301   U. Aschauer and A. Selloni, *Phys. Chem. Chem. Phys.*, 2012, **14**, 16595.

302   E. Maggio, N. Martsinovich and A. Troisi, *J. Phys. Condens. Matter*, 2016, **28**, 074004.

303   C. Di Valentin, G. Pacchioni and A. Selloni, *J. Phys. Chem. C*, 2009, **113**, 20543.

304   P. R. Schreiner, H. P. Reisenauer, F. C. Pickard IV, A. C. Simmonett, W. D. Allen, E. Mátyus and A. G. Császár, *Nature*, 2008, **453**, 906.

305   E. C. C. Baly, I. M. Heilbron and W. F. Barker, *J. Chem. Soc. Trans.*, 1921, 1025.

306   Y. Hori, R. Takahashi, Y. Yuzuru and A. Murata, *J. Phys. Chem. B*, 1997, **101**, 7075.

307   T. Bligaard, J. K. Nørskov, S. Dahl, J. Matthiesen, C. H. Christensen and J. Sehested, *J. Catal.*, 2004, **224**, 206.

308   R. P. Bell, *Proc. R. Soc. London, Ser. A*, 1936, **154**, 414.

309   M. G. Evans and M. Polanyi, *Trans. Faraday Soc.*, 1936, **32**, 1936.

310   A. A. Peterson, F. Abild-Pedersen, F. Studt, J. Rossmeisl and J. K. Nørskov, *Energy Environ. Sci.*, 2010, **3**, 1311.

311   A. Vittadini, A. Selloni, F. P. Rotzinger and M. Grätzel, *J. Phys. Chem. B*, 2000, **104**, 1300.

312 H. Liu, M. Zhao, Y. Lei, C. Pan and W. Xiao, *Comp. Mater. Sci.*, 2012, **51**, 389.

313 M. Shen and M. A. Henderson, *J. Phys. Chem. Lett.*, 2011, **2**, 2707.

314 Q. Cao, S. Berski, Z. Latajka, M. Räsänen and L. Khriachtchev, *Phys. Chem. Chem. Phys.*, 2014, **16**, 5993.

315 S. H. Kim, P. C. Stair and E. Weitz, *Langmuir*, 1998, **14**, 4156.

316 J. Cunningham and A. L. Leahy, *J. Phys. Chem.*, 1972, **76**, 2353.

317 M. E. Pronsato, C. Pistonesi, A. Juan, A. P. Farkas, L. Bugyi and F. Solymosi, *J. Phys. Chem. C*, 2011, **115**, 2798.

318 S. J. Suresh and V. M. Naik, *J. Chem. Phys.*, 2000, **113**, 9727.

319 D. C. Sorescu, W. A. Al-Saidi and K. D. Jordan, *J. Chem. Phys.*, 2011, **135**, 124701.

320 P. V. Kamat, *J. Phys. Chem. Lett.*, 2012, **3**, 663.

321 E. Poli, J. D. Elliott, L. E. Ratcliff, L. Andrinopoulos, J. Dziedzic, N. D. M. Hine, A. A. Mostofi, C. K. Skylaris, P. D. Haynes and G. Teobaldi, *J. Phys. Condens. Matter*, 2016, **28**, 074003.

322 Y. F. Li, U. Aschauer, J. Chen and A. Selloni, *Acc. Chem. Res.*, 2014, **47**, 3361.

323 E. Finazzi, C. Di Valentin, G. Pacchioni and A. Selloni, *J. Chem. Phys.*, 2008, **129**, 154113.

324 M. Cococcioni and S. de Gironcoli, *Phys. Rev. B*, 2005, **71**, 035105.

325 J. D. Goodpaster, A. T. Bell and M. Head-Gordon, *J. Phys. Chem. Lett.*, 2016, **7**, 1471.

326 F. Odobel and Y. Pellegrin, *J. Phys. Chem. Lett.*, 2013, **4**, 2551.

327 A. Nattestad, M. Ferguson, R. Kerr, Y. B. Cheng and U. Bach, *Nanotechnology*, 2008, **19**, 295304.

328 G. Niu, X. Guo and L. Wang, *J. Mater. Chem. A*, 2015, **3**, 8970.

329 H. J. Snaith, *J. Phys. Chem. Lett.*, 2013, **4**, 3623.

330 Y. Zhou and K. Zhu, *ACS Energy Lett.*, 2016, **1**, 64.

331 W. Ma, Y. Jiao and S. Meng, *J. Phys. Chem. C*, 2014, **118**, 16447.

332 N. Madaan, N. R. Shiju and G. Rothenberg, *Catal. Sci. Technol.*, 2016, **6**, 125.