Mechanistic insights into CO_2 valorization via frustrated Lewis pairs, porphyrins, and polyoxometalates : a computational approach

Francesc Penas-Hidalgo

To cite this version:

Francesc Penas-Hidalgo. Mechanistic insights into CO_2 valorization via frustrated Lewis pairs, porphyrins, and polyoxometalates : a computational approach. Catalysis. Sorbonne Université, 2025. English. ⟨NNT : 2025SORUS248⟩. ⟨tel-05362368⟩

HAL Id: tel-05362368

https://theses.hal.science/tel-05362368v1

Submitted on 13 Nov 2025

SORBONNE UNIVERSITÉ

Ecole doctorale 406 - Chimie Moléculaire de Paris

Laboratoire de Chimie des Processus Biologiques - Collège de France

DOCTORAL THESIS

Mechanistic Insights into CO$_2$ Valorization via Frustrated Lewis Pairs, Porphyrins, and Polyoxometalates: A Computational Approach

---

Presented by:

Francesc PENAS-HIDALGO

In front of a jury composed of:

Reviewers:
Jorge J. CARBÓ: Full Professor, Universitat Rovira i Virgili
François-Xavier COUDERT: Directeur de recherche, CNRS, Chimie ParisTech

Examiners:
President of the Jury: Guillaume IZZET: Directeur de recherche, CNRS, Institut Parisien de Chimie Moléculaire, Sorbonne Université
Elsje Alessandra QUADRELLI: Directrice de recherche, CNRS, Université de Lyon, Université Claude Bernard Lyon 1, IRCELYON

Director:
Caroline MELLOT-DRAZNIEKS: Directrice de recherche, CNRS, Collége de France, Sorbonne Université
Co-supervisor:
Albert SOLÉ-DAURA: "la Caixa" Junior Leader, Institute of Chemical Research of Catalonia

SORBONNE UNIVERSITÉ

Abstract

Ecole doctorale 406 - Chimie Moléculaire de Paris
Laboratoire de Chimie des Processus Biologiques - Collège de France

Doctoral Thesis

Mechanistic Insights into CO_2 Valorization via Frustrated Lewis Pairs, Porphyrins, and Polyoxometalates: A Computational Approach

by Francesc PENAS-HIDALGO

The catalytic reduction of carbon dioxide (CO_2) is a central challenge and opportunity in developing sustainable and carbon-neutral technologies. In this context this thesis investigates by means of computational techniques two distinct types of molecular systems capable of promoting CO_2 conversion under mild conditions: Frustrated Lewis Pairs (FLPs) and two polyoxometalate (POM)-containing systems.

Chapter 2 explores the reactivity of molecular FLPs towards CO_2 hydrogenation, focusing on the key step of H_2 activation, aiming to find structure-activity relationships to understand their activity. In that purpose, we built a dataset of 112 intramolecular B/N FLPs with systematically varied electronic and structural features. Using Density Functional Theory (DFT), we computed thermodynamic (ΔG) and kinetic ($\Delta G^{\ddagger}$) parameters associated with H_2 splitting and CO_2 hydrogenation. We used multivariate regression to identify chemically meaningful descriptors, acid/base strength and donor–acceptor distance (d) governing ΔG, and, d and the molecular distortion angle between donor and acceptor centres (λ) governing $\Delta G^{\ddagger}$. These structure–activity models offer insights for designing FLPs, enabling predictive strategies for H_2 activation and CO_2 valorization.

In the subsequent Chapter 3, we focus on the photocatalytic conversion of CO_2 using POM-containing molecular catalysts. These redox-active metal-oxo clusters exhibit high structural versatility and are well-suited for mediating multi-electron transformations. We apply a computational approach to investigate the CO_2 activation and reduction mechanisms on experimentally studied POM-based catalysts systems, in collaboration with the group of Professors Dolbecq and Mialane at the Institut Lavoisier de Versailles. Particular attention is paid to the interplay between the molecular and the electronic structure, supramolecular interactions, and catalytic activity.

Together, these studies offer a dual perspective in the thermal and photocatalytic conversion of CO_2, providing mechanistic understanding and predictive tools for the development of next-generation catalysts.

Acknowledgements

This thesis would not have been possible without the support, guidance, and encouragement of many people, to whom I am deeply grateful.

First, I would like to express my sincere gratitude to my supervisors, Dr. Caroline Mellot-Draznieks and Dr. Albert Solé-Daura, for their invaluable guidance, continuous support, insightful feedback and their patience throughout the course of my doctoral research. Their expertise and mentorship have played a central role in shaping both this thesis and my growth as a researcher. I wish also to thank collaborators in Institut Lavoisier de Versailles, Dr. Anne Dolbecq and Pr. Piere Mialane for inspiring discussions.

I also extend my heartfelt thanks to the members of my thesis monitoring committee, Dr. Héloïse Dossmann and Dr. Julia Contreras-Garcia, for their constructive input, encouragement, and thoughtful discussions at various stages of this project.

I am grateful to my collaborators and colleagues in the Laboratoire de Chimie des Processus Biologiques - Collège de France, especially Dr. Maria Gomez-Mingot and Dr. Lionel Zoubritzky, for their camaraderie, assistance, stimulating conversations and their help.

I gratefully acknowledge the financial support from the French Agence Nationale de la Recherche (ANR) under grant ANR-21-CE07-0028 (project FLIPS) and the HPC national resources from GENCI (CINES /TGCC) under the allocation 20-XX-A01709 07343, which made this research possible.

On a more personal note, I would like to thank my friends for their encouragement, laughter, and grounding presence during both the highs and lows of this journey.

Finally, I am profoundly grateful to my family. To my parents, Mr. Pedro Penas-Bernal and Mrs. Esther Hidalgo-Calvet, and my grandmother Mrs. Rosario Calvet-Seva for their unconditional love, support, and sacrifices.

"Moltes gracies família!"

This thesis is dedicated to all of you.

Contents

List of Figures

List of Tables

List of Abbreviations

BEP	Brønsted-Evans-Polanyi.
BIH	dimethylphenylbenzimidazoline.
BNAH	1-benzyl-1,4-dihydronicotineamide.
btam	N-benzyl-tert-butylamine.
CASSCF	complete active space self-consistent field.
CN	coordination number.
CO_2RR	carbon dioxide reduction reactions.
COF	covalent organic frameworks.
CSVR	Canonical Sampling through Velocity Rescaling.
CV	cyclovoltammogram.
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene.
DCC	dicyclohexylcarbodiimide.
DFT	Density Functional Theory.
DLS	dynamic light scattering.
DMF	N,N-dimethylmethanamide.
DMSO	dimethyl sulfoxide.
EDX	energy dispersive X-ray.
EF	electric field.
ESI	electrospray ionization.
ESP	Electrostatic Potential.
ET	electron Transfer.
Fc	ferrocene.
FEHA	free energy of hydride attachment.
FEPA	free energy of proton attachment.
FLP	Frustrated Lewis Pair.
GCE	glassy carbon electrode.
GGA	Generalized Gradient Approximation.
GPW	Gaussian and Plane Waves.
GTH	Goedecker, Teter, and Hutter.

HER	Hydrogen Evolution Reaction.
HIV	human immunodeficiency virus.
HOMO	highest occupied molecular orbital.
IEF-PCM	Integral Equation Formalism variant of the Polarizable Continuum Model.
ILV	Institut Lavoisier de Versailles.
IR	infrared.
LA	Lewis acid.
LB	Lewis base.
LMCT	ligand to metal charge transfer.
LOOCV	Leave One Out Cross-Validation.
LP	lone pair.
LUMO	lowest unoccupied molecular orbital.
Lut	2,6-lutidine.
LV	latent variable.
MAE	mean absolute error.
MD	molecular dynamics.
MEP	molecular electrostatic potential.
Mes	2,4,6-trimethylphenyl.
MLCT	metal to ligand charge transfer.
MLR	multivariate linear regression.
MO	molecular orbital.
Mo$_6$	$[Mo_6O_{19}]^{2-}$.
MOF	metal-organic framework.
MOP	metal-organic polyhedra.
MP	metalloporphyrin.
MRSA	Methicillin-resistant *Staphylococcus aureus*.
MS	mass spectrometry.
NBO	Natural Bond Orbital.
Ni$_4$	$[Ni_4(H_2O)_2(PW_9O_{34})_2]^{10-}$.
Ni$_4$**Ac**	$[(SiW_9O_{34})Ni_4(CH_3COO)_3(OH)_3]^{8-}$.
Ni$_7$	$[\{(PW_9O_{34})Ni_3(OH)(H_2O)_2(O_3PC(O)(C_3H_6NH_3)PO_3)\}_2Ni]^{14-}$.
Ni$_{14}$	$[(SiW_9O_34)Ni_{14}H_5(O_3PC(O)(C_3H_6NH_3)PO_3)_7(H_2O)_{11}(OH)_7]^{12-}$.
Ni$_{17}$	$Ni[\{(SiW_9O_{34})Ni_4(OH)_3\}_4(OOC(C_4H_2O)COO)_6]^{30-}$.
NMR	nuclear magnetic resonance.
NOCV	natural orbital of chemical valence.
NPA	Natural Population Analysis.

P_2N	bis(triphenylphosphoranylidene)ammonium.
PCA	Principal Component Analysis.
PCET	proton-coupled electron transfer.
PLS	Partial Least Squares.
PMOF	polyoxometalate–metalloporphyrin organic framework.
POM	polyoxometalate.
PS	photosensitizer.
QSAR	quantitative structure–activity relationship.
QTAIM	quantum theory of atoms in molecules.
RMSE	root mean squared error.
$\mathbf{Ru}^{+}$	$[Ru(bpy)_3]^{+}$.
$\mathbf{Ru}^{2+}$	$[Ru(bpy)_3]^{2+}$.
$\mathbf{Ru}^{3+}$	$[Ru(bpy)_3]^{3+}$.
SARS	severe acute respiratory syndrome.
SCE	saturated calomel electrode.
SED	sacrificial electron donor.
SET	single-electron transfer.
SOMO	single occupied molecular orbital.
TBA	tetrabutylammonium.
TD-DFT	time-dependent Density Functional Theory.
TEA	triethylamine.
TEOA	triethanolamine.
TGA	thermogravimetric analysis.
TMP	2,2,6,6-tetramethylpiperidine.
TOF	turn-over frequency.
TON	turn-over number.
TPP	tetraphenylporphyrin.
$TPPS_4$	tetrakis(p-sulfonatophenyl)porphyrin.
TPSS	Tao–Perdew–Staroverov–Scuseria.
TS	transition state.
UFF	universal force-field.
VRSA	Vancomycin-resistant *Staphylococcus aureus*.

Physical Constants

F	The Faraday constant, 23.06 kcal mol^{-1} V^{-1}.
K_b	The Boltzmann constant, 1.381×10^{-23} m^2 kg s^{-2} K^{-1}.
R	The ideal gas constant, 1.987×10^{-3} kcal mol^{-1} K^{-1}.
h	The Planck constant, 6.626×10^{-34} m^2 kg s^{-1}.

Résumé

1 Introduction

Historiquement, le carbone est étroitement intégré aux processus biosphériques naturels. L'industrialisation a provoqué d'importantes perturbations du cycle du carbone: les niveaux de CO_2 atmosphérique sont passés de 280 ppm à l'époque préindustrielle à plus de 415 ppm en 2021.[1] Les océans et les forêts, qui jouaient auparavant le rôle efficace de puits de carbone, ne réabsorbent qu'environ 54% des émissions de CO_2, le reste s'accumulant dans l'atmosphère et nourrissant le changement climatique.

Pour atténuer cette tendance, le CO_2 peut potentiellement être utilisé comme matière première dans des systèmes énergétiques durables. En particulier, le stockage de l'énergie solaire dans les liaisons chimiques de molécules carbonées, selon un processus analogue à la photosynthèse naturelle, représente une stratégie prometteuse pour le stockage d'énergie et la mise en œuvre de la décarbonation. L'hydrogène est souvent cité comme vecteur énergétique propre, mais les défis liés à son stockage (faible densité énergétique, coûts élevés de compression, fragilisation des réservoirs, risques d'explosion) en limitent l'utilisation pratique. En revanche, les carburants à base de carbone (CO, CH_4, méthanol) présentent une densité énergétique plus élevée, une mise à l'échelle industrielle et un stockage plus sûrs, et peuvent être synthétisés à partir de CO_2 en utilisant des sources d'énergie renouvelables.

Le CO_2 est une molécule linéaire et apolaire, possédant des liaisons C=O polaires. Elle possède un moment quadrupolaire lié à la polarisation des deux liaisons C=O, avec un atome de carbone électrophile susceptible d'être attaqué par des nucléophiles. Sa réduction implique le peuplement de l'orbitale atomique (OA) antiliante σ^* des liaisons C=O, qui affaiblit ces liaisons et les rend réactives, permettant la conversion du CO_2 en divers produits réduits tels que le formiate, le monoxyde de carbone, le méthanol ou encore le méthane. Toutefois, la molécule de CO_2 est très peu réactive, en raison de la grande stabilité thermodynamique du CO_2 et de barrières cinétiques importantes. Des catalyseurs efficaces et sélectifs sont nécessaires pour faciliter les processus multi-électroniques et multi-protoniques pour faciliter sa réduction dans des conditions douces.

Ce travail de thèse utilisera trois approches pour la réduction catalytique du CO_2: la catalyse thermique, l'électrocatalyse et la photocatalyse. La catalyse thermique repose sur la température pour déplacer les équilibres chimiques et entraîner la transformation chimique du CO_2 dans le site catalytique. L'électrocatalyse met en œuvre des réactions rédox d'un catalyseur à la surface des électrodes sous l'effet de potentiels appliqués. La photocatalyse exploite l'énergie lumineuse pour exciter les électrons de molécules photosensibles, déclenchant des cascades de transferts d'électrons et de protons à des fins de réduction du substrat.

Dans le contexte de la catalyse thermique, cette thèse s'intéressera aux paires de Lewis frustrées (FLPs) moléculaires pour la conversion du CO_2. Ces molécules constituées d'un acide de Lewis (LA) et d'une base de Lewis (LB) présentent un encombrement stérique qui empêche la formation d'un adduit LA-LB. Cette « frustration » crée un site réactif entre les deux centres de Lewis capable d'activer de petites molécules inertes comme H_2 ou CO_2. Depuis leur première description en 2006 par Stephan et ses collaborateurs, [2] les FLPs sont apparues comme une alternative prometteuse, sans métal, aux catalyseurs classiques d'hydrogénation. Leur capacité à cliver H_2 et à transférer des paires hydrure/proton au CO_2 les rend particulièrement intéressantes pour développer une catalyse durable. Toutefois, leur conception reste complexe, nécessitant de trouver un équilibre délicat entre la force acide/base, l'encombrement stérique associé aux deux centres de Lewis et les propriétés électroniques de la molécule, de façon à favoriser à la fois le clivage de H_2 et l'hydrogénation du CO_2 sans stabiliser excessivement les intermédiaires réactionnels. Ce travail de thèse vise à étudier la famille de FLPs moléculaires à base de bore et d'azote, afin d'en élucider la réactivité pour l'hydrogénation du CO_2 et extraire des critères rationnels de design moléculaire.

Dans le contexte de la photocatalyse, cette thèse se concentrera sur deux grandes familles de catalyseurs moléculaires: les polyoxométallates (POMs) et les porphyrines métalliques. Les POMs sont des clusters anioniques à bases de complexes oxo de métaux de transition. Dans la majorité des cas, ils sont constitués de molybdène (polyoxomolybdate), de tungstène (polyoxotungstate) ou de vanadium (polyoxovanadate), capables de stocker réversiblement plusieurs électrons, ce qui les rend particulièrement adaptés à la médiation électronique dans la réduction du CO_2.[3] Les porphyrines métalliques, quant à elles, sont des complexes métalliques très étudiés pour la réduction photocatalytique du CO_2.[4] Sous forme métalée, à base de fer ou de cobalt, par exemple, les porphyrines combinent une forte affinité pour le CO_2 et des états excités riches en électrons, menant souvent à une conversion $CO_2 \rightarrow CO$ à haute sélectivité et activité.

Cette thèse de doctorat vise à étudier par des méthodes de chimie computationnelle les mécanismes catalytiques en jeu dans la dissociation de H_2 et de réduction du CO_2 par des catalyseurs tels que des paires de Lewis frustrées (FLPs), des polyoxométallates (POMs) et des porphyrines. Elle cherche également à identifier des relations structure-activité qui permettent d'expliquer les tendances réactives observées par calculs ou par l'expérience. Enfin, ce travail ambitionne de développer des modèles prédictifs via des calculs de chimie quantique, des simulations de dynamique moléculaire et des approches guidées par l'analyse de données. L'objectif final est d'établir des principes de conception pouvant orienter le développement rationnel de catalyseurs moléculaires de nouvelle génération, efficaces et durables.

2 Les paires de Lewis frustées B/N pour la dissociation de H_2

2.1 Contexte général

Ce premier chapitre présente notre approche computationnelle pour étudier la réactivité des paires de Lewis frustrées (FLPs) dans l'activation de l'hydrogène et l'hydrogénation du dioxyde de carbone. Dans un premier temps, les FLPs peuvent dissocier H_2 de manière hétérolytique en un proton et un hydrure, qui seront ensuite transférés au CO_2 pour produire de l'acide formique ou des produits apparentés. L'efficacité de ce processus dépend de l'équilibre entre la force acide et la force basique des composants de la FLP. Si la FLP est trop faible, l'activation de H_2 est lente; si elle est trop forte, le transfert au CO_2 devient énergétiquement défavorable.

A ce jour, diverses études computationnelles ont exploré l'influence de la structure moléculaire sur le cycle catalytique, notamment pour déterminer si l'activation de H_2 ou le transfert du proton et de l'hydrure constitue l'étape limitante.[5] Ces études ont proposé des mécanismes séquentiels ou concertés, bien que le mécanisme en deux étapes, activation de H_2 suivie de la réduction du CO_2, soit généralement plus favorable. Un large éventail d'architectures de FLPs a été étudié dans la littérature, et des efforts récents de criblage à haut débit ont commencé à cartographier la relation entre les caractéristiques moléculaires et les performances catalytiques en utilisant des descripteurs thermodynamiques tels que les énergies libres d'attachement du proton et de l'hydrure.[6]

Afin de construire une compréhension plus générale de la réactivité des FLPs, notre travail a porté sur la famille des FLPs moléculaires à base de bore et d'azote, les deux sites de Lewis étant présent au sein de la même molécule. Notre approche vise à définir des descripteurs thermodynamiques transférables et chimiquement significatifs, l'affinité pour le proton pour les bases et l'affinité pour l'hydrure pour les acides, par exemple, basés sur des réactions de référence standardisées, i.e. dissociation de H_2 et hydrogénation du CO_2 en acide formique. Cette approche permet une comparaison cohérente entre différentes molécules et permet le développement de relations structure–activité qui ne sont pas limitées à des géométries spécifiques de FLPs. Le cadre ainsi établi fournit une base pour la conception rationnelle de FLPs et pour le criblage à haut débit dans l'activation de petites molécules.

2.2 Méthodologie

En terme de méthodologies, les calculs DFT (ωB97X-D/def2-TZVP avec solvant PCM dans le toluène) ont été réalisés pour calculer le profil réactionnel de dissociation de H_2, i.e. les paramètres thermodynamiques ($\Delta G_{H_2\ dissociation}$) et cinétiques

($\Delta G^{\ddagger}{}_{H_2\ dissociation}$) de 112 FLPs moléculaires (Figure 2). Dans un premier temps, une procédure automatisée en Python a été développée pour extraire de manière systématique les données structurales et électroniques clés, énergies, géométries, charges, etc., à partir des fichiers de sortie Gaussian, assurant cohérence et passage à l'échelle. Dans un second temps, le script « FLP finder » analyse les données extraites afin d'identifier le motif signature de la FLP, sur la base de la proximité géométrique et de l'identité chimique des centres acide et base de Lewis de chaque molécule. Enfin, des approches de régression ont été utilisées pour établir des relations structure–réactivité en utilisant une variété de descripteurs structuraux et électroniques. La régression linéaire multivariée (MLR) et la méthode des moindres carrés partiels (PLS) ont été combinées avec une validation croisée de type « k-fold » pour construire des modèles prédictifs reliant les descripteurs moléculaires à la réactivité, tant d'un point de vue thermodynamique que cinétique. Une méthode guidée par les données a été appliquée de manière itérative afin d'identifier et sélectionner les descripteurs moléculaires les plus informatifs tout en préservant la qualité des modèles.

2.3 Résultats

Notre analyse s'est concentrée initialement sur la corrélation entre ΔG et $\Delta G^{\ddagger}$, révélant que des énergies libres de réaction faibles ne sont pas nécessairement associées à des barrières d'activation faibles. Les histogrammes des distances H–H à l'état de transition, classés par famille de FLPs, montrent des tendances distinctes, reflétant l'influence des caractéristiques géométriques sur la réactivité.

La régression linéaire multivariée (MLR) a été utilisée comme un outil fondamental pour extraire une compréhension chimique à partir de jeux de données de descripteurs de haute dimension, en construisant des modèles linéaires reliant plusieurs descripteurs moléculaires à une propriété cible. Dans le cadre de ce travail, la MLR est utilisée pour prédire la stabilité thermodynamique (ΔG) des réactions de dissociation de H_2 catalysées par les paires de Lewis frustrées (FLP). Cette méthode permet de construire des modèles interprétables qui quantifient la contribution de chaque descripteur à la propriété étudiée.

Les descripteurs considérés incluent des paramètres géométriques (par exemple, la distance $LA\cdots LB$), des descripteurs électroniques comme FEHA et FEPA, ainsi que des angles structuraux tels que l'angle de distorsion (λ), Φ et l'angle dièdre (γ). Pour éviter la multicolinéarité et le sur-apprentissage, les variables sont normalisées et sélectionnées en fonction de leur pertinence, à l'aide de techniques telles que la validation croisée et l'analyse de corrélation.

Le meilleur modèle obtenu ($r^2 = 0.87$ et $q^2 = 0.85$, Eq. 1) révèle que l'énergie libre de dissociation de H_2, ΔG, dépend uniquement de quelques descripteurs significatifs, les plus influents étant la distance d, FEHA (énergie d'attachement de l'hydrure) et FEPA (énergie d'attachement du proton). La visualisation de l'ensemble de données projetées dans l'espace des descripteurs illustre des tendances nettes, avec des

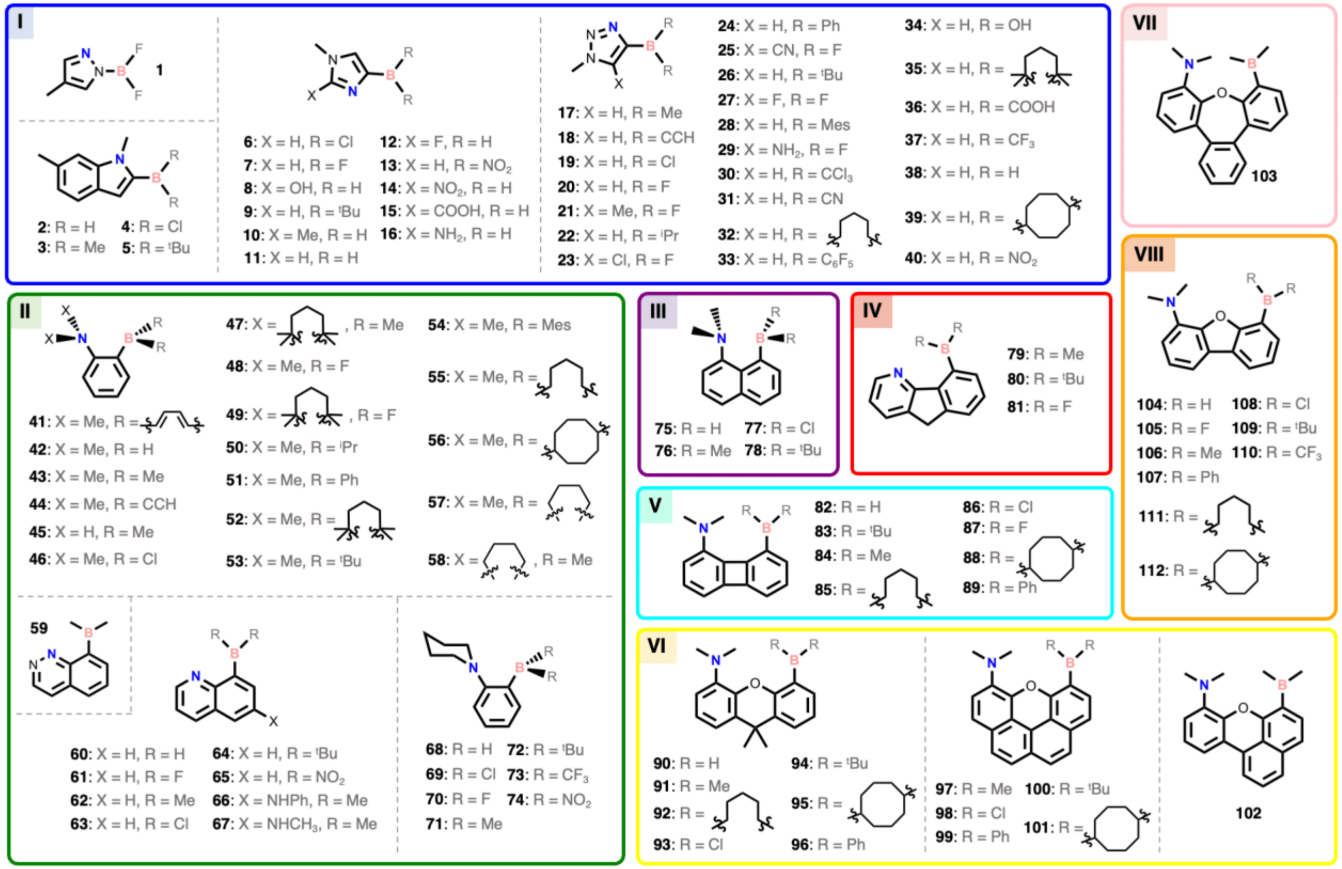


FIGURE 2: Librairie de toutes les FLPs intramoléculaires B/N constituant l'ensemble de données étudiées. La plupart d'entre elles sont inspirées de structures de FLPs analysées dans des travaux antérieurs, expérimentaux ou théoriques.

régions de faible ΔG bien regroupées (Figure 1).

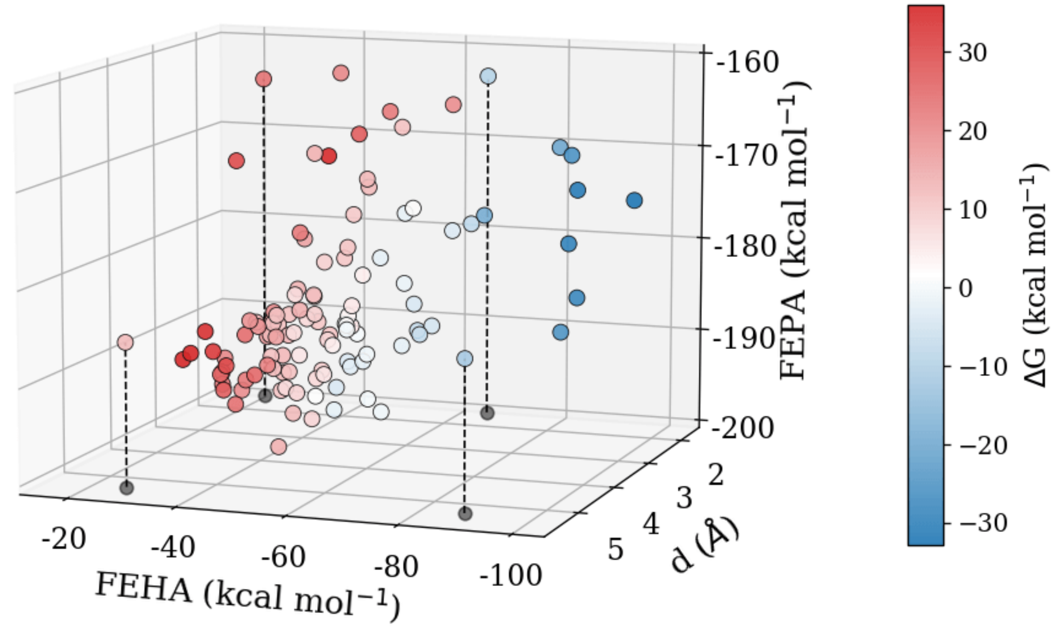


FIGURE 1: Nuage de points des données projetées dans l'espace tridimensionnel défini par d, FEHA et FEPA. Chaque point est coloré en fonction de sa valeur de ΔG calculée par DFT. Quatre points de données sélectionnés sont projetés sur le plan FEHA vs d afin d'illustrer le pouvoir discriminant supérieur de FEHA par rapport aux autres descripteurs.

$$\begin{aligned} \Delta G &= 28.9 \cdot \widetilde{d} + 67.6 \cdot \widetilde{\mathrm{FEHA}} + 24.4 \cdot \widetilde{\mathrm{FEPA}} - 54.4 \\ \Delta G &= 7.37 \cdot d + 0.821 \cdot \mathrm{FEHA} + 0.685 \cdot \mathrm{FEPA} + 151 \\ r^2 = 0.87 \qquad q^2 &= 0.85 \qquad \mathrm{RMSE} = 5.52\ \mathrm{kcal} \cdot \mathrm{mol}^{-1} \end{aligned} \tag{1}$$

Dans l'ensemble, ces observations confirment que la nature des substituants du site de Lewis, affectant directement son acidité, joue un rôle dominant dans la stabilisation du produit de dissociation de H_2. En d'autres termes, une acidité plus élevée est corrélée à des réactions plus exergoniques, comme on l'observe pour des groupements fortement électroattracteurs tels que –NO_2, –CN, –CCl_3 ou –CF_3. Des distances LA$\cdots$LB plus courtes et une basicité plus forte de la base de Lewis favorisent également la dissociation de H_2 d'un point de vue thermodynamique, mais dans une moindre mesure.

Ensuite, nous avons examiné le rôle des caractéristiques structurales et électroniques dans le contrôle de l'énergie libre d'activation ($\Delta G^{\ddagger}$) pour la dissociation hétérolytique de H_2 par des FLP intramoléculaires B/N. L'approche MLR, décrite ci-dessus pour ΔG, n'a pas permis d'identifier les descripteurs les plus importants parmi tous ceux calculés. Afin d'y remédier, nous avons mis au point une autre méthodologie guidée par les données, inspirée de celle de Andries et al.[7] Pour évaluer si un descripteur moléculaire contribue réellement à la prédiction de la réactivité ou s'il contribue au « bruit », i.e. sans contribution spécifique, un ensemble de descripteurs aléatoires est généré. Ces descripteurs aléatoires ont la même distribution que les vrais descripteurs mais sont statistiquement indépendants du système chimique étudié. Ils servent ainsi de référence afin de distinguer le signal pertinent du « bruit ». Une fois les descripteurs aléatoires générés, la performance prédictive des descripteurs réels est comparée à ces derniers à l'aide de la régression par moindres carrés partiels (PLS). Plus précisément, la capacité prédictive de chaque descripteur (mesurée par validation croisée) est évaluée et classée. Un seuil est ensuite défini à partir de la distribution des scores prédictifs des descripteurs aléatoires. Les descripteurs réels doivent dépasser ce seuil pour être considérés comme informatifs. Ce processus garantit ainsi que seuls les descripteurs ayant une contribution statistiquement significative, c'est-à-dire surpassant le bruit, sont retenus pour les modèles ultérieurs.

Les résultats de la régression révèlent que le modèle le plus prédictif pour $\Delta G^{\ddagger}$ repose (r^2 = 0.85 et q^2 = 0.81, Eq. 2) principalement sur deux descripteurs géométriques: la distance d entre les centres acide et base de Lewis, et l'angle de distorsion λ, qui représente la déviation par rapport à la géométrie idéale de la poche réactive de la FLP, représenté par le minimum décrit par le comportement quadratique des descripteurs dans le modèle (Figure 3). Ces deux caractéristiques permettent à elles seules une bonne séparation entre les systèmes à faible et à haute barrière cinétique. Des valeurs plus faibles de d sont généralement corrélées à des barrières d'activation plus faibles, car elles favorisent un meilleur recouvrement des orbitales et une meilleure polarisation de H_2. Cependant, une trop grande proximité peut induire une contrainte stérique ou une formation prématurée d'un adduit. L'angle λ reflète le coût énergétique nécessaire pour déformer le FLP de son état fondamental optimisé vers la géométrie de l'état de transition, influençant ainsi la préorganisation de la FLP pour réaliser la dissociation hétérolytique de H_2.

$$\begin{gathered}
\Delta G^{\ddagger} = 161 \cdot \widetilde{d^2} - 133 \cdot \widetilde{d} + 51.4 \cdot \widetilde{\lambda^2} - 38.8 \cdot \widetilde{\lambda} - 21.9 \cdot \widetilde{\mathrm{FEHA}^2} + 9.32 \cdot \widetilde{\mathrm{FEPA}} \\
-18.2 \cdot \widetilde{q_{lb}(\mathrm{NPA})^2} - 12.4 \cdot \widetilde{\mathrm{EP}(\mathrm{NPA})^2} - 20.0 \cdot \widetilde{\mathrm{EP}(\mathrm{NPA})} + 55.4 \\
\Delta G^{\ddagger} = 5.60 \cdot d^2 - 34.0 \cdot d + 0.00159 \cdot \lambda^2 - 0.222 \cdot \lambda - 0.00228 \cdot \mathrm{FEHA}^2 \\
+0.262 \cdot \mathrm{FEPA} - 30.2 \cdot q_{lb}(\mathrm{NPA})^2 - 57.4 \cdot \mathrm{EP}(\mathrm{NPA})^2 - 45.1 \cdot \mathrm{EP}(\mathrm{NPA}) + 131 \\
r^2 = 0.85 \qquad q^2 = 0.81 \qquad \mathrm{RMSE} = 5.13\ \mathrm{kcal\,mol^{-1}}
\end{gathered} \tag{2}$$

Un résultat clé de cette étude est de montrer que les descripteurs thermodynamiques (FEHA et FEPA), qui influencent fortement les énergies libres de réaction (ΔG), sont moins prédictifs pour $\Delta G^{\ddagger}$. Cela indique que les contrôles cinétique et thermodynamique obéissent à des relations structure–activité différentes. Alors que les descripteurs FEHA et FEPA décrivent la force motrice de l'attachement du proton ou de l'hydrure, la structure de l'état de transition dépend davantage de l'accessibilité

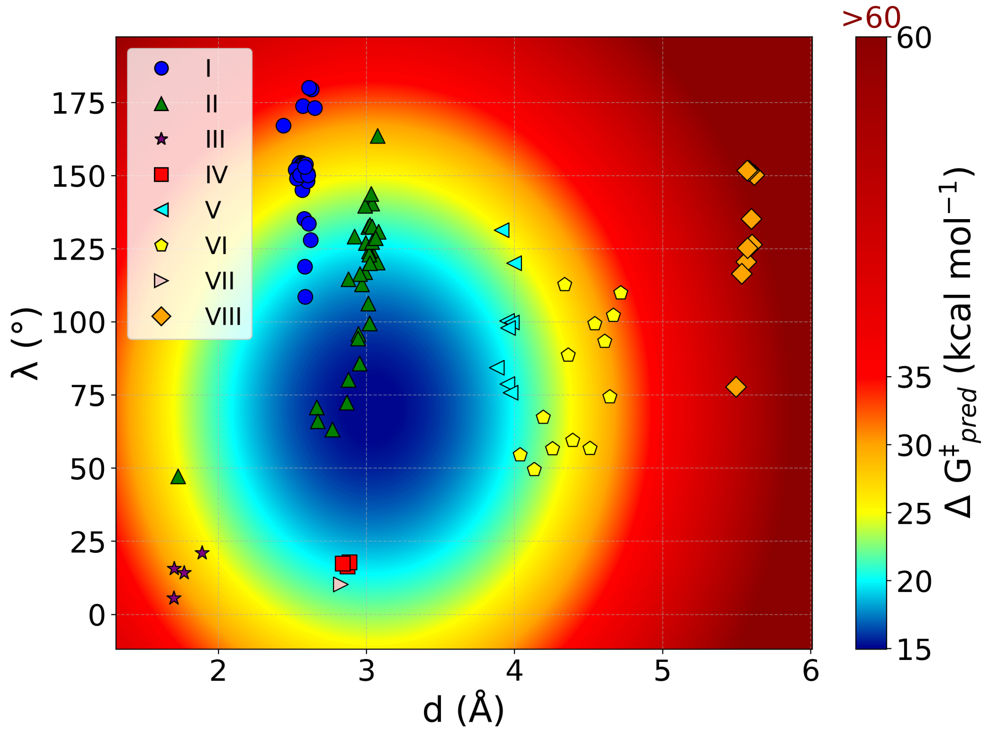


FIGURE 3: Répartition des FLP, classés par familles, superposée à la carte de chaleur de $\Delta G^{\ddagger}$ prédite par l'Éq. 2 en fonction de d et λ. Tous les autres descripteurs ont été fixés à leur valeur moyenne pour générer la surface de réponse.

stérique et de la dynamique conformationnelle. De fait, une corrélation linéaire est observée entre $\Delta G^{\ddagger}$ et l'énergie de distorsion, i.e. l'énergie nécessaire pour déformer la FLP de son état fondamental optimisé vers la géométrie de l'état de transition. Ceci confirme le rôle déterminant de la déformation de la FLP dans l'accès à l'état de transition lors de l'étape de dissociation de H_2. Ce résultat suggère également que les systèmes géométriquement pré-organisés, i.e. déjà proches de la géométrie de l'état de transition, nécessitent un apport énergétique moindre pour l'activation de H_2, et présentent donc des barrières cinétiques plus faibles.

Fait notable, même lorsque les descripteurs de l'acidité et de la basicité intrinsèques, tels que FEHA et FEPA, sont exclus du modèle, ce dernier conserve une forte capacité prédictive ($r^2 = 0.73$ et $q^2 = 0.68$) en s'appuyant uniquement sur des descripteurs géométriques et électroniques. Ce résultat souligne que l'agencement spatial et l'environnement électronique du FLP jouent un rôle plus déterminant dans la réduction de la barrière d'activation que la simple force de ses composants acide et base. En d'autres termes, une préorganisation structurale efficace peut compenser une acidité ou une basicité modérée en réduisant le coût énergétique nécessaire pour atteindre l'état de transition.

Enfin, une procédure de validation « externe » a permis d'évaluer la capacité prédictive des modèles développés en utilisant un ensemble de molécules FLPs non incluses dans l'entraînement du modèle. Les valeurs prédites des énergies libres thermodynamiques et cinétiques ont été comparées aux valeurs de référence calculées par DFT pour ces mêmes molécules « externes ». Cet ensemble de molécules « externes » à l'entrainement du modèle a été choisi afin de comprend des familles de FLP diverses afin d'assurer une représentativité chimique (intramoléculaire B/P, Al/N, Al/P et intermoléculaire B/N FLPs). Cette diversité est nécessaire pour s'assurer que les modèles rendent comptent des relations structure–activité généralisables, plutôt que de s'ajuster à des motifs structuraux spécifiques. Globalement, nos tests de validation « externe » a permis de confirmer la robustesse des modèles pour la prédiction de ΔG et leur utilité pour anticiper les tendances de réactivité dans de nouvelles molécules FLPs non « vues » lors de l'entrainement du modèle. En revanche, concernant la barrière d'activation, $\Delta G^{\ddagger}$, les modèles n'ont pas réussi à reproduire les tendances des nouvelles familles de FLPs, indiquant que d'autres descripteurs sont nécessaires pour déterminer $\Delta G^{\ddagger}$ à travers différentes familles de FLPs.

En utilisant les mêmes descripteurs que pour la coupure de H_2, un modèle a été développé pour prédire la barrière de l'activation du CO_2 ($\Delta G^{\ddagger}{}_{CO_2\ hydro}$). Les meilleures performances (r^2 = 0.82 et q^2 = 0.81; Eq. 3) sont obtenues à l'aide des descripteurs FEHA^2, de la distance LA-LB (d) et de l'angle λ. FEHA^2 était le terme le plus influent, bien que la géométrie joue également un rôle notable.

$$\begin{aligned}
\Delta G^{\ddagger}_{CO_2\ hydro} &= 46.8 \cdot \widetilde{\mathrm{FEHA}^2} - 18.2 \cdot \widetilde{d} - 8.42 \cdot \widetilde{\lambda} - 16.4 \\
\Delta G^{\ddagger}_{CO_2\ hydro} &= 0.00486 \cdot \mathrm{FEHA}^2 - 4.65 \cdot d - 0.0483 \cdot \lambda + 28.6 \\
r^2 = 0.82 \qquad q^2 &= 0.81 \qquad \mathrm{RMSE} = 4.51\ \mathrm{kcal\,mol^{-1}}
\end{aligned} \tag{3}$$

Une analyse plus poussée des géométries des états de transition de l'hydrogénation du CO_2, a révélé une diversité mécanistique selon les molécules FLPs. Alors que certaines molécules présentent un état de transition idéal concerté avec un transfert simultané de H^- et H^+, d'autres montrent un caractère asynchrone, dans lequel les transferts d'hydrure et de proton sont décalés. Dans certains systèmes, un déplacement protonique minimal à l'état de transition suggérait un mécanisme en deux étapes, le transfert de proton survenant plus tard. À l'inverse, d'autres systèmes favorisaient un transfert de proton précoce. Une géométrie unique de transfert d'hydrure perpendiculaire au plan du CO_2 a également été observée, en particulier dans les FLP présentant de grandes distances entre l'acide et la base de Lewis, ainsi que des groupes fortement électroattracteurs sur l'acide de Lewis. Cette variabilité structurale souligne la complexité mécanistique de l'hydrogénation du CO_2 par les FLPs et met en évidence l'importance de modèles complets intégrant à la fois les effets géométriques et électroniques.

En conclusion, ce chapitre présente un cadre fondé sur l'analyse de données pour comprendre et prédire les relations structure–activité dans les paires de Lewis frustrées (FLP) à base de bore/azote pour la dissociation de H_2 et l'hydrogénation du CO_2. Des modèles ont été développés pour prédire les énergies libres de réaction (ΔG) et les barrières d'activation ($\Delta G^{\ddagger}$), en identifiant les descripteurs moléculaires clés, notamment l'acidité (FEHA), la basicité (FEPA) et la distance $LA\cdots LB$ (d), comme les principaux facteurs influençant la thermodynamique, avec un effet prédominant de FEHA. En revanche, $\Delta G^{\ddagger}$ est plus sensible aux caractéristiques géométriques telles que la distance $LA\cdots LB$ (d) et l'angle de distorsion λ, les barrières étant minimisées à des valeurs optimales ($d \approx 3.0$ Å, $\lambda \approx 70°$, Figure 3). Nous observons que nos modèles se généralisent bien pour prédire la thermodynamique de la dissociation (ΔG), mais sont moins transférables pour prédire la cinétique ($\Delta G^{\ddagger}$) selon les types de FLPs considérées. Une approche similaire basée sur les descripteurs a été appliquée à l'hydrogénation du CO_2, produisant un modèle prédictif pour sa barrière d'activation et révélant des trajectoires mécanistiques diverses. Les efforts futurs viseront à intégrer les processus de désactivation et à élargir l'applicabilité des modèles. Dans l'ensemble, cette approche offre des éclairages mécanistiques ainsi que des outils pratiques pour évaluer et concevoir des systèmes FLP efficaces pour l'hydrogenation du CO_2.

3 Réduction photocatalytique du CO_2 assistée par les POMs

Ce chapitre s'inscrit dans le contexte de la réduction catalytique du CO_2 assistée par les polyoxométallates (POMs), en mettant en avant leur potentiel en tant que plateformes moléculaires pour les transformations multi-électroniques. Il souligne d'abord le rôle central des processus photocatalytiques et électrocatalytiques dans le développement des méthodes durables de valorisation du CO_2. Les POMs sont des clusters métal-oxo redox-actifs, dotés d'une diversité structurale exceptionnelle, de propriétés électroniques modulables et d'une grande stabilité en conditions catalytiques. Ces caractéristiques les rendent particulièrement attrayants pour des applications en photosynthèse artificielle et en conversion du CO_2, notamment.[8] Les POMs offrent une variété moléculaire et une modularité permettant un réglage de leur comportement redox et de leur environnement de coordination. De plus, leur capacité à stocker réversiblement plusieurs électrons leur permet de faciliter la réduction complexe et en plusieurs étapes du CO_2 dans des conditions douces. Une attention particulière est accordée au rôle des interactions supramoléculaires et à l'incorporation d'agents actifs supplémentaires (par exemple, des photosensibilisateurs ou des donneurs d'électrons sacrificiels), qui peuvent améliorer l'activité et la sélectivité catalytique de deux systèmes moléculaires à base de POMs. Le chapitre aborde l'étude computationnelle de leur structure électronique, de leurs propriétés redox et des mécanismes réactionnels en jeu dans la réduction du CO_2 en CO, dans le but de dégager des principes de conception généralisables pour des catalyseurs moléculaires de réduction du CO_2 efficaces et sélectifs.

3.1 Ni-substituted POMs as molecular catalysts for CO_2 reduction

3.1.1 Contexte Expérimental

Dans le cadre d'une collaboration avec l'Institut Lavoisier de Versailles, et plus particulièrement dans le contexte de la thèse de doctorat de K. Talbi, cette dernière a exploré l'activité photocatalytique de divers sels de polyoxométallates (POMs) substitués au nickel dans des conditions homogènes, en milieu CH_3CN:H_2O:TEOA saturé en CO_2. Le système étudié utilise $[Ru(bpy)_3]^{2+}$ (Ru^{2+}) comme photosensibilisateur (PS), le TEOA comme donneur d'électrons sacrificiel (SED), et une irradiation par lumière visible ($\lambda > 415$ nm). Le CO et le H_2 sont les seuls produits gazeux détectés, avec une formation négligeable de formiate en phase liquide. Parmi les divers catalyseurs testés à base de POMs, (TBA)Ni_4 a montré la plus forte activité de production de CO, ce qui a conduit à des études d'optimisation supplémentaires. Ses performances optimales ont été obtenues avec un rapport CH_3CN:H_2O:TEOA de 8.8:2:1 et une concentration de 1.95 mM en Ru^{2+}. L'absence d'eau ou l'utilisation d'autres photosensibilisateurs comme $[Ru(phen)_3]^{2+}$ ou $[Ir(dtbbpy)(ppy)_2]^+$ entraîne une réduction significative de la production de CO. L'évolution du CO cesse après environ 4 heures, mais reprend après ajout de PS, indiquant que la dégradation du photosensibilisateur est un facteur limitant. La présence de centres catalytiques de nickel est essentielle à la catalyse, comme le montrent les faibles activités des précurseurs dépourvus de Ni. La sélectivité s'améliore à plus faibles concentrations de catalyseur, et un marquage isotopique confirme que le CO provient du CO_2.

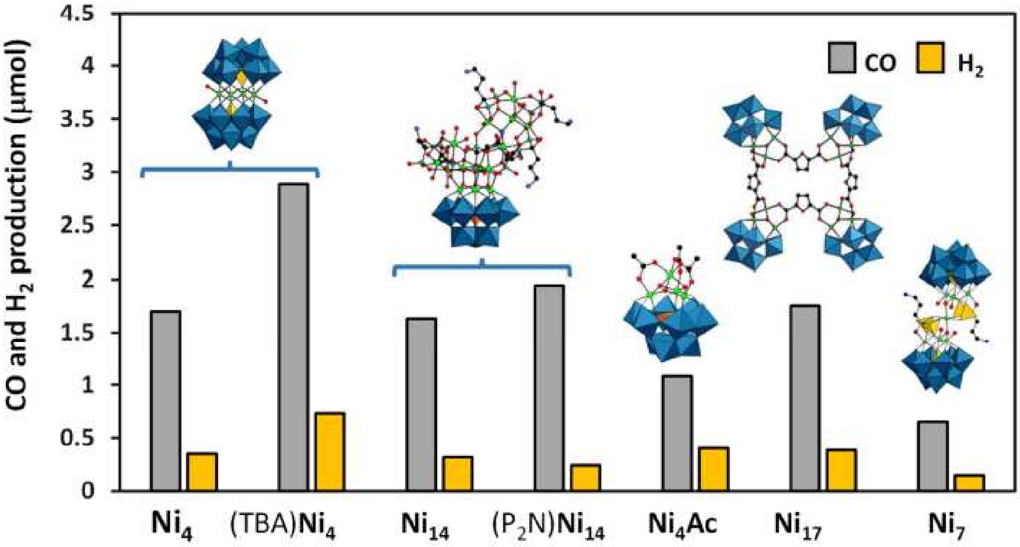


FIGURE 4: Activité de réduction du CO_2 et du H_2 (barres en gris et en orange respectivement) pour toutes les combinaisons catalyseur-contre-ion étudiées. Le tungstène octaédrique est coloré en bleu, le phosphore tétraédrique en jaune ou orange, le carbone en noir, l'azote en bleu clair, l'oxygène en rouge et le nickel en jaune.

Le bullage du CO_2 provoque une séparation de phases révélant un système biphasique, dans lequel le POM se retrouve en phase aqueuse et le cation TBA^+ en phase organique. Cette séparation est essentielle à la catalyse, car l'élimination du CH_3CN entraîne une forte diminution de l'activité. Cet effet bénéfique est attribué à une meilleure séparation des paires ioniques et à des interactions POM–PS favorables à l'interface phase aqueuse/phase organique, où la concentration en CO_2 est plus élevée. En revanche, les contre-ions inorganiques (Na^+, K^+) restent en phase aqueuse et nuisent à l'efficacité catalytique par des interactions plus fortes avec le POM.

L'intégrité structurale et la stabilité de (TBA)Ni_4 ont été confirmées après catalyse par DLS, sans formation de nanoparticules de NiO. Malgré de nombreuses études, les paramètres clés gouvernant la réduction du CO_2 par les POMs restent mal compris, en raison de la complexité des voies rédox et des transferts de protons, ainsi que de mécanismes de dégradation du PS encore incertains. Pour mieux comprendre ces phénomènes, nous avons entrepris une étude computationnelle combinant DFT, TD-DFT et DFT-MD ont été menées sur (TBA)Ni_2 afin de fournir une compréhension mécanistique de la réduction du CO_2 et guider la conception de futurs catalyseurs.

3.1.2 Résultats

Les propriétés rédox de Ni_4 ont d'abord été évaluées dans le cadre de cette thèse. Comprendre les propriétés rédox des polyoxométallates (POMs) nécessite tout d'abord l'évaluation de leurs états de protonation. Dans le cas du POM Ni_4-substitué, une caractérisation acido-basique a été réalisée à l'aide de la forme protonée de la triéthanolamine ($HTEOA^+$), formée lors de la génération d'électrons du TEOA avec le photosensibilisateur sous irradiation. $HTEOA^+$ peut également être générée par transfert de proton à partir de l'eau, en accord avec son pK_a expérimental de 10.75. Le POM Ni_4 contient de nombreux ligands oxo pouvant agir comme bases de Brønsted. Sur la base de la symétrie du catalyseur (axe C_2 et plan miroir), quatre ligands oxo pontants distincts ont été identifiés et évalués pour leur basicité. Ces positions ont été différenciées en types pontants (M–O–M) et terminaux (M=O), les premiers étant plus basiques que les seconds en raison d'interactions métal–oxygène plus faibles. Les énergies libres de protonation calculées montrent une tendance: les ligands oxo pontants entre Ni et W (W–O–Ni) sont les plus basiques (ΔG = –6.8 kcal mol^{-1}), tandis que ceux entre deux atomes de W le sont moins (ΔG = 8.9 kcal mol^{-1}). Cette tendance est confirmée par des cartes de potentiel électrostatique, qui révèlent des régions riches en électrons autour des centres Ni.

Une analyse plus poussée montre que Ni_4 peut spontanément accepter deux protons de $HTEOA^+$, avec des énergies libres favorables de –6.8 et –5.1 kcal mol^{-1}. Ces protons se lient à des sites pontants W–O–Ni symétriques et forment des liaisons hydrogène intramoléculaires stabilisantes. En revanche, une troisième protonation est thermodynamiquement défavorable (4.2 kcal mol^{-1}). Ces résultats constituent une première base pour étudier le comportement rédox de l'espèce doublement protonée H_2Ni_4 et sa réduction par transfert monoélectronique du photosensibilisateur à base de ruthénium.

Avant d'évaluer la réductibilité du polyoxométallate (POM) H_2Ni_4, il est nécessaire d'examiner le pouvoir réducteur du photosensibilisateur (PS), $[Ru(bpy)_3]^{2+}$, et de le triéthanolamine (TEOA), le donneur sacrificiel d'électrons (SED). Ce système peut fonctionner selon deux mécanismes: la désactivation oxydative ou la désactivation réductrice. Dans le mécanisme de désactivation oxydative, le photosensibilisateur excité (PS^*) transfère un électron à H_2Ni_4, puis est régénéré par un transfert d'électron depuis la TEOA. Cependant, cette voie est thermodynamiquement défavorable en raison d'une recombinaison efficace des charges entre le POM réduit et le PS oxydé (64.5 kcal mol^{-1}), ce qui empêche une séparation des charges stable. Les données photophysiques expérimentales confirment cette hypothèse, car aucune espèce à charges séparées de longue durée de vie n'a été observée, rendant le mécanisme oxydatif peu probable.

À l'inverse, la désactivation réductrice, dans laquelle le TEOA réduit le PS^* excité pour former Ru^+, qui réduit ensuite H_2Ni_4, est fortement étayée par des résultats expérimentaux et théoriques. Bien que la réduction initiale du Ru^{2+} excité par le TEOA soit endergonique (10.8 kcal mol^{-1}), elle est rendue thermodynamiquement favorable par la déprotonation exergonique du radical $TEOA^{\bullet+}$ (-11.3 kcal mol^{-1}). De plus, l'espèce Ru^+ est observée expérimentalement comme étant de longue durée de vie et plus réductrice que le radical $TEOA^{\bullet}$. Des études DFT montrent que Ru^+ peut réduire la forme doublement protonée H_2Ni_4 de manière ergoneutre ($\Delta G \approx$ 0.0 kcal mol^{-1}), alors que la réduction de la forme non protonée Ni_4 est bien moins favorable (16.9 kcal mol^{-1}), soulignant ainsi le rôle crucial de la protonation dans la stabilisation des orbitales vacantes du catalyseur.

Dans l'ensemble, les preuves combinées issues des calculs et des expériences photophysiques favorisent fortement un mécanisme de désactivation réductrice, où la protonation du POM ainsi que de possibles effets de paires d'ions renforcent le transfert d'électron vers le POM.

Pour identifier la forme active du catalyseur dans les conditions photocatalytiques, les états de protonation et de réduction du POM Ni_4 ont ensuite été étudiés de manière systématique. Si les réductions initiales du cluster Ni_4 sont fortement endergoniques, sa basicité élevée permet une double protonation spontanée, pour former H_2Ni_4. Cette espèce protonée peut ensuite subir une première réduction ergoneutre, suivie d'une seconde réduction modérément endergonique (7.2 kcal mol^{-1}), conduisant à H_2Ni_4(2e). Une protonation supplémentaire au niveau des ponts Ni–O–W stabilise le système pour former H_4Ni_4(2e), qui présente un profil thermodynamique favorable et permet des réductions supplémentaires accessibles. Toutefois, au-delà de H_4Ni_4(2e), les étapes supplémentaires de protonation et de réduction deviennent de plus en plus défavorables. Une analyse de Boltzmann montre que H_4Ni_4(2e) est l'espèce prédominante (90.7 %) dans les conditions de fonctionnement, avec des populations mineures de H_5Ni_4(3e), H_6Ni_4(4e) et H_4Ni_4(3e). Ces résultats établissent H_4Ni_4(2e) comme la forme active du catalyseur, qui apporte les électrons nécessaires à la conversion de CO_2 en CO.

La réduction du CO_2 en CO catalysée par l'espèce identifiée plus haut, i.e. H_4Ni_4 (2e), a été ensuite étudiée à l'aide de calculs DFT statiques et de simulations de dynamique moléculaire (MD) basées sur la DFT. Thermodynamiquement, la réaction globale, produisant CO et H_2O et régénérant la forme oxydée H_2Ni_4(0e), est légèrement endergonique (4.5 kcal mol^{-1}). Un processus similaire impliquant H_5Ni_4(3e) a été calculé comme étant bien plus endergonique (16.7 kcal mol^{-1}), en raison de la faible stabilité du produit oxydé H_3Ni_4(1e).

L'exploration mécanistique a commencé par un échange de ligand qui consiste en la substitution du ligand aqua sur un site Ni^{II} du POM par une molécule de CO_2. Le calcul DFT statique montre que ce processus est fortement endergonique (53.5 kcal mol^{-1}). Cependant, la prise en compte du couplage antiferromagnétique entre Ni^{II} et l'anion radicalaire $CO_2{}^{-}$ (–$COO^{\bullet -}$) stabilise le produit et réduit le coût énergétique à 34.1 kcal mol^{-1}. Malgré tout, cette barrière reste trop élevée pour une catalyse efficace. Pour surmonter les limites des calculs statiques, des simulations de type path-metadynamics en solvant explicite (DFT-MD) ont été réalisées. Ces dernières ont révélé que le processus d'échange de ligand, impliquant la libération d'eau et l'insertion du CO_2, présente un coût énergétique libre bien plus bas (20.3 kcal mol^{-1}) que celui estimé en solvant implicite, ce qui le rend réalisable dans des conditions expérimentales. La majeure partie de la barrière est associée à la libération de l'eau (10.0 kcal mol^{-1}), plutôt qu'à l'insertion du CO_2 proprement dite (~10.3 kcal mol^{-1}), ce qui met en évidence les limites des modèles statiques à solvant implicite. Les étapes suivantes ont été modélisées à l'aide de calculs DFT statiques. Après la fixation du CO_2 et le transfert d'un électron, une protonation par un site Ni–O(H)–W stabilise l'intermédiaire –$COO^{\bullet -}$ (exergonique de -11.2 kcal mol^{-1}), déclenchant un second transfert d'électron qui forme une espèce Ni–COOH. La déprotonation par le TEOA et la protonation du groupe OH par $HTEOA^{+}$ favorisent ensuite la rupture de la liaison C–OH ($\Delta G^{\ddagger} = 4.3$ kcal mol^{-1}), menant à une espèce Ni–CO et à la libération d'une molécule d'eau. Bien que cette étape apparaisse légèrement endergonique (1.5 kcal mol^{-1}) selon les calculs DFT statiques, la correction tenant compte de l'énergie de libération de l'eau issue des DFT-MD la rend exergonique (-15.0 kcal mol^{-1}). Enfin, la libération du CO et la régénération du catalyseur sont thermodynamiquement favorables, fournissant la force motrice nécessaire à la réaction globale. L'ensemble du mécanisme est illustrée en Figure 5.

Cette section présente une proposition mécanistique pour la réduction du CO_2 par le système photocatalytique Ni_4/Ru^{2+}/TEOA, fondée sur des observations expérimentales et une étude théorique. Le catalyseur accepte deux électrons et quatre protons par un mécanisme de quenching réducteur induit par la lumière, formant une espèce active qui se lie au CO_2 via un des centres Ni^{II} du POM. Des transferts de protons et d'électrons facilitent la rupture de la liaison C–O, produisant

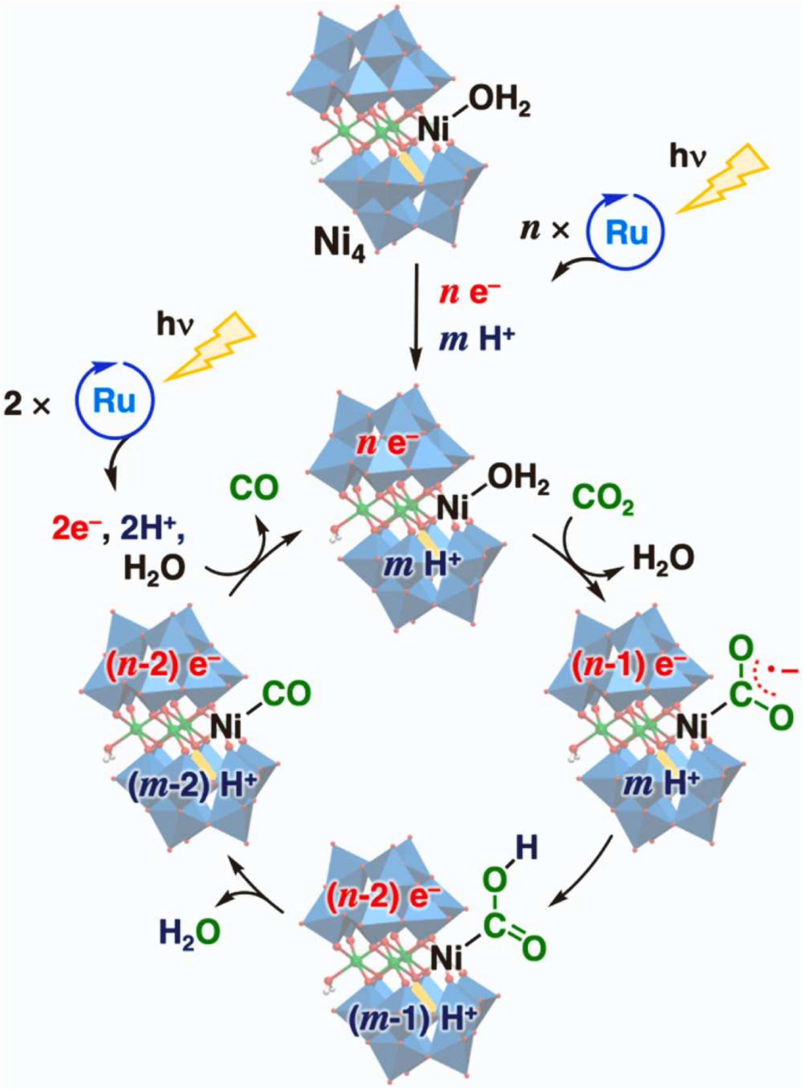


FIGURE 5: Schéma général du mécanisme proposé, avec n = 2 et m = 4 pour H_4Ni_4(2e) et n = 3 et m = 5 pour H_5Ni_4(3e).

du CO et régénérant le catalyseur. Les ligands Keggin trilacunaires, à caractère redox actif, présents dans Ni_4, sont essentiels pour le stockage efficace des électrons et l'accumulation de charge sous irradiation. Les résultats computationnels soulignent la difficulté de modéliser l'échange de ligands avec la DFT conventionnelle, ce qui met en évidence la nécessité d'une solvatation explicite et de techniques d'échantillonnage avancées. L'étude soutient la conception de catalyseurs conservant la capacité redox de Ni_4 tout en minimisant les barrières énergétiques à la coordination du CO_2, offrant ainsi un cadre pour l'amélioration des systèmes de réduction du CO_2.

3.2 Dyades photocatalytiques porphyrine-POM pour la réduction du CO_2

Dans la deuxième partie du Chapitre 3, une dyade covalente porphyrine–POM a été étudiée pour la réduction du CO_2. Les porphyrines sont des macrocycles plans et hétérocycliques, constitués de quatre unités pyrroliques reliées par des ponts méthine, formant ainsi un système π conjugué étendu. Cette structure confère une coloration intense et des caractéristiques d'absorption spécifiques, incluant une forte bande de Soret (~400 nm) et plusieurs bandes Q (500–700 nm) dans le spectre UV-visible, qui sont essentielles à leurs fonctions biologiques dans des systèmes tels que l'hème, la chlorophylle et les cytochromes. La cavité centrale des porphyrines peut coordonner des ions métalliques pour former des métalloporphyrines (MP), qui présentent des propriétés redox, magnétiques et optiques modulables, pertinentes pour la catalyse et la capture de lumière.

Les porphyrines sont accessibles par synthèse via des méthodes telles que les procédés d'Adler–Longo et de Lindsey, et leur robustesse chimique les rend aptes à être utilisées dans des conditions exigeantes, ainsi que comme composants de matériaux avancés tels que les MOFs. Les modifications structurales et l'auto-assemblage en nanostructures élargissent encore leur utilité, en améliorant l'absorption de lumière et la séparation des charges pour des applications en énergie solaire et en thérapie photodynamique. Leur double fonction de ligands et de systèmes photoactifs les rend utiles dans de nombreux domaines, de la biochimie à la science des matériaux et à la catalyse, où la fonctionnalisation contrôlée du noyau et des périphéries permet de concevoir des systèmes porphyriniques sur mesure.

Les métalloporphyrines (MP), inspirées des porphyrines naturelles comme le groupe hème, sont des catalyseurs moléculaires hautement modulables pour la réduction électrochimique et photocatalytique du CO_2 (CO_2RR).[9, 10] Les porphyrines de métaux de transition, en particulier celles incorporant Fe et Co, présentent une grande activité en formant des adduits CO_2–métal et en facilitant les étapes clés de transfert d'électrons. La performance catalytique dépend de la capacité du centre métallique à moduler les forces de liaison et la stabilisation des intermédiaires. Par exemple, CoTPP et FeTPP montrent une interaction optimale avec les intermédiaires de la réduction du CO_2, permettant en électrocatalyse une production efficace de CO avec une haute efficacité faradique et de faibles surtensions.[11] Des modifications structurales telles que la sulfonation ou la modification de la seconde sphère de coordination améliorent l'activité et la sélectivité, tandis que les systèmes hybrides comme les MP incorporées dans des MOF améliorent encore le transfert de charge et la capture du substrat.[12]

En photocatalyse, les MP fonctionnalisées associées à des photosensibilisateurs (PS) montrent de bonnes performances pour la réduction du CO_2 sous lumière visible. Les porphyrines de cobalt et de fer modifiées avec des groupes chargés ou des groupements formant des liaisons hydrogène présentent une excellente sélectivité pour la production de CO. Des études mécanistiques par DFT révèlent des voies contrastées: les porphyrines de fer reposent sur une chimie redox centrée sur le ligand, avec des transferts d'électrons couplés à des protons, menant majoritairement à la formation de monoxyde de carbone (CO).[13] En revanche, les porphyrines de cobalt suivent des mécanismes centrés sur le métal, permettant des réductions plus profondes du CO_2 jusqu'à l'acide formique et au méthane, avec une distribution de produits plus large.[14] Ces résultats soulignent combien la nature du métal et la conception du ligand influencent l'activation du CO_2 et la nature des produits orientant le développement de catalyseurs avancés et sélectifs pour la CO_2RR.

3.2.1 Contexte expérimental

Dans cette optique, le groupe de P. Mialane et A. Dolbecq a récemment développé une série de dyades Mo_6–MP, en particulier les Mo_6–porph(M) avec M = Fe, Co ou Zn (Figure 6). Ces dyades sont synthétisées par réaction du cluster Mo_6 avec des tétraphénylporphyrines (TPP) métallées, en présence de la dicyclohexylcarbodiimide (DCC) dans du DMSO à température élevée. Des analyses spectroscopiques (IR, UV-Vis, RMN ^{1}H, ESI-MS, ATG, EDX) ont confirmé la greffe covalente via des liaisons imido Mo≡N, la préservation des entités Mo_6 et MP, ainsi qu'un effet notable d'attraction électronique du POM sur la MP.

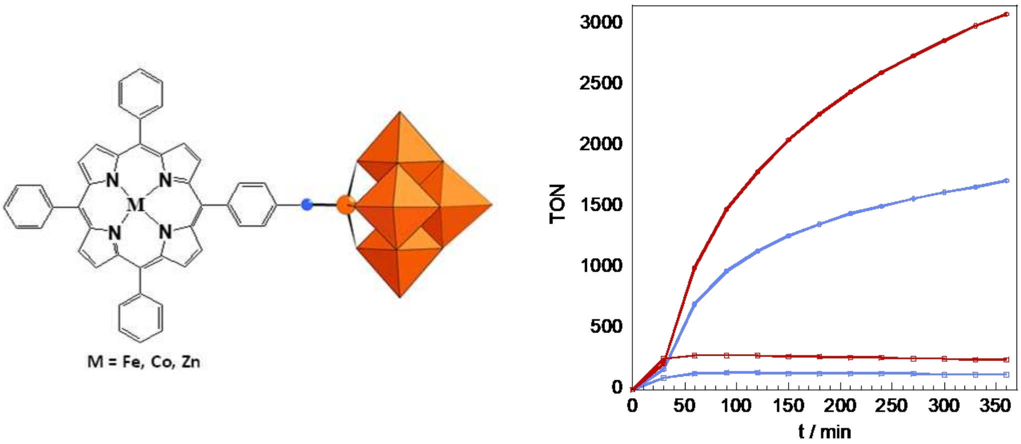


FIGURE 6: Gauche: Représentation des composés Mo_6-porph(M) (M = Fe, Co, Zn), sphère bleue: N, sphère orange: O, octaèdres orange: MoO_6. Droite: TONs de CO et de H_2 pour Mo_6-porph(Fe) (4 μM) et FeTPP (4 μM) en présence de $Ru(dmbpy)_3(PF_6)_2$ (500 μM) et de BIH (48 mM) dans des solutions DMF:TEOA (5:1), obtenus sous irradiation par LED bleue (10 mW cm^{-2}). Cercles rouges: production de CO par Mo_6-porph(Fe); carrés rouges: production de H_2 par Mo_6-porph(Fe); cercles bleus: production de CO par FeTPP; carrés bleus: production de H_2 par FeTPP.

Des tests photocatalytiques en milieu saturé en CO_2 (DMF:TEOA) avec le BIH comme donneur sacrificiel d'électrons et $[Ru(dmbpy)_3]^{2+}$ comme photosensibilisateur ont démontré que Mo_6–porph(Fe) surpasse à la fois le FeTPP libre et les analogues Zn/Co pour la production de CO, avec une suppression de l'évolution de H_2, ce qui indique une sélectivité et une efficacité accrues grâce à l'influence du POM. Des études électrochimiques par voltampérométrie cyclique (CV) ont révélé des déplacements dans les potentiels redox des entités Mo_6 et MP dans la dyade, confirmant une communication électronique mutuelle (Figure 7). Ces déplacements reflètent l'influence du POM doublement négatif et les effets attracteurs électroniques à travers le système conjugué.

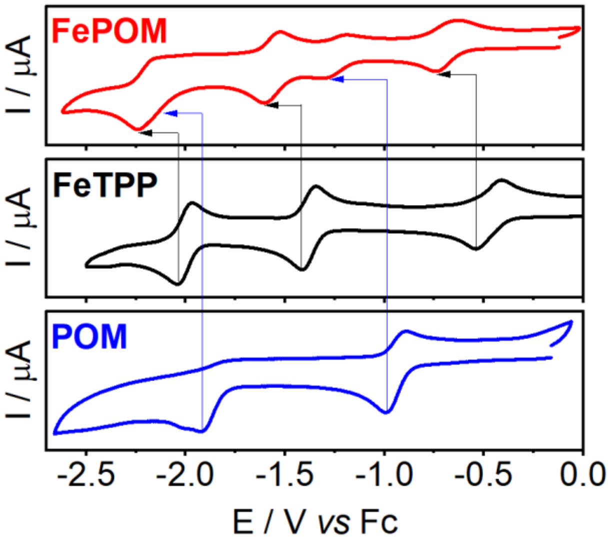


FIGURE 7: Cyclovoltammogramme (CV) de FePOM comparé aux complexes de référence FeTPP et POM [conditions: 0.5 mM de catalyseur, 0.1 M de $(TBA)PF_6$ dans du DMF saturé en Ar, vitesse de balayage de 100 mV s^{-1}]. CVs de Mo_6-porph(Fe) (ligne rouge), Fe(TPP) (ligne noire) et Mo_6 (ligne bleue).

Malgré ces éléments, les interactions électroniques entre le POM et la MP demeurent insuffisamment comprises en raison du recouvrement des événements redox. Pour y remédier, des calculs DFT ont été utilisés afin d'étudier la densité de spin, la distribution orbitale et la localisation des charges dans différents états redox des dyades. Cette analyse théorique vise à clarifier la structure électronique des intermédiaires réduits et à expliquer comment le POM module les propriétés redox et catalytiques de la MP. Les connaissances obtenues ici peuvent guider la conception rationnelle de catalyseurs hybrides modulaires.

3.2.2 Résultats

Caractérisation du complexe entièrement oxydé Mo_6-porph(Fe): Dans un premier temps, nous avons étudié la structure électronique du complexe Mo_6–porph(Fe) entièrement oxydé, dérivé de $[Fe^{III}TPP]Cl$, à l'aide de calculs DFT. Le système présente un centre Fe(III) penta-coordonné avec une géométrie pyramidale carrée, stable quel que soit l'état de spin. Parmi les diverses configurations (doublet, quartet et sextet), le quartet est identifié comme le plus stable, le sextet étant légèrement moins stable de 1.5 kcal mol^{-1}, tandis que le doublet est beaucoup moins favorable ($\Delta G = 14.6$ kcal mol^{-1}). La théorie du champ cristallin et les analyses NBO ont été utilisées pour interpréter les séparations des orbitales d et les interactions ligand–métal. Dans l'état quartet, les orbitales d_{xz} et d_{yz} sont presque dégénérées, et une liaison Fe–Cl plus faible se reflète par une longueur de liaison allongée. Le sextet présente une configuration d^5 avec un transfert de charge minimal. Bien que l'analyse NBO fournisse un aperçu qualitatif de la distribution électronique, elle repose sur des hypothèses de liaisons localisées et sur une description mono-référente issue de la DFT, ce qui limite sa fiabilité pour les systèmes présentant une forte délocalisation électronique ou un caractère multiréférent, où plusieurs états de spin ou configurations électroniques peuvent intervenir, comme c'est le cas des métalloporphyrines. Globalement, les résultats sont en accord avec les études antérieures [15] et suggèrent que le quartet est l'état de spin préféré dans les conditions modélisées, bien que des facteurs environnementaux puissent modifier cet équilibre.

Remplacement du ligand Cl^- dans la dyade par DMF: Une analyse basée sur la DFT a permis d'évaluer la substitution du chlorure par le DMF dans la dyade Mo_6-porph(Fe), révélant un coût en énergie libre endergonique de 3.9 kcal mol^{-1}. Étant donné la concentration bien plus élevée en DMF (12.91 M) comparée à celle de l'espèce Fe(III)-Cl (0.5 mM), l'échange de ligand est favorisé. Un modèle microcinétique, utilisant des constantes de vitesse dérivées de barrières de diffusion supposées et de différences d'énergie libre, a permis de prédire que l'équilibre est atteint en ~8 ns, avec 95 % des dyades coordonnées au DMF. Ce processus met en évidence le déplacement du Cl^- induit par le solvant. Une analyse de densité de spin indique en outre que la coordination du DMF induit des effets électroniques similaires à ceux observés dans le complexe lié au Cl^-.

Première réduction de Mo_6-porph(Fe): Deux régions principales du catalyseur sont susceptibles d'être réduites: le centre Fe de la porphyrine et les atomes de Mo du POM. L'analyse des orbitales moléculaires frontières du système coordonné au DMF montre que la LUMO est majoritairement localisée sur le centre Fe^{III}, indiquant que la première réduction implique la population électronique d'une orbitale de type d sur le Fe, générant une espèce Fe^{II}. Cette réduction augmente la densité électronique sur le site Fe de la porphyrine, affaiblissant la liaison Fe–DMF et facilitant la décoordination du ligand.

La modélisation cinétique, bien qu'intégrant la forte concentration en DMF (12.91 M), prédit que la décoordination reste thermodynamiquement favorable (ΔG = -5.2 kcal mol^{-1}) et se produit rapidement (~150 ps). L'équilibre FeII–DMF $\rightleftharpoons$ FeII + DMF a été modélisé en utilisant des constantes de vitesse directe et inverse dérivées d'hypothèses de contrôle par diffusion, confirmant une dissociation spontanée.

L'analyse des états de spin de l'espèce réduite Fe^{II} montre un état fondamental triplet, avec toutefois un état quintuplet proche en énergie (ΔG = 2.2 kcal mol^{-1}), tandis que le singulet est fortement déstabilisé. L'état triplet présente une densité de spin localisée sur le site Fe de la porphyrine, en accord avec le comportement typique des porphyrines de Fe^{II}. Ces résultats soulignent le rôle central du centre Fe dans l'induction de changements structuraux redox-dépendants permettant des transferts électroniques supplémentaires.

Deuxième réduction de Mo_6-porph(Fe): Pour élucider la structure électronique du complexe Mo_6-porph(Fe) réduit à deux électrons, nous avons examiné ses états de spin potentiels, en nous concentrant sur la localisation du second électron ajouté et l'influence du fragment lié de façon covalente, i.e. le polyoxométalate (POM). En partant de l'espèce triplet réduite à un électron, l'analyse de la LUMO révèle un caractère majoritairement localisé sur un atome Mo du POM, en particulier sur l'atome de Mo coordonné au ligand imido pontant. Cela suggère que le second électron se localise sur le POM, ce qui est confirmé par la β-SOMO de l'espèce réduite à deux électrons, fortement localisée sur le site Mo lié à l'imido.

Parmi les trois états de spin calculés pour le système réduit à deux électrons, le doublet est le plus stable, avec le centre Fe conservant une configuration semblable à celle de l'espèce triplet réduite à un électron. L'état sextuplet, seulement 1.6 kcal mol^{-1} plus haut, reflète un centre Fe à spin élevé avec un électron non apparié supplémentaire sur le POM. L'état quadruplet, en revanche, est nettement moins stable (10.9 kcal mol^{-1} au-dessus du doublet), probablement en raison d'un réarrangement défavorable des orbitales d centrées sur le Fe. Ces résultats mettent en évidence le rôle central des interactions Fe–POM dans la modulation des énergies d'état de spin et de la structure électronique de ce catalyseur hybride.

Troisième réduction de Mo_6-porph(Fe): Dans le complexe Mo_6-porph(Fe) réduit à trois électrons, l'analyse des états de spin révèle que le triplet est l'état fondamental, présentant un centre Fe à spin intermédiaire, un électron α localisé sur le POM (au niveau du Mo lié à l'imido), et un électron β dans une orbitale π^* centrée sur la porphyrine—ce qui conduit à une configuration globale triplet. L'état quintuplet, 2.9 kcal mol^{-1} plus haut, conserve la même distribution électronique mais avec un centre Fe^{II} à spin élevé, ce qui le rend moins favorable, tandis que l'état singulet, situé à 38.1 kcal mol^{-1} au-dessus du triplet, impose une configuration fermée avec deux électrons sur le POM. Ces résultats soulignent la flexibilité redox du système hybride, le troisième électron se localisant sur la porphyrine et permettant la délocalisation du spin et de la charge à travers les composants Fe, porphyrine et POM.

Quatrième réduction de Mo_6-porph(Fe): La quatrième réduction du système Mo_6-porph(Fe) se localise principalement sur le fragment polyoxométalate (POM), comme en témoigne le caractère fortement d du Mo dans la LUMO. Parmi les états de spin étudiés, doublet, quadruplet et sextuplet, le doublet apparaît comme l'état fondamental. Il est significativement plus stable (respectivement de 3.8 et 9.5 kcal mol^{-1}) que les états sextuplet et quadruplet. Dans la configuration doublet, l'analyse de la densité de spin montre que l'électron ajouté est délocalisé sur le POM, tandis qu'un électron β réside sur l'orbitale π^* de la porphyrine, et que le centre Fe conserve un caractère triplet local, similaire à celui de l'espèce réduite à trois électrons. En revanche, les états de spin plus élevés impliquent des schémas de couplage différents, avec des électrons non appariés supplémentaires sur le POM menant soit à un doublet déformé Fe–porphyrine (état quadruplet), soit à un électron localisé non couplé sur la porphyrine (état sextuplet). Les tentatives de convergence de configurations de spin encore plus élevées ont échoué. La HOMO de l'état doublet est localisé sur la porphyrine, ce qui indique son rôle potentiel dans le transfert électronique catalytique, notamment dans la réduction du CO_2. Ainsi, nous avons conclu que la configuration de spin doublet, combinant un triplet local sur Fe, un électron π^* sur la porphyrine et un singulet ouvert sur le POM, est la plus stable et la plus pertinente sur le plan catalytique.

Cinquième réduction de Mo_6-porph(Fe): Lors de la cinquième réduction du système Mo_6-porph(Fe), l'électron ajouté se localise principalement sur le ligand porphyrine, complétant les deux réductions précédentes centrées sur la porphyrine et les deux sur le POM. L'analyse des états de spin révèle que le quintuplet est l'état fondamental, avec le triplet et le singulet ouvert respectivement à 1.3 et 2.1 kcal mol^{-1}. Ce faible écart énergétique suggère que plusieurs états de spin peuvent être accessibles thermiquement. Dans l'état singulet, le centre Fe adopte une configuration triplet locale couplée à deux électrons non appariés sur la porphyrine, tandis que le POM reste apparié. L'état triplet présente une délocalisation du spin à travers les espèces Fe, porphyrine et POM. Le quintuplet, malgré un couplage plus faible Fe–porphyrine, est le plus stable, suggérant que la stabilisation est principalement gouvernée par la configuration des électrons d du Fe. Les états de spin plus élevés, comme le septuplet, sont fortement défavorisés (18.6 kcal mol^{-1}), probablement en raison d'une déstabilisation des orbitales d du Fe. Globalement, le système montre une flexibilité de spin, mais le Fe reste le déterminant central des énergies d'état de spin.

Prédiction des potentiels de réduction: Les potentiels de réduction de Mo_6-porph (Fe) ont été calculés par DFT à partir des énergies libres des configurations de spin les plus stables et référencés à l'électrode de ferrocène pour comparaison avec les valeurs expérimentales obtenues par voltamétrie cyclique (CV). Alors que les deux premières réductions montrent un bon accord avec les valeurs expérimentales, les réductions ultérieures présentent des divergences croissantes. Cela met en évidence les limites des méthodes à déterminant unique, telles que la DFT, et souligne le besoin potentiel de recourir à des méthodes quantiques multiréférence, comme le méthode CASSCF, [16, 17] malgré leur coût computationnel élevé. De plus, les effets de paires d'ions impliquant les contre-cations TBA, qui influencent le comportement rédox, n'ont pas pu être modélisés de manière fiable avec les méthodes DFT statiques, suggérant la nécessité de simulations dynamiques ou de techniques hybrides QM/MM pour des prédictions plus précises.

Validation expérimentale et proposition mécanistique: Parallèlement aux investigations théoriques, une étude expérimentale menée par le Prof. Leibl et Dr. Gotico (Université Paris-Saclay et CEA) a utilisé des techniques électrochimiques et spectroélectrochimiques pour explorer le comportement rédox du complexe Mo_6-porph (Fe). Leurs résultats ont corroboré l'identification théorique des sites rédox-actifs et ont mis en évidence des vagues catalytiques lors des troisième et quatrième événements rédox. Ces observations ont permis de proposer un mécanisme réactionnel détaillé, mettant en évidence que l'activation du CO_2 se produit non seulement à l'état réduit à trois électrons, mais de manière plus efficace aux états réduits à quatre et cinq électrons (I_3, I_4 et I_5 Figure 8). Une analyse plus approfondie de l'insertion du CO_2 sur ces intermédiaires a montré que les états de spin les plus stables sont le quintet pour I_3 et le quartet pour I_4. Dans les deux cas, un transfert monoélectronique (SET) de la porphyrine vers le ligand CO_2 est confirmé, accompagné de réarrangements structuraux dans la sphère de coordination du Fe lors de la liaison du CO_2. Sur le plan thermodynamique, les énergies libres calculées ($\Delta G = 19.6$ kcal mol^{-1}) sont comparables à celles du système de référence Fe-TPP, indiquant que la liaison du CO_2 en soi n'est pas l'étape limitante. Ce seraient plutôt les étapes de protonation ultérieures qui pourraient contrôler la vitesse globale de production de CO, une hypothèse actuellement en cours d'investigation.

Cette sous-section présente donc une étude computationnelle détaillée du comportement rédox et de la structure électronique de la dyade hybride Mo_6-porph(Fe), révélant un mécanisme de réduction du CO_2 séquentielle débutant au niveau du centre Fe de la porphyrine, suivi par la réduction du POM et le ligand porphyrine. En utilisant la DFT, l'analyse NBO, la cartographie de densité de spin et la modélisation microcinétique, cette étude met en évidence une localisation cohérente des électrons ajoutés et une préférence pour des états de spin intermédiaires tout au long des étapes de réduction, malgré les limites inhérentes à la DFT. Les résultats montrent que la distribution électronique et les énergies des états de spin sont fortement influencées par l'environnement des ligands et l'ordre de réduction, avec des étapes clés telles que l'échange de ligands qui s'avèrent thermodynamiquement et cinétiquement viables. La liaison imido pontant renforce le couplage électronique entre les unités porphyrine et POM, améliorant ainsi le pouvoir réducteur global du système. Ces résultats offrent une base mécanistique pour la conception de catalyseurs rédox porphyrine–POM aux propriétés modulables, notamment pour la réduction du CO_2.

4 Conclusions et perspectives

Cette thèse de doctorat présente une étude computationnelle de catalyseurs moléculaires pour l'activation de petites molécules chimiquement inertes, telles que H_2 et

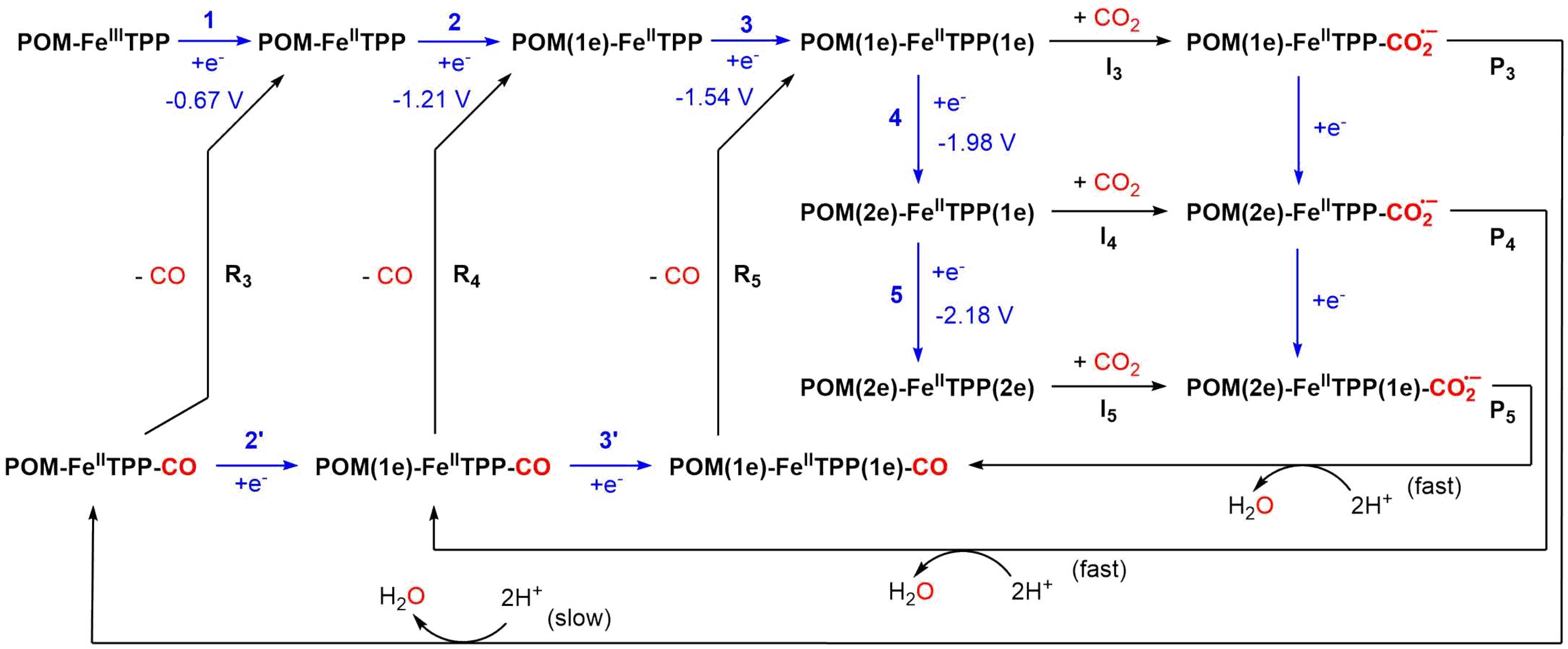


FIGURE 8: Mécanisme réactionnel proposé pour le Mo_6-porph(Fe) (POMFeTPP), incluant les cinq processus de réduction explorés ainsi que d'éventuelles voies concurrentes. Les processus de réduction sont indiqués en bleu. Les atomes de CO_2 sont surlignés en rouge. Les observations expérimentales sont indiquées entre crochets.

CO_2, avec un accent particulier sur les réactions de dissociation de H_2 et de réduction de CO_2. Les systèmes catalytiques étudiés dans cette thèse couvrent des architectures moléculaires allant de paires de Lewis frustrées (FLPs) moléculaires à base de bore et d'azote (Chapitre 2) à des systèmes hybrides plus complexes à base de polyoxométallates (POMs) (Chapitre 3). Notre investigation computationnelle s'est appuyée sur des calculs de la théorie de la fonctionnelle de la densité (DFT), des approches guidées par l'analyse dedonnées et l'entrainement de modèles, ainsi que sur des simulations avancées de dynamique moléculaire, afin de i) établir des relations structure–activité dans les FLPs B/N pour la dissociation de H_2 et l'hydrogénation du CO_2, et ii) fournir une compréhension mécanistique des systèmes catalytiques à base de POMs. Nos principaux objectifs étaient à la fois d'approfondir la compréhension théorique de ces systèmes et d'identifier, dans la mesure du possible, des principes directeurs pour la conception de catalyseurs efficaces. Nous mettons ci-dessous en évidence les principaux résultats liés à ces deux familles distinctes de catalyseurs, avant de proposer des perspectives possibles pour les travaux futurs dans ces domaines.

Dans le chapitre 2, nous avons développé des modèles multivariés pour prédire la thermodynamique (ΔG) et la cinétique ($\Delta G^{\ddagger}$) de la coupure de H_2 par des paires de Lewis frustrées (FLP) intramoléculaires à base de B/N, en intégrant des calculs DFT à une modélisation fondée sur les données. Un pipeline Python personnalisé a été mis en œuvre pour extraire systématiquement 38 descripteurs structuraux et électroniques à partir des fichiers de sortie de Gaussian. Des régressions linéaires multivariées (MLR) et des régressions par moindres carrés partiels (PLS), vérifiées par validation croisée, ont révélé que ΔG est principalement gouverné par des facteurs électroniques (notamment FEHA et FEPA) ainsi que par la distance $LA \cdots LB$, tandis que $\Delta G^{\ddagger}$ est majoritairement influencé par des descripteurs géométriques tels que d et λ, tous deux présentant des dépendances quadratiques optimales. Le modèle de ΔG a démontré une forte précision prédictive et une bonne généralisabilité au-delà des systèmes B/N, tandis que le modèle de $\Delta G^{\ddagger}$ reste fiable principalement au sein des FLP de type B/N. Des modèles préliminaires pour l'hydrogénation du CO_2, utilisant le même jeu de descripteurs, ont identifié FEHA comme une caractéristique prédictive clé, soulignant sa pertinence mécanistique plus large.

Ces résultats fournissent des indications exploitables sur les relations structure--réactivité et offrent des outils prédictifs pour orienter la conception rationnelle des FLP. Cependant, des voies compétitives, y compris des mécanismes de désactivation et la formation de produits alternatifs, peuvent affecter significativement les résultats réactionnels. La classification des états de transition (TS) en types « précoces » et « tardifs » suggère en outre une dépendance structurale dans l'activation de H_2 qui mérite une investigation plus approfondie. L'intégration de méthodes statistiques ou d'apprentissage automatique (ML) plus avancées, telles que «random forest» pourrait améliorer notre capacité à prédire les aspects cinétiques (états de transition) et la diversité mécanistique.

La performance des modèles fondés sur les données dépend également de manière critique de la diversité et de la taille du jeu de données. L'étude du chapitre 2, fondée sur 112 FLP intramoléculaires de type B/N, illustre que la génération de données reste laborieuse, non pas à cause du coût des calculs DFT, mais en raison de la nécessité d'une supervision manuelle. Pour garantir une applicabilité plus large et une meilleure généralisation des modèles, il est essentiel d'élargir le jeu de données pour inclure une gamme plus étendue de chimies de FLP (par exemple B/P, Al/N, FLP intermoléculaires). Les stratégies d'apprentissage actif offrent une voie prometteuse pour réduire l'effort manuel tout en permettant la construction efficace de jeux de données plus vastes et représentatifs.

Le Chapitre 3 s'intéresse aux catalyseurs à base de polyoxométallates (POM) pour la réduction du CO_2, en se concentrant sur deux systèmes représentatifs: un POM de type Keggin « sandwich » contenant quatre centres nickel (Ni_4), actif en conditions photocatalytiques, et un complexe hybride porphyrine–POM, Mo_6-porph-(Fe). Des études computationnelles démontrent que les POM agissent comme centres rédox actifs, facilitant des transformations multiélectroniques et multiprotoniques essentielles à la conversion du CO_2. Dans le système chargé Ni_4, l'activation du CO_2, rendue possible par photoaccumulation, progresse via un échange de ligands et une coupure de liaison C–O, les simulations par méta-dynamique chemin abordant la surestimation des effets de solvatation souvent rencontrée dans les calculs DFT statiques. Dans le complexe Mo_6-porph(Fe), le transfert électronique et la dynamique des ligands sont élucidés par des calculs DFT et une modélisation cinétique, mettant en lumière le rôle des équilibres d'états de spin et du couplage électronique inter-fragments. Ensemble, ces études soulignent la multifonctionnalité des POM et renforcent l'importance des techniques computationnelles avancées pour la conception de catalyseurs rédox actifs de nouvelle génération destinés à la réduction du CO_2.

Afin d'identifier les mécanismes les plus favorables, plusieurs voies réactionnelles possibles sont explorées. Cette tâche est particulièrement complexe dans des systèmes tels que les POM, où les exigences computationnelles sont élevées. Bien qu'une analyse plus exhaustive, incluant des processus secondaires comme la formation de H_2, serait idéale, le coût prohibitif des méthodes avancées (par exemple, post-Hartree–Fock) limite leur application courante. Par conséquent, le développement de stratégies de modélisation à la fois précises et abordables en termes de coût computationnel demeure un enjeu majeur.

Enfin, les effets supramoléculaires, incluant les interactions avec le solvant et les contre-ions, influencent fortement la réactivité dans des systèmes tels que Ni_4 et Mo_6–porph(Fe). Une modélisation explicite de l'environnement est importante, mais coûteuse. Les potentiels dérivés par apprentissage automatique et les outils, tels que DESC, [18] offrent des alternatives prometteuses aux approches DFT traditionnelles pour capturer ces effets environnementaux et effectuer des échantillonnages améliorés. Les travaux futurs incluront des simulations de dynamique moléculaire classique pour évaluer l'influence de l'environnement sur la réactivité des FLPs et des POMs.

Chapter 1

General introduction

1.1 Human impact on Earth carbon cycle

Humanity primarily relies on fossil fuels for energy, [19] depleting Earth's thermodynamic disequilibrium and contributing to greenhouse gas emissions and global warming.[1] The concept of diminishing use of fossil fuels has become central to sustainable development strategies. Even renewable energy use (e.g., biomass, wind, ...) can weaken natural Earth processes if overexploited, forming the basis of the "Limits to Growth" concept, [20] though human technology can enhance Earth's energy production capacity.

Carbon is the element with the most complex chemistry on Earth. It can form either organic and inorganic compounds through covalent bonds of shared electrons between atoms. Its ability to form large molecules and its unique properties make carbon-containing molecules the main chemical energy carrier of solar energy and the base element of the biosphere. Besides the central role of carbon in the biosphere, it also has played a fundamental role to access energy in human civilization, from the earliest use of fire to the present-day industrial era. Human activities have increasingly influenced the Earth's carbon cycle, transforming it from a relatively closed natural loop to a system heavily impacted by anthropogenic emissions (Figure 1.1). During primitive civilization, human activities remained largely integrated within the biosphere. Limited in number and with modest energy consumption based primarily on biomass, the carbon dioxide (CO_2) emitted from fire use could be readily buffered by Earth's natural carbon reservoirs.

However, the advent of agricultural civilization marked a turning point. The expansion of farming practices and rising populations led to widespread deforestation and intensive firewood use. The carbon that would have remained stored in terrestrial biomass for centuries was prematurely released into the atmosphere. In addition, the deforestation weakened the forest carbon pool's storage capacity, initiating a slow but noticeable rise in atmospheric CO_2 concentrations around 7000 years ago. The Industrial Revolution in the 18th century brought about a profound shift. Large-scale exploitation of fossil fuels released vast amounts of ancient organic carbon into the Earth system, changing its pre-human configuration. With the forest carbon sinks in decline, only approximately 31% of human-emitted CO_2 could

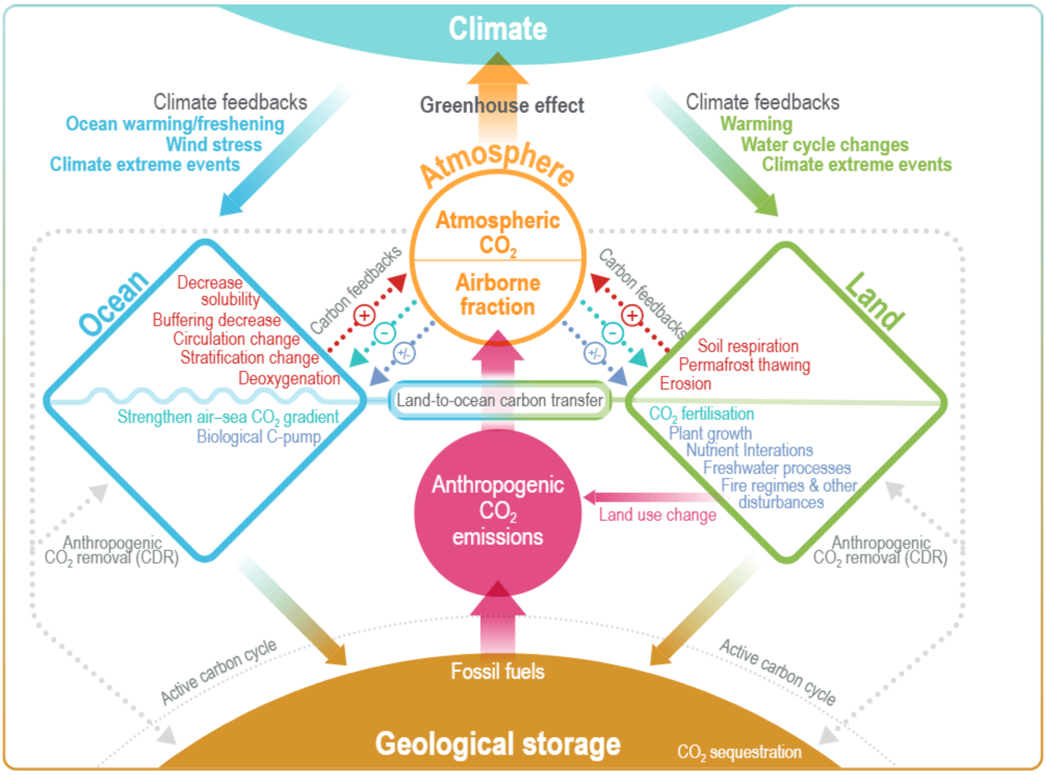


FIGURE 1.1: Key processes regulating CO$_2$ levels involve negative feedbacks distributing anthropogenic emissions among the ocean, land, and atmosphere. Positive feedbacks from ocean and land processes are driven by carbon–climate and carbon-concentration interactions. Human CO$_2$ removal and additional biosphere processes are also included. Reprinted from ref. 1.

be absorbed by terrestrial vegetation. A similar portion, 23%, was absorbed by the oceans, which simultaneously experienced acidification that reduced their carbon uptake efficiency. The remaining 46% of anthropogenic CO$_2$ emissions accumulated in the atmosphere, leads to a rapid increase in its concentration: from 280 ppm before 1760 to 415 ppm in 2021.[1] Over the past 70 years, the rate of atmospheric CO$_2$ increase has been approximately 100 times greater than during the end of the last glacial period. This unprecedented rise is now a primary driver of global warming and climate-related feedback loops (Figure 1.1).

1.2 The value of CO$_2$ in sustainable energy systems

To recover Earth system optimal conditions for human live while being compatible with sustainable human civilization, the later needs to gain more control of the planet, through technological advances. Solar energy is the principal source of low-entropy energy and the main cause of other energy sources on Earth (wind, biomass, etc.). Chemistry is crucial to develop technologies to store solar energy into long-time stable carbon forms of chemical energy, using the excess of CO_2 in the atmosphere, like biosphere do, but keeping or even improving the efficiency of the related processes.

Hydrogen is increasingly recognized as a clean, carbon-free energy carrier with significant potential for supporting sustainable energy systems.[21] It can be produced through water electrolysis or photocatalytic water splitting and integrated into energy systems consisting of three main components: hydrogen generation (e.g., electrolysers), storage, and energy conversion units such as fuel cells. However, storing hydrogen poses critical challenges: as a gas it has very low volumetric energy density at ambient conditions, requiring either high-pressure tanks, cryogenic liquid storage, or complex solid-state materials, all of which face issues such as hydrogen embrittlement of metal-based high pressure tanks, low weight-to-volume ratios, high energy requirements for compression or liquefaction, and poor reversibility or slow kinetics in metal/complex hydrides.[22]

The intermittent nature of most renewable energies makes finding alternatives for fossil fuels challenging, as our energy consumption is not intermittent. Storing excess energy produced during favourable weather periods is a key issue. Carbon, and specially CO_2, is ideal for chemical energy storage, overcoming H_2 main limitations as a source for generating liquid and gaseous products (such as CO, methane, methanol, etc.) with very stable and energetic covalent bonds, increasing their energy density (0.0107 MJ l^{-1} for hydrogen vs 22.2 - 34.6 MJ l^{-1} for liquid carbon based products).[23] Such chemical storage process can be observed in nature where plants, algae and bacteria can harvest light and use water during the light phase of photosynthesis to stock energy and release O_2, then use the stored energy to transform CO_2 into sugars, cellulose or starch during the dark phase of photosynthesis. Transforming CO_2 in the atmosphere into chemically valuable products is an attractive strategy: carbon dioxide is largely produced from the combustion of fossil fuels and using it as a C_1 carbon feedstock to store energy in chemical bonds would help establish a circular CO_2 economy.

1.3 CO_2 properties

CO_2 is a linear molecule with two polar C=O bonds despite its overall non-polar character ($^{\delta-}O = C^{\delta+} = O^{\delta-}$). This structure imparts a quadrupole moment,[24] making the carbon atom electrophilic and thus susceptible to nucleophilic attack, while the oxygen atoms can be targeted by electrophiles. Theoretically, CO_2's lowest unoccupied molecular orbital (LUMO) is a symmetric σ^* orbital of both C=O bonds, facilitating electron uptake, whereas its highest occupied molecular orbital (HOMO) consists on a π non-bonding orbital of lone pairs on oxygen.[25]

CO_2 activation involves occupying the σ^* LUMO weakening both C–O bonds, favouring its bending. The resulting radical anion $COO^{\bullet-}$ is bent and presents the extra electron delocalized across the O–C–O π system. Stabilization of $COO^{\bullet -}$, via protonation, further reduction or solvation clustering makes these processes highly sensitive to pH, solvent environment, and the number of electrons and protons involved. Proton-assisted, multielectron reduction steps are typically more favourable, particularly under acidic conditions and in solvents that stabilize charged intermediates, leading to more stable molecules, when compared to the radical anion, such as carbon monoxide (CO), formic acid (HCOOH), formaldehyde (H_2CO), methanol (CH_3OH), methane (CH_4) or products containing more than one carbon atoms (C_{2+}) (see Table 1.1).[26]

TABLE 1.1: Equations and potentials of CO_2 reduction (vs. NHE, at pH 7).

Entry	**Equation**	**E (V)**
1	$CO_2 + e^- \rightarrow COO^{\bullet -}$	-1.90
2	$CO_2 + 2H^+ + 2e^- \rightarrow HCOOH$	-0.61
3	$CO_2 + 2H^+ + 2e^- \rightarrow CO + H_2O$	-0.52
4	$CO_2 + 4H^+ + 4e^- \rightarrow H_2CO + H_2O$	-0.51
5	$CO_2 + 6H^+ + 6e^- \rightarrow CH_3OH + H_2O$	-0.38
6	$2CO_2 + 12H^+ + 12e^- \rightarrow C_2H_4 + 4H_2O$	-0.34
7	$CO_2 + 8H^+ + 8e^- \rightarrow CH_4 + 2H_2O$	-0.24

Compared to other common gases, such H_2, N_2 or O_2, CO_2 is more soluble in water, being in equilibrium with carbonic acid (H_2CO_3), the main reason for ocean acidification (Figure 1.1). CO_2 is even more soluble in organic solvents like acetone, acetonitrile, or toluene, an interesting property for design CO_2 reducing systems.[27] Still, the chemical reduction of CO_2 faces two main challenges: i) the difficulty of capturing CO_2 due to its dilution in ambient air, making it an expensive process; ii) the chemical reduction of CO_2 in energy-dense molecules (CH_4, ethanol, methanol, ethylene, etc.) is fundamentally challenging due to the high thermodynamic stability of CO_2 and the endergonic nature of its reduction. (see Table 1.1).

On top of the previously mentioned thermodynamics challenges underlying CO_2 reduction, kinetic factors must thus also be considered. Multiple electrons and protons are necessary to chemically reduce CO_2 inducing high energy barriers and kinetic limitations as more electrons and protons are needed to produce more hydrogenated C_1 molecules (CO, HCOOH HCOH, CH_3OH, CH_4). Thus, efficient and selective catalysts are needed to facilitate CO_2 chemical reduction process into valuable chemicals.

Current methods for catalytic CO_2 reduction can be grouped based on the type of energy used to drive the reaction: heat (thermal), light (photocatalysis), or electricity (electrocatalysis). In thermal catalysis, the kinetic energy of molecules is leveraged to overcome associated reaction barriers, whereas thermodynamic limitations are offset by strongly favourable side processes, such as acid-base events between the reaction products and external agents. Photocatalysis uses light, i.e. the energy of photons, to excite molecules into high-energy, highly unstable and thus reactive states. These excited states can trigger a series of electron and proton transfer processes that eventually lead to CO_2 being reduced into useful products. Photocatalytic systems often include photosensitizers (PSs) and sacrificial electron donors (SEDs).[28] These help collect and transfer electrons to catalysts, which are not capable of absorbing light themselves. Electrocatalysis relies on the application of an electric potential, which allows the catalyst to be reduced at the electrode-solvent interface. In analogy to photocatalytic approaches in the presence of SED, this forms highly reducing species, which can store electrons and spontaneously transfer them to CO_2 through a range of mechanisms, involving either the formation of catalyst-CO_2 bonds, or the reduction of protons to hydrides which are then used to hydrogenate CO_2.

Besides the development of molecular catalysts for CO_2 conversion, it is worth noting that there is also increasing interest in heterogeneous electrocatalytic and photocatalytic systems for CO_2 reduction. Despite these are generally rather active, more robust and recyclable, they typically offer less control over the product selectivity compared to molecular complexes, in which secondary coordination spheres can be tuned to modulate selectivity.

As explained in more detail below, this thesis presents computational studies on molecular catalysts for CO_2 reduction involving thermal catalysis (Chapter 2) and photo- and electro-catalysis (Chapter 3). Our goal is to better understand how these reactions work and identify the key factors that affect their performances, in order to help develop clear design guidelines and create new, improved catalysts.

1.4 CO_2 reduction by Photocatalysis

A photocatalytic system designed for CO_2 reduction comprises three key components: a catalyst, a light harvesting photosensitizer (PS), and a sacrificial electron donor (SED). These components work together to facilitate the conversion of CO_2 into useful products. Light-driven catalytic routes can be categorized into two types: oxidative quenching and reductive quenching, depending on the interaction between the photosensitizer and the electron acceptor. By understanding and manipulating the interplay between the catalyst, PS, and SED, researchers aim to develop efficient photocatalytic systems that can effectively use light energy to convert CO_2 into valuable products.

Photocatalysis present exclusive factors related to the usage of photons to generate electrons. Most of photocatalytic systems use PSs, systems that have strong visible-light absorption with enough long excited-state (PS^* in Figure 1.2) lifetimes in order to allow electron transfer processes of the excited PS to the catalyst. The most common PSs ase $[Ru(bpy)_3]^{2+}$ (Ru^{2+}) (bpy = 2,2′-bipyridine), $[Ir(ppy)_3]$ (ppy =

2-phenylpyridine), and $[Os(5dmb)_3]^{2+}$ (5dmb = 5,5′-dimethyl-2,2′-bipyridine).[29] Photo-degradation processes are still problematic, exhausting the concentration of the photoactive forms of PSs. Another fundamental agent in a photocatalytic system is the sacrificial electron donor (SED) (D in Figure 1.2), which electrons are consumed to regenerate the PS photoactive form. The most common SEDs consist in amines such triethylamine (TEA) and triethanolamine (TEOA), ascorbate ($AscH^-$), BNAH, or benzimidazole derivates (BIH, BI(OH)H). [30] Some SEDs also exhibit decomposition processes that affect the activity of the photocatalytic system. Photocatalytic systems, in contrast with electrochemical systems, provide electrons one-by-one via PS excitation, making one-electron intermediates more relevant. This slower process allows for for alternative reaction pathways that affects product distribution and activity deactivation.

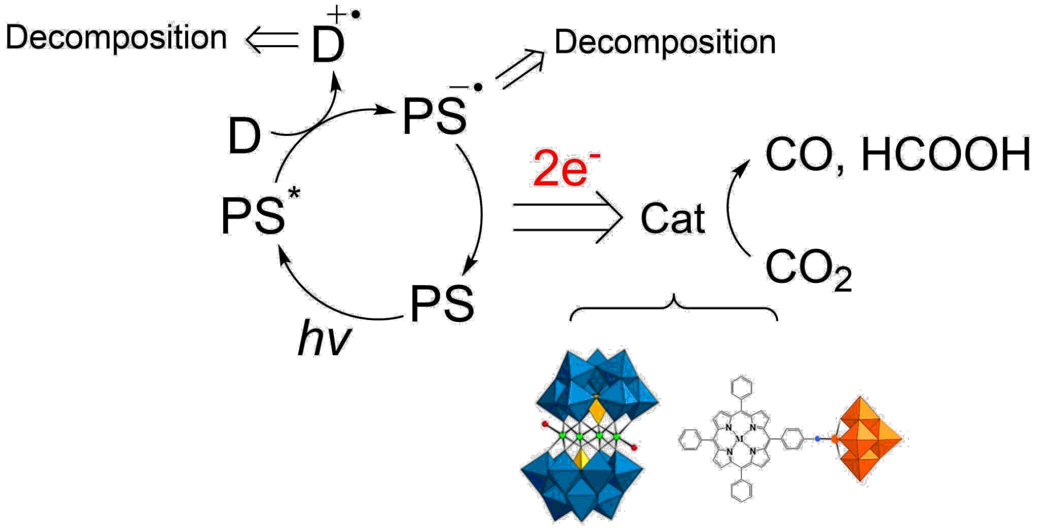


FIGURE 1.2: Schematic reaction mechanism for photocatalytic multi-electron CO_2-reduction reactions. Catalysts studied in this PhD thesis are illustrated.

In the search for effective catalysts for photochemical CO_2 reduction, organometallic complexes have emerged as leading candidates. Their well-defined structures, featuring a distinct active site and surrounding environment, make them ideal for in-depth mechanistic studies using various spectroscopic and computational chemistry methods. By modifying the ligands, researchers can fine-tune the redox properties of the metal centre, allowing for precise control over CO_2 reduction potential. These distinctive features have driven the development of a broad array of molecular catalysts aimed at converting CO_2 into "solar fuels".

In the context of this PhD work (Chapter 3), we have focused on two families of such molecular catalysts: i) polyoxometalates (POMs), which constitute a vast family of anionic metal-oxide clusters and that exhibit exceptional redox flexibility and electron storage capacity.[31–33] These molecular metal oxides, composed primarily of early transition metals in high oxidation states (W^{VI}, Mo^{VI}, V^{V}), can undergo reversible multi-electron reduction processes, making them ideal components for mediating complex multi-electron transformations such as CO_2 reduction.[3] Their modular structure and tunable properties have positioned them as versatile platforms for designing hybrid catalytic systems. ii) Porphyrin-based complexes are one of the most active and extensively studied families of molecular catalysts for photocatalytic CO_2 conversion. Porphyrin-based molecular catalysts incorporating transition metals, typically Fe and Co, have proved effective in reducing CO_2 to solar fuels. They have a high affinity for carbon dioxide thanks to the formation of a CO_2-metal bond, while in the excited state they can accept the electrons needed to reduce the substrate.[9, 10] As a seminal example, Bonin et al. [4] evidenced in 2014 that an iron porphyrinic compound was able to perform carbon dioxide reduction reactions (CO_2RR) in homogeneous conditions under visible-light irradiation in presence of an organic PS and a SED, with 100% selectivity for CO production and no deactivation of the catalyst after long-time periods. One of the fundamental advantages of metalloporphyrins (MPs) is that they can be custom functionalised.

In Chapter 3, we will focus on this topic investigating two different systems, a Nickel-based POM, $[Ni_4(H_2O)_2(PW_9O_{34})_2]^{10-}$ (**Ni**$_4$), which was studied under photocatalytic conditions for CO_2 conversion into CO, and a covalent iron-porphyrin-POM dyad, noted **Mo**$_6$-porph(Fe), which produces CO from CO_2 under both electro- and photocatalytic conditions.

1.5 CO_2 catalytic conversion with Frustrated Lewis Pairs

Frustrated Lewis Pairs (FLPs) are organic molecules made of a Lewis acid (LA) and a Lewis base (LB) that are frustrated due to geometric and/or steric constrains that preclude the formation of the LA–LB adduct. This "frustration" creates a reactive space between the LA and LB centres where small and stable molecules such as H_2 and CO_2 can be activated and transformed. These FLPs architectures are either intra- or intermolecular, and have emerged as promising metal-free catalysts for the heterolytic cleavage of H_2 and the subsequent hydrogenation of CO_2.[34, 35]. FLPs, often called a "Holy Grail" of catalysis, have opened up a new path for more sustainable and eco-friendly chemistry. In particular, they allow the transition metal-free activation of small molecules, demonstrating significant potential in hydrogen activation and double bond reduction in fine chemicals, including CO_2 hydrogenation. This unique reactivity positions FLPs as a promising greener and metal-free alternative to traditional, often costly and low abundance transition metals-based complexes. The splitting of H_2 by FLPs was introduced in the seminal work by Stephan and co-workers, [2] who reported the reversible cleavage of H_2 using a boron-based LA and phosphorus-based LB molecule. Since then, FLP chemistry has expanded to include various LA/LB combinations (B/N, Al/P) across inter- and intramolecular architectures. The heterolytic splitting yields a zwitterionic species where the proton

(H^+) binds to the LB site and the hydride (H^-) binds to the LA centre. These zwitterionic products can transfer both hydride and proton to unsaturated molecules including CO_2, [36] underscoring FLPs' potential as metal-free hydrogenation catalysts. The reversibility of H_2 splitting has been shown to depend on specific molecular characteristics, such the electronic and geometric nature of LA's substituents, [37–41] offering promising applications in chemical energy storage of H_2. Computational chemistry has been instrumental in advancing the understanding of FLP-mediated H_2 splitting and CO_2 reduction.[5] While the LA and LB strengths have been recognized as key factors governing H_2 splitting, their relative contributions, as well as the possible roles of additional parameters, is still a matter of intense debate, with no clear or quantitative consensus. Hence, although important reactivity trends have been identified, comprehensive and quantitative structure–activity relationships across a broad range of FLP families with diverse structures and properties are still lacking. This lack of general insight hinders both a complete understanding of the factors governing FLP-promoted H_2 splitting and the development of general design rules. Chapter 2 aims at bridging this gap through a data-driven approach of FLP's reactivity for H_2 splitting.

1.6 Thesis Objectives

This PhD thesis applies a range of computational tools and techniques to gain unprecedented insights into various CO_2 reduction processes. The ultimate goal is to push the boundaries of current knowledge and inspire the development of more efficient catalytic systems by achieving a deep understanding of the reaction mechanisms and by revealing structure–activity relationships that govern catalytic efficiency. This thesis presents a comprehensive computational study aimed at elucidating the fundamental mechanisms underlying CO_2 reduction by both FLPs and POMs, with particular emphasis on establishing quantitative structure-activity relationships that can guide catalyst optimization. Through the application of state-of-the-art Density Functional Theory (DFT) calculations, molecular dynamics simulations, and statistic and data-driven approaches, we seek to bridge the gap between molecular-level understanding and macroscopic catalytic performance.

Chapter 2 presents the development of predictive models for H_2 splitting by B/N Frustrated Lewis Pairs (FLPs), with the main goal of establishing a quantitative framework that links simple molecular descriptors to FLP reactivity. The chapter pursues several specific objectives: i) constructing multivariate regression models to uncover chemically meaningful descriptors and general structure–activity relationships; ii) identifying the molecular features that most strongly govern thermodynamic and kinetic parameters underlying this reactivity; and iii) generating predictive models for both the reaction free energies (ΔG) and activation barriers ($\Delta G^{\ddagger}$) associated with H_2 cleavage and subsequent CO_2 reduction. Collectively, these efforts aim to deliver mechanistic insights and quantitative tools that can inform the rational design of FLP systems with enhanced performance in H_2 activation and CO_2 hydrogenation.

Chapter 3 then explores the role of POMs as electron mediators in photocatalytic CO_2 reduction, focusing on two molecular systems: $[Ni_4(H_2O)_2(PW_9O_{34})_2]^{10-}$ ($\mathbf{Ni_4}$) and $\mathbf{Mo_6}$-porph(Fe). Using a combination of static DFT calculations, microkinetic modelling and advanced molecular dynamics simulations, we investigate how POMs serve as dynamic electron reservoirs to support efficient multi-electron CO_2 reduction under mild conditions. For the $\mathbf{Ni_4}$ catalyst, we aim to elucidate detailed mechanistic pathways, examine the impact of protonation states and redox properties, and identify key factors governing catalytic performance under photocatalytic conditions. In the $\mathbf{Mo_6}$-porph(Fe) hybrid catalyst, we characterize the electronic structure and sequential redox behaviour of the catalyst, explore the nature of electronic communication between porphyrin and POM moieties, and determine how the POM scaffold modulates the catalytic properties of the porphyrin unit.

This PhD thesis applies a range of computational tools and techniques to gain unprecedented insights into various CO_2 reduction processes, developing comprehensive understanding that can guide rational catalyst design and inspire development of more efficient catalytic systems for sustainable CO_2 conversion.

Chapter 2

Frustrated Lewis Pairs for H_2 splitting and CO_2 hydrogenation

2.1 Introduction

Frustrated Lewis Pairs (FLPs) are molecular architectures comprising a Lewis acid (LA) and a Lewis base (LB) whose direct association into a classical Lewis adduct (Figure 2.1a) is hindered or prevented. This frustration arises from steric hindrance between bulky substituents or geometric constraints that preclude close interaction of the LA and LB centres. Depending on the spatial arrangement, FLPs can be either intramolecular (Figure 2.1b), in which both reactive centres are tethered within a single molecular scaffold, or intermolecular (Figure 2.1c), where LA and LB reside on distinct molecules.

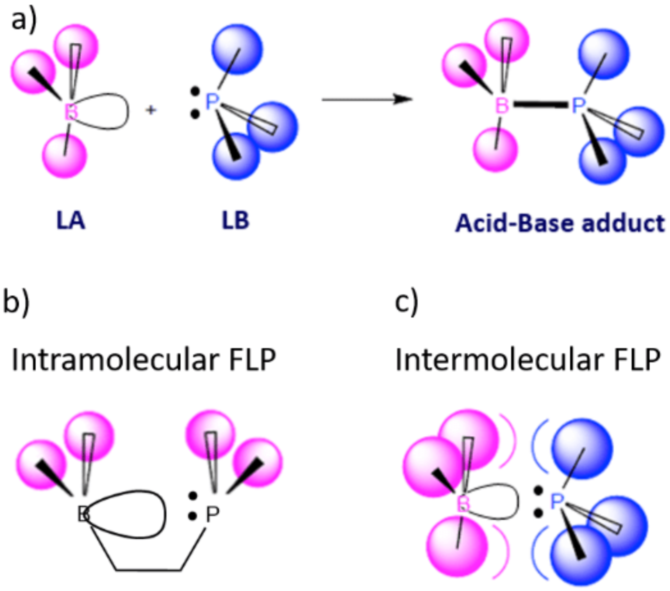


FIGURE 2.1: Representation of a) an acid-base adduct formation, b) an intramolecular FLP and c) an intermolecular FLP.

The region of space situated between the Lewis acidic and basic centres in an FLP serves as a reactive pocket, enabling the activation of otherwise inert molecules such as H_2, CO_2, or N_2.[34, 35] This unique reactivity profile underscores the promise of FLPs as metal-free hydrogenation catalysts and as tools for molecular activation in synthetic chemistry. The discovery of FLPs capable of splitting dihydrogen (H_2) heterolytically without the assistance of transition metals marked a milestone in the development of sustainable chemical transformations. The splitting of H_2 by FLPs was introduced in the seminal work by Stephan and co-workers, who reported the reversible cleavage of H_2 using a molecule (Figure 2.2) that features a boron-based LA and a phosphorus-based LB.[2] Since then, this field exhibits an ever-expanding range of molecular LBs and LAs partners in intra- or inter-molecular fashions.[42, 43] They demonstrated their large potential for the activation of H_2 and the subsequent reduction of double bonds in fine chemicals,[44, 45] including CO_2 hydrogenation.[46–48] across a broad range of molecular architectures, spanning both inter- and intramolecular scenarios.

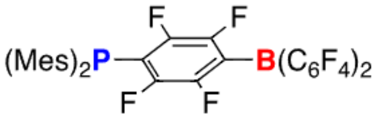


FIGURE 2.2: System reported by Stephan *et. al.* named p-(Mes_2P)-$C_6F_4(B(C_6F_5)_2)$ (Mes = 2,4,6-trimethylphenyl).

2.1.1 Molecular FLPs for CO_2 hydrogenation

Our project fits in the specific context of the computational investigation of molecular FLPs for CO_2 hydrogenation. The hydrogenation of CO_2 by FLPs has often been shown to require the initial generation of an active hydrogenating species. FLPs can first heterolytically split H_2 into a hydride (H^-) and a proton (H^+), which are accommodated on the LA and LB centres forming the FLP, respectively, and can then be transferred to CO_2 in a second step to release formic acid, as outlined in path (i) in Figure 2.3. However, competition between H_2 and CO_2 for binding to FLPs complicates the reaction mechanism. FLPs may often bind CO_2 more strongly than H_2 (path (i) vs (ii) in Figure 2.3), potentially inhibiting the first activation step unless alternative pathways are available.[49]

Initial computational studies examined how FLPs can mediate both H_2 splitting and CO_2 hydrogenation. Wen et al. studied in 2013 an intermolecular FLP using 2,2,6,6-tetramethylpiperidine (TMP) as LB and $B(C_6F_5)_3$ as LA (Figure 2.4a), showing that the ion pair $[TMPH]^+[HB(C_6F_5)_3]^-$ could transfer H^+/H^- to CO_2 through a concerted transition state, yielding formic acid, which is deprotonated to form a formato-borate.[50] Fontaine and Stephan analysed in 2015 B/N intramolecular FLPs (Figure 2.4b) computationally and showed that they undergo H_2 activation and splitting followed by CO_2 hydrogenation via low-energy concerted pathways.[47] They also identified feasible protodeborylation processes after H_2 activation for such systems and calculated feasible energy barriers for subsequent CO_2 reduction. Jiang

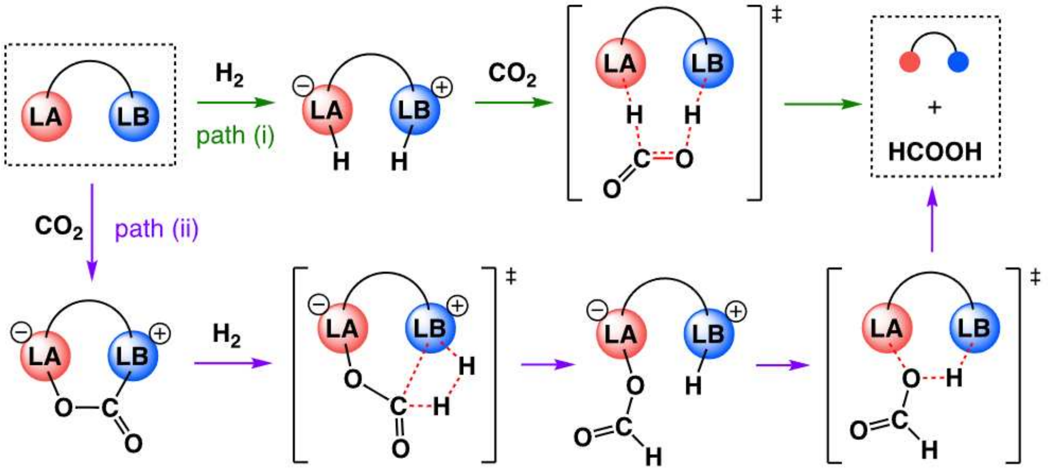


FIGURE 2.3: Proposed pathways for CO_2 hydrogenation with intramolecular FLPs.

et al. further proposed in 2018 two competing mechanisms for CO_2 hydrogenation (Figure 2.3): the standard sequential path (i), involving H_2 activation first, and a reverse path (ii), where CO_2 first binds to the FLP.[51] While path (i) was generally favoured for the studied systems (Figure 2.4c), some specific FLPs preferred path (ii) such as systems **i** and **iv**. Ghara and Chattaraj later considered simultaneous activation of H_2 and CO_2, but this concerted mechanism was associated with high energy barriers (37–50 kcal mol^{-1}), making it kinetically less favourable.[52] These characteristics make the concerted H_2 + CO_2 activation mechanisms less competitive than the stepwise pathway depicted in Figure 2.3, path (i), as later confirmed by the same authors (Figure 2.4d).[53]

Designing effective FLPs for CO_2 hydrogenation thus requires optimizing both H_2 activation and hydride/proton transfer to CO_2. Liu *et al.* examined five B/N FLPs (Figure 2.4e), combinations of $B(C_6F_5)_3$ and p-$B(C_6F_2H_3)_3$ LAs with TMP, N-benzyl-tert-butylamine (btam) and 2,6-lutidine (Lut), by means of DFT calculations. They found that while stronger acid/base pairs reduced the barrier for H_2 splitting, they often raised the barrier for the CO_2 hydrogenation step due to overly stable ion pairs.[54] Overall, the studies discussed above indicate that optimal FLPs for CO_2 hydrogenation must achieve a delicate balance between efficient H_2 activation and effective hydride/proton transfer from the zwitterionic LB^+H/LAH^- intermediate to CO_2. This balance is crucial to ensure that both steps proceed with accessible free-energy barriers. FLPs that are too weak may face slow H_2 splitting kinetics (Figure 2.5, purple arrow), whereas overly strong FLPs can make the hydride/proton transfer to CO_2 the rate-limiting step (Figure 2.5, red arrow), a balanced pathway towards CO_2 hydrogenation into formate being illustrated in the green path (Figure 2.5).

Corminbœuf and co-workers recently conducted in 2022 a comprehensive computational investigation of intermolecular B/N and B/P frustrated Lewis pairs (FLPs) for the hydrogenation of CO_2 into formate [6]. Their study involved the virtual

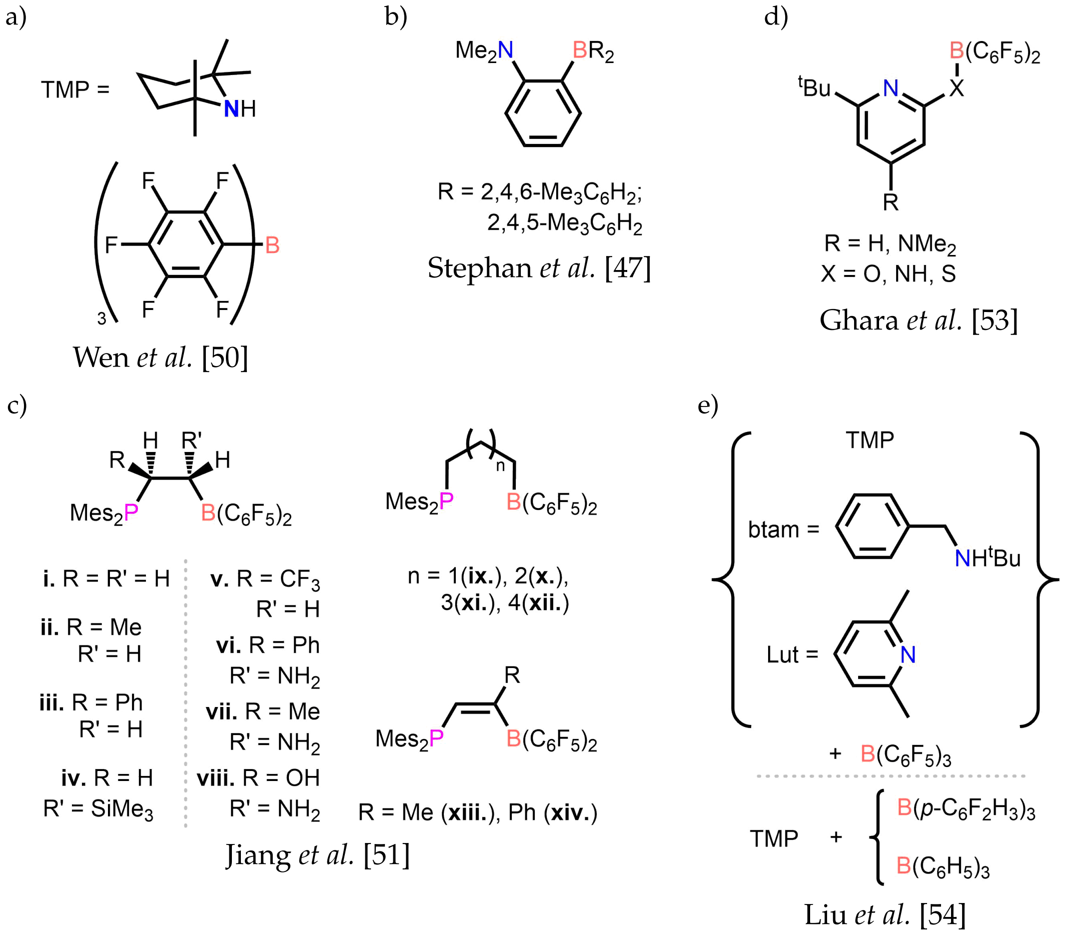


FIGURE 2.4: Compilation of computationally studied FLP structures for CO_2 hydrogenation with H_2.

screening of 3840 FLP candidates constructed by combining a library of 60 nitrogen- and phosphorus-based LBs with 64 triarylborane LAs. Each combination was assessed along a proposed catalytic cycle, computing the relative free energies of key intermediates and transition states (TSs), assuming the LA at catalytic concentrations and the LB in excess to drive product formation.

Two sub-libraries were extracted from this large dataset by fixing one Lewis partner and varying the other, enabling an independent analysis of LA and LB effects. To correlate structure with activity, linear free-energy scaling relationships were established using two thermodynamic descriptors: the free energy of proton attachment (FEPA) (Eq. 2.1) to the LBs and the free energy of hydride attachment (FEHA) (Eq. 2.2) to the LAs, that serve as direct indicators of basicity and acidity, respectively, and are particularly relevant to H_2 activation in FLP systems.

$$\mathrm{LB} + \mathrm{NH}_4^+ \rightarrow [\mathrm{LBH}]^+ + \mathrm{NH}_3 \tag{2.1}$$

$$\mathrm{LA} + \mathrm{H}^- \rightarrow [\mathrm{LAH}]^- \tag{2.2}$$

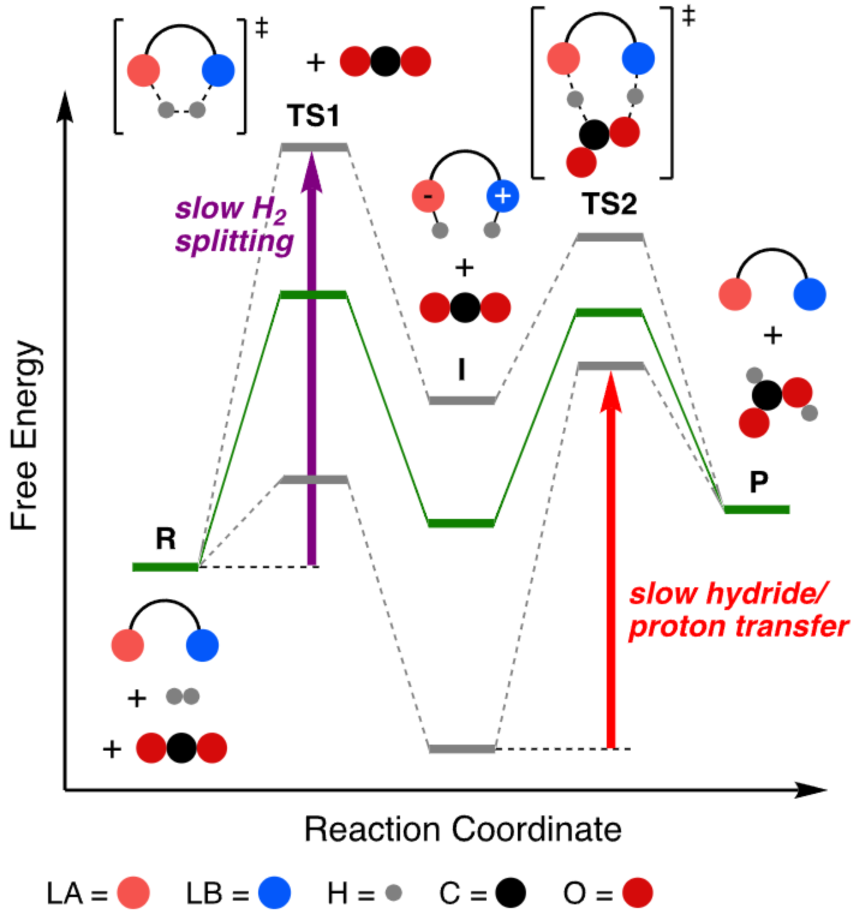


FIGURE 2.5: Schematic free-energy profile for the two-step hydrogenation of CO_2 catalyzed by FLPs. The green profile illustrates an optimal situation where both H_2 splitting are kinetically accessible. Conversely, grey-dashed profiles represent sub-optimal situations limited either by a slow H_2 splitting step, for which the barrier is highlighted with a purple arrow; or by a slow hydride/proton transfer step, in which the associated barrier is highlighted with a red arrow.

Volcano plots relating turn-over frequency (TOF) to FEHA and FEPA highlighted optimal and suboptimal reactivity regions, depending on the nature of the rate-determining step. For example, with a weak LA (BMe_3), efficient H_2 activation was only achieved with strong bases like 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU), while strong LAs paired with weak bases (e.g., pyridine) led to over-stabilization of the boroformate intermediate, hindering catalytic turnover. By plotting FEHA and FEPA against TOF (Figure 2.6), a clear optimal red-colored catalytic performance zone emerged, surrounded by regions of poor activity caused by either overly stabilized intermediates or high-energy TSs. This seminal work not only demonstrates the importance of an acid-base balance in designing effective FLPs for CO_2 hydrogenation but also establishes a powerful framework for predicting an FLP's reactivity using simple thermodynamic descriptors.

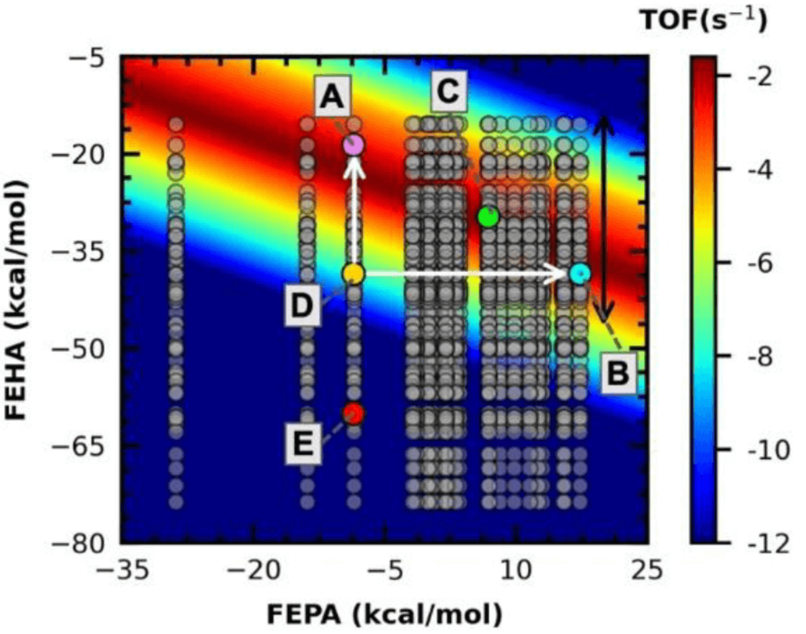


FIGURE 2.6: Map correlating the TOF (log scale) for CO_2 hydrogenation into formate to FEPA and FEHA descriptors for the 1664 B/N intermolecular FLPs studied by Corminbœuf and coworkers. The highlighted systems correspond to **A**: $B(C_6Me_4COOH)_3$/DBU, **B**: $B(C_6(OMe)_2F_2COOH)_3$/pyrazine, **C**: $B(C_6H_2Me_2COOH)_3$/4-OMe-py, **D**: $B(C_6(OMe)_2F_2COOH)_3$/DBU and **E**: $B(C_6F_4COOH)_3$/DBU. Reprinted from reference 6.

Computational studies have shown that FLPs can engage in side reactions that may hinder CO_2 hydrogenation. A key issue is the competition between H_2 and CO_2 for FLP binding sites, as FLPs may readily form stable LB–C(O)O–LA adducts. If CO_2 binds faster and more strongly than H_2, it can inhibit catalysis by raising the effective barrier for H_2 splitting ($\Delta G^{\ddagger}_{overall,1}$, Figure 2.7, light-blue lines). In some cases, however, such as systems following path (ii) (Figure 2.3), CO_2-bound FLPs may still proceed to H_2 activation via an alternative mechanism, although the factors dictating this path remain unclear. To enhance selectivity for H_2 over CO_2 binding (assuming path (i), Figure 2.3), studies have explored structural modifications of FLPs. Tuning LA/LB substituents affects both gas similarly, but shortening the LA$\cdots$LB distance selectively strengthens H_2 binding.[55] Electron-conjugation effects also play a role: increased aromaticity lowers barriers associated to CO_2 binding,[56] while reduced anti-aromaticity aids H_2 activation.[57] These insights offer strategies for designing more selective and efficient FLPs. Product inhibition is another challenge, as formic acid can form stable adducts with FLPs (Figure 2.7, red lines), potentially stalling the cycle. FLP-HCOOH complexes have been identified as the most stable, compared to the separate species, in several systems.[6] Using bulky bases like DBU can trap formic acid as [Hbase][HCOO] salts, avoiding FLP deactivation. Additionally, moisture sensitivity and protodeborylation [47] can degrade FLPs or alter their reactivity, forming H_2O–LA adducts in the presence of H_2O or braking LA–substituent

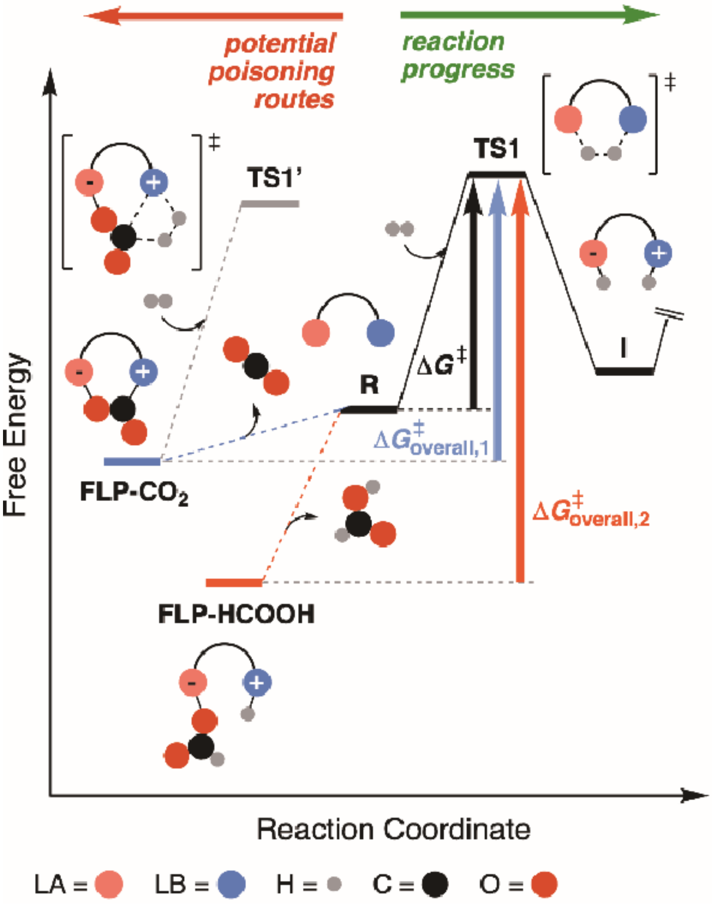


FIGURE 2.7: Schematic free energy profile for the early steps of CO_2 hydrogenation, highlighting the impact of possible stabilizing interactions between the bare FLP and CO_2 (light-blue dashed lines) or formic acid (red dashed lines) on the kinetics of H_2 splitting. Black, light-blue and red arrows represent the individual free-energy barrier for the H_2 splitting step ($\Delta G^{\ddagger}$), and overall barriers for the same process measured from stable FLP–CO_2 ($\Delta G^{\ddagger}_{overall,1}$) and FLP–HCOOH ($\Delta G^{\ddagger}_{overall,2}$) complexes, respectively. TS1′ accounts for possible hydrogenation of CO_2 through path (ii) of Figure 2.3.

bonds in the case of protodeborylation. DFT studies suggest that increasing steric bulkyness around boron sites can suppress such degradation, highlighting the need to account for these effects in the catalyst's design.

Overall, efficient FLP catalysts for CO_2 hydrogenation require a careful balance between H_2 activation, effective H^+/H^- transfer to CO_2, and the mitigation of poisoning pathways. Achieving this involves tuning the relative affinities of FLPs for CO_2 and H_2, as well as ensuring that FLP-HCOOH complexes can readily release the product to sustain catalytic turnover.

2.1.2 Molecular FLPs for H_2 heterolytic cleavage

A key prerequisite for catalytic activity is the ability of FLPs to split H_2 under mild conditions. Since this step initiates the entire CO_2 hydrogenation cycle, understanding the mechanistic and structural factors that govern H_2 splitting is essential. These aspects, such as the influence of the geometry of LA–LB pair, its electronic properties, and the transition state stabilization, are the focus of this chapter, which aims at providing a detailed examination of the H_2 splitting step and its central role in determining the FLP's reactivity and efficiency. As illustrated in Figure 2.8a, the heterolytic splitting of H_2 by FLPs involves a zwitterionic species, where the resulting proton (H^+) is accommodated on the LB site while the hydride (H^-), which retains both electrons from the original σ(H–H) bond, binds to the electron-deficient LA centre. These zwitterionic products have demonstrated the ability to transfer both the hydride and the proton to unsaturated molecules such as alkenes, imines, enamines, or even CO_2, [36] underscoring the potential of FLPs as powerful, metal-free hydrogenation catalysts. Moreover, the reversibility of H_2 splitting has been shown to strongly depend on LA and LB substituents or their interatomic distance of each FLP.[37–41]

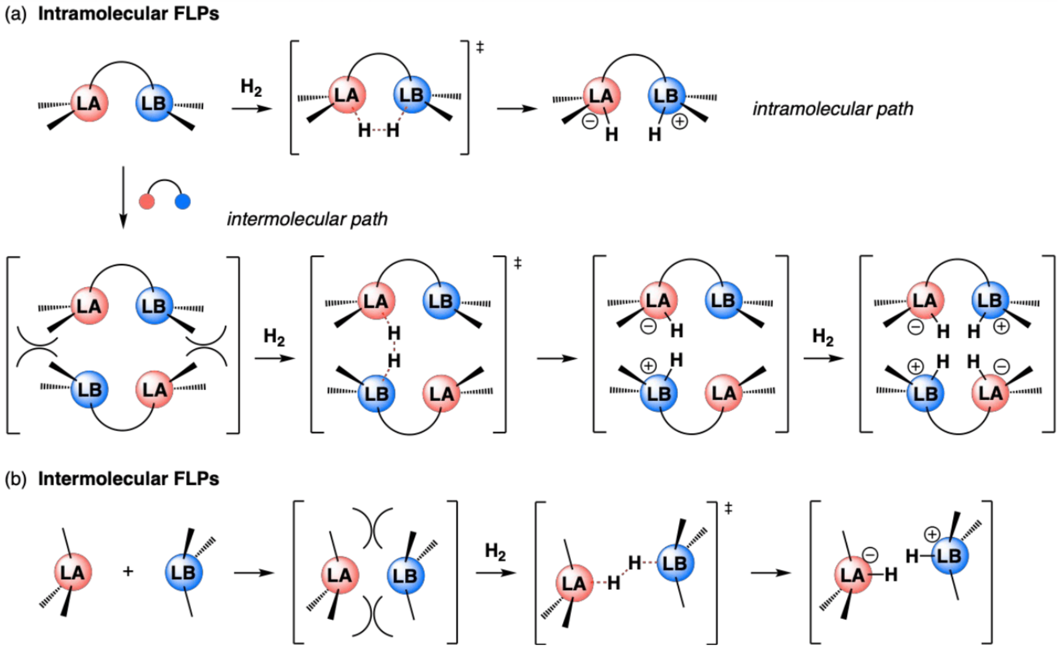


FIGURE 2.8: Overview of the main reaction mechanisms characterized by DFT calculations to govern the H_2 splitting by (a) intramolecular and (b) intermolecular FLPs.

Mechanistic origins and atomistic characterization

The very first example of the reversible splitting of H_2 by an FLP system was reported by Stephan and co-workers in 2006,[2] who demonstrated that a sterically encumbered phosphine-borane pair, *p*-(Mes_2P)-C_6F_4-($B(C_6F_5)_2$) (Mes = $C_6H_2(CH_3)_3$, containing both boron LA and phosphorous LB centres, could reversibly split H_2 at room temperature. The resulting product, $Mes_2P^+H(C_6F_4)B^-H(C_6F_5)_2$, a zwitterionic phosphonium borate, releases H_2 when heated to 150 °C, regenerating the original FLP. This discovery was particularly notable because the $P\cdots B$ distance in the molecule is about 6.2 Å, far too long to permit the direct and concerted heterolytic H_2 cleavage on P and B centres simultaneously. To rationalize this behaviour, Stephan proposed two possible multi-step intramolecular pathways: the hydride migration mechanism and the proton migration mechanism which are illustrated in Figure 2.9. Both involve initial H_2 addition either to the P–C or to the B–C bond, followed by

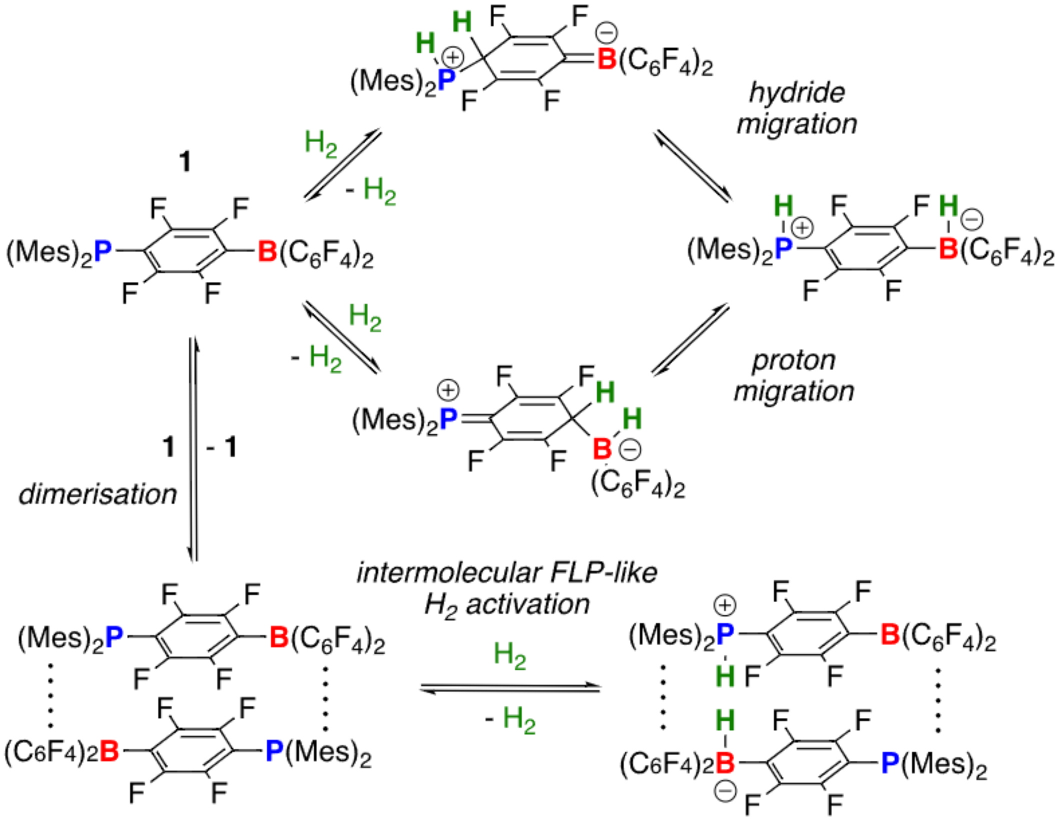


FIGURE 2.9: Overview of the main reaction mechanisms characterized by DFT calculations to govern the H_2 splitting by (a) intramolecular and (b) intermolecular.

migration of H^- or H^+ to the respective LA or LB sites. These findings motivated the first computational exploration of the H_2 splitting mechanism by Guo and Li in 2008.[58] Using DFT calculations, they found that both the hydride and proton migration mechanisms carried prohibitively high free-energy barriers: 69.1 and 54.7 kcal mol^{-1}, respectively. Instead, they proposed a third pathway involving the formation of a non-covalent head-to-tail dimer between two FLP molecules (see Figures 2.8 and 2.9).

This intermolecular dimer allows the formation of a shorter effective P$\cdots$B distance of 5.5 Å, enabling the concerted activation of H_2 by the P and B centres of two different molecules through a significantly lower free-energy barrier of 33.7 kcal mol^{-1} when compared to the hydride and proton-migration pathways. This revised model suggests that even for apparently intramolecular FLPs like Mes_2P-$(C_6F_4)B(C_6F_5)_2$, the operative mechanism for H_2 activation might involve dimerisation and an intermolecular mechanism. Further studies on related intramolecular P/B FLPs reinforced this type of conclusions. Spies *et al.* investigated systems like $Mes_2PCHRCH_2B(C_6F_5)_2$ (R = H, Me), where H_2 splitting occurs readily, whereas bulkier analogues such as $Mes_2PCH(Ph)CH_2$- $B(C_6F_5)_2$ showed no reactivity.[39, 59, 60] For the ones splitting H_2, early computational studies showed that the most stable "closed" conformer (Figure 2.10) displaying a weak P–B bond can be easily opened (with an energy cost of only ca. 7.0 kcal mol^{-1}) allowing for the subsequent activation of H_2.[40, 59, 61, 62] However, some years later, Vankova et al. revisited the reactivity of these systems using DFT calculations and found that FLPs capable of dimerizing into a reactive, stacked configuration could activate H_2, while those that are too sterically hindered could not.[63] These findings revealed that the previously assumed mechanistic scenario, in which a single FLP molecule activates H_2 in a concerted, intramolecular fashion (Figure 2.8a), cannot account for the observed variations in reactivity as a function of substituent bulkiness. In contrast, the authors proposed an alternative, intermolecular mechanism (also shown in Figure 2.8a), where H_2 is split by a stacked dimer of "open" FLP conformers arranged in a head-to-tail configuration, similar to what Guo and Li proposed for the $Mes_2P(C_6F_4)B(C_6F_5)_2$ FLP.[58] This mechanism was fully characterized and, unlike the intramolecular pathway, explains the inertness of sterically demanding FLPs, which cannot dimerize due to steric clashes that prevent formation of a suitable reactive pocket. Still, this intermolecular mechanistic picture cannot be generalized to all intramolecular or linked FLPs.

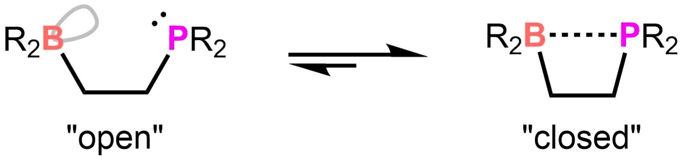


FIGURE 2.10: Open-closed displaced equilibrium of B/P FLP.

More recent computational studies have confirmed that not all intramolecular FLPs require dimerisation by disclosing numerous examples of FLPs being able to split H_2 through an intramolecular pathway with no need of FLP dimerisation. In many cases, especially with appropriate LA$\cdots$LB distances (typically 3–4 Å), a concerted heterolytic cleavage of H_2 can occur between the LA and the LB centres within a single molecule. These pathways often proceed via a single concerted TS, directly producing a zwitterionic adduct that supports a hydride at the LA and a proton at the LB.[56, 64–66] Turning to genuine intermolecular FLPs, Pápai and co-workers applied DFT calculations to analyse the H_2 splitting mechanism promoted by the combination of sterically hindered boranes and phosphines.[67] In such *intermolecular* FLPs, the LA and LB partners are separated into two distinct and different molecules as illustrated in Figure 2.8b. The authors selected the experimentally tested case of $B(C_6F_5)_3$ and $P(tBu)_3$, which was reported in one of the seminal works by Welch and Stephan.[68] The authors characterised a plausible mechanism that starts with the energetically favourable formation of a $P(tBu)_3\cdots B(C_6F_5)_3$ adduct ($\Delta E_{binding}$ = -11.5 kcal mol^{-1}) in which the two FLP partners interact via C–H$\cdots$F hydrogen bonds and dispersion forces. This association was found to lead to a reactive pocket between the P and the B centres, which are placed 4.2 Å apart. In the latter, H_2 can be heterolytically activated through a concerted TS whereby the H–H bond cleavage occurs concomitantly with the formation of B–H(δ^-) and P–H(δ^+) bonds, overcoming a calculated energy barrier from $P(tBu)_3\cdots B(C_6F_5)_3$ + H_2 to the TS of only 10.4 kcal mol^{-1}.

Despite ongoing debate over whether H_2 splitting by intra- and intermolecular FLPs follows a concerted[69, 70] or stepwise mechanism,[71–75] this dichotomy has been extensively explored through explicit-solvent DFT-based molecular dynamics (DFT-MD) simulations. Notably, Liu et al. revisited the archetypal $tBu_3P/B(C_6F_5)_3$ system using metadynamics simulations, reconstructing the free-energy landscape along a multidimensional reaction coordinate that included the H–H distance and the coordination numbers of B and P by H atoms.[75] Their results revealed a subtle intermediate near the transition state, indicative of a stepwise pathway, in contrast to the single transition state predicted by conventional static DFT calculations. The simulations showed that H_2 cleavage proceeds via hydride transfer to the Lewis acid, followed by proton capture by the base, with the former being rate-limiting, underscoring the dominant role of the acid component in controlling reaction kinetics. More recent DFT-MD studies on intramolecular P/N and B/N FLPs suggest concerted pathways,[69, 70] reflecting the complexity of the mechanistic landscape. Nonetheless, static DFT approaches, particularly those employing dispersion-corrected hybrid or meta-GGA functionals such as ωB97X-D,[76] TPSS-D3,[77–80] B3LYP-D3,[80, 81] M05-2X,[82] or M06-2X,[83] in combination with implicit solvation models [84] and triple-ζ basis sets [85] with polarization [86, 87] and diffuse functions, [88, 89] remain reliable for accurately predicting geometries, reaction free energies, and activation barriers, as supported by benchmarking against experimental and higher-level computational data.[90–92]

Overall, the above findings constitute a well-established mechanistic picture for the activation of H_2 by FLPs that is illustrated in Figure 2.8 and can be summarized as follows. On the one hand, *intermolecular* FLPs proceed via the formation of a non-covalent adduct between the LA and LB partners that is usually referred to as “encounter complex”, followed by the insertion of H_2 into the reactive pocket and its subsequent splitting to form a hydrogenated $[LB–H]^+\cdots[LA–H]^-$ ion pair. On the other hand, molecules that contain both LA and LB sites and can thus be classified as *intramolecular* FLPs may either activate H_2 by themselves or dimerize to operate as an intermolecular FLP. Although the competition between intra- and intermolecular pathways in intramolecular FLPs has not been systematically investigated for a whole range of systems to set clear conclusions, we can envisage that the prevalence of each of these paths might depend on the equilibrium LA$\cdots$LB distance within the intramolecular FLP structure and on the ability of FLPs to dimerize into stable non-covalent adducts. Up to now, this mechanistic knowledge has been successfully applied to rationalize the H_2 splitting activity of a wide variety of FLPs, including mostly P/B and N/B pairs,[38, 90, 93] and less often carbene/B,[94] and N/TM (TM = Ni, Pt) pairs,[95] or even that of heterogeneous systems such as hydroxylated indium oxide surfaces.[96–98]

H_2 polarization and cleavage

Two complementary models have been proposed to rationalize the origin of H_2 polarization and cleavage in FLPs: the Electron Transfer (ET) model,[67, 99] and the electric field (EF) model (Figure 2.11).[100] In the ET model, H_2 splitting is described as a synergistic charge-transfer process: the p-type lone pair of the LB donates to the σ^* orbital of H_2, while the σ orbital back-donates to the vacant p-type orbital of the LA. This push-pull interaction weakens the H–H bond and facilitates cleavage. Conversely, the EF model emphasizes the polarization of H_2 within the electric field existing between the LA and LB. Grimme and co-workers showed that such fields can elongate the H–H bond even in the absence of direct orbital interactions. The energy cost of H_2 insertion into the FLP pocket correlates with the magnitude of the electric field and may become barrierless at sufficiently high field strengths.

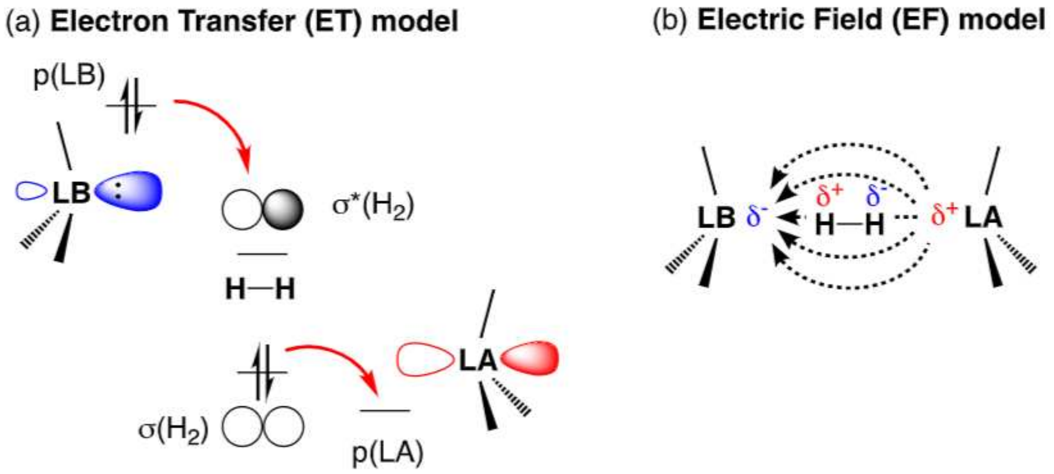


FIGURE 2.11: Schematic representation of the ET (a) and the EF (b) models proposed to explain the physicochemical foundations for H_2 splitting by FLPs.

Using DFT-MD simulations in combination with the Metadynamics enhanced-sampling technique, Liu *et al.* reinvestigated the splitting of H_2 by the archetypal tBu_3P/ $B(C_6F_5)_3$ FLP and revealed that both mechanisms can operate simultaneously.[75] The simulations showed that at long distances from the reactive centres (>2.5 Å), H_2 is polarized by the EF existing between LA and LB partners, displaying a slightly elongated bond with respect to the equilibrium distance (0.80 Å vs. 0.74 Å). By contrast, at shorter distances, charge transfer dominates, as evidenced by measurable charge density transfer from LB to H_2 and from H_2 to LA, supporting that the ET model governs at short FLP$\cdots H_2$ distances, although the EF model can also contribute to ease the H_2 splitting.

Skara *et al.* reported a classification of FLP's TSs which further supported this duality, by scrutinizing the reactivity of six Lewis pairs towards H_2 employing DFT calculations.[101] These include three classical Lewis pairs: Me_3P–BF_3, Me_3P–$B(C_6F_5)_3$ and Lut–$B(C_6F_5)_3$ (Lut represented in Figure 2.4e); and three FLPs: carb/$B(C_6F_5)_3$, $tBu_3P/B(C_6F_5)_3$, Mes_3P/ BPh_3 (carb = N,N′-tBu disubstituted imidazole-based carbene). This investigation identified two different types of TSs for H_2 splitting, which differ both in structure and in energy demand and were classified accordingly. Reactive systems exhibiting low-energy TSs exhibited short H–H bond lengths ($\sim$0.80 Å) with geometries resembling the reactants, thus referred to as "early" TSs, whereas unreactive systems (e.g., Me_3P–BF_3 and Mes_3P–BPh_3) showed high-energy TSs characterized by more product-like geometries and elongated H–H bonds ($\sim$0.95–0.98 Å), thus referred to as "late" TSs. These findings support the existence of two different mechanisms: one dominated by EF activation and another governed by ET. The authors concluded that the ET model applies to high-energy, ie “late” TSs. On the contrary, the EF was proposed to play a more critical role in the splitting of H_2 by systems that operate through “early” TS structures, in which no sign of base $\rightarrow \sigma^*$ donation was detected. While this classification advances our mechanistic understanding, predicting whether a given FLP will follow an "early" or "late" TS pathway remains challenging and likely requires broader systematic analysis of reactivity data and machine learning approaches to establish generalizable classification.

Structure–activity relationships in FLP reactivity

In addition to characterizing atomistic pathways of H_2 splitting by FLPs, computational efforts have aimed to uncover structural and electronic factors governing this process. Indeed, the first key requirement for an active FLP is a suitable LA$\cdots$LB distance. Illustratively, by performing constrained potential-energy surface scans along the P$\cdots$B distance of the $P(tBu)_3/B(C_6F_5)_3$ FLPs, Vankova *et al.* showed that the optimal P$\cdots$B distance for H_2 activation lies in the 3-5 Å range.[102] For B/N FLPs, Corminbœuf *et al.* found for H_2 splitting an ideal B$\cdots$N distance of $\sim$2.9 Å, with an effective activation over a larger 2.6-3.5 Å range based on a Morse potential fit.[103]

Notably, experimental studies by Pápai *et al.* [40], have demonstrated that the electronic properties of substituents on the LA and LB centres significantly influence H_2 splitting by FLPs. Their partition analysis of the Gibbs free energies for H_2 splitting promoted by a series of intra- and intermolecular FLPs yielded several key insights: (i) unreactive FLPs for H_2 splitting typically exhibit unfavourable thermodynamics, highlighting that the stabilization of the product is an essential requirement; (ii) the thermodynamics of H_2 splitting correlates with the cumulative acid–base strengths of the FLP components; (iii) intramolecular FLPs experience a reduced entropic penalty along the reaction coordinates when compared to intermolecular FLPs, and thus require lower cumulative acid–base strengths to be active; (iv) acid–base strengths can be quantified using proton and hydride attachment energies; and (v) in intramolecular FLPs, the reaction free energy correlates linearly with the inverse of the LA$\cdots$LB distance, indicating that shorter donor–acceptor distances favour more exergonic H_2 splitting. Although accessible kinetics are also necessary, the Brønsted-Evans-Polanyi (BEP) principle [104, 105] implies that favourable thermodynamics often ensure low activation barriers, as subsequently confirmed. [54]

By analyzing a series of intramolecular B/N FLPs (Figure 2.12), Yepes *et al.* [64] reported that the height of the reaction barrier is highly sensitive to the electronic nature of the substituents of both the LA and LB sites, as these modulate the cooperative LB $\rightarrow \sigma^*(H_2)$ and $\sigma(H_2)\rightarrow$LA charge transfers as well as the electrostatic interactions along the reaction path. In particular, electron-withdrawing or soft/poor π-donor substituents on the LA were found to lower the energy barrier. Vankova *et al.* [102] extended these findings by studying a series of experimentally tested intermolecular FLPs with systematically varied acid-base strengths, showing that both strong LBs and strong LAs enable irreversible H_2 activation, while combinations of weak LA/LB fail to activate H_2. Intermediate strength FLPs exhibited reversible H_2 splitting [2, 106], with calculated Gibbs free energies correlating with cumulative acid-base strength [102]. Strong cumulative strengths led to highly exergonic and irreversible H_2 activation, while weak pairs showed endergonic processes. Neu *et al.* [107] experimentally confirmed this correlation, and the concept has since guided the design of FLPs for CO_2 hydrogenation. Despite consensus on the importance

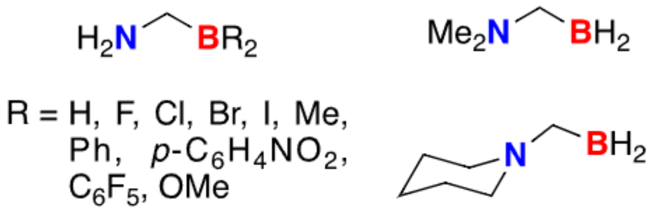


FIGURE 2.12: Intramolecular B/N FLPs studied computationally for heterolytic H_2 splitting by Yepes *et. al.*[64]

of LA and LB strengths, their relative influence on kinetics and thermodynamics remains underexplored. Liu *et al.* [38, 75] used metadynamics to show that in tBu_3P/ $B(C_6F_5)_3$, the hydride transfer to the LA is rate-limiting, while the subsequent protonation of the LB drives the thermodynamics of the reaction. Thus, the authors concluded that LA acidity controls kinetics and contributes to thermodynamics, while LB basicity chiefly determines the latter. In contrast, experimental work by Berke and co-workers [108] suggested that the LA acidity is the primary factor in the reversibility of the H_2 splitting, with the LB basicity playing a secondary role. In light of this, it can be inferred that there is no clear consensus on how FLP parameters affect their reactivity.

Even though some qualitative trends relating to the H_2 splitting ability of FLPs and their molecular structure had been previously reported, Ye and Johnson highlighted a lack of clear structure–activity relationships in this area.[109] In their computational study on CO_2 hydrogenation using FLPs grafted onto the organic linkers of the metal–organic framework UiO-66, they focused specifically on the H_2 activation step, as well as on the subsequent CO_2 hydrogenation step. The study focussed on a series of eight B/N FLPs incorporating pyrazole–BR_2 motifs (R = CH_3, F, H, Cl, Br, CN, CF_3, NO_2), which were chosen for their ability (at the experimental level) to be covalently grafted onto the terephthalate linkers of the MOF (Figure 2.13a). This work aimed to establish the first structure–activity relationships between the FLPs' structure and H_2 activation and, ultimately, CO_2 hydrogenation under a gas stream of H_2 and CO_2. Here, we focus on their findings concerning the initial H_2 splitting step. The selected FLPs form a series of intramolecular B/N pairs of increasing Lewis acid strength, allowing the authors to compute electronic energy profiles and probe for Brønsted-Evans-Polanyi (BEP) relationships between reaction energetics (reaction energies and barriers) and molecular descriptors. These descriptors included the strength of the LA and the LB, estimated as the hydride and proton attachment free energies, respectively; electronegativity was calculated from the ionisation potential and electron affinity of the FLP using the Mulliken formula,[110] whereas hardness and softness were also derived employing the Parr and Pearson's formula.[111]

Electronic energy profiles revealed a linear correlation between the H_2 dissociative adsorption energy and the hydride attachment free energy so that, as the LA strength increases, so does the H_2 dissociative adsorption energy (Figure 2.13b), but no correlation with the proton attachment energy, consistent with Berke's results.[108] However, the computational work of Ye and Johnson addressed the impact of the substituents at the LA boron centre exclusively, resulting in a rather narrow range of proton attachment energies. In other words, the analysed FLPs do not significantly differ in terms of LB basicity, and therefore, the impact of this parameter remained uncharted. Still, a BEP relationship was found between the H_2 splitting barrier and the LA's chemical hardness (Figure 2.13c). Increasing the hardness of the LA thus decreases the barrier for H_2 split in accordance with Pearson's

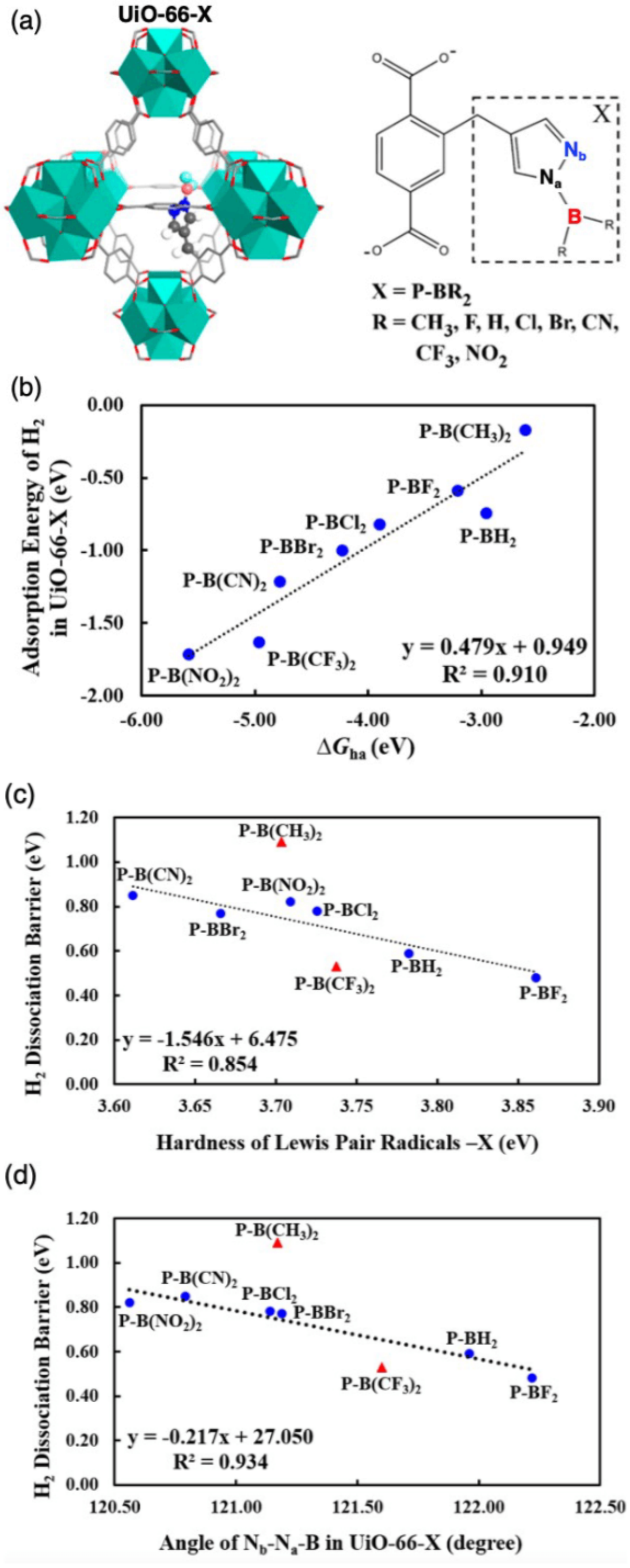


FIGURE 2.13: (a) Octahedral cage of the UiO-66 metal-organic framework bearing an organic linker functionalized with a N/B Lewis Pair (X). (b) Dissociative adsorption energy of H_2 in UiO-66-X plotted against the free-energy of hydride attachment to the B centre of the Lewis pair. BEP relationships between the H_2 dissociation barrier on X and the chemical hardness of the Lewis pairs (c) and the N_b-N_a-B angle in the bare UiO-66-X structure (d). Reprinted from ref. 109.

theory, as this is a process along which the LA binds a hydride. We wish to stress here that, unlike the H_2 adsorption energy, this barrier was found not to correlate with the acidity of the LA. This can be interpreted by the fact that the BEP principle does not fully apply to this process, given the absence of a linear correlation between reaction energies and barriers; and that some other factors affect the kinetics but not the thermodynamics of the reaction. Another BEP relationship was found between the H_2 splitting barrier and the bond angle N_b-N_a-B formed by N_b, N_a and B centres (Figure 2.13d) in the bare FLP, which was ascribed to the fact that larger angles induce more strain in the FLP, thus lowering the barrier for H_2 splitting. The height of the reverse barrier for H_2 release, however, was found to be inversely proportional to the H_2 dissociative adsorption energy and to the hydride attachment energies, which as mentioned above, correlate linearly to one another. It is also relevant that the MOF environment was not found to significantly affect the splitting energy of H_2 on the FLP, as similar reaction energies were obtained from both isolated and MOF-hostedFLP models. Later, Ye *et al.* extended this to four FLP families (B/N and B/P), finding strong correlations within individual FLP sub-families between H_2 splitting and hydride attachment energies ($r^2 > 0.95$), but weaker correlations when considering all FLP families ($r^2 = 0.83$).[55] A multivariate relationship was derived (Eq. 2.3), involving the LA$\cdots$LB distance, d_1, the hydride attachment energy on the LA site (ΔG_{ha}), and $\Delta\theta$, which stands for the variation of the angles concerning the LA/LB sites and the two consecutive atoms of the scaffold in the direction of the FLP partner upon H_2 binding. This evidences that besides hydride attachment free energies, there are other significant factors that influence the splitting of H_2 (ΔE).

$$\Delta E = (0.184 \cdot d_1 + 0.273) \cdot \Delta G_{ha} + (0.990 \cdot d_1 + 0.010 \cdot \Delta\theta - 0.400) \tag{2.3}$$

Zhang *et al.* [112] studied the splitting of H_2 for four similar B/N FLPs, including inactive Piers's ansa-aminoborane [113] and active derivatives reported by Repo's, [114, 115] which differ in the substituent's electronic properties on both the B and N centres. The authors concluded that the acidity of the LA has an important effect on the reaction free energies of the H_2 splitting. The latter correlate with the Hammett constants of the substituents for a series of compounds bearing a -B(p-$C_6F_4R)_2$ LA moiety whereby R is modified systematically, as shown in Figure 2.14, again highlighting LA acidity's influence on thermodynamics. However, the reaction barrier was found to depend more of the LB strength than of the LA acidity. Also, increasing the solvent's polarity was also found to stabilize the zwitterionic product, improving in turn the thermodynamics of H_2 splitting. Tussing *et al.* used similar descriptors to relate FLP-catalyzed imine hydrogenation rates to substituent effects.[116] Their work further demonstrated that, while P-based LB groups form strong zwitterionic adducts, their larger size promotes effective product separation, a feature crucial for catalytic turnover. By contrast, oxygen-based LBs resulted in more tightly bound products, impeding further reactivity.

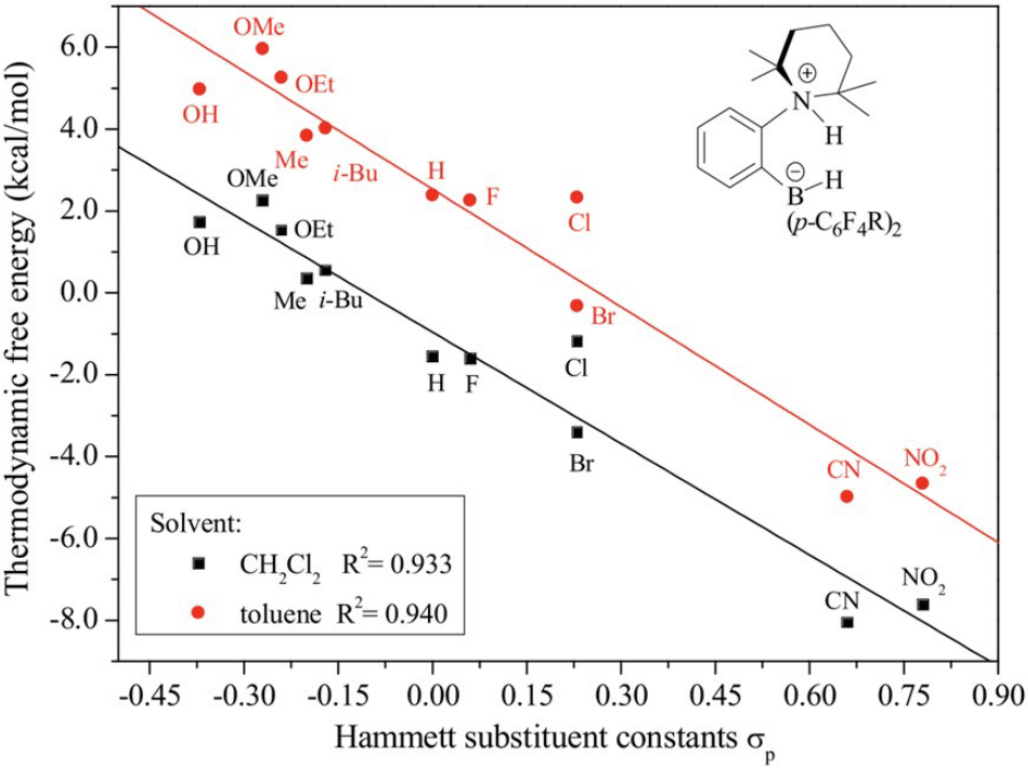


FIGURE 2.14: Relationship between the reaction Gibbs free-energy for H_2 splitting by an intramolecular FLP (shown in the inset) and the Hammett constants of the substituents. Reprinted from ref. 112.

In 2020, Heshmat and Ensing performed a broad computational screen of 75 intermolecular FLPs which was aimed at disclosing the impact of the electronic and steric effects of both LA and LB components on the kinetics and thermodynamics of H_2 splitting.[117] The array of analysed FLPs consisted of combinations of 12 experimentally reported borane-derived LAs with two families of LB partners, including strong phosphines (tBu_3P and Me_3P) and less basic ethereal solvents (terahydrofuran and Ph_2O). Their study confirmed that modifying the LA electrophilicity and the LB nucleophilicity allows to adjust the thermodynamics of H_2 splitting, consistent with the cumulative acid-base strengths model.[40, 102] O-based LBs yield strong ion-pair interactions and short $H^{-}\cdots H^{+}$ distances, hindering H_2 dissociation, while P-based LBs give longer distances and more exergonic reactions due to stronger P–H bonds and bulkier geometry. Concerning H_2 splitting kinetics, they found that: i) strong LA and LB partners reduce energy barriers; ii) bulky LBs promote early transition states, reducing steric clashes and barriers (as also shown by Skara *et al.* [101]); iii) LA bulkiness increases barriers by restricting B access. The authors noted that BEP principles apply well to these intermolecular FLPs, as barriers correlated with reaction energies. These features contrast with those reported by Ye and Johnson on intramolecular FLPs,[109] suggesting fundamental differences between intra- and intermolecular systems.

Overall, these computational findings establish that FLP's reactivity is indeed governed by a combination multiple factors, including acid-base strengths, LA$\cdots$LB distance, electronic hardness, and steric environment. Still, several points still lack of a quantitative and unambiguous answer. For instance, the exact structural and electronic parameters that determine whether an intermolecular FLP operates through an intra- or an intermolecular mechanism, the relative impact of electric field polarisation and orbital overlap on H_2 splitting and how they depend on the FLP's nature and structure, are aspects of the H_2 splitting by FLPs that remain controversial or unclear nowadays. Despite the above progresses, generalizable and quantitative structure-activity relationships remain still underdeveloped.

2.1.3 Motivation and scope of this work

To overcome these limitations and establish transferable generic models, we aim at developing multivariate regression models to identify and quantify general structure-activity relationships between FLPs' molecular features, one the one hand, and their reactivity towards H_2 splitting, on the other hand. Our ultimate goal is to be able to predict the reactivity of a given FLP molecule for H_2 splitting on the basis of simple molecular features or descriptors of this FLP molecule. In that purpose, we have constructed a dataset of 112 intramolecular B/N FLPs possessing systematically varied electronic and structural features, i.e. covering various backbones and chemical substituents for LA and LB centres. Using dispersion-corrected DFT calculations, we computed for each FLP molecule the related reaction free energy (ΔG) and free-energy barrier ($\Delta G^{\ddagger}$) for H_2 splitting, together with a set of chemically meaningful molecular descriptors, described in section 2.3.

We then developed multivariate regression models aiming at correlating these molecular descriptors with both the thermodynamic (ΔG) and kinetic metrics ($\Delta G^{\ddagger}$) of H_2 splitting (Figure 2.15), thus aiming at identifying the key molecular features governing FLPs' reactivity for H_2 splitting. These models aim to provide new insights into how an FLP's structure dictates its reactivity towards H_2 splitting and to offer *predictive* capabilities for structurally and chemically related FLP systems. Our approach aims to bridge the gap between mechanistic understanding and quantitative prediction, enabling rational design and prescreening of effective FLPs candidates for optimal H_2 activation and splitting, requirement for further transformations.

An important long-term objective of this work is to move beyond the sole reliance on quantum chemical calculations by leveraging predictive models to estimate H_2 splitting reactivity metrics directly from FLP's molecular features. The use of such predictive simple models would in principle allow for faster screening and design of efficient FLPs for H_2 splitting, bypassing the computational cost of full DFT evaluations.

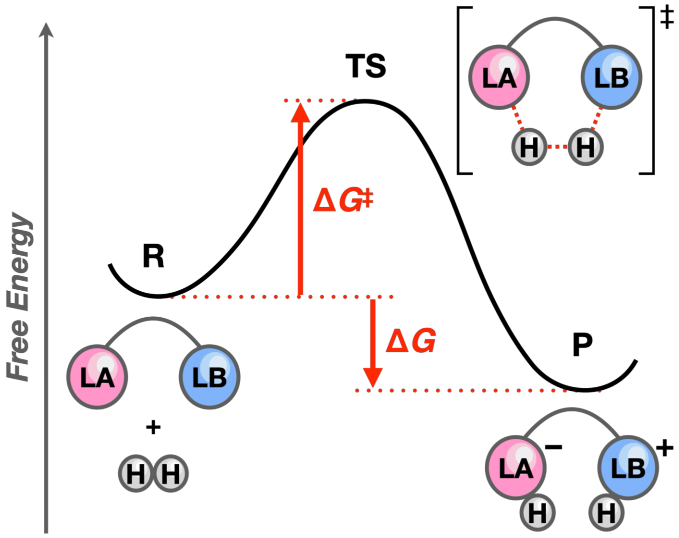


FIGURE 2.15: Schematic free-energy profile illustrating the heterolytic splitting of H_2 promoted by intramolecular FLPs. LA and LB stand for Lewis acid and Lewis base, respectively. The main energy parameters discussed in this work (i.e., reaction free energy, ΔG; and free-energy barrier, $\Delta G^{\ddagger}$) are highlighted with red arrows.

This project aims not only at elucidating which descriptors best predict an FLP's reactivity for H_2 splitting, but also how steric and electronic features jointly modulate barrier heights and reaction energies. Importantly, the ultimate goal is to integrate these findings to guide the design of FLPs for tandem processes such as CO_2 hydrogenation, where H_2 splitting must be obviously balanced with subsequent reactivity.

2.2 Computational methods

In order to create a representative sample of the target chemical space, the H_2 splitting free energy profile as illustrated in Figure 2.15, of a library of 112 B/N FLPs (Figure 2.23) was computed and characterized by means of DFT calculations. Regression methods, such multivariate linear regression (MLR) and Partial Least Squares (PLS), were used to build mathematical expression models based on molecular descriptors of FLPs in order to predict their catalytic performances for H_2 splitting, i.e. the reaction free-energy (ΔG in Figure 2.15), that captures the thermodynamic stabilization of the process; and the free-energy barriers ($\Delta G^{\ddagger}$ in Figure 2.15), which captures the kinetics of the process (Figure 2.16). We expect being able to assess correlations to i) identify key FLP's molecular descriptors at play in the H_2 splitting reactions, ii) predict in a quantitative fashion the reactivity of FLPs using simple molecular descriptors, while iii) gaining new chemical understanding of FLP's reactivity which will be of valuable use for further design. The creation and exploration of different predictive models based on a variety of descriptors (structural, chemical

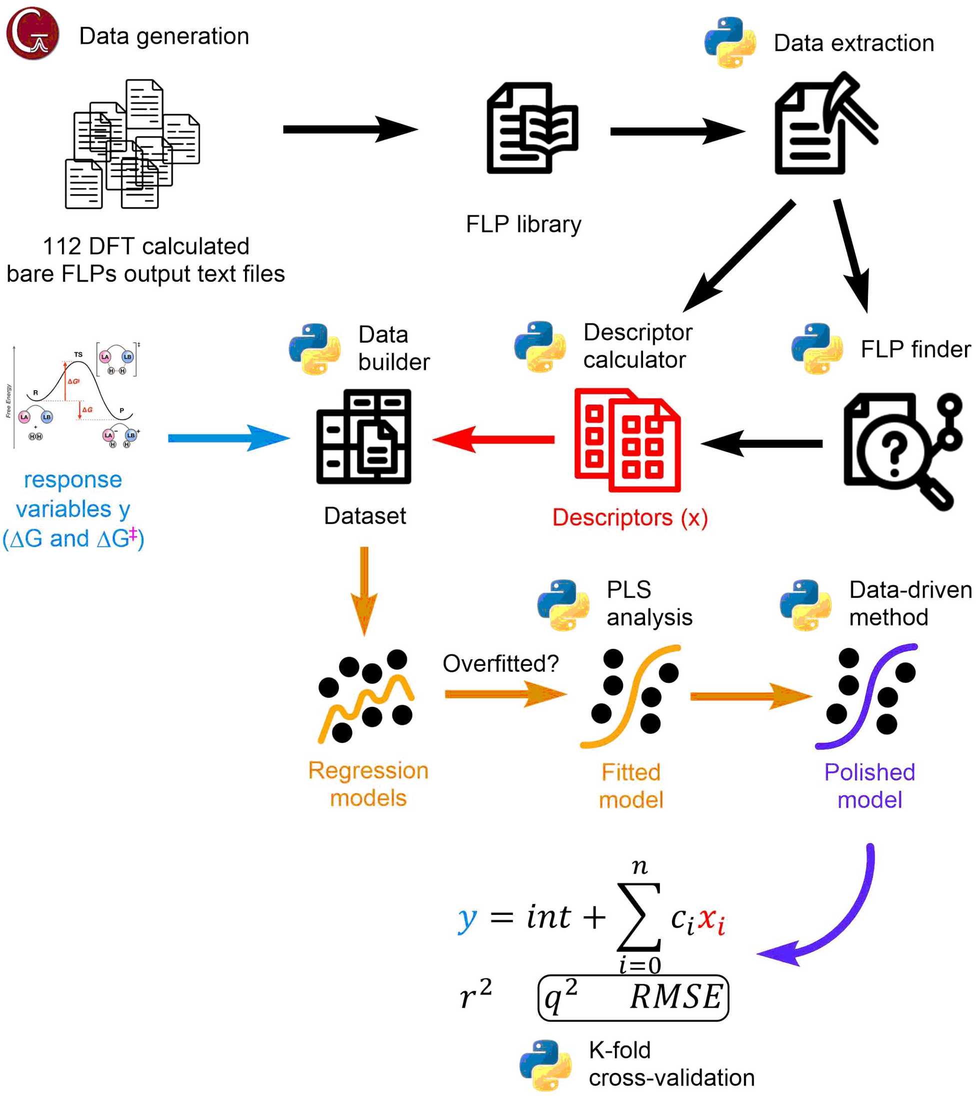


FIGURE 2.16: Schematic representation of the methodology followed to create all the models showed in this chapter. In blue the response variables (ΔG and ΔG‡) in Figure 2.15). In red the descriptors detailed in section 2.3.1. In orange the model fitting methods. Finally, in purple the data-driven model polishing method.

or electronic) of FLPs should allow us to identify the underlying key features that control their reactivity towards H_2 splitting. The proposed approach is meant to be generic and applicable to the further step of CO_2 hydrogenation. All the python codes and files mentioned in the following sections can be found in the corresponding GitHub repository (https://github.com/FrancescPenas/FLPs_for_H2_splitting .git).

2.2.1 Data generation from DFT calculations

The free-energy profile of each FLP-containing molecule of the library (Figure 2.23), mentioned above, were evaluated by means of Density Functional Theory (DFT) calculations at the ωB97X-D/6-311++G(d,p)[76, 85, 87–89, 118] level using the Gaussian 16 (rev C.02)[119] quantum chemistry software. Such a level of theory has been successfully applied before to model FLP chemistry.[90, 92, 101] Solvent effects of toluene were accounted both in geometry optimizations and energy calculations through the IEF-PCM implicit solvent model.[120] The nature of all stationary points on the potential energy surface, minima and TSs, was confirmed by frequency calculations. The orbital occupancy was studied using the Natural Bond Orbital (NBO) analysis version 3.1 already implemented in Gaussian. All the information of each DFT-optimized FLP, TS and H_2 splitting product (R, TS and P, respectively in Figure 2.15) are stored in plain text ".log" output files.

A change of +1.9 kcal mol^{-1} was applied to the free energies of all species to account for the free energy variation in going from the reference state of an ideal gas at 1 atm assumed in Gaussian's thermochemistry calculations, to a standard state of 1 mol L^{-1} in solution at 298.15 K. These corrected energies were used to calculate ΔG and $\Delta G^{\ddagger}$ of the H_2 splitting process (Figure 2.15).

2.2.2 Automated data extraction from DFT calculations

An in-house developed Python code (data_extractor.py) has been used to retrieve in a systematic fashion the information contained in the Gaussian output files of the bare molecule (R in Figure 2.15) of all the studied FLPs. As shown in the flowchart (Figure 2.17), the code first collects the name of all the Gaussian output files in the input directory, sorting them alphanumerically. It then creates empty lists that are filled with the extracted data of all the Gaussian output files. The code then loops through each text output file calling the "gaussian_reader" function that transforms the text file into a list of strings corresponding to each line of the Gaussian output file. For each file, several subroutines are called to identify desired sections of the list of text lines containing the corresponding information (data types), checking if the information exists in the file, or detecting possible errors in the original Gaussian output file. The mentioned subroutines extract optimized cartesian coordinates (gaussian_coordinates_extractor.py), Natural Bond Orbitals (NBOs) (gaussian_nbo_extractor.py), connectivity (gaussian_connectivity_extractor.py), natural charges derived from the Natural Bonding Analysis (gaussian_nat_charges_extractor.py) and ChelpG charges (gaussian_esp_charges_extractor.py). All the extracted information is saved in the previously created lists and the possible generated errors. After looping through all the Gaussian output files, the extracted data is compiled in python directories. Finally, the data is saved in the "input_data.pkl" binary file and the errors, if any, in the "data_extractor_error_list.csv" file. Notable, the code can be easily tuned to adapt it to other kinds of files types and to make it compatible with other kinds of software.

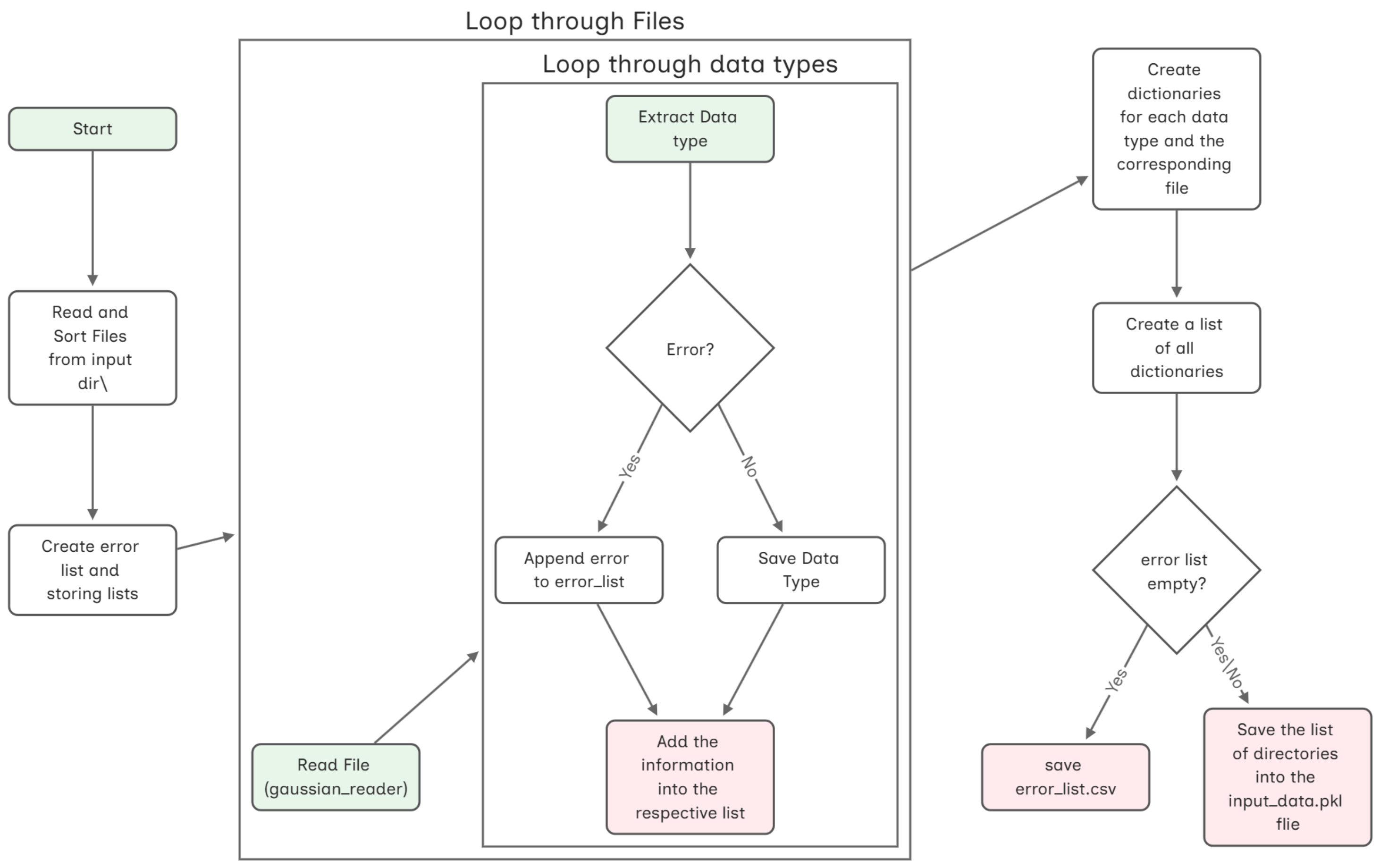


FIGURE 2.17: Flowchart illustrating the structure and functionality of the data_extractor.py script.

2.2.3 FLP finder

The following code is aimed to detect both LA and LB centres for each FLP-containing system (flp_detector.py), for which flowchart is shown in Figure 2.19. First, the code checks the existence of the extracted data and loads the information used to detect both Lewis partners, i.e. the coordinates, NBO and the connectivity of each FLP. Then, it creates empty lists to store the generated information. The code loops through all the FLP-containing molecules, calling two subroutines that identify the LA ("gaussian_la_detector.py") and the LB ("gaussian_lb_detector.py"), respectively. Importantly, for each FLP, the "gaussian_la_detector.py" code looks for the lowest energy-lying unoccupied orbital (labelled as LP* in NBO analysis) of the boron atom using the NBOs information. If such an orbital does not exist, it returns an error message. The "gaussian_lb_detector.py" code then identifies the LB by searching for the occupied lone pair (LP) orbital of the nitrogen atom. When more than one N atom with a LP (N_1, N_3 and N_4 of Figure 2.18) are present in the FLP molecule, the code calculates their respective distance from the identified LA, and takes the closest one that is non-directly bound to the LA (N_3 of Figure 2.18). Finally, the Gaussian indexes corresponding to the Lewis partners identified above and the calculated LA$\cdots$LB distance are compiled in the python dictionaries and saved in the "data_flp.pkl" binary file.

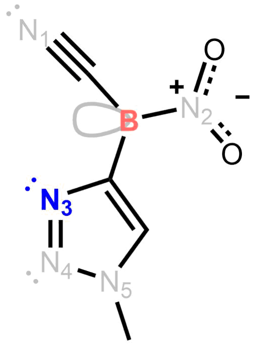


FIGURE 2.18: Example of an FLP containing one LA (B) and several LB candidates containing a lone pair (LP) of electrons (N_1, N_3 and N_4), without a LP (N_5) and directly bonded to the LA (N_2).

2.2.4 Regression methods

Using the data extracted from the Gaussian output files (Section 2.2.2) and the Lewis partners identified above (Section 2.2.3), a set of selected molecular descriptors (x) was further calculated using this information (Section 2.3.1). These descriptors were selected to cover both structural and electronic properties of the FLPs and will be described in more detail in section 2.3 and in Table 2.1. They constitute the dataset which will be used to create mathematical expression models (Eq. 2.4) in order to predict ΔG and $\Delta G^{\ddagger}$ (Figure 2.15), also referred in this work as response variables

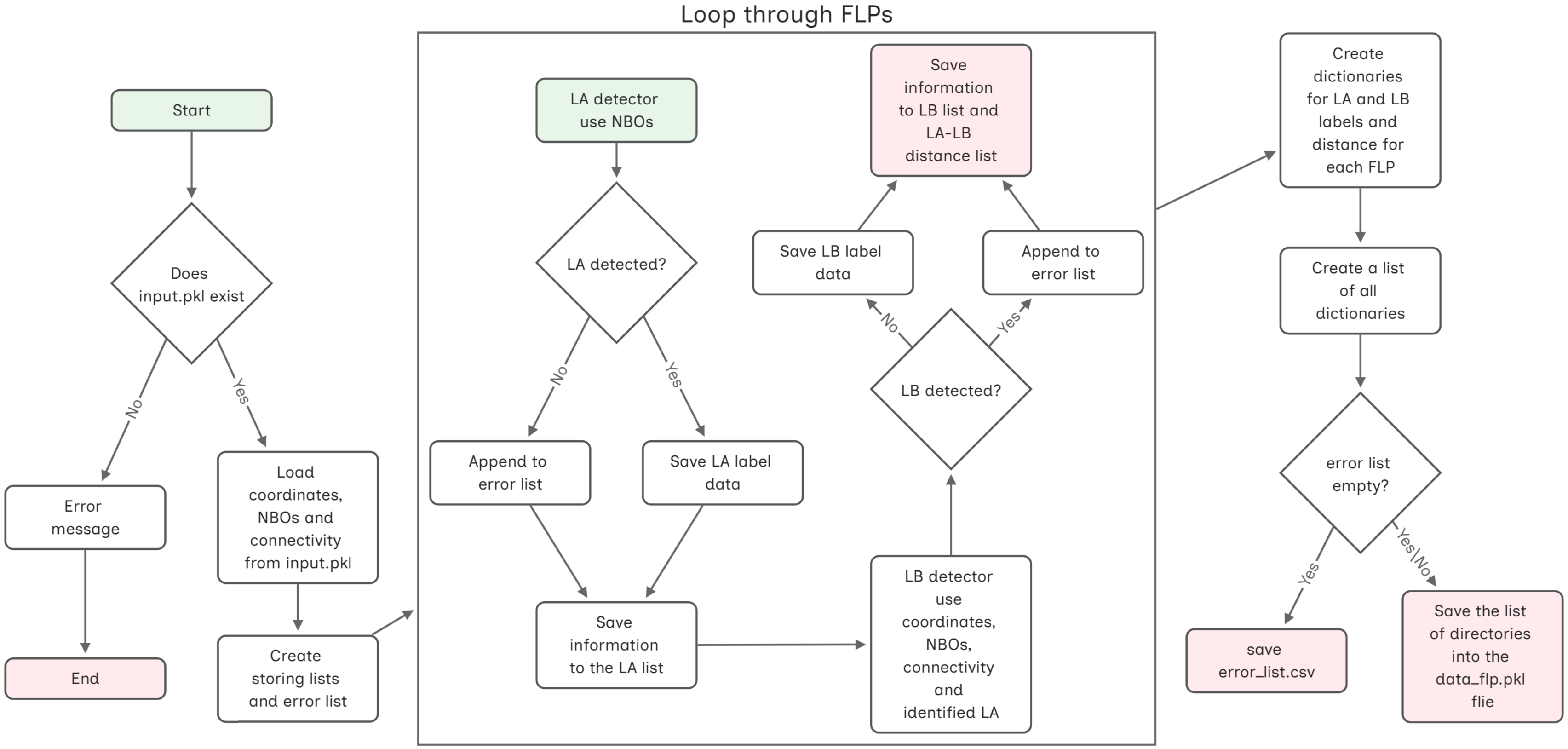


FIGURE 2.19: Flowchart illustrating the structure and functionality of the flp_detector.py script.

(y). The mathematical models (Eq. 2.4) consist in a linear combination of n descriptors (x) weighted by coefficients (c), the coefficients (c) thus directly reflecting the contribution of the descriptors to the model.

$$y = \sum_{i=1}^{n} c_i x_i + int \tag{2.4}$$

First of all, the data needs to be normalized. This procedure guarantees that various descriptors contribute equally to the analysis, especially when these descriptors differ in units and scales, as is common with chemical, structural, or electronic properties. For that purpose, we used the "MinMaxScaler" function, imported from the `sklearn` python package,[121] to reshape the range of the descriptors, also referred to as original variables (x), from 0 to 1. The normalized variables ($\widetilde{x}$) are calculated subtracting their corresponding minimum value ($min(x)$) divided by their range value ($max(x) - min(x)$), according Eq. 2.5.

$$\widetilde{x} = \frac{x - min(x)}{max(x) - min(x)} \tag{2.5}$$

The regression models are evaluated using the correlation coefficient r^2 to check the model's fitting following Eq. 2.6, using the response variable value (y_i) of the FLP (i), the predicted value using the model ($\hat{y}_i$) and the mean value of the response variable of all FLPs ($\overline{y}$).

$$r^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \overline{y})^2} \tag{2.6}$$

For r^2, the best possible score is 1.0 and it can be negative, due to the fact that the model can be arbitrarily worse than simply predicting the mean of the target variable for all observations.

K-fold cross-validation

In order to evaluate the predictive ability of the fitted models, we employed the cross-validation parameter q^2 (Eq. 2.7), which is calculated using the same strategy as r^2 (Eq. 2.6), but using cross-validation methods. The q^2 coefficient is calculated from leave-one-out cross-validation, where $\hat{y}_i$ is the predicted value of y_i when the i-th data point is left out during the model training.

$$q^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \overline{y})^2} \tag{2.7}$$

$$RMSE = \sqrt{\sum(y_i - \hat{y}_i)^2} \tag{2.8}$$

According to Tropsha, [122] models with q^2 values gives an estimation of the prediction ability of the model: from 1.0 to 0.9, the model can be considered as presenting an excellent precision; from 0.9 to 0.7, a good precision; from 0.7 to 0.5, a screening ability; and under 0.5, it fails to predict the response variable, y. Using the same values than for q^2 (y_i and $\hat{y}_i$), the root mean squared error (RMSE) is also calculated as an indicator of the precision of the models but keeping the original scale and units of the response variable (Eq. 2.8). The q^2 parameter can be obtained from the Leave One Out Cross-Validation (LOOCV) methodology, which consists in withdrawing each one datapoint of the dataset used to train the model, rebuild a model with the remaining ones and finally predict the excluded datapoint. While this is feasible in manageable datasets, the LOOCV makes the calculation of q^2 computationally more expensive as datasets become bigger and when the number of models to fit and evaluate increases significantly. Thus, to save computational resources, the K-fold methodology has been used. This consists in randomly divide the dataset into *k* equally sized subsets of datapoints or “folds”. The methodology of the LOOCV is used by keeping each "fold" aside and training a model with the remaining folds. Each data point of each separated fold is then predicted with its related trained model, and q^2 is recalculated using those predicted values. Figure 2.20 shows the evaluation of q^2 and the time required to calculate it, using different number of folds generated using the "k-fold_opt.py" code. The q^2 value is practically conserved as the number of folds is reduced until 60, when the q^2 value starts to fluctuate, meaning that the probability of generating biased models increases. All the models created in this chapter were evaluated using 60 folds that reduces the computation time to the half, barely loosing quality.

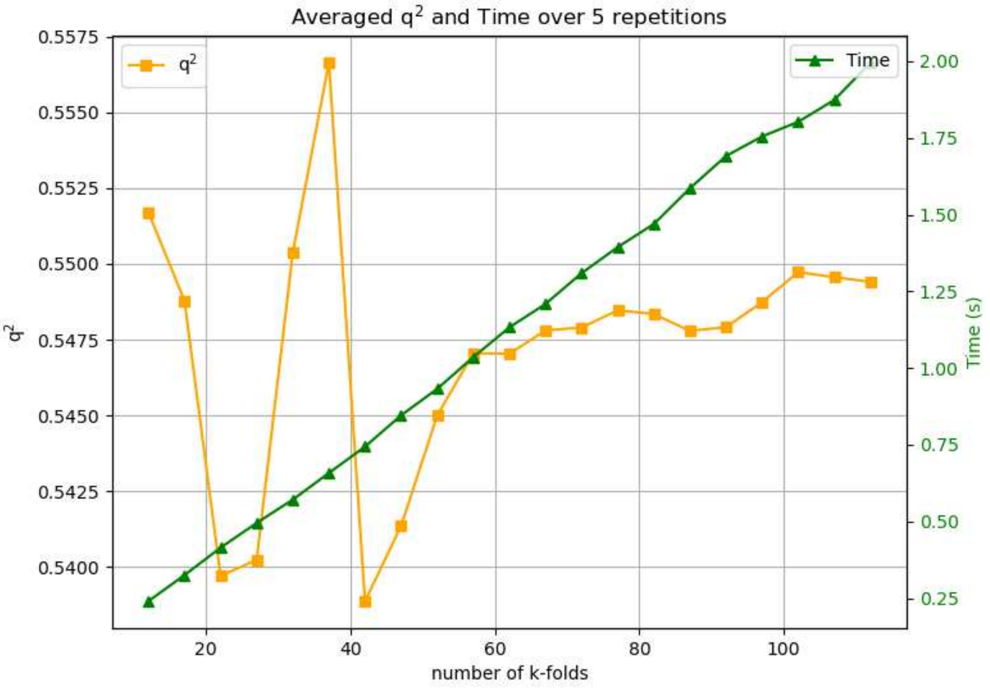


FIGURE 2.20: Average value of q^2 and the computation time using different folds for the evaluation of the model predicting $\Delta G^‡$ using all descriptors.

Partial least squares analysis

Partial Least Squares (PLS) modelling [123] is a statistical method that builds predictive models by finding latent variables, latent variable (LV), new, unobserved variables that are linear combinations of the original predictors, which best explain the variation in both the independent variables (our descriptors) and the response variables (ΔG and $\Delta G^{\ddagger}$ for H_2 splitting). It is especially useful when there are many collinear variables (I.e. descriptors that are highly correlated with each other), which can make traditional regression methods unstable. Unlike Principal Component Analysis (PCA), which focuses only on the variance in the descriptors, PLS selects components that maximize the covariance between the descriptors and the response variable, thus enhancing the prediction accuracy. This makes it ideal for analysing high-dimensional, noisy datasets, such as those commonly found in chemometrics, bioinformatics, and spectroscopy.

The PLS analysis is coded in "pls_analysis.py" to create predictive models, importing all tools from the `sklearn` python package.[121] Each data point is predicted to calculate the correlation parameter, r^2. The k-fold cross validation is carried out using the "cross_val_predict" function, which allows calculating q^2 and the RMSE. The coefficients of the PLS model are extracted to evaluate their impact; and the intercept (*int*) value is calculated following Eq. 2.9. $\overline{y}$ is the average value of the response variable, c_i the coefficient of each descriptor and $\overline{x_i}$ the mean value of the corresponding descriptor across all the dataset points.

$$int = \overline{y} - \sum_{i=1}^{n} c_i \overline{x_i} \tag{2.9}$$

Then, coefficients are transformed into the non-normalized ones in order to create the non-normalized model that predicts the response variable using directly the original (non normalized) descriptor values. This is done with the "convert_to_non_normalized_coefficients" function. Finally, a table is created that contains the real, predicted and error values for each data-point for data analyses. The main advantage of the PLS method is the transformation of the original set of variables, here the descriptors, into a new set of variables called latent variables, latent variables (LVs), which are linear combination of the original ones, and sorted according their contribution to the response variable. This fact allows to reduce the number of LV, sacrificing the minimum amount of data representation and overcome possible overfitting effects due to the limited size of the training dataset. Nevertheless, when re-converting the LV into the original variables (here descriptors), their number remain unreduced without providing any information about the uninformative descriptors.

Data-driven method

In order to discriminate uninformative descriptors, we adopted the data pretreatment protocol inspired by Andries *et al.* [7] This procedure, coded in informative_var_filter_PLS_iter.py, which flowchart is shown in Figure 2.22, involves augmenting the original set of N descriptors (including both linear and squared, normalized terms) with N randomly generated, intentionally non-informative variables. To avoid artifacts emanating from randomness, this process was repeated 10 times, each with a different set of random variables. For each of these 10 extended datasets, we built a series of PLS models with the number of latent variable (LV) ranging from 1 to 2N. For each dataset, we selected the model that simultaneously minimizes the RMSE. The coefficients of all real variables (both descriptors and random variables) were averaged across the 10 selected models. The highest average coefficient among the random variables was used as a threshold (see Figure A.1), so that descriptors with mean coefficients above this threshold were considered informative, while those below it were discarded (see Figure A.2). This process was iteratively repeated using the classified informative descriptors of previous iteration and new sets of random variables until no further original descriptor was classified as non-informative.

In the end, only 23 descriptors out of the initial 38, either linear or quadratic terms, passed this filtering process and were subsequently used to model relationships with the $\Delta G^{\ddagger}$. The coefficients of the final model including all 23 informative descriptors can be used to sort them according their relevance to the response variable. Using our final sorted set of 23 informative descriptors, we then performed a combinatorial analysis to generate all possible models. This involved creating ev-

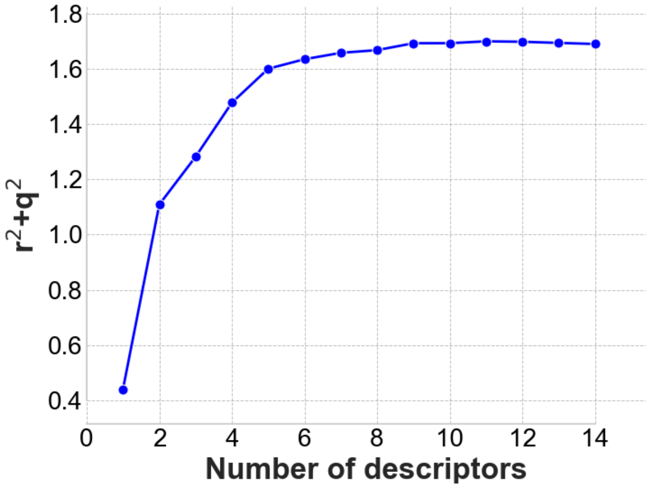


FIGURE 2.21: Quality of the best $\Delta G^{\ddagger}$ model of certain number of informative descriptors (r^2+q^2) vs each number of descriptors, from 1 to 14.

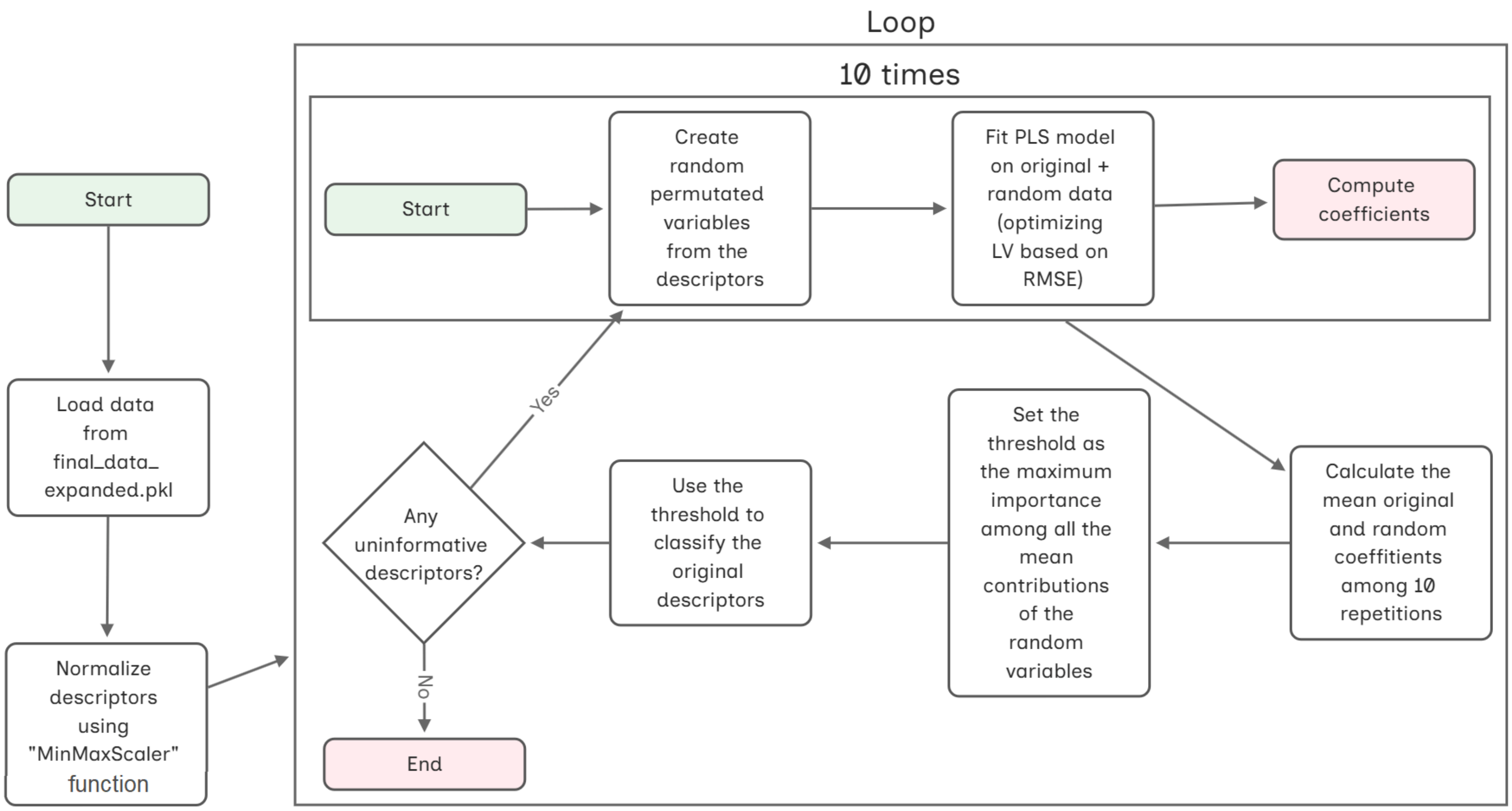


FIGURE 2.22: Flowchart illustrating the structure and functionality of the informative_var_filter_PLS_iter.py script.

ery combination of 2 to 23 descriptors and, for each, building MLR models, if the number of LV are equal to the number of original variables used, as well as PLS, if the number of LV are less than the original. Starting from combinations of two descriptors, this analysis was carried out by systematically increasing the number of included descriptors, following the sorted informative descriptors order, until the performance of the best model began to drop (Figure 2.21). The array of resulting models were ranked based on the sum of their r^2 and q^2 values, used as a combined metrics of the model's quality and stored in opt_barr_models.pkl.

2.3 Results and discussions

2.3.1 Construction and analysis of the dataset

Computational sampling of the chemical space was carried out, focusing on intramolecular FLPs featuring a boron atom as the LA and a nitrogen atom as the LB. This subclass of molecular FLPs represents one of the broadest and most extensively studied families of FLPs reported to date, owing to their rich structural diversity, well-documented reactivity, and synthetic accessibility,[5, 90, 124] making it an ideal platform for uncovering general structure–activity relationships. Figure 2.23 illustrates the library of the 112 FLP-containing molecules selected in this work to form our dataset, which serves as a representative subset of the target chemical space. A diverse set of molecular backbones was used to explore geometric constraints, resulting in LA$\cdots$LB distances spanning from 1.70 Å to 5.62 Å. We also investigated various substituents on both Lewis partners to examine a range of geometric and electronic environments surrounding the reactive sites. The FLPs were grouped into sub-families, each uniquely defined by the number of atoms in the shortest path between LA and LB, as well as by the relative orientations of these sites imposed by the molecular framework. Most of the core structures were derived from FLPs that have been either experimentally reported or previously investigated computationally.

We initially performed DFT calculations to determine the reaction Gibbs free energies (ΔG) and free-energy barriers ($\Delta G^{\ddagger}$) associated to the H_2 splitting promoted by each of the FLPs in Figure 2.23. As shown in Figure 2.24, our library covers a broad range of reaction free energies, from approximately –30.0 to +35.0 kcal mol^{-1}, encompassing FLPs that promote both highly favourable and unfavourable H_2 splitting. Notably, a significant number of data points cluster around $\Delta G \approx 0$, a potentially desirable feature for designing FLPs with tunable H_2 splitting reversibility, which is particularly relevant for applications in energy storage and hydrogenation catalysis. In terms of free-energy barriers, the dataset also captures a wide range of transition-state stabilities, spanning from low barriers of 10 – 15 kcal mol^{-1} to moderate-high (up to 35.0 kcal mol^{-1}) and even prohibitively high barriers in the 35 – 60 kcal mol^{-1} range. Therefore, the selected library constitutes a representative subset of the chemical space in terms of the response variables, encompassing FLP

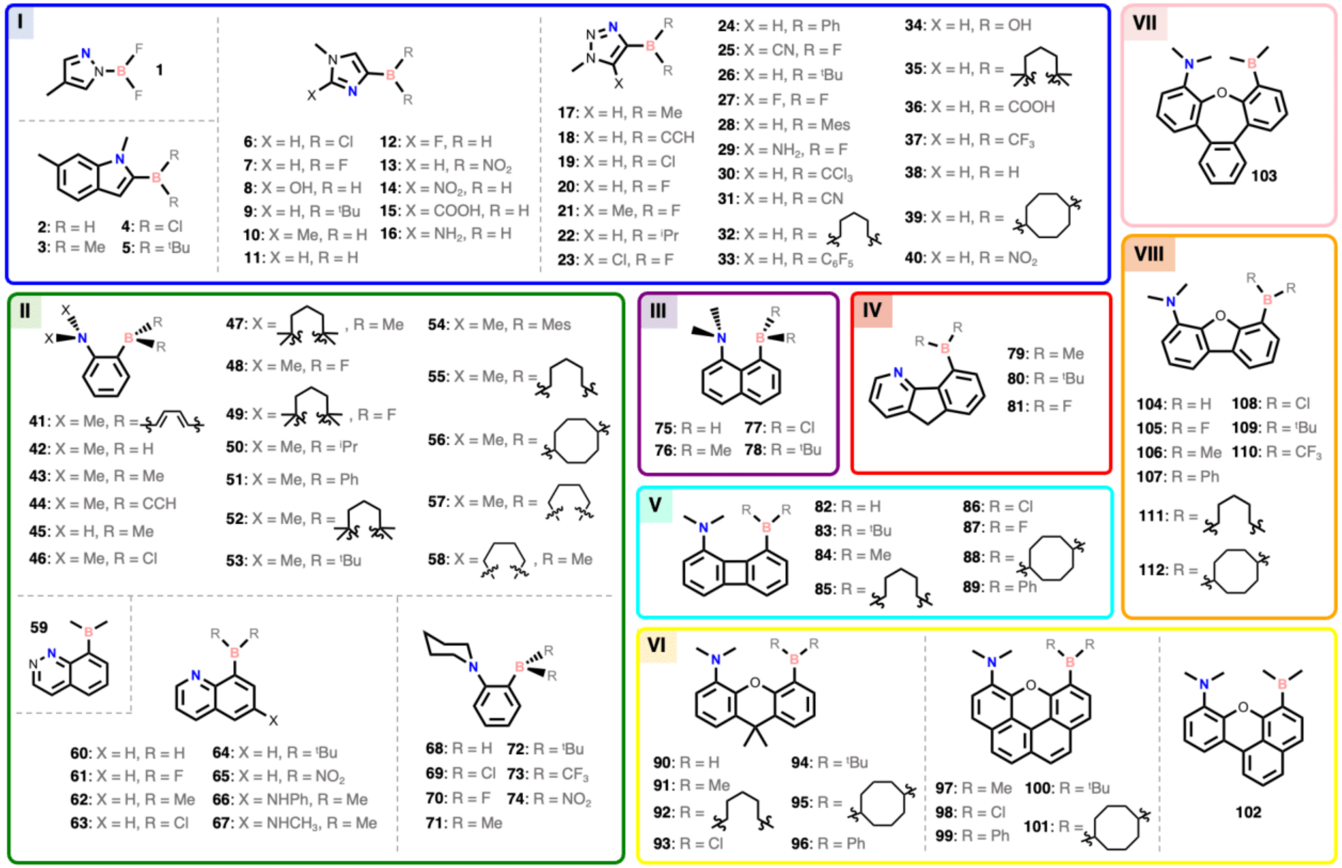


FIGURE 2.23: Library of all the intramolecular B/N FLPs constituting the dataset. Most of them were inspired by FLP structures analyzed in previous works. FLP **1**:[55, 109], **42**:[47], **41-54**:[55, 114], **46**:[125], **46** & **48**: [55], **47** & **49**:[112, 115, 126], **61** & **81**:[103], **60-65**:[127], **75-78**:[128, 129], **82-89**:[130], **90-96**:[131–133], **97-103**:[134] and **104-112**:[131]

structures that were predicted to be active, as well as inactive ones limited by either kinetic or thermodynamic factors.

Interestingly, Figure 2.24 also reveals a weak correlation between ΔG and $\Delta G^{\ddagger}$ (r^2= 0.37), indicating that the Brønsted-Evans-Polanyi (BEP) principle, [135, 136] correlation between ΔG and $\Delta G^{\ddagger}$ of the same chemical process, does not hold across the entire dataset. This suggests that different factors influence the thermodynamics and kinetics of H_2 splitting in distinct ways. However, when evaluating the BEP principle within individual families (Figure 2.25), significant correlations emerge. For example, among the geminal FLP family I (r^2 = 0.85) and the dibenzofuran-based family VIII (r^2= 0.91). In contrast, other families, such as II, containing anilines and quinolines, show little to no correlation (r^2= 0.15). This observation recalls to earlier results of Ye and Johnson[109], who reported that linear relationships between the binding energy of H_2 (ΔE) on FLPs and FEHA held within individual FLP sub-families (with r^2> 0.95), while the correlation weakened substantially (r^2= 0.83) when data from all families were combined, supporting the idea that meaningful structure–reactivity relationships are more readily captured within structurally related groups. Interestingly, as shown in Figure 2.26, the families that exhibit strong

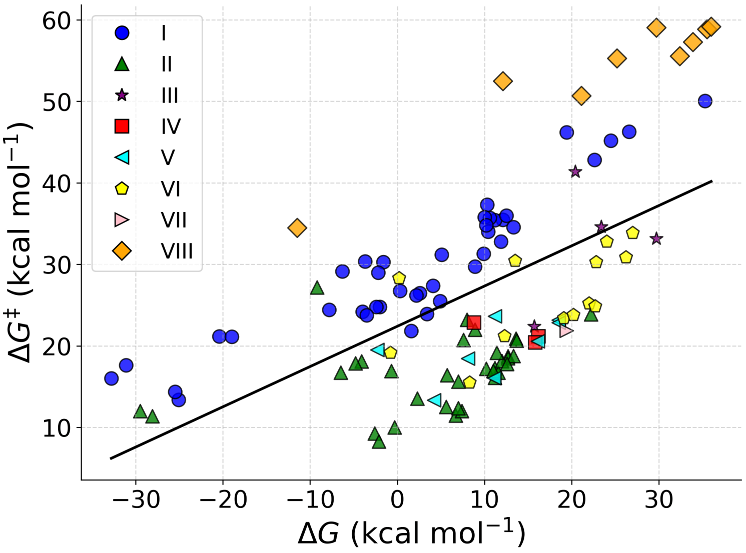


FIGURE 2.24: Scatter plot of reaction free-energies (ΔG) versus free-energy barriers ($\Delta G^{\ddagger}$) for H_2 splitting across the FLP library. Each point represents a distinct FLP and is colored according to the family it belongs to (see Figure 2.23). The black line represents the linear regression fit across the entire dataset (BEP correlation), with a r^2 = 0.37.

BEP relationships are characterized by having the H_2 splitting TSs in the dataset with H–H distances around 0.95 Å. This indicates that the TS is structurally close to the product and referred to as "late" TS, which according to Hammond's postulate, is also energetically close. These findings suggest that the nature of the molecular backbone plays a critical role in determining whether a given FLP proceeds via an

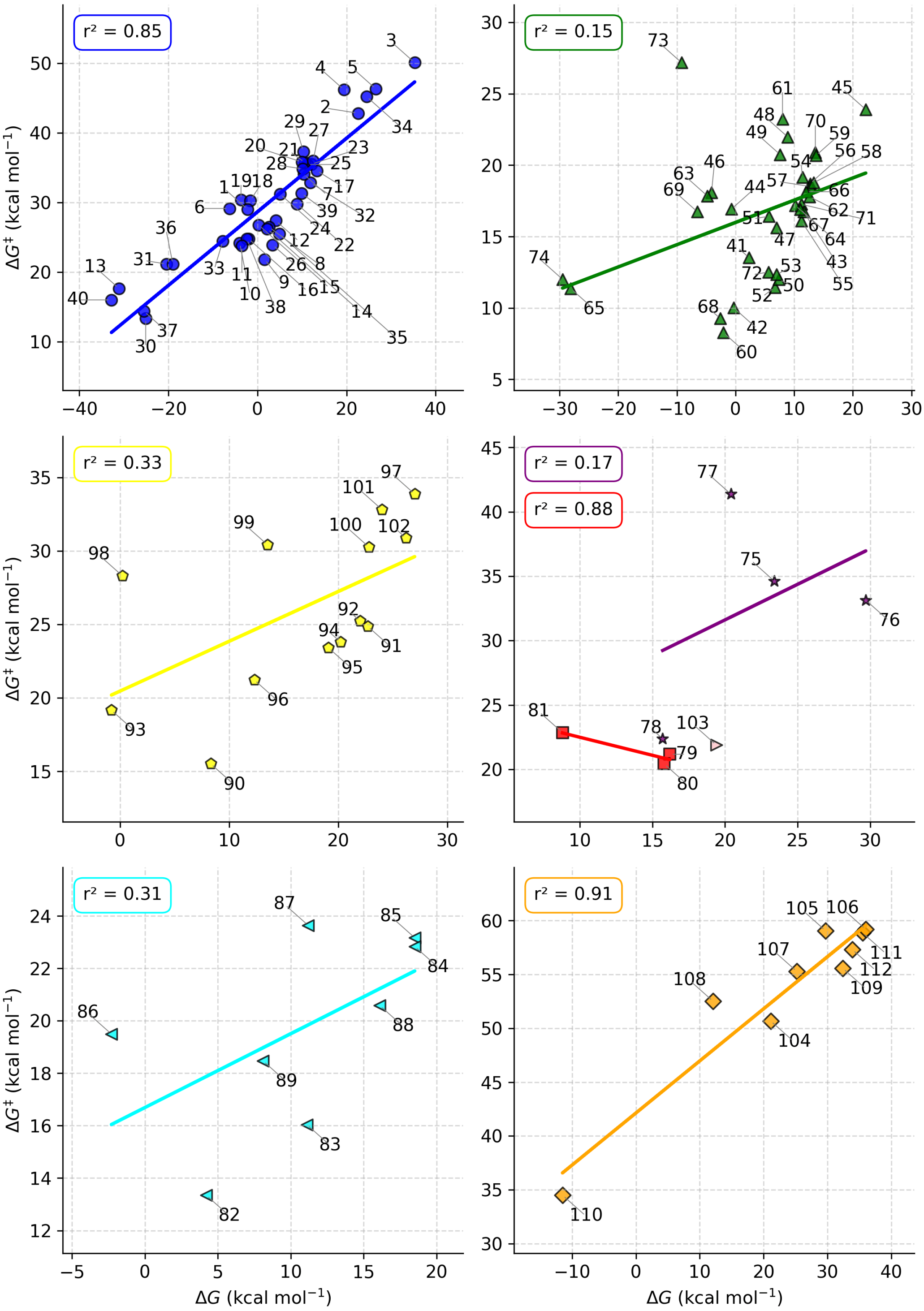


FIGURE 2.25: Correlation of ΔG vs ΔG‡ for the BEP of the FLPs included in the library by families.

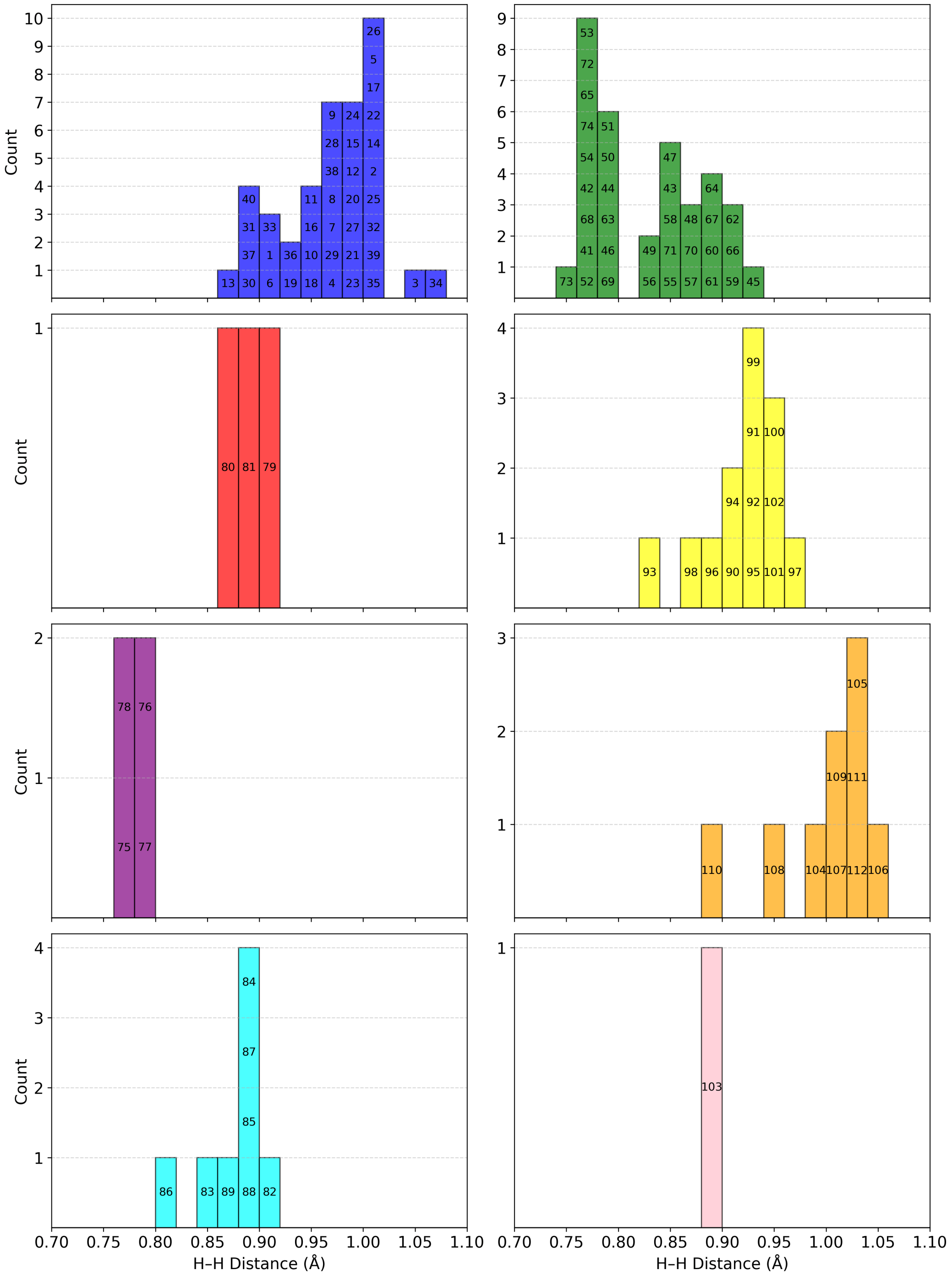


FIGURE 2.26: Histograms of the H–H distances of the converged transition states for each of the molecules of the library by families.

"early" (close to the reactant) or "late" (close to the product) TS and, consequently, whether the BEP principle holds or not. Note that Skara *et* al. were also able to distinguish between "early" and "late"-type TSs for H_2 splitting as function of the FLP features.[101]

Thus far, the structural and electronic properties of FLPs have been shown to strongly influence their reactivity, as further exemplified by the broad variability in H_2 splitting reaction energies and barriers observed in our library (Figure 2.24). However, there is still no clear consensus regarding which parameters govern the thermodynamics and kinetics of H_2 splitting, or the extent to which each contributes. Aiming to bridge this gap, we defined a set of chemically meaningful molecular

TABLE 2.1: List of descriptors used in this study along with their detailed descriptions. See Figure 2.28 for histograms showing the distribution of all descriptor values across the FLP library.

Descriptor	Description
$d^{[a]}$	Distance between LA and LB centres.
$\phi^{[a]}$	Angle between the vectors defining the direction of the p-type orbitals of both Lewis partners. Illustrated in Figure 2.27a.
γ	Dihedral angle between the directions of the p-type orbitals of each Lewis partner ($\vec{v}_B$–B–N–$\vec{v}_N$), shown in Figure 2.27b.
λ	Sum of the angles between each p-type orbital direction and the LA–LB vector (α and β in Figure 2.27a).
$E_{p(LA)}$	Orbital energy of the lowest unoccupied p-type orbital of LA, obtained from Natural Bond Orbital (NBO) analysis.
$E_{p(LB)}$	Orbital energy of the highest occupied p-type orbital of LB, obtained from NBO analysis.
$q_{la}(NPA)$	Atomic charge of LA calculated using Natural Population Analysis (NPA).
$q_{lb}(NPA)$	Atomic charge of LB calculated using NPA.
$EF_{la}(NPA)$	Electric field generated by $q_{la}(NPA)$.
$EF_{lb}(NPA)$	Electric field generated by $q_{lb}(NPA)$.
$EP(NPA)$	Electrostatic potential at the LA–LB midpoint, generated by NPA charges.
$q_{la}(ESP)$	Atomic charge of LA derived from the molecular electrostatic potential (CHelpG method).
$q_{lb}(ESP)$	Atomic charge of LB derived from the molecular electrostatic potential (CHelpG method).
$EF_{la}(ESP)$	Electric field generated by $q_{la}(ESP)$.
$EF_{lb}(ESP)$	Electric field generated by $q_{lb}(ESP)$.
$EP(ESP)$	Electrostatic potential at the LA–LB midpoint, generated by ESP (CHelpG) charges.
FEPA[a]	Free energy of proton attachment.
FEHA[a]	Free energy of hydride attachment.
M_w	Molecular weight.

[a] Previously used molecular descriptors in the field of FLP chemistry. [6, 55, 103, 109]

descriptors, readily obtainable from simple DFT calculations, to capture both the structural and electronic properties of FLPs.

These descriptors were subsequently applied in the following sections to establish clear relationships with the key energetic metrics of the H_2 splitting process, namely, ΔG and $\Delta G^{\ddagger}$. These descriptors are listed in Table 2.1.

Descriptors calculator

Before calculating any descriptor, the text format of the Gaussian output files is converted into a data format that can further be processed (see Section 2.2.2) and both LA and LB are automatically identified (see Section 2.2.3). The "desc_gen.py" script utilizes the data produced by the workflows described in Section 2.2.2 and Section 2.2.3 to systematically compute a set of descriptors (Table 2.1) for each FLP molecule in the library, including the B–N distance previously determined in Section 2.2.3.

The descriptors related to the conformation of the FLPs use the the vectors $\overrightarrow{v_B}$ and $\overrightarrow{v_N}$, which describe the orientation of the LB's lone pair and the LA's unoccupied p-type orbital, respectively (see Figure 2.27a). Using the coordinates and the connectivity, the first vector is calculated as the perpendicular vector to the plane defined by the coordinates of the three atoms directly bound to the B centre ($\overrightarrow{v_B}$ in Figure 2.27a). Then, the second is calculated the same way if N has three substituents. If the N centre has two substituents, the code calculates the vector perpendicular to the line defined by the two substituents of N in the plane that contains the N atom ($\overrightarrow{v_N}$ in Figure 2.27a). The angle ϕ corresponds to the angle between these two vectors. In addition, we propose two new descriptors, γ and λ, designed to evaluate the FLP molecules from a structural standpoint, accounting for the alignment and directionality of the p-type orbitals of the FLP partners, respectively. λ is calculated as the sum of the angle α, between $\overrightarrow{v_B}$ and $\overrightarrow{v_{BN}}$ (in black in Figure 2.27a), and β, between the vectors $\overrightarrow{v_N}$ and $\overrightarrow{v_{NB}}$. The dihedral angle (γ) is calculated between the plane containing the vectors $\overrightarrow{v_B}$ and $\overrightarrow{v_{BN}}$ and the plane containing the vectors $\overrightarrow{v_N}$ and $\overrightarrow{v_{NB}}$ (see Figure 2.27b).

To further account for the electronic properties of the LA and LB centres, we drew inspiration from two main reactivity models developed to explain FLP-mediated H_2 splitting explained in Section 2.1.2. Accordingly, the energy levels of these p-type orbitals ($E_{\mathrm{p(LA)}}$ and $E_{\mathrm{p(LB)}}$) serve as relevant descriptors for FLP reactivity. Here, using the identified LA and LB, the energies of the orbitals involved in the reactivity, are directly taken from the NBO information, implemented in Gaussian 16. In this context, the electric fields generated by the LA and LB centres may correlate with reactivity. We therefore considered, in first instance, the atomic charges on the LA and LB as potential descriptors. These were obtained using two methods: Natural Population Analysis (NPA), which assigns charges based on wavefunction localization, and Electrostatic Potential (ESP), which fits atomic charges to reproduce the molecular electrostatic potential. Using these atomic charges, we computed the electric

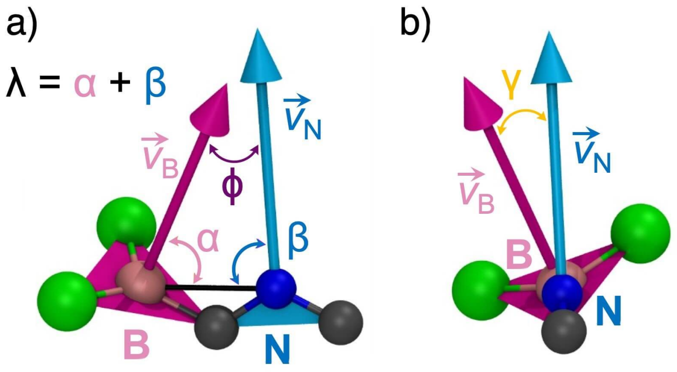


FIGURE 2.27: Graphical representation of the λ, Φ (a) and γ (b) geometric descriptors. B is colored in pink, with its orbital direction ($\overrightarrow{v_B}$) in magenta. N is colored in blue, with its orbital direction ($\overrightarrow{v_N}$) in cyan. Backbone atoms are colored in dark grey and substituents of the B centre in green. $\overrightarrow{v_B}$ is defined as the perpendicular vector to the plane formed by the three substituents of B that crosses the B centre; while $\overrightarrow{v_N}$ is the vector included in the plane formed in sp^2 N. In FLPs containing a sp^3 N, $\overrightarrow{v_N}$ is calculated as $\overrightarrow{v_B}$. λ is the sum of α and β angles; Φ corresponds to the angle between $\overrightarrow{v_B}$ and $\overrightarrow{v_N}$; and γ accounts for the dihedral angle formed between $\overrightarrow{v_B}$, B, N and $\overrightarrow{v_N}$.

field (EF) and Electrostatic Potential (EP) at the midpoint of the LA$\cdots$LB segment via Coulomb's law. The EF generated at the midpoint between the LA and the LB sites (r being thus 1/2 of the LA$\cdots$LB distance) is considered using the charges (Q) extracted previously from the DFT outputs of the FLP library (Eq. 2.10). The constant K was set to 1 for simplicity, any other value would only shift the data and the models were built using normalized descriptors anyway. Hence, the EF descriptor was calculated with Eq. 2.11.

$$EF = K \cdot \frac{Q}{r^2} \qquad (2.10) \qquad\qquad EF_{desc} = \frac{Q}{r^2} \qquad (2.11)$$

The electrostatic potential (EP) descriptor accounts for the Coulombic interaction between two charged centres (Q_{LA} and Q_{LB}) and a particle placed at the midpoint between these two (hence at distance r of both) obeying Eq. 2.12. Again, as done for the EF descriptor, the constant K is set to 1 so that the EP equation is simplified to Eq. 2.13.

$$EP = K \cdot \frac{Q_{LA} \cdot Q_{LB}}{r} \qquad (2.12) \qquad\qquad EP_{desc} = \frac{Q_{LA} \cdot Q_{LB}}{r} \qquad (2.13)$$

This gives rise to the q, EF, and EP descriptors listed in Table 2.1. The attachment free-energies (FEHA and FEPA) were calculated by means of DFT as the free-energy difference of the processes showed in Eq. 2.14, being R and R′ the rest of the molecule, and added externally to the final dataset of descriptors.

$$\begin{aligned} &\textbf{FEHA: } \text{R-LA} + \text{H}^- \rightarrow \text{R-LAH}^- \\ &\textbf{FEPA: } \text{R}'\text{-LB} + \text{H}^+ \rightarrow \text{R}'\text{-LBH}^+ \end{aligned} \tag{2.14}$$

Finally, we also considered the molecular weight of the FLP (M_w) as a potential 0-D descriptor, given that translational entropy depends on molecular weight and can therefore influence the height of free-energy barriers in bimolecular processes. Using the atomic masses of the `mendeleev` python module,[137] the atom types of the coordinates are used to calculate the molecular weight of the molecule.

The histograms of all the computed descriptors (Figure 2.28), confirm the variability of the explored chemical space, essential to perform data-driven analysis (see section 2.2.4).

All the computed descriptors, the response variables, the name of the files and the labelling used are stored in "final_data.csv" and "final_data.pkl" files.

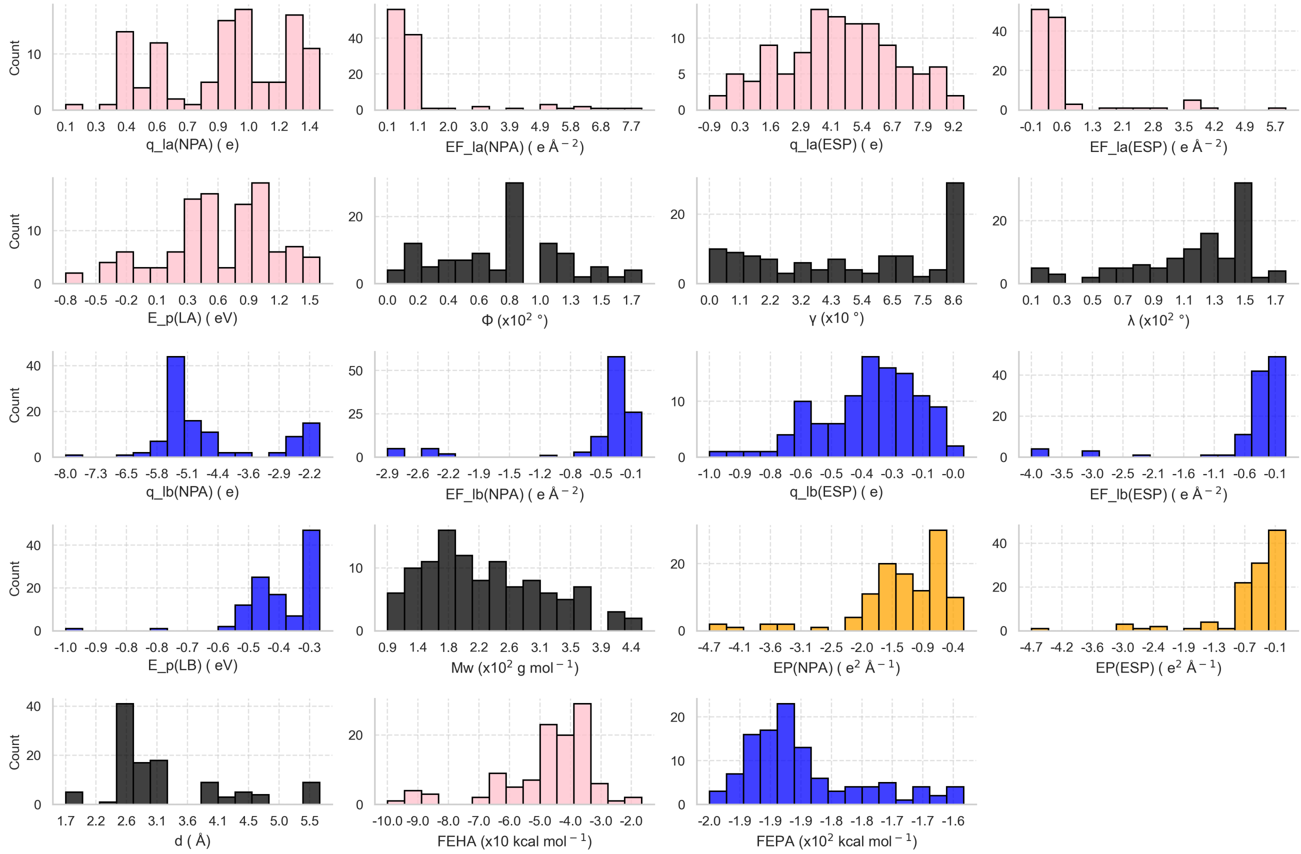
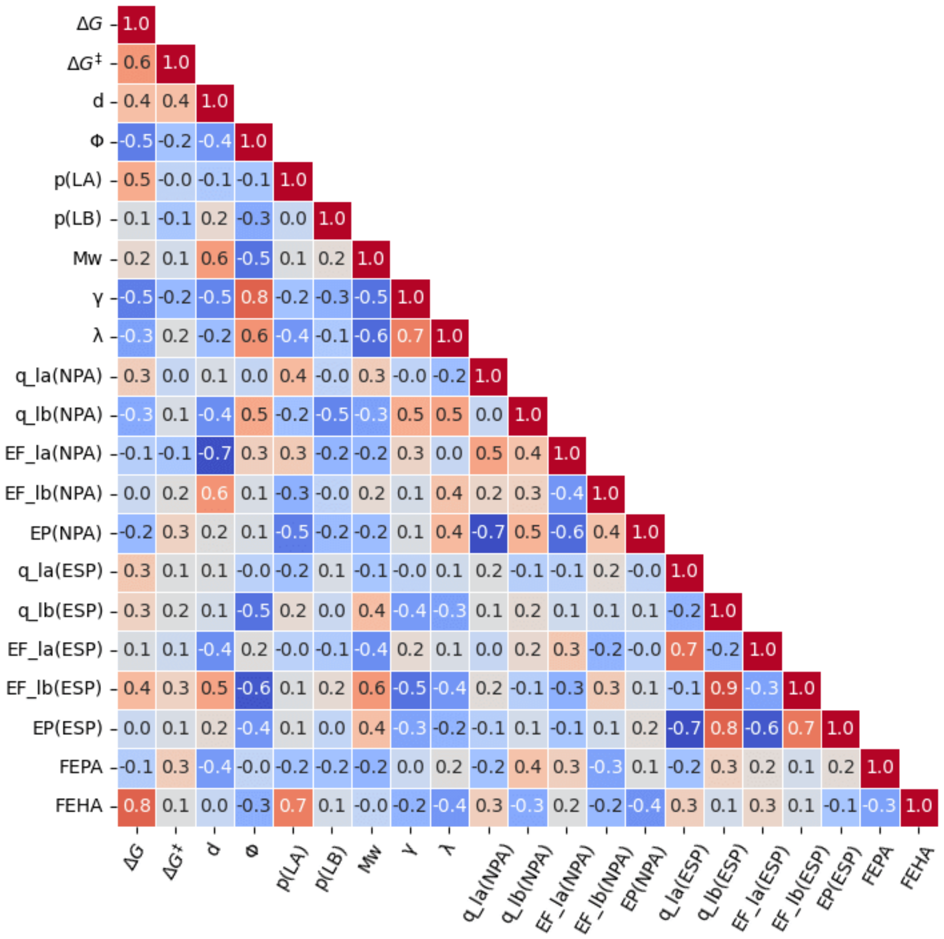


FIGURE 2.28: Histograms of the value of all the descriptors for all the FLP containing molecules of the library. LA related electronic descriptors in pink, structural descriptors in grey, LB related electronic descriptors in blue and electrostatic potential descriptors in orange.

2.3.2 Structure-activity relationships governing product thermodynamic stability (ΔG)

The correlation matrix shown in Figure 2.29 reveals weak individual correlations between the computed descriptors and both free-energy values, with the notable exception of FEHA, which shows a strong correlation with ΔG. This preliminary analysis suggests that the acidity of the LA, which influences the stabilization of the hydride formed upon heterolytic H_2 cleavage, plays a more significant role in determining the thermodynamics of H_2 splitting than the basicity of the LB. This observation is consistent with previous findings based on smaller datasets, [55, 108, 109, 112] confirming that the trend remains valid when extended to a broader and more diverse set of FLP families. As a first and simplest approach to predicting ΔG, we fitted a linear regression model using FEHA as the sole descriptor. The model's performance was evaluated using three metrics: the coefficient of determination (r^2), which measures the goodness of fit; the q^2 coefficient, which assesses predictive ability, and the root mean squared error (RMSE), which reflects the average prediction



	ΔG	ΔG‡	d	Φ	p(LA)	p(LB)	Mw	γ	λ	q_la(NPA)	q_lb(NPA)	EF_la(NPA)	EF_lb(NPA)	EP(NPA)	q_la(ESP)	q_lb(ESP)	EF_la(ESP)	EF_lb(ESP)	EP(ESP)	FEPA	FEHA
ΔG	1.0																				
ΔG‡	0.6	1.0																			
d	0.4	0.4	1.0																		
Φ	-0.5	-0.2	-0.4	1.0																	
p(LA)	0.5	-0.0	-0.1	-0.1	1.0																
p(LB)	0.1	-0.1	0.2	-0.3	0.0	1.0															
Mw	0.2	0.1	0.6	-0.5	0.1	0.2	1.0														
γ	-0.5	-0.2	-0.5	0.8	-0.2	-0.3	-0.5	1.0													
λ	-0.3	0.2	-0.2	0.6	-0.4	-0.1	-0.6	0.7	1.0												
q_la(NPA)	0.3	0.0	0.1	0.0	0.4	-0.0	0.3	-0.0	-0.2	1.0											
q_lb(NPA)	-0.3	0.1	-0.4	0.5	-0.2	-0.5	-0.3	0.5	0.5	0.0	1.0										
EF_la(NPA)	-0.1	-0.1	-0.7	0.3	0.3	-0.2	-0.2	0.3	0.0	0.5	0.4	1.0									
EF_lb(NPA)	0.0	0.2	0.6	0.1	-0.3	-0.0	0.2	0.1	0.4	0.2	0.3	-0.4	1.0								
EP(NPA)	-0.2	0.3	0.2	0.1	-0.5	-0.2	-0.2	0.1	0.4	-0.7	0.5	-0.6	0.4	1.0							
q_la(ESP)	0.3	0.1	0.1	-0.0	-0.2	0.1	-0.1	-0.0	0.1	0.2	-0.1	-0.1	0.2	-0.0	1.0						
q_lb(ESP)	0.3	0.2	0.1	-0.5	0.2	0.0	0.4	-0.4	-0.3	0.1	0.2	0.1	0.1	0.1	-0.2	1.0					
EF_la(ESP)	0.1	0.1	-0.4	0.2	-0.0	-0.1	-0.4	0.2	0.1	0.0	0.2	0.3	-0.2	-0.0	0.7	-0.2	1.0				
EF_lb(ESP)	0.4	0.3	0.5	-0.6	0.1	0.2	0.6	-0.5	-0.4	0.2	-0.1	-0.3	0.3	0.1	-0.1	0.9	-0.3	1.0			
EP(ESP)	0.0	0.1	0.2	-0.4	0.1	0.0	0.4	-0.3	-0.2	-0.1	0.1	-0.1	0.1	0.2	-0.7	0.8	-0.6	0.7	1.0		
FEPA	-0.1	0.3	-0.4	-0.0	-0.2	-0.2	-0.2	0.0	0.2	-0.2	0.4	0.3	-0.3	0.1	-0.2	0.3	0.2	0.1	0.2	1.0	
FEHA	0.8	0.1	0.0	-0.3	0.7	0.1	-0.0	-0.2	-0.4	0.3	-0.3	0.2	-0.2	-0.4	0.3	0.1	0.3	0.1	-0.1	-0.3	1.0

FIGURE 2.29: Scatter matrix of the correlation coefficient (R) of all the descriptors, ΔG and ΔG‡.

error in the units of the response variable (kcal mol^{-1}). The last two metrics are obtained via K-fold cross-validation as explained in Section 2.2.4. The resulting model yielded r^2, q^2, and RMSE values of 0.62, 0.60, and 8.90 kcal mol^{-1}, respectively. This confirms that while FEHA can qualitatively account for the variation in ΔG, a more accurate prediction likely requires incorporating multiple descriptors.

Multivariate linear regression (MLR)

Multivariate linear regression (MLR) models were next used to evaluate how multiple molecular descriptors jointly influence the response variable (ΔG). MLR assigns weights to each descriptor in a linear equation, optimizing the fit by minimizing the error between predicted and measured values. Initially, all descriptors listed in Table 2.1 were included to improve the model's accuracy. Since descriptors vary in scale, each of them was normalized to a 0 – 1 range by subtracting its minimum value and dividing the result by the maximum value in the range. The resulting normalized dataset was then used to train the MLR model and therefore, the coefficients of the normalized descriptors in the model can be interpreted as the relative influence that the descriptor has on the response variable (Figure 2.30).

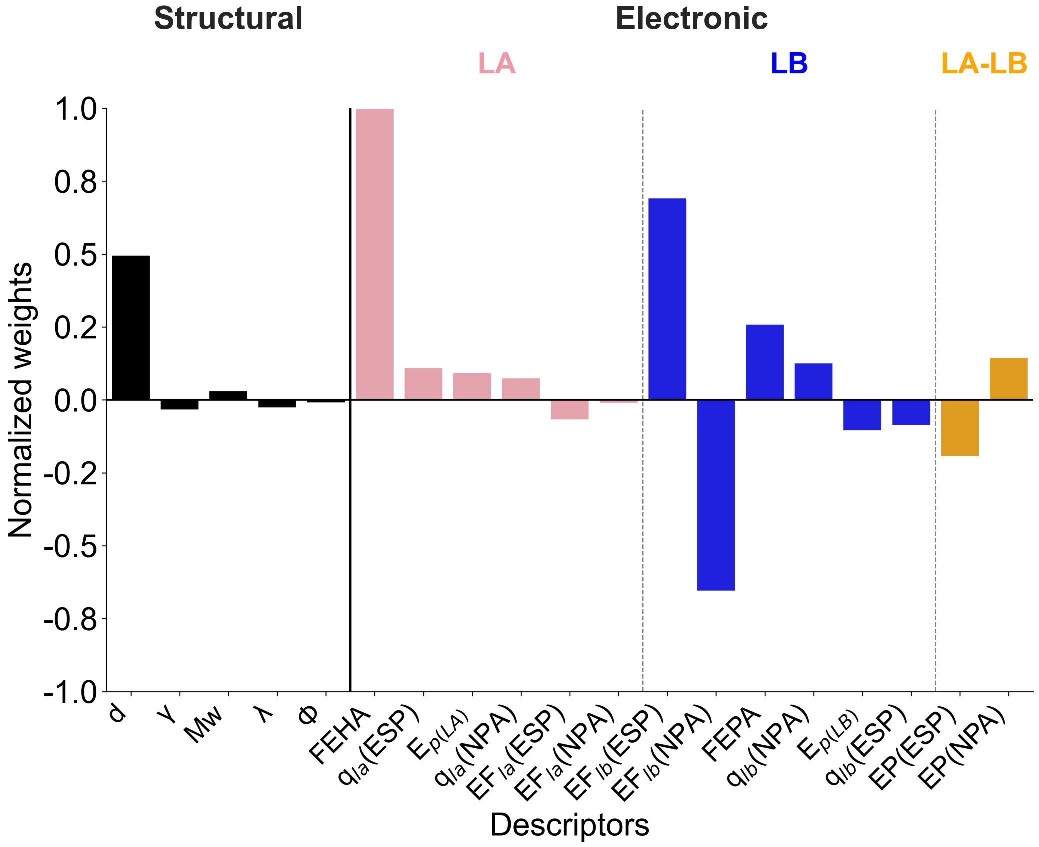


FIGURE 2.30: Coefficients of the normalized MLR model using all the descriptors for ΔG. $r^2 = 0.93$, $q^2 = 0.84$ and RMSE = 5.50 kcal mol^{-1}.

The improved quality metrics (caption of Figure 2.30) demonstrate that MLR offers significantly better fitting and predictive performance than the single-descriptor model. As expected, FEHA emerged as the most influential descriptor, followed by the LA$\cdots$LB distance (d) and EF_{lb}. Notably, the EF descriptors calculated using either NPA or CHelpG charge schemes appear in the model with similar weights but opposite signs, effectively cancelling each other and complicating their chemical interpretation. Based on this, we next built a "simplified" MLR model using only d, FEPA and FEHA, aiming to maintain representativeness while reduce overfitting, thereby improving predictive ability. The resulting model is given in Eq. 2.15. The first line, with tilded descriptors, corresponds to the model built using normalized descriptors, where the weights (coefficients) can be meaningfully compared. In the second line, where the descriptors are not tilded, the model uses the original, non-normalized values, allowing direct calculation of the response variable. The bottommost line presents the fitting and validation metrics. The coefficients of the normalized variables indicate that all three descriptors contribute comparably to the reaction energy, with FEHA having the greatest influence, as expected.

$$\begin{aligned}\Delta G &= 28.9 \cdot \tilde{d} + 67.6 \cdot \widetilde{\mathrm{FEHA}} + 24.4 \cdot \widetilde{\mathrm{FEPA}} - 54.4 \\ \Delta G &= 7.37 \cdot d + 0.821 \cdot \mathrm{FEHA} + 0.685 \cdot \mathrm{FEPA} + 151 \\ r^2 = 0.87 \qquad q^2 &= 0.85 \qquad \mathrm{RMSE} = 5.52\ \mathrm{kcal} \cdot \mathrm{mol}^{-1}\end{aligned} \tag{2.15}$$

The positive coefficients of FEHA and FEPA are consistent with the idea that weaker hydride and proton affinities (i.e., more positive values), make the H_2 splitting less favourable (more positive ΔG), due to lower acidity and basicity of the FLP sites. The positive coefficient of d may reflect the contribution of attractive electrostatic interactions between LA–H^- and $^+$H–LB groups in the zwitterionic product, which become stronger at shorter distances, thereby stabilizing the H_2 splitting product and lowering ΔG. These findings are consistent with previous observations by Pápai and co-workers, who showed that H_2 splitting becomes more exergonic with increasing acidity and basicity of the FLP partners, and with decreasing donor–acceptor distance.[40] However, Eq. 2.15 provides the first near-quantitative model that explicitly assesses the individual contribution of each parameter to the reaction free energy. In terms of model performance, the simplified model loses correlation (r^2 of 0.93 vs 0.87 for simplified and non-simplified models, respectively) while conserving the exact same predictability (q^2 of 0.85 for both models), indicating that the excluded descriptors added noise rather than value, potentially leading to overfitting. The MLR model presented in Eq. 2.15 not only highlights the relative contributions of key FLP parameters, acidity, basicity, and the LA$\cdots$LB distance, to their H_2 activation ability, but also functions as a simple and practical quantitative structure–activity relationship (QSAR) model, with a reasonable RMSE at 5.0 kcal mol^{-1}.

Notably, models constructed using only two of these descriptors perform substantially worse than the three-dimensional model shown in Eq. 2.15.

$$\begin{aligned}\Delta G &= 64.4 \cdot \widetilde{\text{FEHA}} + 12.3 \cdot \widetilde{\text{FEPA}} - 36.8\\ \Delta G &= 0.781 \cdot \text{FEHA} + 0.346 \cdot \text{FEPA} + 110\\ r^2 &= 0.66 \qquad q^2 = 0.63 \qquad \text{RMSE} = 8.62\ \text{kcal mol}^{-1}\end{aligned} \tag{2.16}$$

$$\begin{aligned}\Delta G &= 20.7 \cdot \widetilde{d} + 58.1 \cdot \widetilde{\text{FEHA}} - 36.3\\ \Delta G &= 5.28 \cdot d + 0.705 \cdot \text{FEHA} + 24.9\\ r^2 &= 0.75 \qquad q^2 = 0.73 \qquad \text{RMSE} = 7.37\ \text{kcal mol}^{-1}\end{aligned} \tag{2.17}$$

Still, when FEHA is paired with either FEPA (Eq. 2.16) or d (Eq. 2.17), the resulting models still show reasonable correlation (r^2 up to 0.75) and screening power (q^2 up to 0.73). In contrast, models using only FEPA and d (Eq. 2.18) fail to capture any meaningful relationship, further highlighting the key role of FEHA in determining the reaction thermodynamics.

$$\begin{aligned}\Delta G &= 24.3 \cdot \widetilde{d} + 4.58 \cdot \widetilde{\text{FEPA}} - 2.87\\ \Delta G &= 6.21 \cdot d + 0.129 \cdot \text{FEPA} + 12.2\\ r^2 &= 0.16 \qquad q^2 = 0.09 \qquad \text{RMSE} = 13.52\ \text{kcal mol}^{-1}\end{aligned} \tag{2.18}$$

This is consistent with the scatter plot in Figure 2.31 and Figures 2.32, 2.33 and 2.34, which show that only FEHA can qualitatively categorize FLPs based on their ability to stabilize the H_2 splitting product (i.e., their ΔG values), whereas the other descriptors cannot.

With the insights gained from Eq. 2.15, the computed ΔG values for the FLPs in our dataset can be rationalized based on the magnitudes of their associated molecular descriptors. As shown in Figures 2.31 and 2.32, a group of FLPs featuring strongly electron-withdrawing substituents on the LA fall within FEHA values ranging from –100 to –80 kcal mol^{-1}, leading to highly favourable H_2 splitting. These include –NO_2 substituents in molecules **13**, **40**, **65**, and **74**; –CN in **31**; –CCl_3 in **30**; and –CF_3 in **37** and **110**. FLPs with FEHA values between –80 and –60 kcal mol^{-1} mainly contain moderately electron-withdrawing groups, resulting in slightly less favourable, but still exergonic, H_2 splitting. Examples include –COOH in **36**, or –Cl in molecules **6**, **19**, **46**, **63**, **69**, **86**, **93**, and **98**. Still, some structures with strong electron-withdrawing groups, such as in **33** and **73**, containing –C_6F_5 and –CF_3, respectively, were also found to fall within this range. The remaining molecules, with FEHA values between –60 and –20 kcal mol^{-1}, typically feature aliphatic substituents and cluster in a region characterized by endergonic H_2 splitting. Notably, molecule **110**, despite lying within the highly stabilizing FEHA range (–100 to –80 kcal mol^{-1}), shows a

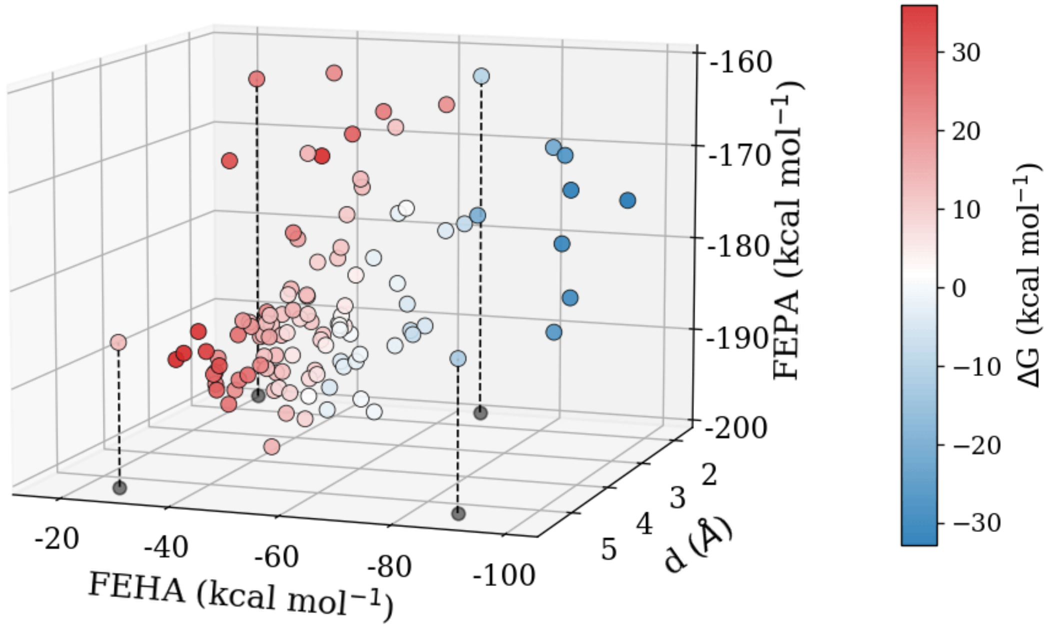


FIGURE 2.31: Scatter plot of the data points projected on the three-dimensional space of d, FEHA and FEPA. Each point is colored according to its DFT-calculated ΔG value. Four selected data points are projected on the FEHA vs d plane to illustrate the stronger discrimination power of FEHA compared to the other descriptors.

higher ΔG than its neighbors in FEHA/FEPA space due to its longer LA$\cdots$LB distance imposed by the molecular framework (see Figure 2.33c,e), which reduces electrostatic stabilization of the zwitterionic product. Similarly, although molecules **4** and **73** have comparable FEHA and FEPA values (Figure 2.32c), **73** exhibits a more exergonic H_2 splitting due to a shorter LA$\cdots$LB distance (Figure 2.33e). Another interesting example is the difference in ΔG between molecules **4** (ΔG = +19.4 kcal mol^{-1}) and **19** (ΔG = –3.7 kcal mol^{-1}), which exhibit similar FEHA and d values (Figure 2.33e). This disparity can be ascribed to the stronger basicity of the LB (more negative FEPA) in the FLP **19** (Figure 2.32c). Overall, these observations confirm that the nature of the LA substituents, directly affecting its acidity, plays the dominant role in stabilizing the H_2 splitting product; that is, higher acidity correlates with more exergonic reactions. Shorter LA$\cdots$LB distances and stronger LB basicity also favour H_2 splitting from a thermodynamic perspective, albeit to a lesser extent.

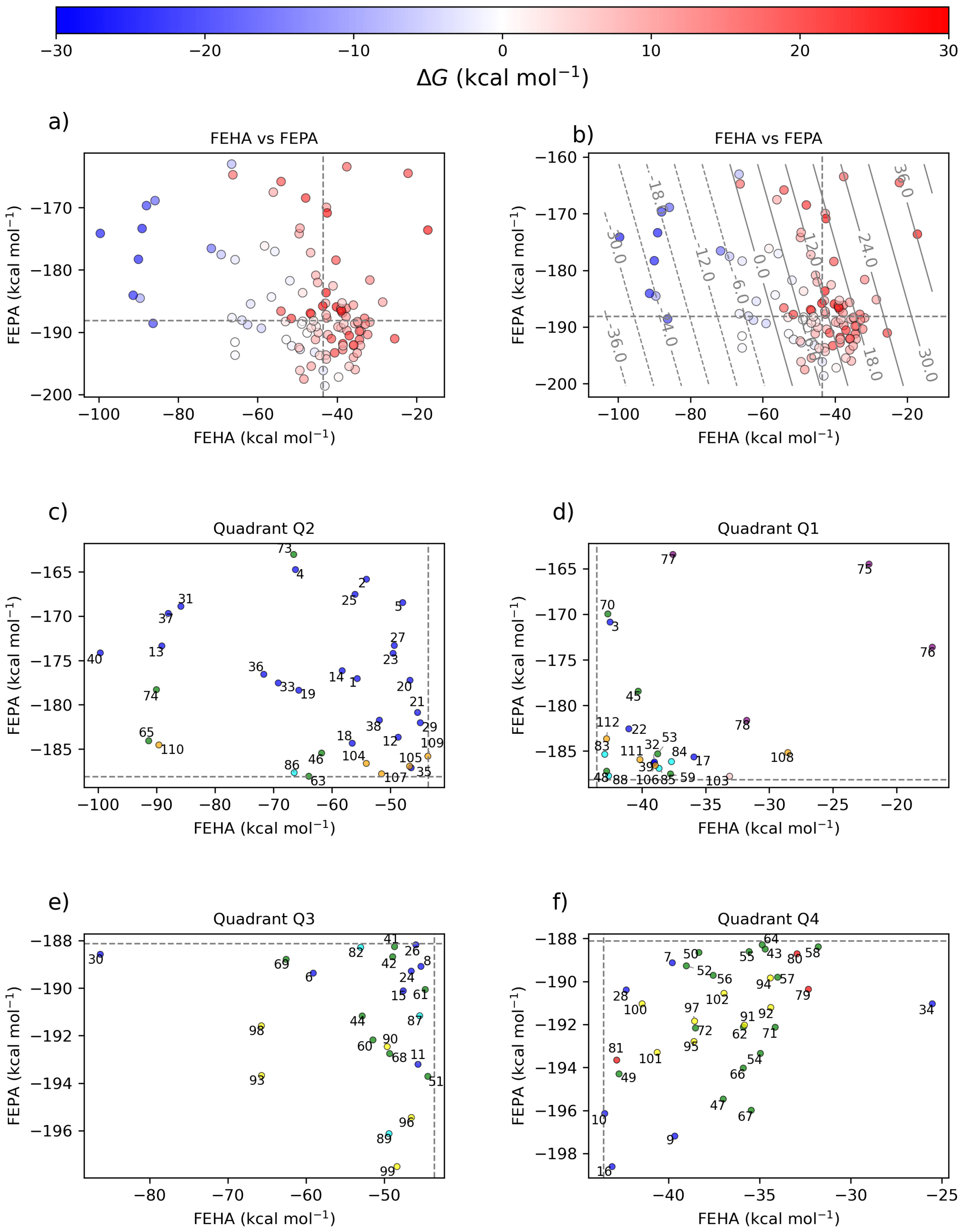


FIGURE 2.32: a) Scatter plot of the dataset projected on the FEHA vs FEPA space. Data points are coloured according to their ΔG values. b) The same showing isovalues of the predicted response variable (kcal mol^{-1}) holding d to the mean value of the dataset. c-f) Zoomed quadrants of graph a) labelled according to Figure 2.23.

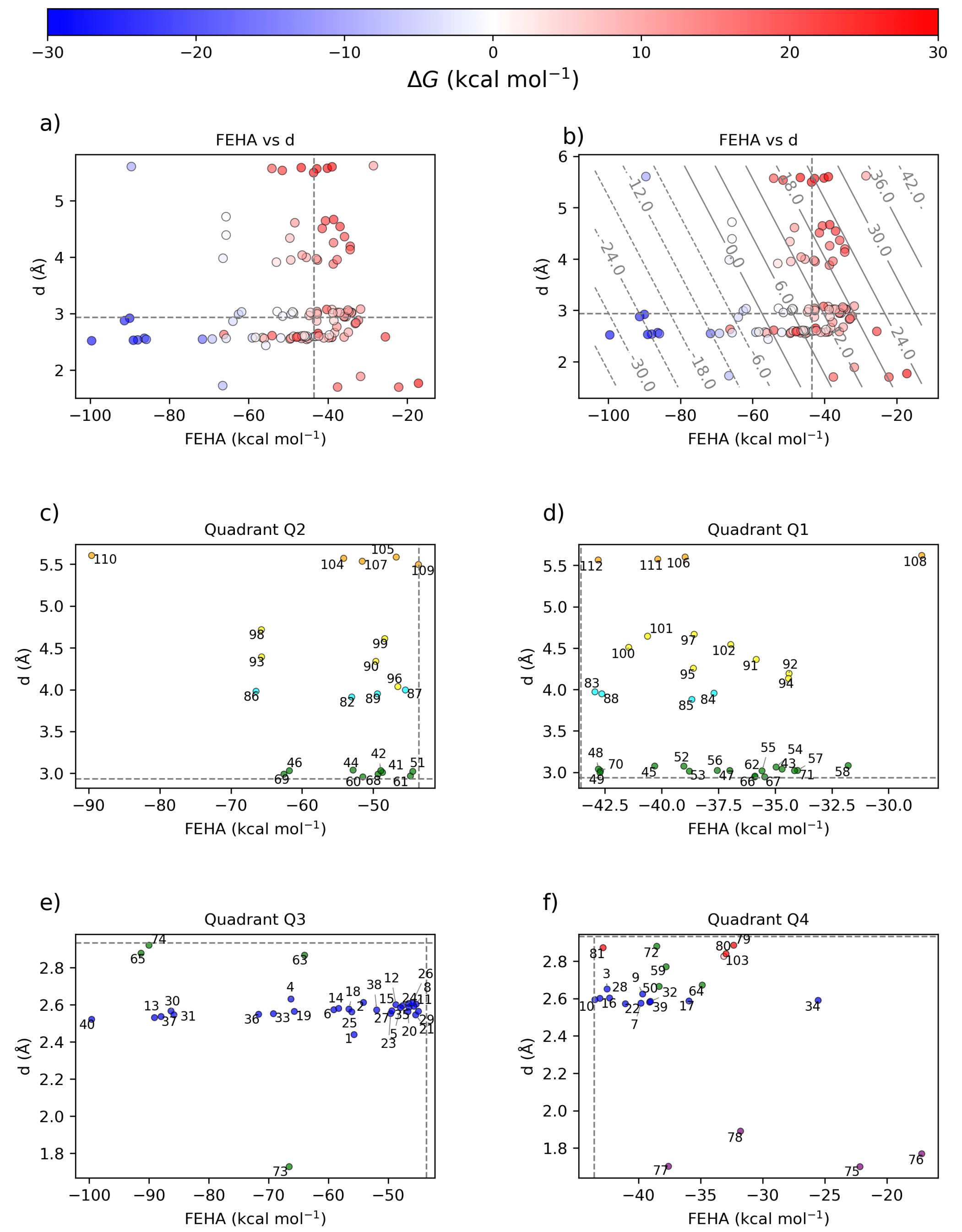


FIGURE 2.33: a) Scatter plot of the dataset projected on the d vs FEHA space. Data points are coloured according to their ΔG values. b) The same showing isovalues of the predicted response variable (kcal mol^{-1}) holding d to the mean value of the dataset. c-f) Zoomed quadrants of graph a) labeled according to Figure 2.23.

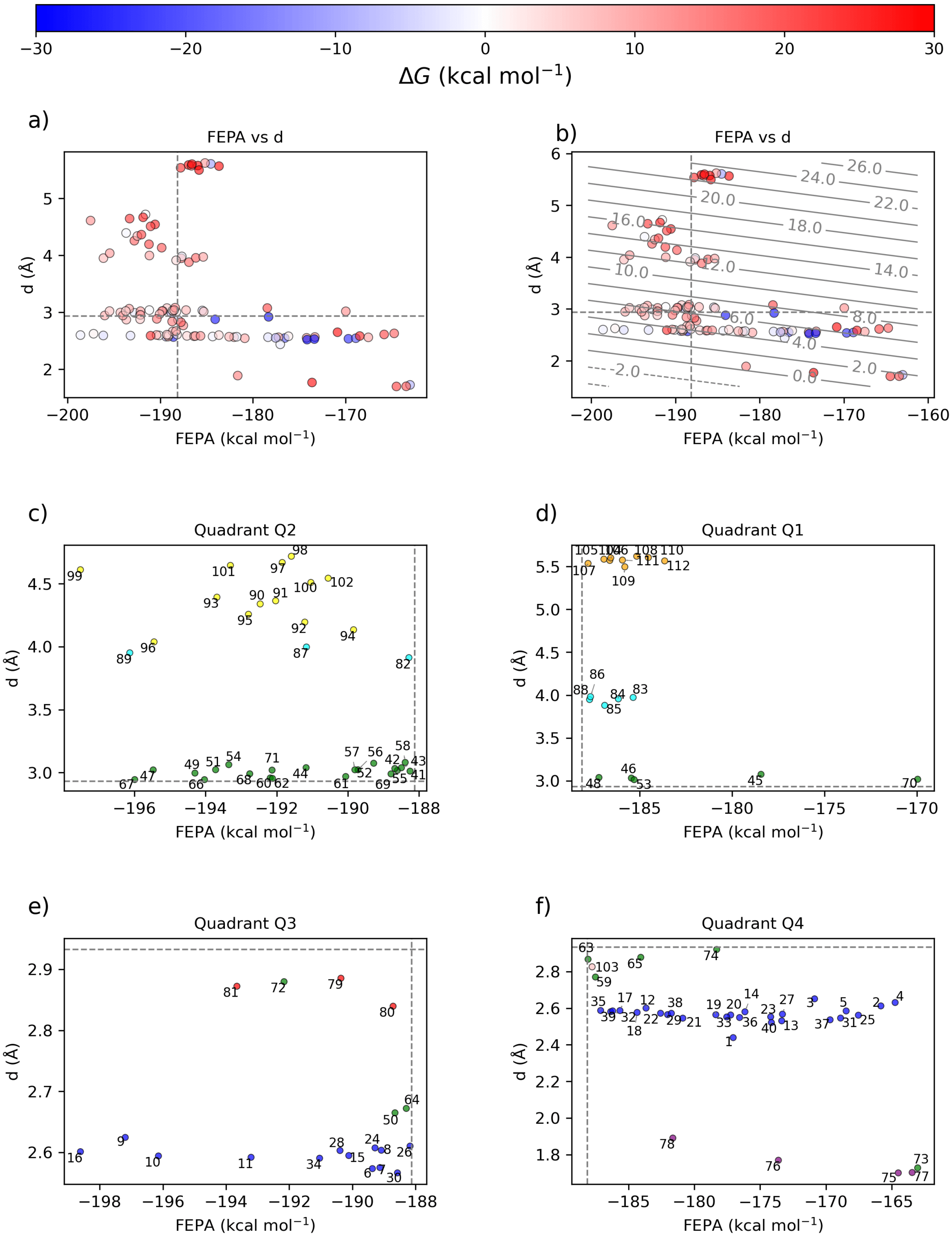


FIGURE 2.34: a) Scatter plot of the dataset projected on the d vs FEPA space. Data points are coloured according to their ΔG values. b) The same showing isovalues of the predicted response variable (kcal mol^{-1}) holding d to the mean value of the dataset. c-f) Zoomed quadrants of graph a) labelled according to Figure 2.23.

2.3.3 Structure–activity relationships governing activation free-energy ($\Delta G^{\ddagger}$)

In addition to an accessible reaction free energy, FLP-promoted H_2 splitting must also proceed through surmountable free-energy barriers ($\Delta G^{\ddagger}$) to be kinetically feasible. We therefore next aimed to develop mathematical models relating the descriptors listed in Table 2.1 with the calculated H_2 splitting barriers, in order to identify which features of FLP structure most strongly influence the reaction kinetics. None of the individual descriptors exhibit a clear correlation with $\Delta G^{\ddagger}$ (Figure 2.29), suggesting that no single parameter can independently account for the activation barriers across the FLP dataset. To gain initial insights, we constructed an MLR model using all normalized descriptors, following the same approach as for ΔG. However, the resulting model, whose coefficients and quality metrics are given in Figure 2.35, shows a lower accuracy and predictive power than the corresponding model for ΔG. This initial model can only capture linear relationships between the descriptors and $\Delta G^{\ddagger}$. However, the true dependence may involve non-linear trends, including the existence of optimal values for certain descriptors. Previous studies have identified such behaviour for the LA···LB distance (Vankova *et al.* [102] and Corminbœuf *et*

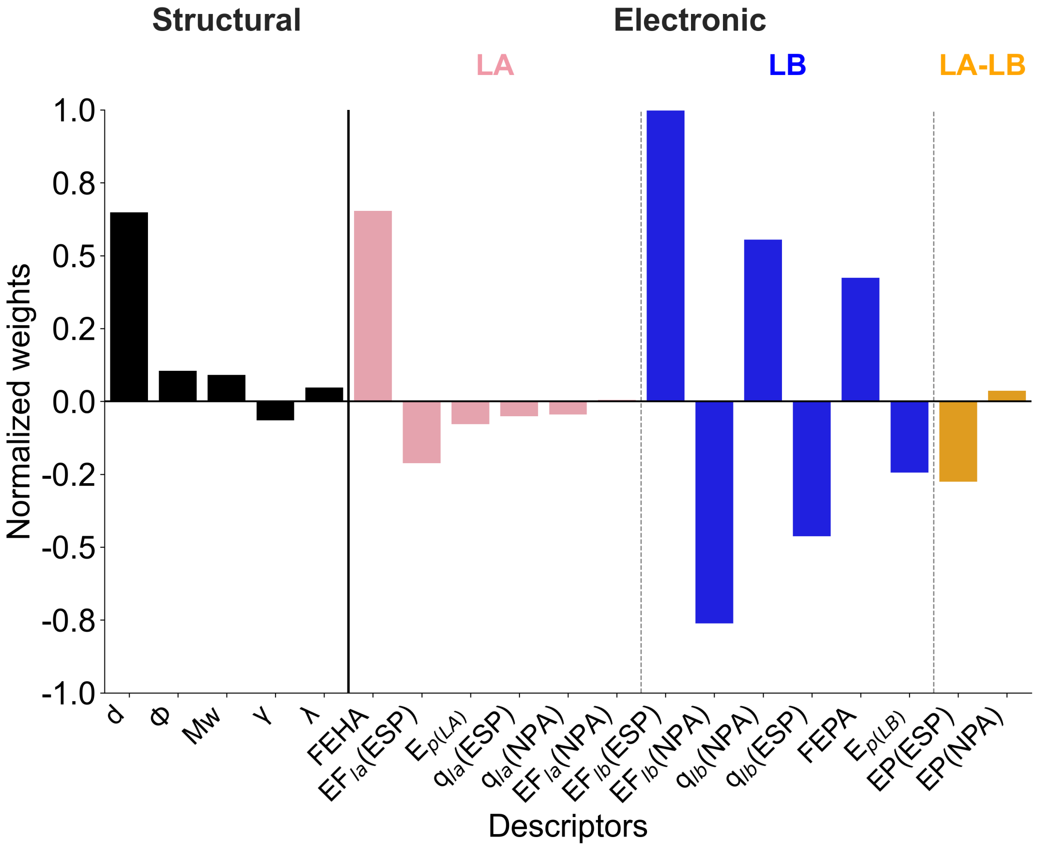


FIGURE 2.35: Normalized weights of the MLR model for predicting $\Delta G^{\ddagger}$ built using all the descriptors. Quality metrics: r^2 = 0.75, q^2 = 0.43 and RMSE = 8.92 kcal mol $^{-1}$.

al. [103], see Section 2.1.2), finding minimized H_2 splitting barriers at distances between 3 and 5 Å. These findings suggest that more sophisticated, non-linear models may be necessary to accurately capture the factors influencing $\Delta G^{\ddagger}$. To this end, we incorporated the squared values of the molecular descriptors as an additional set of features, enabling the model to account for potential quadratic relationships, thus providing it with more flexibility. The data_expander.py code is created to calculate the square values of the descriptors in final_data.pkl and append them into the new file final_data_expanded.pkl. Nevertheless, further increasing the number of variables certainly increases the risk of overfitting, given the size of the dataset used to train these models.

To maintain the necessary model flexibility while minimizing overfitting, we adopted a data-driven approach that combines the Partial Least Squares (PLS) method, [123] automated selection of informative descriptors and combinatorial analysis (see Section 2.2.4) to systematically explore dimensionality reduction. The resulting best-performing model, given in Eq. 2.19 represents a MLR model build with 9 descriptors, which exhibits reasonably good fitting, along with predictive ability provided a q^2>0.8 and a RMSE in barrier heights below 6.0 kcal mol^{-1}, obtained from cross validation.

$$
\begin{aligned}
&\Delta G^{\ddagger} = 161 \cdot \widetilde{d^2} - 133 \cdot \widetilde{d} + 51.4 \cdot \widetilde{\lambda^2} - 38.8 \cdot \widetilde{\lambda} - 21.9 \cdot \widetilde{\mathrm{FEHA}^2} + 9.32 \cdot \widetilde{\mathrm{FEPA}} \\
&-18.2 \cdot \widetilde{q_{lb}(\mathrm{NPA})^2} - 12.4 \cdot \widetilde{\mathrm{EP(NPA)}^2} - 20.0 \cdot \widetilde{\mathrm{EP(NPA)}} + 55.4 \\
&\Delta G^{\ddagger} = 5.60 \cdot d^2 - 34.0 \cdot d + 0.00159 \cdot \lambda^2 - 0.222 \cdot \lambda - 0.00228 \cdot \mathrm{FEHA}^2 \\
&+0.262 \cdot \mathrm{FEPA} - 30.2 \cdot q_{lb}(\mathrm{NPA})^2 - 57.4 \cdot \mathrm{EP(NPA)}^2 - 45.1 \cdot \mathrm{EP(NPA)} + 131 \\
&r^2 = 0.85 \qquad q^2 = 0.81 \qquad \mathrm{RMSE} = 5.13\ \mathrm{kcal\,mol^{-1}}
\end{aligned}
\tag{2.19}
$$

Interestingly, the weights of the descriptors in this model (illustrated graphically in Figure 2.36) indicate that the energy barrier is primarily governed by the geometry of the FLP (black bars), rather than its electronic properties. The largest coefficients are related with the LA$\cdots$LB distance, being positive for its squared value and negative for its linear one. This agrees with previous reports indicating that there is an optimal distance that minimizes the height of the barrier; as too short or too long distances may lead either to strained TS geometries or to poor orbital overlap, respectively. The second most influential descriptor is λ, which also exhibits a similar quadratic relationship with the barrier height (Figure 2.36). This reveals that, in addition to an optimal LA$\cdots$LB distance, an optimal orbital directionality maximizing the interaction with H_2, is also required to lower the H_2 activation barrier. Importantly, these two features, and thus the reaction kinetics, are mainly dictated by the structure of the FLP backbone, i.e., the molecular framework in which both FLP sites are embedded. Electronic effects, captured by FEHA^2, FEPA, q_{lb}(NPA)^2, EP(NPA) and EP(NPA)^2 in our model, were found to play a secondary role, suggesting that efforts to accelerate H_2 splitting in FLP systems should prioritize the design of the scaffold, rather than focusing on tuning the electronic properties of the LA and LB centres through substituent screening. Also, the fact that the $\Delta G^{\ddagger}$ is mostly governed by structural features, whereas ΔG is more strongly influenced by electronic effects might explain the lack of general BEP relationships, as observed in Figure 2.24.

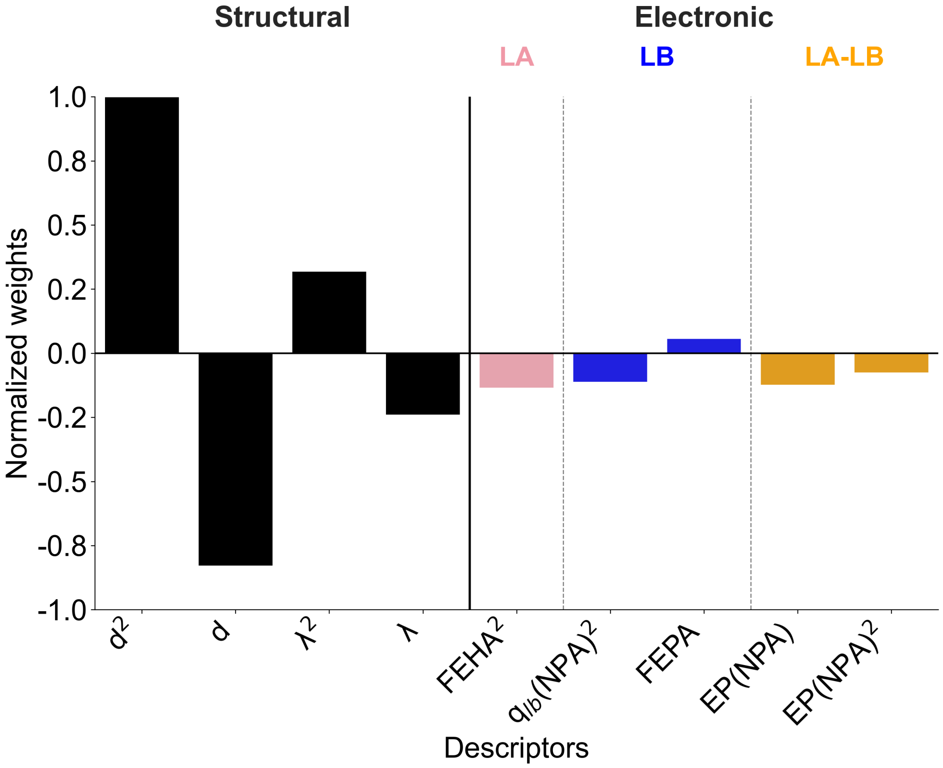


FIGURE 2.36: Normalized weights of the descriptors influencing the H_2 splitting barrier according to the MLR model given in Eq. 2.19. The signs of the weights indicate whether the relationships are directly or inversely proportional.

As shown in Figure 2.37, projecting our data points ($\Delta G^{\ddagger}$ values) onto the d vs. λ two-dimensional space reveals clustering by FLP families, as defined in Figure 2.23. This analysis shows that, regardless of the electronic properties of the LA and LB sites, families II, IV, and V consistently exhibit the lowest barriers due to optimal values of both d vs. λ. In fact, all five FLP systems with the lowest calculated barriers are, in order of increasing values, **60**, **68**, **42**, **65**, and **52**, all them belonging to the same family (see Figure 2.23). Conversely, although the FLPs in family I possess an optimal LA$\cdots$LB distance, they display systematically higher barriers than FLPs with similar electronic properties belonging to the aforementioned families due to a suboptimal orbital directionality (λ), which forces significant structural rearrangement to reach the TS structure. At the other extreme, families VI and VIII present favourable or near-optimal λ values but suffer from excessively long LA$\cdots$LB distances, introducing strain and poor orbital overlap into the TS geometry. Plotting the response variable predicted by the model against d and λ, while holding all other descriptors at their mean values across the dataset (Figure 2.37), the blue-colored region reveals that the barrier height is indeed minimized at optimal d values around

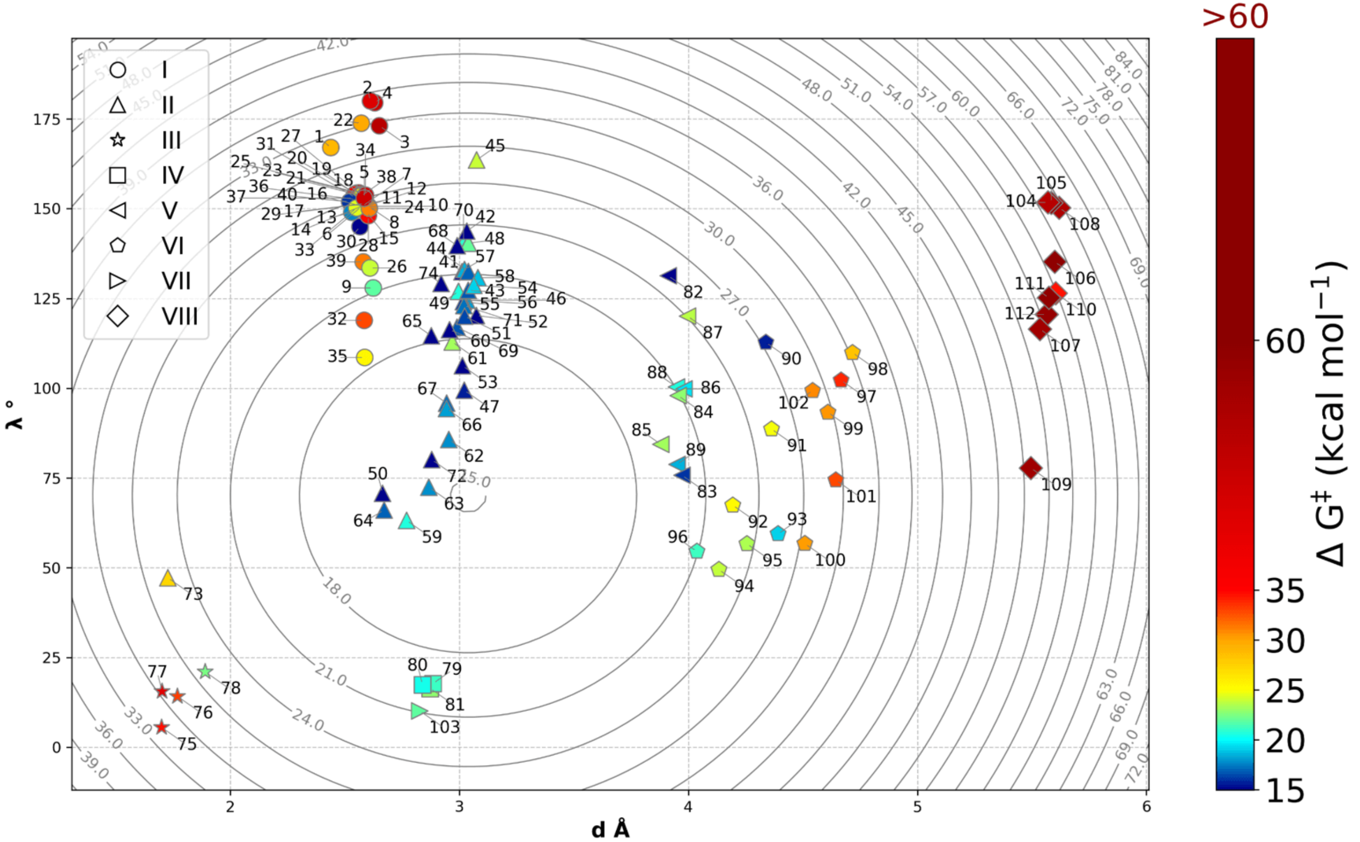


FIGURE 2.37: Scatter plot of the dataset projected on the d vs λ space, being the most important descriptors of the optimized model predicting $\Delta G^{\ddagger}$. Isovalue lines represent the $\Delta G^{\ddagger}$ as predicted by the model upon holding all the other descriptors to their mean value in the dataset. FLPs classified according their family as shown in the legend and colored according their DFT-calculated $\Delta G^{\ddagger}$ shown in the color-bar at the right.

3.0 Å, in excellent agreement with previous reports,[103] and at λ values around 70°. Together, these structural features can be considered practical design rules that may significantly aid future FLP screening efforts. Nonetheless, beyond these structural effects, the electronic properties of the LA and LB sites do play a secondary role (Figure 2.36). Indeed, the influence of electronic descriptors is further evidenced by the broad range of barrier heights observed within individual FLP families (Figure 2.24). Therefore, electronic properties should also be considered for fine-tuning FLP reactivity, becoming particularly important to modulate the reaction thermodynamics, as discussed in the previous section.

The above observations indicate that the main factor influencing the height of the H_2 splitting barrier is the energy penalty associated with the structural distortion an FLP must undergo to reach the TS configuration from its equilibrium geometry. This conclusion is supported by the strong correlation observed between distortion energies and H_2 splitting barriers (Figure 2.38). Typically, the distortion energy is lower and thus, the reaction barriers tend to be smaller, for FLPs bearing small substituents on the boron centre, as these ease the rotation of the B-containing group to align with the LB partner in the TS structure. This trend is exemplified by the three FLPs exhibiting the lowest barriers (systems **60**, **68**, and **42**), all of which contain the minimally hindered –BH_2 group. Also, FLPs with bulky substituents on boron, such as tBu or iPr (e.g., systems **50**, **53**, and **72**), tend to adopt equilibrium configurations in which the p-type orbital of boron is already oriented toward the LB site. This prealignment reduces the structural reorganization required to reach the TS, thereby

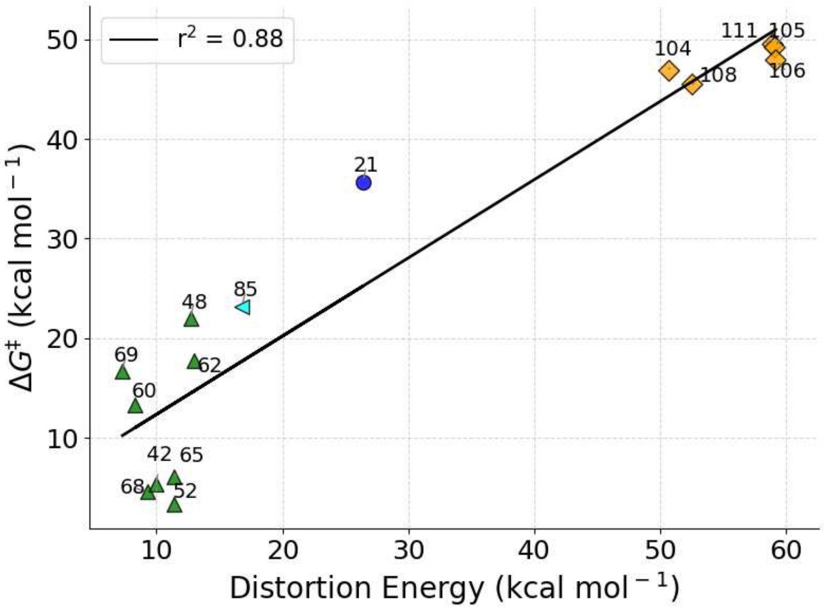


FIGURE 2.38: Distortion energy of the bare FLP required for reaching the TS structure from its equilibrium geometry plotted against $\Delta G^{\ddagger}$ for some randomly selected FLP systems. The black line represents fitting of the data to a linear regression model.

lowering the energy barrier.

Prompted by residual impact of FEHA and FEPA on the $\Delta G^\ddagger$ model, we also developed a multivariate model for predicting $\Delta G^\ddagger$ using the same methodology, but excluding FEHA and FEPA from the initial pool of descriptors. This approach allows barrier predictions to be made using descriptors derived solely from a single DFT calculation of the bare FLP structure, thus bypassing the need for additional DFT calculations on the protonated and hydrated analogues required to compute the attachment energies, further easing the practical application of the model. The resulting model is presented in Eq. 2.20, and the weights of the descriptors are graphically illustrated in Figure 2.39. Excluding the attachment energies results in only a modest decline in model performance, consistent with their secondary role in the original model. Specifically, r^2 decreases from 0.85 to 0.73 and q^2 from 0.81 to 0.68. The descriptor weights in this simplified model remain largely consistent with those in the full model. These results suggest that H_2 splitting barriers can be reliably estimated using descriptors exclusively derived from a DFT calculation on

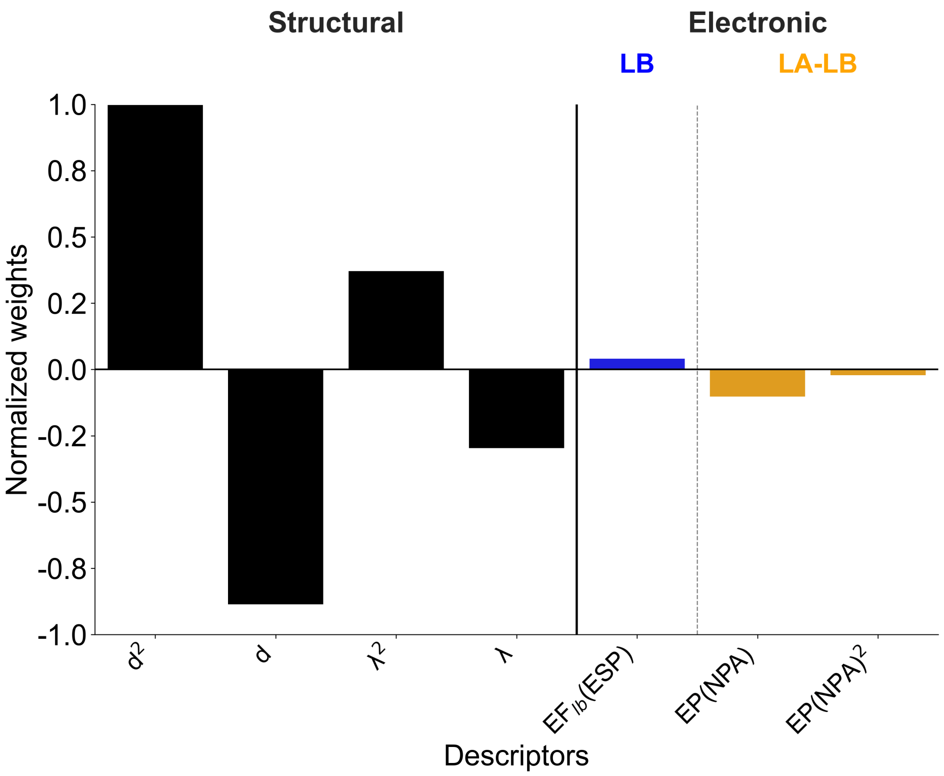


FIGURE 2.39: Normalized weights of the best obtained model in terms of $r^2 + q^2$ for predicting $\Delta G^\ddagger$ without using attachment energies FEHA and FEPA. Quality metrics: $r^2 = 0.73$, $q^2 = 0.68$ and RMSE = 6.65 kcal mol $^{-1}$.

the optimized structure of the FLP, without requiring any additional computation.

$$\begin{aligned}
&\Delta G^{\ddagger} = 181 \cdot \widetilde{d^2} - 160 \cdot \widetilde{d} + 67.4 \cdot \widetilde{\lambda^2} - 54.0 \cdot \widetilde{\lambda} + 7.74 \cdot \widetilde{EF_{lb}(\mathrm{ESP})} \\
&-4.35 \cdot \widetilde{\mathrm{EP(NPA)}}^2 - 18.9 \cdot \widetilde{\mathrm{EP(NPA)}} + 46.3 \\
&\Delta G^{\ddagger} = 6.30 \cdot d^2 - 40.9 \cdot d + 0.00208 \cdot \lambda^2 - 0.309 \cdot \lambda + 19.1 \cdot \mathrm{EF_{lb}(ESP)} \\
&-20.1 \cdot \mathrm{EP(NPA)}^2 - 42.5 \cdot \mathrm{EP(NPA)} + 87.2 \\
&\quad r^2 = 0.73 \qquad q^2 = 0.69 \qquad \mathrm{RMSE} = 6.65\ \mathrm{kcal\,mol^{-1}}
\end{aligned} \tag{2.20}$$

2.3.4 Validation of the models

Finally, we sought to further validate the presented models by performing two additional tests coded in model_validation.py. First, the response variable of the models was shuffled, assigning a random free energy value to each FLP containing molecules, in order to verify that the methodology does not fit good quality models systematically. All the models presented in this work completely lose their fitting quality and predictive ability upon shuffling the response variable. Secondly, we explored the application of our models to predict the H_2 splitting ΔG and $\Delta G^{\ddagger}$ values for an external set of FLPs previously unseen by the models (Figure 2.40), and generated using the same codes used to generate the training set. This includes eight new B/N intramolecular FLP molecules (**E1**-**E8**) stored in intra_B_N_data.pkl, five intermolecular B/N FLPs (**E9**-**E13**) stored in inter_B_N_data.pkl, and two groups of intramolecular FLPs varying the chemical nature of either the LA or the LB. These consist of three B/P FLPs (**E14**-**E16**) stored in intra_B_P_data.pkl and three Al/N FLPs (**E17**-**E19**) stored in intra_Al_N_data.pkl, inspired from previous works.[6, 138, 139] The set of structural molecular descriptors for intermolecular FLPs (**E9**-**E13**) were determined from the optimized structure of their encounter complex (i.e., the non covalent adduct between LA and LB molecules).

As shown in Table 2.2, the qualitative trends in ΔG across all families in the external validation set are correctly captured by our model. Quantitatively, the model performs well for intramolecular B/N FLPs, with a mean absolute error (MAE) of 4.9 kcal mol^{-1}, and the largest deviations remaining below 9 kcal mol^{-1}. Remarkably, ΔG values are also reasonably well reproduced for the other subsets, despite their more pronounced structural differences from the molecules in the training set. The largest discrepancies were observed for **E13**, **E14**, and **E19**, which correspond to an intermolecular B/N, a B/P, and an Al/N FLP, respectively. These systems are not structurally related to those included in the training set, and thus, accurate predictions for them were not initially expected. Thus, although certain parameters such as the LA$\cdots$LB distance may exhibit distinct optimal values depending on the specific nature of the FLP, our model is still able to capture general reactivity trends, providing reasonable reaction energy estimates even for FLP types different than intramolecular B/N systems.

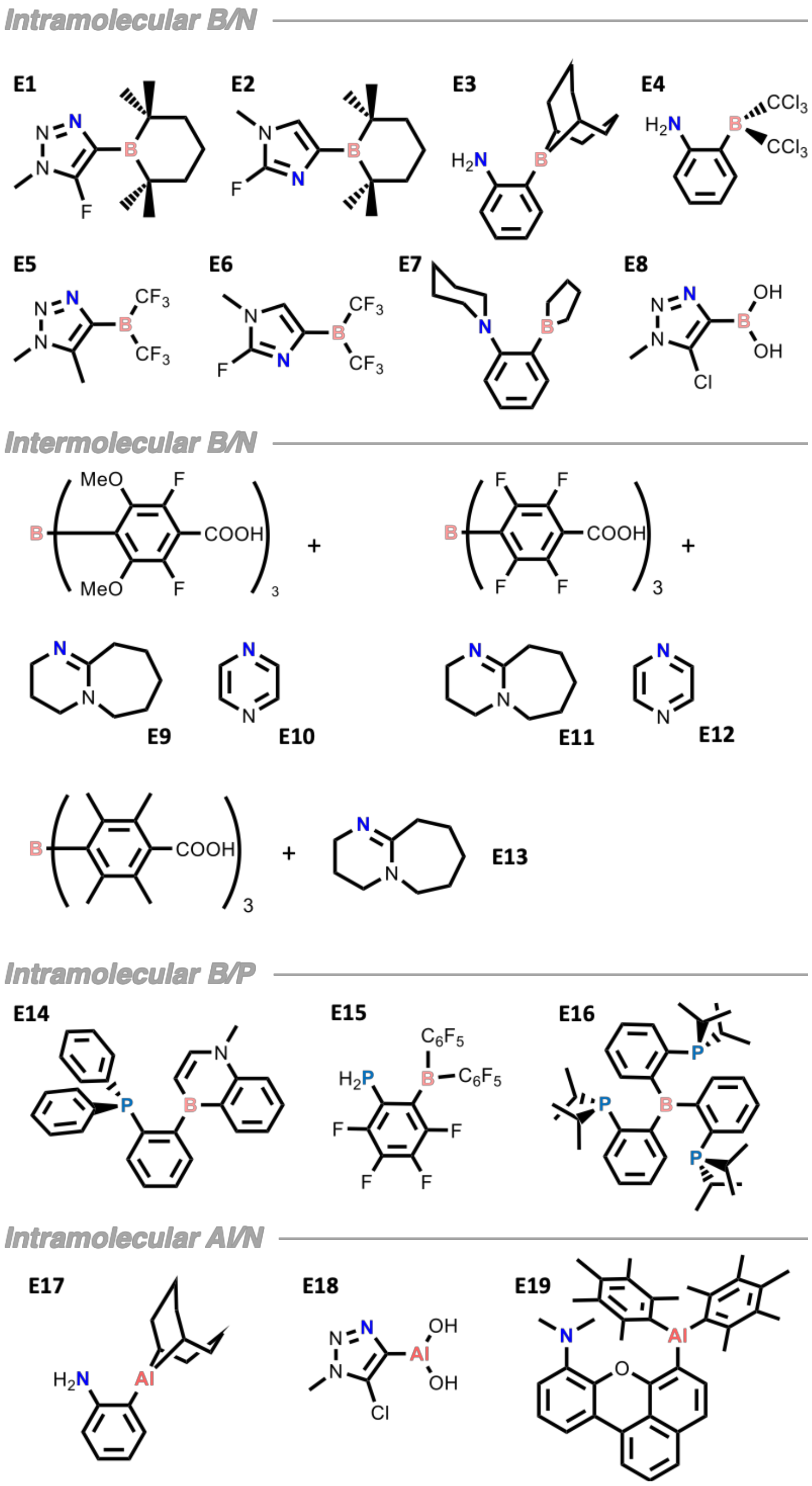


FIGURE 2.40: FLPs included in the external validation set classified by families according to their composition.

TABLE 2.2: Comparison of calculated and predicted reaction free energies (ΔG) and free-energy barriers ($\Delta G^{\ddagger}$) for the external test set.

Entry	ΔG			$\Delta G^{\ddagger}$		
	DFT	**Pred.**	**Error**	**DFT**	**Pred.**	**Error**
Intramolecular B/N FLPs						
E1	8.2	7.3	0.9	27.7	36.5 [*32.1*]	8.8 [*4.4*]
E2	12.0	8.9	3.1	30.1	30.2 [*34.9*]	0.1 [*4.8*]
E3	21.3	16.6	4.7	23.2	19.9 [*34.4*]	3.3 [*11.2*]
E4	0.8	−7.6	8.4	29.0	16.2 [*40.0*]	12.8 [*11.0*]
E5	−28.6	−21.2	7.4	11.3	27.9 [*31.7*]	16.6 [*20.4*]
E6	−19.4	−16.7	2.7	17.3	27.3 [*32.2*]	10.0 [*14.9*]
E7	11.7	7.9	3.8	18.4	19.7 [*23.0*]	1.3 [*4.6*]
E8	29.6	21.1	8.5	45.5	44.3 [*36.5*]	1.2 [*9.0*]
Intermolecular B/N FLPs						
E9	−22.7	−23.1	0.4	12.6	6.1 [*19.9*]	6.5 [*7.3*]
E10	11.4	20.6	9.2	19.8	53.8 [*50.5*]	34.0 [*30.7*]
E11	−34.0	−34.0	0.0	8.4	0.4 [*15.3*]	8.0 [*7.0*]
E12	−4.9	−6.0	1.1	11.4	19.1 [*20.6*]	7.7 [*9.2*]
E13	3.2	20.1	16.9	22.1	25.9 [*30.0*]	3.8 [*7.8*]
Intramolecular B/P FLPs						
E14	34.1	16.0	18.1	40.9	39.3 [*38.3*]	1.6 [*2.6*]
E15	7.8	1.1	6.7	20.7	8.4 [*4.4*]	12.3 [*16.3*]
E16	9.2	2.1	7.1	20.3	-67.3 [*-32.9*]	87.6 [*53.2*]
Intramolecular Al/N FLPs						
E17	16.8	11.8	5.0	17.5	8.0 [*44.1*]	9.5 [*26.6*]
E18	5.8	2.8	3.0	20.2	41.9 [*41.9*]	21.7 [*21.7*]
E19	19.4	62.1	42.7	21.0	54.8 [*43.0*]	33.8 [*22.0*]

All values are in kcal mol^{-1}. Errors are absolute deviations. Bracketed values correspond to predictions from the simplified model excluding FEHA and FEPA (Eq. 2.20)

Regarding free-energy barriers ($\Delta G^{\ddagger}$), the predictions are generally less accurate than those obtained for reaction energies, in line with the lower quality metrics of the corresponding models. Still, the $\Delta G^{\ddagger}$ of intramolecular B/N FLPs are, in general, well reproduced by the full model (Eq. 2.19) with a MAE of 6.8 kcal mol^{-1}. As expected, the simplified model (Eq. 2.20) provide worse quantitative, although still reasonable predictions, with a MAE of 10.0 kcal mol^{-1}. The largest deviations were observed for systems **E4**–**E6** and **E8**, which contain –CCl$_3$, –CF$_3$ and –OH substituents on the boron centre, features that might be underrepresented in the original training set. Prediction errors for barrier heights increase in going from intra- to intermolecular B/N FLPs. Even so, the models can still differentiate between low (**E9**, **E11**, **E12**) and high (**E10**, **E13**) barriers, although the latter tend to be overestimated. Lastly, predictions for FLPs other than B/N exhibit larger errors in general, even leading to a non-physical negative barrier for system **E16**. These results indicate that the applicability of the $\Delta G^{\ddagger}$ models is currently limited to B/N FLPs, and predictions of intermolecular pairs should be interpreted with caution. Thus, while the model developed for ΔG prediction exhibits a certain degree of transferability to other types of FLPs, the same does not hold true for $\Delta G^{\ddagger}$. This strongly suggests that the molecular features governing the height of the barrier are more sensitive to the specific nature of the FLP, making them more difficult to capture and thereby complicating the development of general predictive models. Consequently, extending the applicability of $\Delta G^{\ddagger}$ models to other FLP types will most likely require the incorporation of additional descriptors that can account for the impact of both the nature of LA and LB centres and the geometric nuances introduced by non-covalently linked donor–acceptor architectures.

2.3.5 Towards CO_2 hydrogenation

The potential catalytic activity of FLPs for CO_2 hydrogenation to formate has been both experimentally evidenced and theoretically studied.[5] Following the commonly favoured path (i) in Figure 2.3 that involves initial H_2 activation by the FLP, followed by a concerted transfer of hydride and proton to a CO_2 molecule, the latter step was addressed using the same methodology than that applied to H_2 splitting. We have thus computed the second transition state (TS2) for the reaction, corresponding to the formation of formic acid (see Figure 2.5). Inspired by our structure-property approach used for H_2 splitting, we conducted a preliminary attempt to build a numerical model to predict the energy barrier associated to TS2 ($\Delta G^{\ddagger}{}_{CO_2\ hydro}$) for CO_2 activation based on the same descriptors than for H_2 splitting (Table 2.1). A three-variable model was obtained, shown in Eq. 2.21, incorporating the square of FEHA descriptor, the LA$\cdots$LB distance (d), and the angle λ.

$$
\begin{aligned}
\Delta G^{\ddagger}_{CO_2\ hydro} &= 46.8 \cdot \widetilde{\mathrm{FEHA}^2} - 18.2 \cdot \tilde{d} - 8.42 \cdot \tilde{\lambda} - 16.4 \\
\Delta G^{\ddagger}_{CO_2\ hydro} &= 0.00486 \cdot \mathrm{FEHA}^2 - 4.65 \cdot d - 0.0483 \cdot \lambda + 28.6 \qquad (2.21) \\
r^2 = 0.82 \qquad & q^2 = 0.81 \qquad \mathrm{RMSE} = 4.51\ \mathrm{kcal\,mol^{-1}}
\end{aligned}
$$

Among these descriptors, FEHA^2 appears to contribute most significantly to the model's predictive power, possibly reflecting that $\Delta G^{\ddagger}{}_{CO_2\ hydro}$ is in many cases governed by hydride motion; the stronger the hydride is attached to the LA, the more difficult it is to make it react with CO_2. A simplified one-variable model based solely on FEHA^2 achieves r^2 = 0.63 q^2 = 0.61, and RMSE = 6.54 kcal mol^{-1}. While this suggests a notable correlation, inclusion of the geometric descriptors d and λ improves the model fit, pointing to a non-negligible influence of FLP geometry on $\Delta G^{\ddagger}{}_{CO_2\ hydro}$.

Among all the TSs calculated by means of DFT, types were identified. Ideally is expected a H^-/H^+ concerted TS, in which both are transferred simultaneously to the physisorbed CO_2 moiety, as observed for system **81**. However, many TS geometries appear intermediate between this concerted case and more asynchronous variants, where proton displacement is minimal at the transition structure. This may indicate that proton transfer occurs in a subsequent step, potentially representing the true rate-limiting transition state in such systems like **106**. The other way around, TSs in which the proton displacement is the predominant, where identified for systems like **38**. Finally, in certain systems, we also observed a distinct perpendicular TS geometry, where hydride transfer occurs out of the CO_2 plane. This motif seems to be more frequent in FLPs characterized by large LA$\cdots$LB distances and strong electron-withdrawing substituents on the LA such system **110**.

Overall, this analysis offers an initial quantitative perspective on the factors influencing $\Delta G^{\ddagger}{}_{CO_2\ hydro}$. The diversity of transition state motifs observed further highlights the mechanistic complexity of CO_2 hydrogenation by FLPs, emphasizing the need for comprehensive modeling strategies that integrate both electronic and geometric contributions. Future efforts may benefit from incorporating dynamic or solvent effects to capture more accurately the energetics of this multifaceted process.

2.4 Conclusion

This work presents a robust data-driven framework to uncover structure–activity relationships in B/N-based FLPs for H_2 splitting. Practical multivariate models were developed to predict both reaction free energies (ΔG) and activation barriers ($\Delta G^{\ddagger}$), while identifying and quantifying, for the first time, the influence of key molecular descriptors.

Our analysis reveals that the thermodynamics of the reaction is primarily controlled by the acidity and basicity of the FLP sites, captured by the free energies of hydride and proton attachment (FEHA and FEPA, respectively) and the LA$\cdots$LB distance. Stronger acidity/basicity and shorter LA$\cdots$LB separations lead to more favourable H_2 splitting energies. While these factors were previously acknowledged to influence reactivity, their relative contributions remained unclear thus far. Here, we demonstrate that FEHA has the dominant effect, with FEPA and the LA$\cdots$LB distance having similar, but approximately three times lower, impact. In contrast, barriers were found to be more sensitive to structural features, especially the LA$\cdots$LB distance (d) and the λ parameter, which measures the directionality between the LB lone pair and the LA's unoccupied p-type orbital, both showing a quadratic dependence. Additional descriptors related with the electronic properties of LA and LB sites play a secondary, albeit non negligible role in determining $\Delta G^{\ddagger}$. Thus, barrier heights are minimized at optimal values of d and λ around 3.0 Å and 70°, respectively, assuming average values for the remaining descriptors across the dataset. This optimized minimum and the captured trend provides useful insights in the preselection or design of future operational FLPs for h_2 splitting. External validation suggests that the ΔG model generalizes well to structurally diverse FLPs, whereas predictions of $\Delta G^{\ddagger}$ are reasonable for intramolecular B/N FLPs although less transferable to other FLPs types, reflecting the greater sensitivity of transition-state energetics to the detailed nature of the FLP. In this regard, ongoing efforts aim to broaden the model's applicability by expanding the training set incorporating new descriptors that better capture geometric characteristics of intermolecular FLPs and the diversity of different LA/LB combinations.

Building upon these insights, we extended our analysis to the hydrogenation of CO_2, a key downstream transformation following H_2 activation. Using the same descriptor-based approach, a preliminary multivariate model was developed to predict the activation barrier for hydride/proton transfer to CO_2 ($\Delta \mathrm{G}^{\ddagger}{}_{CO_2\ hydro}$). A three-variable model involving FEHA^2, d, and λ yielded a strong correlation ($r^2 = 0.82$, $q^2 = 0.81$, RMSE = 4.51 kcal mol^{-1}), indicating that similar descriptors govern this subsequent step, with FEHA^2 exerting the dominant influence. Complementary analysis of TS geometries revealed a diversity of mechanistic motifs beyond the idealized concerted H^-/H^+ transfer pathway. This mechanistic variety underlines the complexity of CO_2 hydrogenation and the need for refined models that account for dynamic, geometric, and electronic subtleties.

Additional deactivation pathways, such as FLP poisoning processes discussed in Section 2.1.1 (see Figure 2.7), should be incorporated into future modeling efforts to enhance accuracy and predictive power. Including these side processes is essential for developing more realistic and comprehensive models of overall reactivity, as they can significantly impact the global free energies of reaction and activation ($\Delta G_{\mathrm{global}}$ and $\Delta G^{\ddagger}_{\mathrm{global}}$). Integrating both H_2 splitting, CO_2 hydrogenation, and potential deactivation events into a unified framework will provide deeper mechanistic insight and offer more robust guidance for experimental FLP design.

Overall, this work delivers actionable structure–reactivity insights and computational tools to guide the design and screening of efficient FLP-based systems not only for H_2 activation but also for downstream transformations such as CO_2 hydrogenation. By identifying the molecular features that most strongly impact both steps, and by setting the foundation for models that can incorporate side processes like FLP deactivation, this approach enables the rational targeting of FLPs with tailored thermodynamic and kinetic profiles. Such comprehensive models hold promise for advancing chemical energy storage, CO_2 valorization, and the broader field of sustainable catalysis.

Chapter 3

Catalytic CO_2 reduction assisted by polyoxometalates

3.1 Introduction

Polyoxometalates (POMs), are anionic molecular metal oxide clusters formed by transition metals of groups 5 and 6 such as W, Mo, V, *etc.*, in their highest oxidation state (d^0) that form MO_x ($x = 5, 6$) polyhedral units. [31–33] These transition-metals are in a distorted octahedral environment and are linked to one another through bridging oxo ligands forming edge-sharing and corner-sharing patterns. However, other environments like square or pentagonal pyramid have been also reported for V and Mo, respectively, in mixed-valent compounds where the metals are not found in their highest oxidation state. [140, 141] POMs are usually synthesized from their mononuclear oxo-anions governed by the equilibrium, showed in Eq. 3.1, that are strongly dependent upon concentration, pH, ionic strength, the nature of the countercation or temperature. The transition metal oxides such V_2O_5, MoO_3 or WO_3, dissolve at high pH to give orthometalates (${VO_4}^{3-}$, ${MoO_4}^{2-}$ or ${WO_4}^{2-}$). As the pH is lowered, orthometalates protonate to give oxide-hydroxide compounds that condense losing water and forming M–O–M linkages.

$$q{MO_4}^{n-} + rH^+ \rightleftharpoons [M_qO_y]^{m-} + \frac{r}{2}H_2O \tag{3.1}$$

$$q{MO_4}^{n-} + t{XO_4}^{k-} + rH^+ \rightleftharpoons [X_tM_qO_y]^{m-} + \frac{r}{2}H_2O \tag{3.2}$$

Introducing other oxides like silicates (${SiO_4}^{4-}$) or phosphates (${PO_4}^{3-}$) following Eq. 3.2, the resulting POMs can also incorporate p- and d-block atoms, named heteroatoms (X). POM can thus be divided into two sub-families, isopolyanions and heteropolyanions.[142] Representative examples of POM structures are shown in Figure 3.1. The Lindqvist are compact clusters with a general formula $[M_6\mathbf{O}_{19}]^{n-}$ without heteroatoms (Figure 3.1 w). Their structure is composed of six MO_6 octahedra that form an almost spherical shape. Lindqvist-type POMs are valuable in redox chemistry due to their multiple accessible oxidation states and high stability in various solvents that makes them a good platform for metal substitution and incorporation into larger supramolecular assemblies. They are often used in the design of

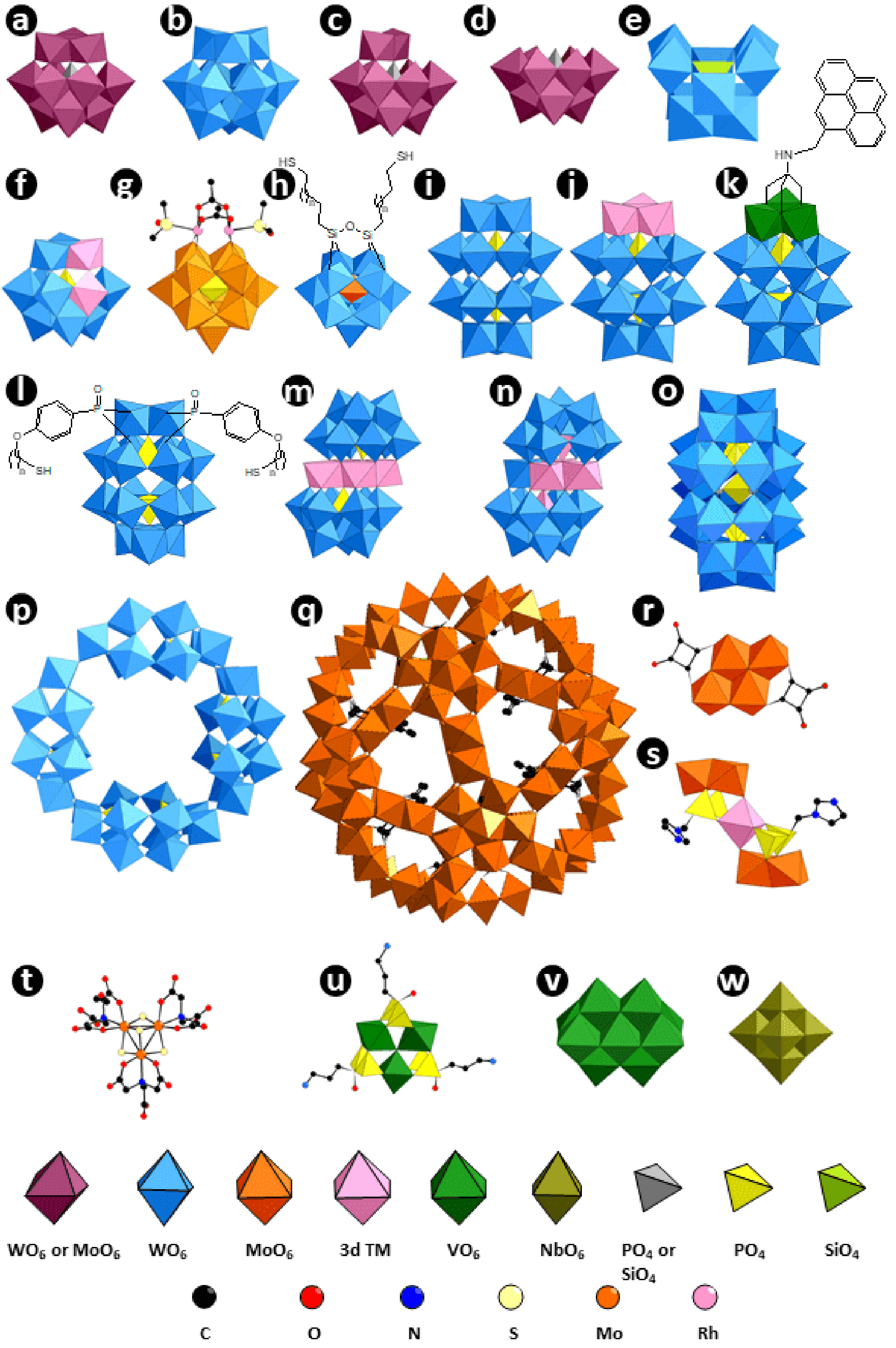


FIGURE 3.1: Polyhedral representations of various POM structures. The figure includes Lindqvist-type (w), Keggin-type (a), Wells-Dawson-type (i), Preyssler-type (o), lacunary structures (b–e), substituted and sandwich-type POMs (f, j, m, n), hybrid organic-inorganic POMs (g, h, k, l), keplerate structures (p, q), and other structural motifs (r–v).

organometallic complexes, battery electrodes,[143, 144] and molecular electronics. The Keggin-type POMs, $[XM_{12}O_{40}]^{n-}$, are the most well-known and widely studied polyanions (Figure 3.1 a). They structure consists of a central XO_4 tetrahedron surrounded by twelve edge-sharing MO_6 octahedra, forming a nearly spherical, highly symmetric anion. The Wells-Dawson-type POMs, $[X_2M_{18}O_{62}]^{n-}$, are larger and structurally more complex than Keggin-type POMs, and are composed of two central tetrahedral heteroatoms embedded within a framework of eighteen MO_6 octahedra, offering a greater number of redox-active sites and a higher surface area (Figure 3.1 i). The Wells-Dawsons are particularly valued as molecular precursors to functionalized hybrids. Their structural complexity allows for easier functionalization at specific sites, enabling fine-tuning of their electronic and chemical properties for targeted applications. Preyssler-type POMs (Figure 3.1 o) are large symmetric clusters with a formula $[X_5M_{30}O_{110}]^{n-}$. Their "crown-shaped" structure makes it the smallest POM with internal cavity that confer them an extra ability to encapsulate certain cations. The most prominent families offer distinct architectures that influence their chemical behaviour, highlighting the adaptability and use of POMs as molecular building blocks for advanced technologies. On top of it, POMs present several isomers that further increase their structural variability and affecting their properties.[145] Furthermore, the metal positions of POMs can be substituted by other metals[146], the controlled addition of a base, increases the pH that leads to the dissociation of certain metal oxo units, generating lacunary POMs (Figure 3.1 b-e) that serves as polydentate ligands for the substituting metal that comes from the dissolved corresponding salt (Figure 3.1 f, j, m and n) or allows their clusterization to form bigger structures (Figure 3.1 p and v). Nevertheless, not only the metal positions can be substituted, but the oxo-ligands can also be substituted by other ligands that can contain organic molecules forming organic-inorganic POM-based hybrid platforms[147–149] (Figure 3.1 g, h, k and l) exhibiting enhanced stability, control on the number and relative orientation of the organic and inorganic components, tuning of the steric and/or electronic features and possible increase of synergetic effects, where both the organic and inorganic moieties mutually alter their properties in response to each other's presence.

By combining all the possible structural modifications, it becomes feasible to design an almost infinite variety of species tailored for specific roles in chemical processes (Figure 3.1 r–u). For instance, certain structures can be engineered as scaffolds to host specific chemical reactions (Figure 3.1 q), or modified to enhance the dispersion of POMs in polymers, enable their immobilization on surfaces, or fine-tune their supramolecular properties. Such control over their behaviour in solution paves the way for developing highly efficient heterogeneous catalysts.[150–158]

Properties

Thermal distortion of POMs involves structural changes at elevated temperatures, such as reversible geometric distortions of metal–oxygen bonds, loss of structural groups, and formation of lacunary sites upon removal of metal-oxo units, often preceding decomposition above 350–500°C for stable tungsten-based Keggin- and Dawson-type POMs, while molybdenum- and lacunary POMs degrade earlier. POMs can also undergo multiple redox processes without suffering strong distortion thanks to the delocalization of negative charges across the entire POM structure that enhances electronic stability, and their entirely inorganic composition, which helps avoid common decomposition pathways typical of organic compounds. Thermal distortion is monitored by thermogravimetric analysis, differential scanning calorimetry, and in-situ spectroscopy, while strategies to minimize distortion include using tungsten-based POMs, embedding them in protective matrices, functionalizing for stability, and controlling the heating atmosphere to prevent unwanted redox reactions.

Commonly, the traditional POMs exhibit diamagnetism as all their electrons are paired. However, transition-metal substituted POMs can present unpaired electrons making them exhibit paramagnetism and other magnetic behaviour.

The most accessible oxygen sites in POMs are typically the surface bridging (M–O–M) and terminal (M=O) oxygens at their surface, both of which exhibit significant basic character due to the high electron density on the O^{2-} anions. In specific structures, such as the Preyssler-type POM, even the internal oxygen atoms bonded to the central heteroatom (X–O–M) within the molecular cavity can be exposed and are among the most basic sites. These anionic oxygens, particularly in the bridging positions, act as strong Brønsted bases, a behaviour supported by both crystallographic data and computational studies, which consistently show a preference for these sites in host–guest and protonation interactions.[159]

The redox properties of POMs[145] comes from the presence of many metal atoms in high oxidation states, which make POMs prone to accept electrons. Extra electrons occupy non-bonding low-energy molecular orbitals of the POM, consisting on linear combinations of d-type orbitals centred on metal centres. The ability of POMs to be reduced and therefore the oxidizing power, can be generalized by considering the overall negative charge and the number of metal atoms. The ratio between the charge and the metals strongly determines their redox behaviour being a simplified molecular charge density evaluation of POMs. Furthermore, solvation effects also play an important role, precisely highly negatively charged POMs electronic nature depend on how solvent molecules are able to stabilize the charge, having a strong impact to their properties. The multiple available low-energy metal orbitals make POMs prone to be reduced multiple times, which in turn, increases their Brønsted basicity. Protonation of reduced POMs offsets to some extent their increased charge density, decreasing the charge/metal ratio, enhancing their ability to be further reduced[160]. This synergetic behaviour is very promising for their application as electron storage [161] and relay agents or as hydrogen evolution promoters.[162, 163]

Applications

Owing to their structural versatility, which allows fine tuning of their properties, POMs are used in many areas like in biochemistry and biomedicine which POMs have demonstrated antitumoral, antiviral, and antibacterial activities.[164] For instance, certain titanium- and vanadium-substituted POMs exhibit potent antiviral properties, including activity against human immunodeficiency virus (HIV) and SARS viruses, as well as antibacterial effects against strains like MRSA and VRSA. POMs also play a role in protein crystallography. Their high negative charge and ability to form multiple hydrogen bonds make them effective additives in protein crystallization processes,[165] aiding in the formation of high-quality crystals for X-ray diffraction studies. In enzymatic catalysis, POMs have been explored as mimetics, particularly in hydrolytic processes.[166] Their structural resemblance to certain enzyme active sites allows them to mimic enzymatic functions, offering insights into catalytic mechanisms.

POMs are renowned for their versatile redox properties, enabling them to participate in a variety of chemical reactions and energy-related applications. Their ability to undergo reversible multi-electron and proton-coupled electron transfer processes makes them valuable in both oxidation and reduction reactions. POMs can act as direct oxidants, accepting electrons in reactions such as the dehydrogenation of alcohols and dienes. They also serve as co-catalysts in oxidation reactions, facilitating electron transfer processes and can be regenerated using molecular oxygen, producing water as a byproduct.[166] Moreover, POMs can participate in Mars–van Krevelen-type oxidations, transferring oxo ligands during the reaction cycle.[167] Beyond oxidation, POMs have shown promise in energy storage applications due to their ability to undergo reversible multi-electron redox processes. This characteristic makes them suitable for use in batteries and other energy storage systems.[168] For instance, certain POMs can reversibly store up to 24 electrons, highlighting their potential in high-capacity energy storage devices.[169] Their capacity for multi-electron and proton-coupled electron transfer processes also positions POMs as attractive candidates in multi-electron/proton catalysis, including applications in solar energy conversion processes like water oxidation and proton reduction. Studies have demonstrated the use of tetraruthenium POM in photochemical water oxidation, showcasing their ability to facilitate multiple oxidative proton-coupled electron transfer events.[170]

More relevant to the present chapter is the capability of POMs to promote reduction reactions, including proton reduction for hydrogen evolution,[171] nitrogen fixation,[172] and carbon dioxide conversion.[168] Their ability to reversibly accept and store multiple electrons makes them excellent redox-active ligands or catalytic platforms for multi-electron transfer reactions.[3] For example, POM-based materials have been explored for electrocatalytic nitrogen reduction reactions, where they act as electron reservoirs to facilitate the activation and reduction of N_2 molecules.

The reduction of N_2 to NH_3, a critical step in nitrogen fixation, has gained attention in the context of sustainable chemistry. Traditional Haber–Bosch processes are energy-intensive, but POM-based catalysts offer a low-temperature, solution-phase alternative. Recent work has demonstrated that certain metal-substituted POM systems, M–POM (M = Mo, Fe, etc.), can serve as biomimetic models for nitrogenase enzymes. For example, an iron-substituted POM cluster has shown promising electrocatalytic activity for N_2 reduction under ambient conditions, producing NH_3.[172] Furthermore, hybrid architectures involving POMs and transition metal complexes, such as Fe–POM assemblies, have also demonstrated potential for photocatalytic nitrogen fixation. These systems mimic key features of the nitrogenase active site, including the requirement for coupled proton and electron transfers. A recent study reported that a Fe-functionalized Dawson-type POM enabled visible-light-driven N_2-to-NH_3 conversion, highlighting the potential of POM-based systems in artificial nitrogen fixation.[173] POMs have been integrated into CO_2 reduction systems, enhancing the selectivity and efficiency of converting CO_2 into valuable hydrocarbons.[3] In electrocatalytic systems, POMs have been used in combination with molecular complexes or conductive surfaces to facilitate CO_2 reduction. For instance, copper nanocubes modified with Mo_8-type POMs generated acetate with high selectivity (48.7% Faradaic efficiency) and current density, highlighting the role of Cu–O–Mo interfaces in steering product selectivity.[3] POM-based materials have also been explored for photocatalytic CO_2 reduction. A manganese-modified POM photocatalyst was reported by Zhang *et al.* to exhibit a high CO evolution rate with 90.3% selectivity over H_2, enabled by efficient charge separation and electron transfer.[174] Another study demonstrated the selective reduction of CO_2 to methane using a Zn-based POM in conjunction with an Ir-based photosensitizer under visible light irradiation, achieving high selectivity due to favourable redox alignment and multi-electron capability.[175]

Finally, polyoxometalate–metalloporphyrin organic frameworks (PMOFs) have been developed to optimize directional electron flow and promote multi-electron catalytic cycles. In one example, a Co–PMOF achieved nearly quantitative Faradaic efficiency for CO_2-to-CO electroreduction,[176] and analogous frameworks are being extended to explore N_2 reduction pathways as well.

These studies collectively underscore the versatility of POMs in promoting key redox processes, including CO_2 and N_2 conversion. Their ability to stabilize high-energy intermediates and serve as electron reservoirs makes them valuable as tunable platforms for designing catalysts with high efficiency and selectivity for reduction reactions of critical small molecules.

POMs for electro- and photocatalytic CO_2 reduction

Regardless of whether the approach is photocatalytic or electrocatalytic, the central goal in using POMs for carbon dioxide reduction is to inject electrons into the POM framework, which then acts as a redox-active reservoir to mediate the multi-electron reduction of CO_2 into value-added products such as CO, formate, or methanol. Fully-oxidized POMs do not typically absorb light in the visible range. Thus, using POMs in photocatalytic schemes requires external photosensitizers, such as transition metal complexes or semiconductors, that harvest photons and transfer excited-state electrons to the POM. These photosensitizers enable the stepwise reduction of POMs under illumination, allowing them to participate in downstream CO_2 activation and reduction. In contrast, during electrocatalysis, the POMs are directly reduced at the electrode surface due to an externally applied potential. This electrochemical input allows for precise control over the number of electrons introduced into the POM, which can then be channeled toward CO_2 transformation without the need for light or additional sensitizing agents.

The state of the art in POM-mediated CO_2 reduction reveals a broad and evolving landscape of molecular design and hybrid material strategies. Studies such as those by Wang *et al.* (2021)[8] have demonstrated how redox-tunable POMs can be selectively reduced and coupled to catalytic moieties for efficient CO_2 activation. Fabre *et al.* (2022)[177] highlighted how surface-functionalized POMs enhance electron transfer properties when immobilized on conductive supports, improving catalytic turnover and selectivity. Gu *et al.* (2021)[178] reviewed various POM-based nanostructures, including hybrids and single-cluster systems, and emphasized their ability to balance redox storage with structural stability under electrochemical or photochemical conditions. Cao *et al.* (2019)[179] provided key insights into how the composition and nuclearity of POMs affect their performance and electron transfer dynamics in photocatalysis, while Song and Tsunashima (2012)[180] explored the versatility of POMs as scaffolds that can be easily functionalized with metal centres or organic ligands to enhance activity and selectivity. The latest contributions by Zang and Wang (2022)[181] offer a comprehensive perspective on how structural and electronic modifications in POMs, such as incorporating transition metals or forming organic–inorganic hybrids, can fine-tune their catalytic behaviour, pushing the boundaries of efficiency and selectivity in both photo- and electrocatalytic CO_2 reduction.

3.2 Ni substituted POMs for CO_2 reduction

This section was written based on of the paper: *Photocatalytic CO_2 reduction by Ni-substituted polyoxometalates: Structure-activity relationships and mechanistic insights*, by Khadija Talbi, Francesc Penas-Hidalgo, Amanda L. Robinson, Philipp Gotico, Winfried Leibl, Pierre Mialene, Maria Gomez-Mingot, Marc Fontcave, Albert Solé-Daura, Caroline Mellot-Draznieks, and Anne Dolbecq, Applied Catalysis B: Environment and Energy, **2024**, 345, 123681. [182] F.P.-H. contribution: computational investigation, formal analysis, and visualization.

3.2.1 Introduction

Nickel(II)-substituted POMs proved highly active for photocatalytic proton reduction and subsequent H_2 evolution.[183–189] Also, recent reports demonstrate their activity in photocatalytic CO_2 reduction,[190–192] although the application of Ni-POMs in the field of CO_2 reduction is still significantly underexplored compared to that in H_2 evolution. Inspired by these findings, our experimental collaborators at Institut Lavoisier de Versailles synthesized a series of POMs with vacant SiW_9 or PW_9 units connected to Ni clusters (Figure 3.2) and tested them for the photocatalytic reduction of CO_2 within the context of the PhD of Khadija Talbi. The set of selected Ni-substituted POMs included structures that exhibit different Ni contents, sizes, charges and shapes, and were synthesized using reported protocols. The smallest system, **Ni_4Ac**, conformed by a cluster of four Ni centres coordinated by a trilacunary Keggin and three acetate ligands, was synthesized from the trilacunary Keggin and nickel acetate acidified with HCl.[193] A tetra-nickel-containing sandwich Keggin-type POM was synthesized using $Na_2WO_4{\cdot}2H_2O$ and Na_2HPO_4, followed by the slow addition of nickel acetate, resulting in the formation of **Ni_4**.[183] By modifying the synthetic conditions and starting from $NiCl_2{\cdot}6H_2O$, a bisphosphonate ligand, and an excess of trilacunary Keggin units, a system composed of two subunits was obtained. In this system, the bisphosphonate groups form two tetrahedral phosphorus centres that coordinate with the nickel centres, leading to the formation of **Ni_7**.[194] Such species were shown to be highly stable in the pH 7-10 range, and derivates with different organic groups or coordination complexes were further synthesized.[195] Using a similar strategy and adjusting

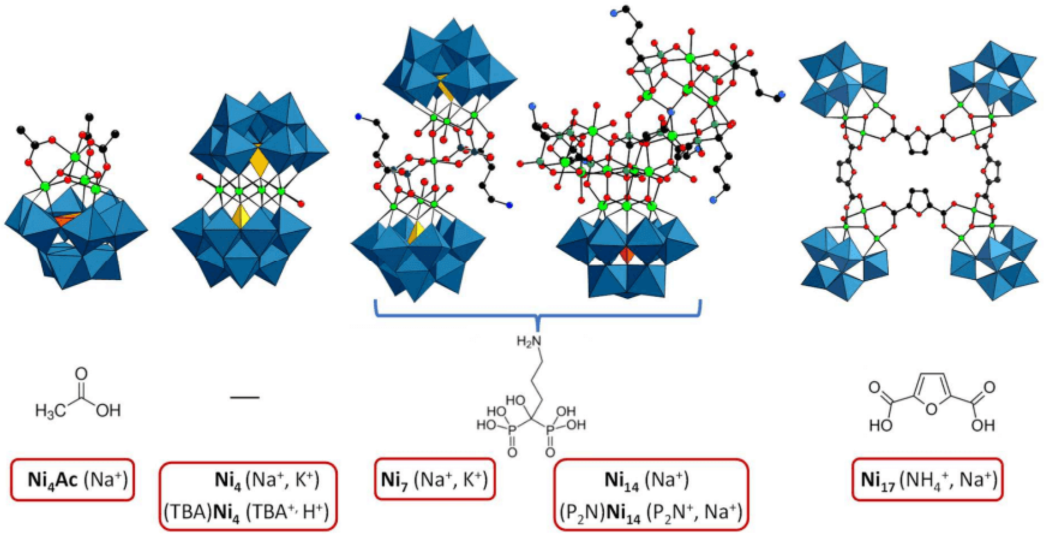


FIGURE 3.2: Ni containing systems studied specifying the counterions used and the formula of the organic ligands. Octahedral W colored in blue, tetrahedral P in yellow or orange, C in black, N in light blue, oxygen in red and Ni in yellow.

the concentration of the salts and setting the pH at 9.1, a **Ni_{14}** system was synthesized, made of a trilacunary Keggin ligand capped by a tetradecanuclear Ni^{II} cluster.[187] Substituting the bridging acetate ligands of system **Ni_4Ac** by rigid dicarboxilates generates **Ni_{17}** which corresponds to a metal-organic polyhedra (MOP) due to the acute angle of 72° between its three acetate ligands that avoid the formation of possible metal-organic frameworks (MOFs).[192] As it is shown in Figure 3.2, two organic counter-ions have been studied, tetrabutylammonium (TBA) and bis(triphenylphosphoranylidene)ammonium (P_2N) in order to evaluate their potential effect on the catalyst's activity.

3.2.2 Experimental background

All the POM salts represented in Figure 3.2 were studied under photocatalytic homogeneous conditions in a mixture of CH_3CN:H_2O:TEOA solution saturated with CO_2, using TEOA as SED, H_2O as proton source and $[Ru(bpy)_3]^{2+}$ (**Ru^{2+}**) as PS. The solution was irradiated with visible light at $\lambda > 415$ nm for 2 h with a 280 W Xe lamp. Analyzing both gas and liquid phases, negligible traces of formate were detected in the liquid phase and the concentration of CO and H_2 were quantified in the gas phase as the only products. The catalytic activity of the set of POMs has been evaluated and is expressed in Figure 3.3 using optimized conditions in terms of water and PS content, as well as in terms of the nature and concentration of the PS. As (TBA)**Ni_4** salt proved to be the most efficient catalyst, the experimental conditions were further optimized for this system. In short, the catalytic activity of (TBA)**Ni_4** was found to maximize in a reaction mixture consisting in a CH_3CN:H_2O:TEOA volumetric ratio of 8.8:2:1. Modifying the volume of TEOA decreased the production of CO. Also the presence of water proved to be beneficial with a significant decrease in CO production in the absence of water. The concentration and different types of PS were screened showing that **Ru^{2+}** exhibits the best performance with 1.95 mM concentration, while $[Ru(phen)_3]^{2+}$ affords systematically lower activities and no activity was observed with $[Ir(dtbbpy)(ppy)_2]^+$. The production of CO and H_2 was measured for the seven salts of Ni-substituted POMs using the above optimized conditions. **Ni_7** and **Ni_4Ac** exhibited the lowest photocatalytic activities, while the catalyst (TBA)**Ni_4** exhibited the highest one. The other catalysts exhibit similar suboptimal performances. Notably, the activity of **Ni_4** significantly decreased using a fully inorganic salt consisting of a mixture of Na^+ and K^+ cations compared to the analogous TBA^+ salt. The same effect is observed for **Ni_{14}** using Na^+ as counter-ion instead of P_2N^+. Control experiments proved that the production of CO becomes negligible in the absence of any agent of the catalytic system, CO_2, light, PS, SED or the catalyst itself. The CO production increased linearly for the first 2 h but leveled off after 4 h. The production restarted after addition of 8 mg (10.5 μmol) of fresh $[Ru(bpy)_3]Cl_2$, indicating that the photocatalytic plateau can be attributed to the deactivation of the photosensitizer, as commonly observed.[196, 197] The production of CO and H_2 of the $Na_8H[PW_9O_{34}]$ precursor measured using the same conditions was found very low, evidencing that the Ni centres of the POM play a crucial role in the catalytic process. Noticeably, it has been experimentally observed that the selectivity of the catalyst increases as its concentration decreases. Through isotope tracing experiments the origin of the CO product has been attributed to the initial CO_2. Using mass spectrometry (MS) techniques, the solutions after the photocatalytic processes were analyzed confirming the previously observed oxidized product of the TEOA, the aldehyde product,[12, 198] supporting that TEOA acts indeed as a SED during the catalysis. Finally, the stability of the catalyst is obviously an important parameter in photocatalytic CO_2 reduction. First, EDX measurements performed on the precipitate obtained by addition of $[Ru(bpy)_3]Cl_2$ after photocatalysis have shown that the Ni\W and P\Ni ratio are the ones expected for $[Ni_4(H_2O)_2(PW_9O_{34})_2]^{10-}$ POM. Analyzing the same precipitate using infrared (IR) spectroscopy, confirmed the preservation of the POM during the reaction, being identical to the IR spectrum of the pristine POM in the W–O, W=O and P–O vibration regions. Dynamic light scattering (DLS) experiments did not detect any nanoparticles higher than 1 nm, suggesting that no nickel oxide nanoparticle was found during the catalytic event. All these observations support that (TBA)**Ni**$_4$ is stable under the described catalytic conditions.

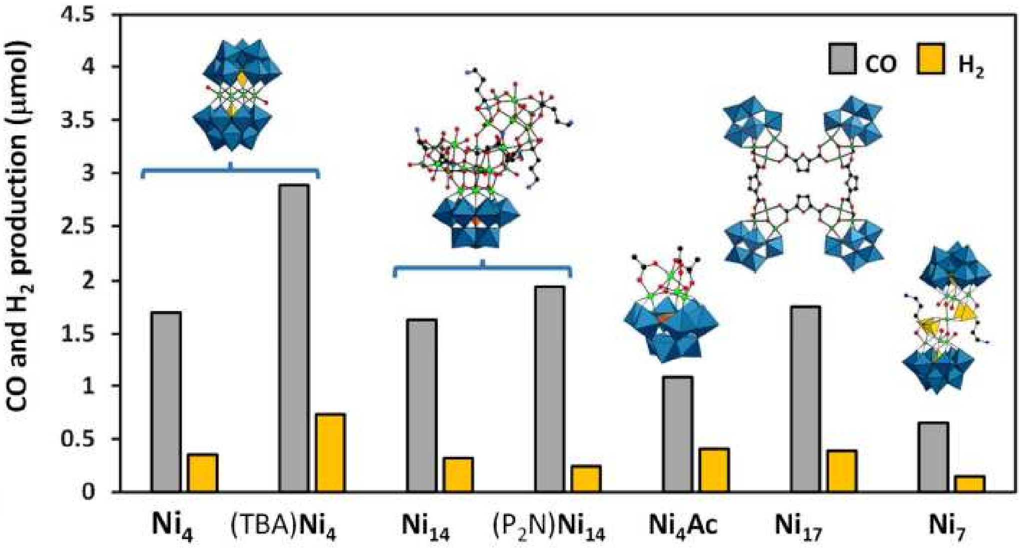


FIGURE 3.3: CO_2 and H_2 reduction activity (bars in gray and orange respectively) of all the catalyst-counterions combinations studied. Octahedral W colored in blue, tetrahedral P in yellow or orange, C in black, N in light blue, oxygen in red and Ni in yellow.

For all the photocatalytic experiments mentioned above, a phase separation has been observed triggered by the bubbling of CO_2 in the mixture, explained by a salt-induced liquid-liquid phase separation in CH_3CNH_2O mixtures.[199] The saturation of CO_2 molecules enhance the formation of TEOA-CO_2 adducts[200], carbamate, $HTEOA^+$ and $HCO_3{}^-$ species, which in turn enhances the water solvation of the ions by breaking the CH_3CN-H_2O interactions that allows the miscibility of both liquids.[201] Interestingly, adding a few drops of N,N-dimethylmethanamide (DMF) recovers an homogeneous phase which is maintained upon further bubbling of CO_2. It can be proposed that such addition of DMF regenerates the interactions between solvent molecules, preventing the liquid-liquid phase separation. ^1H Nuclear magnetic resonance (NMR) measurements on both phases verified that the POM was present exclusively in the aqueous phase while the counter-cation TBA^+ was in the organic phase. Surprisingly, the experiments showed that the biphase formation is essential for the catalysis as the activity drastically drops in the absence of CH_3CN in the catalytic medium. This can be attributed to the fact that the organic counter ion remains in the organic phase of the biphasic system and that the catalytic reaction takes place in the water phase at its interface with the CH_3CN phase, whereby the separation of the POM from its TBA counter-ions has a beneficial impact on the resulting photocatalytic activity. This explanation is in line with the higher catalytic activity of (TBA)$\mathbf{Ni_4}$ over (Na^+/K^+)$\mathbf{Ni_4}$. In the latter the alkaline cations present in the aqueous phase have strong interactions with the POM[202] and interfer with the $\mathbf{Ru^{2+}}$–$\mathbf{Ni_4}$ interactions[203] thus hindering the electron transfer from the PS to the POM and consequently, the CO_2 reduction process. Along that line, the behaviour of a Keggin-type POM was studied in 2009 in an organic-aqueous interphase in the presence of different types of counterions (Cs^+, $NBu_4{}^+$, H_3O^+, $H_5O_2{}^+$ and Eu^{3+}) by means of molecular dynamics (MD) simulations.[204] The results of these simulations further support that POMs remain in the aqueous phase. Moreover, in the case of TBA^+ salts, water-solvated POMs were found to accumulate close to the interface with the organic solvent, where TBA^+ cations reside due to its hydrophobic nature. This supramolecular behaviour may be responsible for the superior activity of (TBA)$\mathbf{Ni_4}$ compared to other salts based on inorganic cations by: (1) enhancing ion-pair separation, which favours POM–PS interactions; and (2) positioning the catalyst closer to the organic phase, where the concentration of CO_2 is higher due to its greater solubility in acetonitrile than in water. The latter effect may facilitate interaction between the active from of the catalyst and CO_2, bypassing diffusion limitations and thus positively impacting the catalysis.

The diverse range of properties associated with POMs has been extensively investigated over the last decades. However, understanding their behaviour under specific conditions can become extremely challenging due to the multitude of potential processes they can undergo, often involving redox processes, proton transfer events or even structural changes. Despite substantial research into their structure and reactivity, clear evidence regarding the structural and compositional parameters essential for the catalytic activity in CO_2 to CO conversion of POMs is still lacking. This uncertainty is compounded by the complex interactions within photocatalytic systems, whereby the behaviour of the photosensitizer may differ based on whether it undergoes reductive or oxidative quenching pathways. These processes are not understood in a systematic fashion, hindering the development of efficient photocatalytic systems for CO_2 reduction. The lack of a comprehensive understanding of the critical factors driving the CO_2 reduction process in POMs highlights the need for a more detailed exploration of the fundamental interactions involved. Furthermore, the involvement of complex intermediates and the dynamics of charge transfer processes, such as the quenching of the photosensitizer, require a deeper investigation.

In response to these challenges, our computational study aims to bridge the gaps in understanding by analyzing the interactions between the principal components of the photocatalytic system, providing a complete atomistic description of the reaction mechanism and elucidating the parameters that affect the catalytic performance. Focusing on the **Ni**$_4$ catalyst as a representative example, we combined static Density Functional Theory (DFT) and time-dependent Density Functional Theory (TD-DFT) calculations with DFT-based molecular dynamics (DFT-MD) simulations, to elucidate the nature of the chemical processes and mechanisms governing the photocatalytic CO_2 to CO conversion. By doing so, we aim to identify the critical features that facilitate optimal catalytic performance, thus paving the way for the design of more efficient photocatalysts for CO_2 reduction.

3.2.3 Computational methods

The initial structure of the **Ni**$_4$ catalyst was retrieved from the ioChem-BD repository.[205] This initial geometry and all the subsequent intermediates, as well as other components of the catalytic system (PS, reactants, products, SED...) were optimized by means of DFT adopting the ωB97X-D functional[76] using Gaussian16 (rev C.02) quantum chemistry software[119]. The LANL2DZ basis set[206] and associated pseudopotentials[207] supplemented by Frenking's f-type polarization functions[208], were used to describe the heaviest elements such Ni, W and Ru. An all-electron 6-31g(d,p) Pople-type basis set[86, 209, 210] was used for the remaining atoms. As the POM was experimentally observed to reside exclusively in the aqueous phase, solvent effects of water were included by means of Integral Equation Formalism variant of the Polarizable Continuum Model (IEF-PCM) implicit solvent model.[120] The nature of all stationary points on the potential energy surface was confirmed by frequency calculations. The standard-state correction of +1.9 kcal mol^{-1} (from the 1 atm reference state assumed in Gaussian calculations to the standard state of 1 mol L^{-1} in solution at 25 °C) was applied to the free energy of all the species with the exception of water molecules. For the latter, which are part of the solvent, the standard-state correction accounts for +4.3 kcal mol^{-1} according to their concentration of 55.3 M. UV-vis spectra have been simulated using TD-DFT calculations at the same level of theory. The **Ni**$_4$ system has two unpaired electrons per Ni^{II} centre, making the overall spin multiplicity equal 9 in a high-spin configuration, which is supported by previous studies.[211]

DFT-MD simulations were performed using the CP2K code (2022.1 release)[212, 213] at the BLYP level of theory[214, 215], augmented with the Grimme's D3 dispersion correction [80] using the Gaussian and Plane Waves (GPW) method, whereby the wave function is expanded using a localized Gaussian basis set and the density is expanded using an auxiliary plane-waves basis set.[216] Valence orbitals were described by the split-valence DZVP-MOLOPT Gaussian basis set [217] of double-ζ quality and one set of polarization functions in conjunction with Goedecker, Teter, and Hutter (GTH) pseudopotentials.[218, 219] The auxiliary plane-wave basis set was cutoff at 400 Ry. Simulated systems contain a $H_4Ni_4(2e)]^{8-}$ anion and a CO_2 molecule embedded in a 3D-periodic cell of 18 × 18 × 20 $Å^3$ containing 170 water molecules. Hydrogen atoms were replaced by Deuterium atoms for the sake of computational efficiency. During the DFT-based molecular dynamics simulations, positions and momenta were propagated within the Born–Oppenheimer scheme with a time step of 0.5 fs in the canonical (NVT) ensemble. The Canonical Sampling through Velocity Rescaling (CSVR) thermostat [220] was used to keep constant the temperature of the system at 300 K. Following 500 ps relaxations of the solvent around fixed solutes at the universal force-field (UFF) level,[221] the system was equilibrated for 3 ps of DFT-MD. Then, biased DFT-MD simulations were carried out in combination with the path-metadynamics enhanced sampling technique,[222, 223] in order to study firstly the release of an aqua ligand from the coordination sphere of a Ni^{II} centre; and secondly, the binding and activation of CO_2 to the ligand vacancy of Ni^{II}. This was done in a sequential manner, as a configuration generated in the first simulation, in which the Ni bears no aqua ligands served as a starting point for the second one. In the first simulation, set to study the decoordination of the aqua ligand, the path collective variable was built from a non-linear combination of the $Ni\cdots O$(aqua) distance and the coordination number (CN) of Ni by O atoms of solvent water molecules, defined by a $r_0 = 2.8$ Å and exponents of $n = 12$ and $m = 24$ to the numerator and denominator, respectively. In the second one, aimed at investigating the binding of CO_2, the reaction path was defined by the $Ni\cdots C(CO_2)$ distance and the O–C–O angle. In order to prevent the binding of other water molecules to the Ni centre, an upper potential wall with a force constant of 150 kcal mol^{-1} was applied to the CN of Ni by O atoms of solvent water molecules, which becomes active at $CN \geq 0.2$; the latter being defined by $r_0 = 3.0$ Å and exponents of $n = 12$ and $m = 24$.

3.2.4 Results and discussions

Exploration of the basicity and protonation state of Ni_4

As it is well known that the redox properties of POMs depend on their protonation state[160], the acid-base properties of the $\mathbf{Ni_4}$ system were evaluated in a first step. As it is mentioned in the experimental background, the mechanism in which TEOA acts as a SED, also involves proton releasing and the generation of the protonated version of the amine ($HTEOA^+$), as illustrated in Figure 3.4. The latter is the product of the spontaneously deprotonation of $TEOA^{\bullet +}$ radical cation by a second TEOA molecule after the first step of the reductive quenching of the photoexited $\mathbf{Ru}^{2+}$.[12, 198, 224–228] Furthermore, $HTEOA^+$ can be also spontaneously generated in the

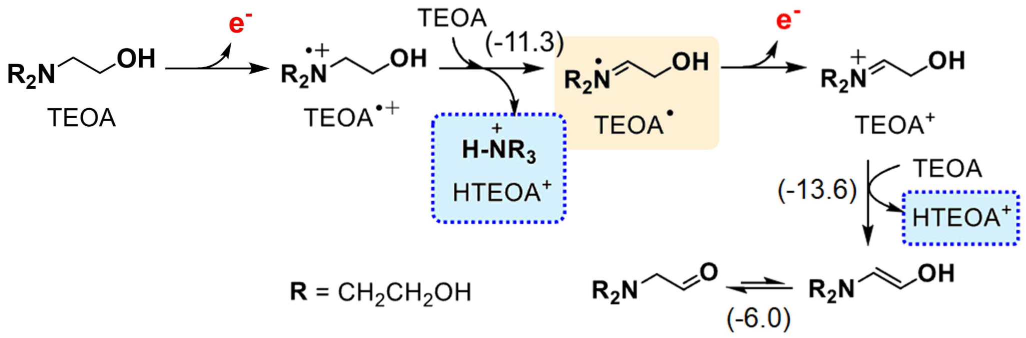


FIGURE 3.4: Radical chain decomposition mechanism of TEOA acting as proton source and SED.

presence of water via protonation of TEOA by a water molecule, according to the experimental pKa value of 10.75 determined for $HTEOA^+$. For this reason, $HTEOA^+$ was used here as the protonation agent to evaluate the basicity of the **Ni_4** POM. Its polyanionic makes all oxo ligands susceptible to behave as Brønsted bases. Still, the symmetry of the catalyst allows identifying equivalent oxo ligands (see Figure 3.6). Among all the non-equivalent positions, two main types of oxo-ligands can be distinguished, those occupying bridging positions (O−M−O) and those in the terminal positions (M=O). Bridging positions present two σ interactions with the metal and the terminal only one σ interaction together with an extra π interaction with the same metal. Previous computational studies have demonstrated that bridging oxygens are typically more basic than terminal ones.[145] This can be ilustrated in the simple case of the Keggin ion, for which the molecular electrostatic potential (MEP) is represented in Figure 3.5. The red-colored region represents the most electron-rich regions of the POM, which correspond to the parts of the surface occupied by the bridging oxo-ligands, whereas the terminal O^{2-} positions are less basic as they interact more strongly with the metal, making them less electron rich, in other words, less basic. When compared to a Keggin-type POM, **Ni_4** presents even more potentially basic bridging oxo-ligand positions, as illustrated in Figure 3.6. Interestingly, the system under study presents some symmetry that makes this evaluation simpler. The **Ni_4** system presents a C_2 symmetry axis and a perpendicular mirror plane (m) as shown in Figure 3.6. The symmetry axis make each trilacunary Keggin moiety equivalent. Furthermore, the symmetry plane divides both trilacunary Keggin moieties, making the oxo ligands inside the highlighted region unique in terms of electronic environment. Knowing that the bridging positions are the most basic ones due to their electronic particularities, the protonation energy of four bridging oxo ligands have been computed using $HTEOA^+$ as Brønsted acid (positions 1–4 in Figure 3.6). The corresponding protonation free energies for these four positions are shown in the table of Figure 3.6, ranked from less to more favourable sites, revealing

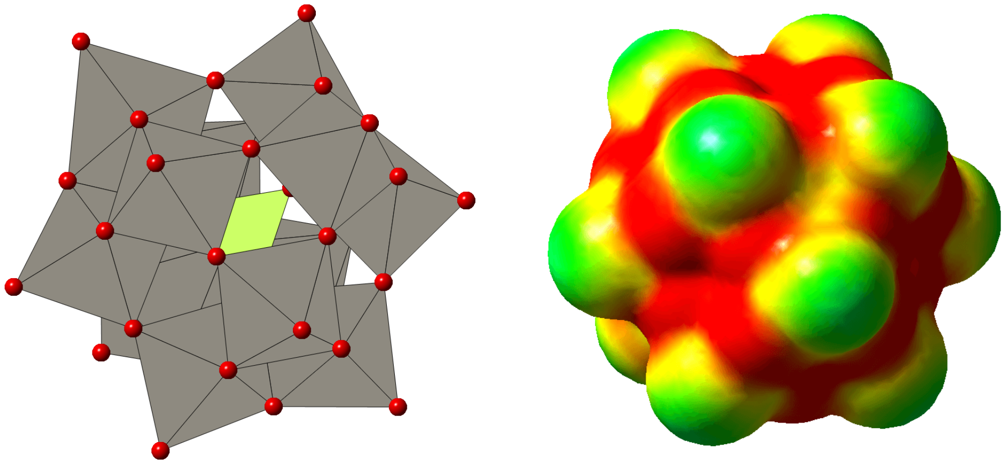
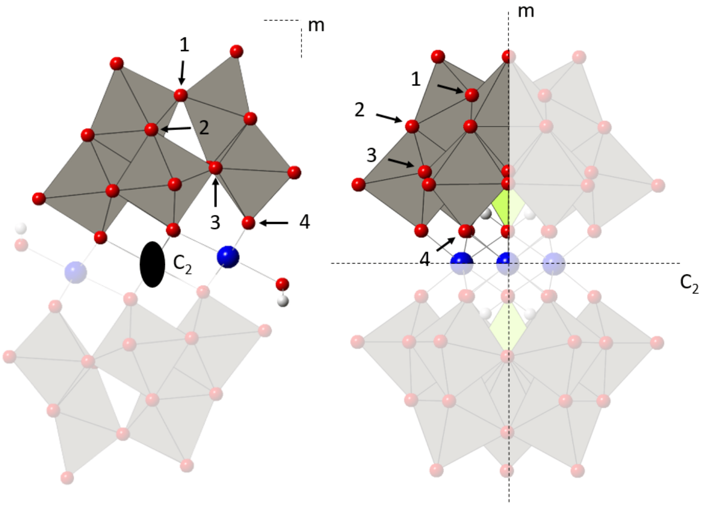


FIGURE 3.5: Left: the polyhedral representation of a Keggin, W octahedra in grey, P tetrahedra in yellow and O as red spheres. Right: molecular surface of electrostatic potential of the same orientation of the Keggin being the energy difference between red and green -11.0 kcal mol^{-1}. The red-colored area represents the most basic regions of the system.

that W–O–Ni are those exhibiting the strongest basicity. This behaviour can be ascribed to the fact that the four Ni^{II} are richer in electrons compared to the W^{VI} ones, meaning that they are weaker Lewis acids, which renders O atoms directly bound to them more electron rich than those connecting two W^{VI} ions. Regarding the positions 1–3, even if the nature of the neighboring metals are the same, the closer to the Ni^{II} sites, the more basic they are. This trend is in line with the corresponding ESP surface of the $\mathbf{Ni_4}$ system illustrated at the left of Figure 3.7 in which the red-colored region, representing the most basic area of the system, is concentrated around the Ni^{II} centres. The other bridging oxo-ligands also comprise more basic areas compared to the terminal positions, in a similar fashion than in the Keggin system at the left of Figure 3.5.

These basic positions of $\mathbf{Ni_4}$ were found to spontaneously abstract up to two protons from $HTEOA^+$, the reaction Gibbs free-energies for the first and second proton transfer events being -6.8 and -5.1 kcal mol^{-1}, respectively. The protons in the doubly protonated anion ($H_2\mathbf{Ni_4}$) are accommodated on two Ni–O–W oxo-ligands (at the right of Figure 3.7) which corresponds to symmetric sites of the position 4 in Figure 3.6. Furthermore, intramolecular hydrogen bonds with neighboring O sites of the POM further favours the protonation process to occur at these positions. By symmetry, two more equivalent sites 4 are still available, but the evaluation of a third protonation turned to be unfavourable, being endergonic by about 4.2 kcal mol^{-1}. With this in hand, we next explored the redox properties of $H_2\mathbf{Ni_4}$ and its potential reduction by the employed Ru-based PS via single-electron transfer (SET) steps. Throughout this process, we also considered additional protonation events of the reduced form of the POM, owing to its increased negative charge density.



	ΔG_{prot} (kcal mol^{-1})
1	8.9 (W^{VI}-O-W^{VI})
2	2.4 (W^{VI}-O-W^{VI})
3	1.6 (W^{VI}-O-W^{VI})
4	-6.8 (W^{VI}-O-Ni^{II})

FIGURE 3.6: At the left the symmetry elements on the C_2 symmetry axis orientation of the polyhedral representation on $\mathbf{Ni_4}$ coloring the octahedral W in grey, the tetrahedral P in yellow, the O atoms in red, the Ni atoms in blue and H atoms in white; indicating the calculated basic positions from 1 to 4. At the centre the m plane orientation. At the right a table of the free protonaton energies of the positions explored.

Exploration of the electronic structure and redox properties of the $H_2\mathbf{Ni_4}$ species

A first exploration of the redox properties of a system involves the calculation of the molecular orbitals (MOs) diagram in order to identify the nature of the lowest energy virtual orbitals involved in the reduction processes of the system. The calculated MOs for the $\mathbf{Ni_4}$ POM indicate that the lowest unoccupied molecular orbital (LUMO) and the following ones, correspond to d-type orbitals of the W centres (Figure B.1). Noticeably, the d-type orbitals of Ni are inserted into the "band" of W centred d-type orbitals. This kind of MO distribution is common in POMs whereby the frontier MOs correspond to the occupied p-type orbitals of the oxo-ligands and the virtual d-type orbitals belonging to the metals. Recovering the concept of the ESP surface of $\mathbf{Ni_4}$ (Figure 3.7), the trilacunary Keggin moiety of the $\mathbf{Ni_4}$ system, presents the most positive region. In line with this fact, the MOs diagram indicates that the lowest-energy unoccupied orbitals correspond to the W centres most distant from the Ni^{II} region. The diagrams of the mono- and the bis-protonated system ($H_2\mathbf{Ni_4}$) (Figure B.2 left and right side respectively) show that all the orbitals are systematically lower in energy when compared to those of $\mathbf{Ni_4}$, supporting that protonating the POM make all W centres easier to reduce.[160] Such calculated virtual orbitals of a system are indeed useful to anticipate its potential reduction sites. However, they are just a prediction derived from the optimized occupied orbitals distribution that do not account for possible interactions between the incoming electrons and the electronic structure of the POM. In order to evaluate the stabilization of these extra electrons of the system, the reduced versions of the $H_2\mathbf{Ni_4}$ POM were calculated

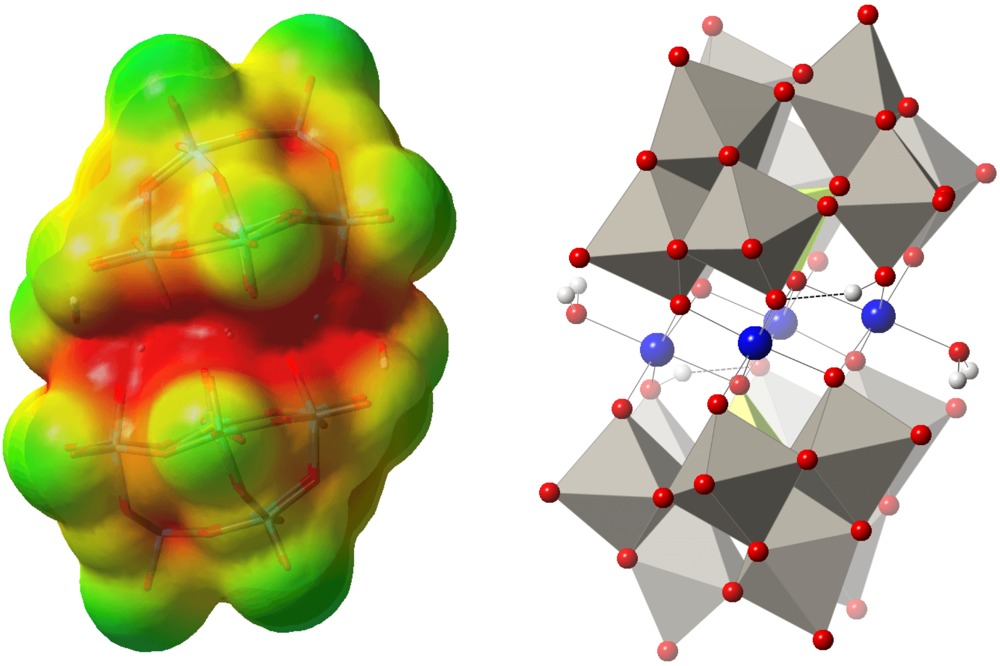


FIGURE 3.7: At the left the ESP energy surface overlapped with the tube representation of $\mathbf{Ni_4}$ being the energy difference between red and green -90.9 kcal mol^{-1}. At the right, the polyhedral representation of $H_2\mathbf{Ni_4}$ colored the octahedral W in grey, the tetrahedral P in yellow, O atoms in red, Ni atoms in blue and H atoms in white. Hydrogen bonds are represented as dashed lines.

next. Figure B.3 represents the spin density of $H_2\mathbf{Ni_4}$(1e) with a spin multiplicity of 10, that is, assuming that the extra electron is magnetically decoupled from spin-up electrons located on Ni^{II} centres. Due to the virtual orbitals of the initial system indicates that the most accessible positions involve spin down or beta electrons on the d-type orbitals of the W centres, the spin multiplicity of 8 of $H_2\mathbf{Ni_4}$(1e) was also explored. This was found to be nearly degenerate with the high-spin solution, with a free energy difference of only 0.5 kcal mol^{-1}, indicating that even if the spin-down electron is localized in a different W centre than the spin-up suggesting a possible coupling with Ni^{II}, it barely affects the stabilization of the system. Thus, in the following, we assumed for simplicity that all reduction events of the POM increase the multiplicity of the system. In other words, that the added electrons are magnetically decoupled from the others.

Identifying the nature of the electron-donor species

Before exploring the feasibility of the $H_2\mathbf{Ni_4}$ to be reduced, it is crucial to evaluate the reducing ability of the electron source, in this case the photocatalytic system conformed by the PS and the SED. As mentioned in the experimental section, a photocatalytic system can operate via either reductive or oxidative quenching pathways (Figure 3.8). The oxidative quenching pathway involves a SET process from the triplet metal to ligand charge transfer (MLCT) exited state of $\mathbf{Ru}^{2+}$ to $H_2\mathbf{Ni}_4$ followed for a SET from the SED to the resulting $\mathbf{Ru}^{3+}$. On the other hand, the reductive quenching involves first the reduction of the photoexcited state of the PS by the SED followed by the reduction of $H_2\mathbf{Ni}_4$ by the reducing $\mathbf{Ru}^{+}$ species. As shown

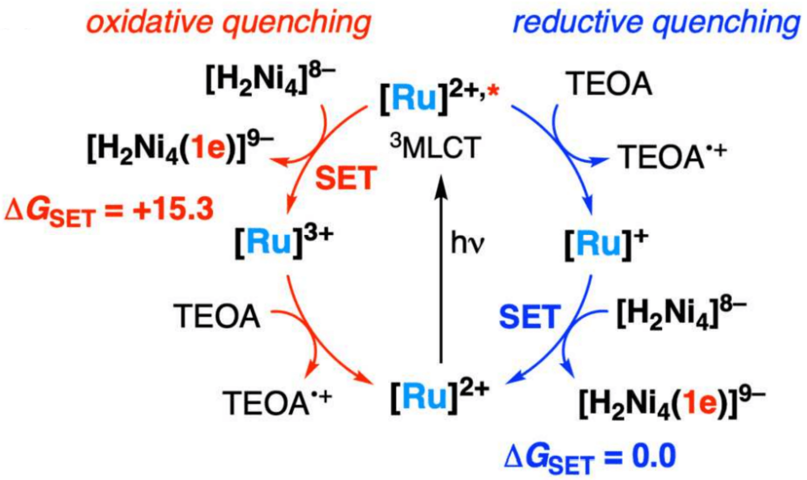


FIGURE 3.8: Reductive and oxidative quenching scheme of $[Ru(bpy)_3]^{2+}$ ($\mathbf{Ru}^{2+}$) as PS.

in Figure 3.8, the first step of oxidative quenching pathway involves the SET from the triplet MLCT state of $\mathbf{Ru}^{2+}$ to $H_2\mathbf{Ni}_4$, which has been calculated as endergonic (+15.3 kcal mol^{-1}). Even though this might be accessible at room temperature, the charge recombination between the reduced POM and the oxidized PS was found to be strongly favourable (by 64.5 kcal mol^{-1}), thus preventing the efficient formation of the reduced POM. In fact, experimental photophysical studies revealed that although the presence of the POM can quench the excited state of the PS, the charge-separated state cannot be detectable on a nano-seconds time scale. This strongly suggest that the $[Ru(bpy)_3]^{3+}$ species formed after electron injection into the POM is unable to escape the association complex before charge recombination, rendering the oxidative quenching pathway unlikely. Moreover, experiments in the presence of TEOA revealed the formation of a long-lived $[Ru(bpy)_3]^{+}$ species (reduced form of the PS), which decay is accelerated upon increasing the POM concentration. Hence, photophysical studies clearly point to a reductive quenching pathway.

Our DFT calculations indicate that the reduction of the exited PS by TEOA is endergonic (by 10.8 kcal mol^{-1}), although the spontaneous deprotonation of the as-formed $TEOA^{\bullet +}$ by a second molecule of TEOA, which is present in large excess in the reaction mixture, is exergonic by 11.3 kcal mol^{-1} (see Figure 3.4), providing the thermodynamic driving force for the formation of $\mathbf{Ru}^{+}$. As shown in Figure 3.8, the reduction of $H_2\mathbf{Ni}_4$ by $\mathbf{Ru}^{+}$ was estimated to be ergoneutral. Therfore, these results strongly indicate that the $\mathbf{Ni}_4\backslash\mathbf{Ru}^{2+}\backslash$TEOA photocatalytic system operates through a reductive quenching mechanism (Figure 3.8, blue arrows), further supporting the mechanistic picture inferred from photophysical studies. It is worth noting that as previously reported in the literature, the oxidation of TEOA leads to the formation of aldehyde groups from the alcohol ones for each substituent of the amine.[12] The alcohol groups need to transfer two electrons and two protons to be oxidized to aldehydes. Mechanistically, this implies the formation of a neutral $TEOA^{\bullet}$ radical (Figure 3.4).[228] The latter is a stronger reducing agent compared to TEOA, producing the correspondent cation $TEOA^{+}$ upon one-electron oxidation, which further evolves to an enol via deprotonation and finally tautomerizes into the aldehyde. Hence, it is also likely that $TEOA^{\bullet}$ radicals involved in the chain decomposition mechanism of TEOA also participate in the formation of $\mathbf{Ru}^{+}$. However, according to experimentally determined oxidation potentials, $\mathbf{Ru}^{+}$ is a stronger reducing agent compared to $TEOA^{\bullet}$ (-1.33[229] and -1.12[230] V vs SCE, respectively and therefore, we used $\mathbf{Ru}^{+}$ to evaluate subsequent reduction processes of the POM catalyst. It is also relevant here that protonation of the $\mathbf{Ni}_4$ is essential to allow efficient charge accumulation in the catalyst. As a matter of fact, the SET from $\mathbf{Ru}^{+}$ to the non-protonated $\mathbf{Ni}_4$, which bears an overall charge of -10, is much less favourable compared to that to the doubly-protonated $[H_2\mathbf{Ni}_4]^{8-}$ anion, with a reaction free-energy of +16.9 kcal mol^{-1} vs 0.0 kcal mol^{-1} (see Figure 3.9). This can be ascribed to the polyanionic nature of the POM, as reducing its overall negative charge via protonation lowers the energy of its unoccupied MOs, favouring in turn reduction events. Also, possible ion-pairing effects involving the POM, which are not explicitly captured in our models; are expected to ease SET processes due to the same principle, although to a lesser extent than protonation. Hence, it cannot be ruled out that the SET processes from the PS to the POM are actually somewhat more exergonic than what computed values suggest.

Exploration of the redox and protonation states accessible to the POM under photocatalytic conditions

Once we explored the protonation ability of the POM and the reduction ability of the photocatalytic system, we aimed to identify the nature of the active form of the catalyst under photocatalytic conditions. This is crucial to understand the interaction between the catalyst and the reactant, CO_2 in this case. Due to the synergy between the reduction and the protonation of POMs[160], multiple protonated and reduced states of $\mathbf{Ni}_4$ have been explored (Figure 3.9). The first and second reduction of the $\mathbf{Ni}_4$ catalyst were computed to be endergonic by +16.9 and +20.8 kcal mol^{-1}, respectively. However, we showed that the catalyst is basic enough to be spontaneously protonated twice ($H_2\mathbf{Ni}_4$) by -6.8 and -5.1 kcal mol^{-1} for the first and the second protonation processes, respectively. Then, the first reduction process of the doubly protonated catalyst was evaluated to be ergoneutral generating $H_2\mathbf{Ni}_4$(1e) and the second one endergonic by 7.2 kcal mol^{-1} producing $H_2\mathbf{Ni}_4$(2e). The protonation reduces the overall negative charge of the POM, thus lowering the energy of its unoccupied MOs, and further favouring the reduction events, as previously discussed. Despite this last process was computed to be slightly endergonic, the subsequent

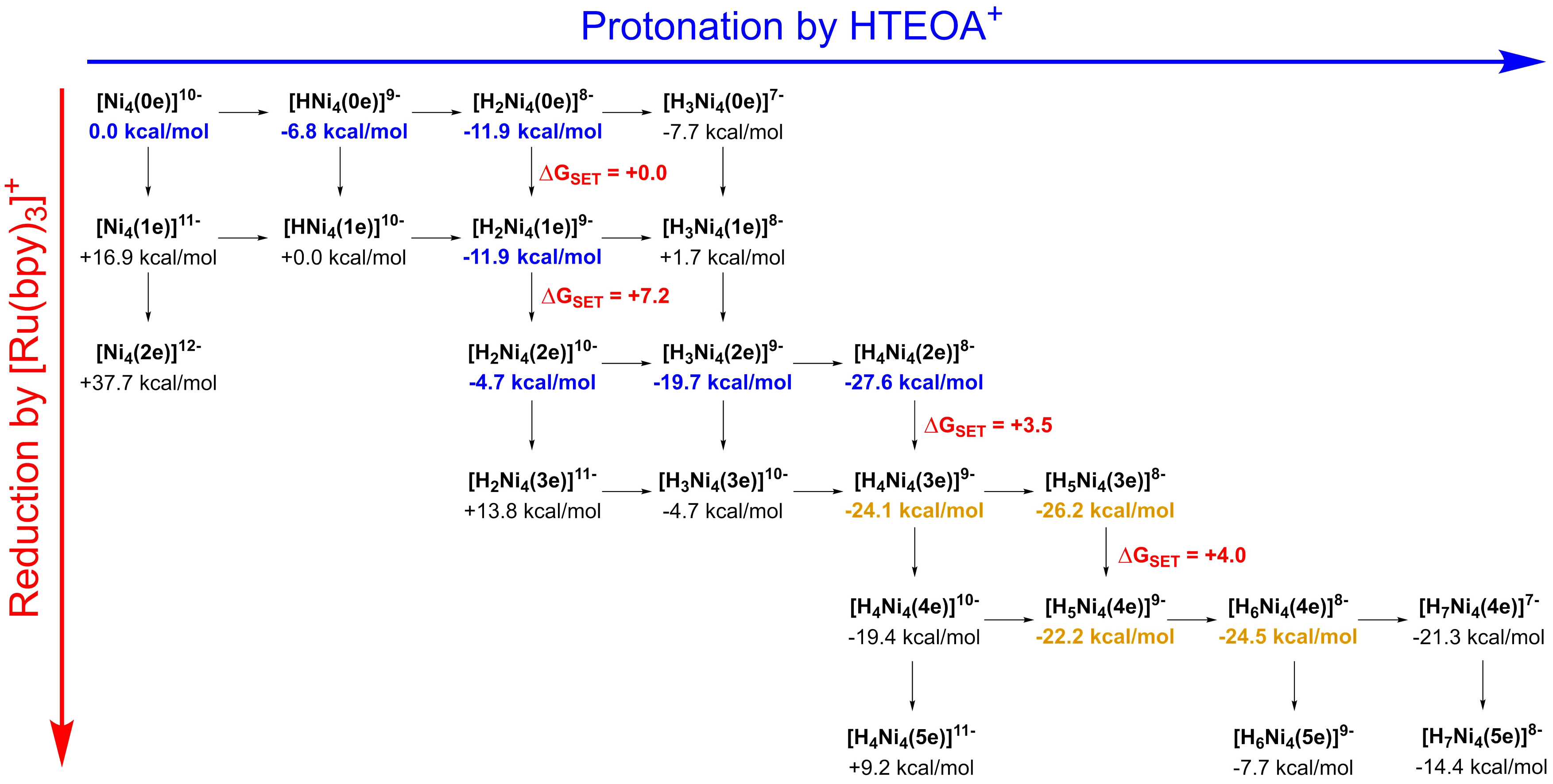


FIGURE 3.9: Gibbs free-energy diagram of all the protonation and reduction processes explored for $\mathbf{Ni_4}$ in blue the less energetic path for the resting state of the catalyst and in orange the path for feasible other intermediates.

spontaneous association of two additional protons to the remaining bridging Ni–O–W positions (-15.0 and -7.9 kcal mol^{-1}), provides the thermodynamic driving force to the second SET process while preventing undesired back electron transfer from the reduced POM to the PS. It is worth mentioning here that the protonation of the most basic region drives the localization of the electrons to the W positions neighboring Ni centres. The resulting 4-fold protonated and two-electrons reduced POM, $H_4\mathbf{Ni_4}$(2e), exhibits an accessible third reduction in standard conditions by 3.5 kcal mol^{-1}. However, in this case, its protonation is exergonic by only 2.1 kcal mol^{-1}, being thus unable to compensate the cost of its reduction. Thus, from $H_4\mathbf{Ni_4}$(2e) onwards, all further reduction\protonation events proceed uphill (see Figure 3.9).

Overall, Figure 3.9 summarizes the free-energies associated to all the explored processes, highlighting in blue the most likely reduction\protonation path leading to $H_4\mathbf{Ni_4}$(2e). Still, some species were found to be close to that one in energy (orange path), opening the possibility for some of them to coexist in a significant concentration. In order to evaluate the concentration of each potential intermediate, a Boltzmann distribution was calculated to quantify the probability (P_i) of each species (*i*) with energy (E_i) of being populated at thermal equilibrium in standard conditions over the contribution of all the participating species (*j*), following Eq. 3.3, using the ideal gas constant (*R*).

$$\mathrm{P}_i = \frac{e^{-E_i/(RT)}}{\sum_{j=1}^{N} e^{-E_j/(RT)}} \tag{3.3}$$

TABLE 3.1: Gibbs free energy, population probability (calculated using Boltzmann distributions), and aqua ligand substitution energy by CO_2 for the four most stable catalyst intermediates.

System	**Free energy (kcal mol^{-1})**	**Probability (%)**	**$H_2O \rightarrow CO_2$ substitution energy (kcal mol^{-1})**
$H_4\mathbf{Ni_4}$(2e)	-27.6	90.72	34.1
$H_4\mathbf{Ni_4}$(3e)	-24.1	0.25	–
$H_5\mathbf{Ni_4}$(3e)	-26.2	8.54	27.0
$H_6\mathbf{Ni_4}$(4e)	-24.5	0.48	31.4

As shown in Table 3.1, the 90.72% of the catalyst is estimated to be in the form of $H_4\mathbf{Ni_4}$(2e) under photocatalytic conditions, followed by $H_5\mathbf{Ni_4}$(3e) (8.54%), $H_6\mathbf{Ni_4}$(4e) (0.48%) and $H_4\mathbf{Ni_4}$(3e) (0.25%). The concentration of the other intermediates appearing in Figure 3.9 were estimated to be negligible. Hence, subsequent steps of the reaction mechanism were analyzed focusing on $H_4\mathbf{Ni_4}$(2e). Of note, this species already hosts the number of extra electrons required to promote the conversion of CO_2 into CO, as observed experimentally.

Towards CO_2 insertion and CO production

Next, the mechanism of the experimentally observed reduction of CO_2 to CO was explored. From a thermodynamic point of view, the reduction of CO_2 to CO promoted by $H_4\mathbf{Ni_4}$(2e), giving a water molecule as a byproduct and regenerating the fullyoxidized $H_2\mathbf{Ni_4}$ anion (Eq. 3.4) was computed to be slightly endergonic by 4.5 kcal mol^{-1}.

$$H_4\mathbf{Ni}_4(2e) + CO_2 \rightarrow H_2\mathbf{Ni}_4(0e) + CO + H_2O \tag{3.4}$$

Out of curiosity, the same energy balance was calculated for $H_5\mathbf{Ni_4}$(3e), generating the oxidized $H_3\mathbf{Ni_4}$(1e) system (Eq. 3.5), being endergonic by 16.7 kcal mol^{-1}. This energy difference comes from the fact that $H_3\mathbf{Ni_4}$(1e) is less stabilized than $H_2\mathbf{Ni_4}$(0e).

$$H_5\mathbf{Ni}_4(3e) + CO_2 \rightarrow H_3\mathbf{Ni}_4(1e) + CO + H_2O \tag{3.5}$$

Inspired by the reaction mechanisms reported for other molecular CO_2 reduction catalysts containing redox-active ligands, [231–234] we initially explored the replacement of the aqua ligand by CO_2 in the coordination sphere of one of the accessible Ni^{II} centres of $H_4\mathbf{Ni_4}$(2e). The spin density distributions of the resulting species (Figure 3.10) indicates that indeed, such ligand-exchange process induces the transfer of one electron from the POM framework to the incoming CO_2 ligand, which coordinates to the Ni^{II} site through the carbon atom and adopting a bent configuration.

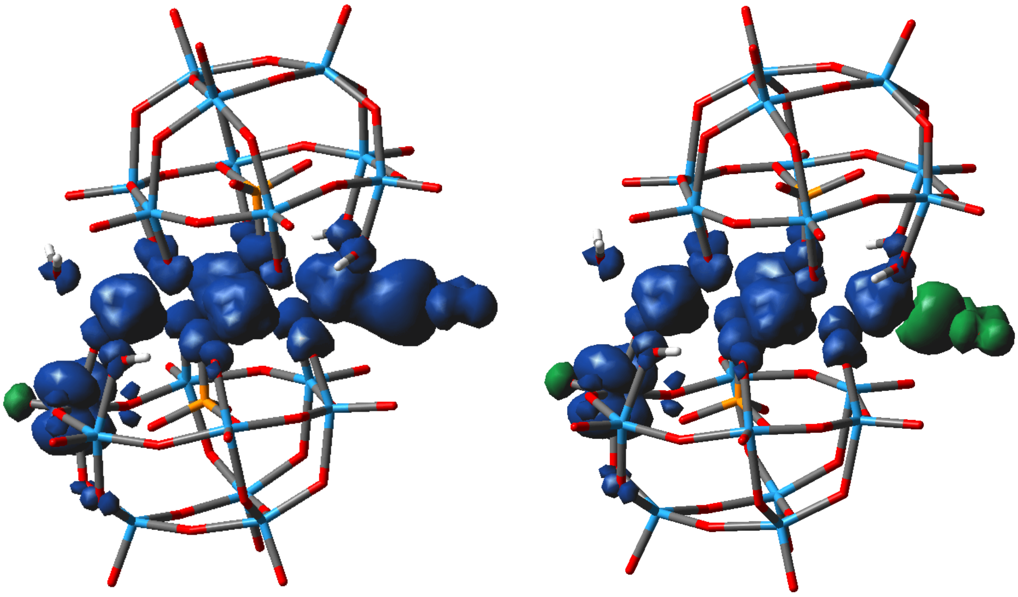


FIGURE 3.10: Spin densities of $H_4\mathbf{Ni_4}$(1e)-$COO^{\bullet-}$. At the left, the spin up state where the electron of -$COO^{\bullet-}$ is not coupled to those of Ni^{II}, at the right the spin down configuration in which the extra electron on the CO_2 moiety is coupled with Ni^{II}.

Assuming no variation of the total spin multiplicity (left of Figure 3.10) this process was computed to be prohibitively endergonic by about 53.5 kcal mol^{-1}. However, exploring the possible antiferromagnetic coupling between the available spin up electrons of the Ni^{II} and the transferred electron to the -$COO^{\bullet-}$ ligand (right of Figure 3.10) resulted in a strong stabilization of the system, reducing the free-energy cost of the ligand exchange (and associated CO_2 activation) from 53.5 to 34.1 kcal mol^{-1}. Notably, despite the stabilizing effect of the coupling, this energy cost is still too high to afford efficient CO_2 reduction at the experimental conditions. However, as shown in Table 3.1, the calculated free energy costs considering electronic coupling are all quite alike, suggesting that the protonation and redox state of the POM does not have a strong impact on the overall $H_2O \rightarrow CO_2$ ligand substitution energy. Other commonly employed correlation-exchange functionals (B3LYP-D3BJ or PBE0) provided a very similar free-energy cost of +34.3 kcal mol^{-1}, suggesting that such a high energy cost does not arise from a methodological artifact coming from the nature of the employed density functional.

Splitting the ligand substitution process into two elementary steps, (I) the decoordination of the aqua ligand followed by (II) the insertion of CO_2, reveals individual free energy costs of 26.5 and 7.6 kcal mol^{-1}, respectively, indicating that the most limiting step corresponds to the release of the aqua ligand to the solvent bulk to generate a ligand vacancy on the Ni^{II} centre where CO_2 can bind. On these grounds, we developed two possible hypotheses to explain the experimentally observed reactivity:

- **Hypothesis 1**: As CO_2 reduction takes place under constant light irradiation, the formation of a $W^V \rightarrow Ni^{II}$ excited state in $H_4\mathbf{Ni_4}$(2e) may facilitate H_2O release, allowing subsequent CO_2 binding and activation.

- **Hypothesis 2**: Due to the complexity of aqueous solutions, characterized by hydrogen-bond networks and cavitational effects, our static DFT calculations, combined with implicit solvent models, may not accurately capture the free-energy change associated with releasing a water molecule into the bulk solvent.

To probe Hypothesis 1, we carried out time-dependent Density Functional Theory (TD-DFT) calculations to investigate the photochemical properties of $H_4\mathbf{Ni_4}$(2e). While, $\mathbf{Ni_4}$ does not absorb in the visible range, the presence of d(W) electrons in $H_4\mathbf{Ni_4}$(2e) may allow excitations at higher wavelengths, as commonly observed in reduced classical POMs.[144] Indeed, the simulated UV-vis spectrum of $H_4\mathbf{Ni_4}$(2e) (see Figure 3.11) reveals electronic transitions in the visible range. However, these are exclusively associated to $W^V \rightarrow W^{VI}$ transitions and do not involve Ni^{II} centres. Hence, there are no indications that the formation of $H_4\mathbf{Ni_4}$(2e)-centred excited states could assist the Ni-aqua dissociation.

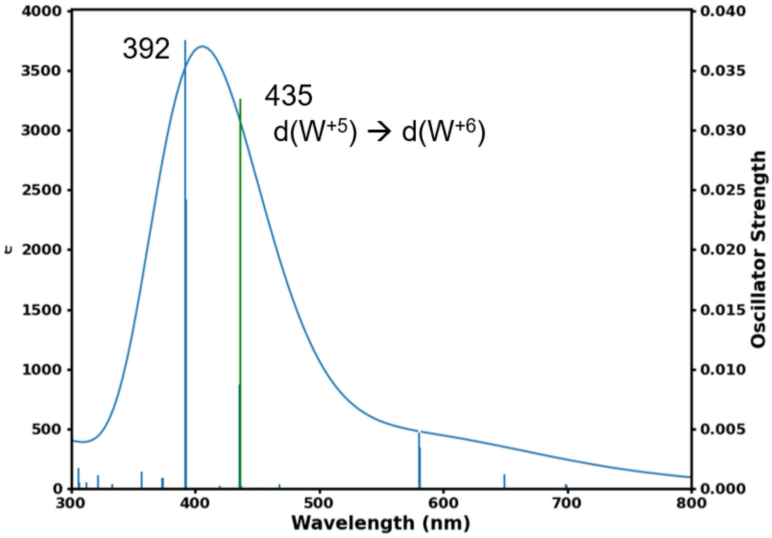


FIGURE 3.11: Simulated UV-vis spectrum of H_4**Ni**$_4$(2e).

Probing Hypothesis 2 requires a significantly more complex computational setup. To investigate the ligand-exchange process within a more realistic and dynamic environment, we carried out DFT-based MD simulations on the system represented in Figure 3.12a. With the currently available computational power, DFT-MD simulations can be only performed in the picoseconds time scale. Hence, the Hamiltonian was augmented with a bias potential using the path-metadynamics technique, [222, 223] which allows the simulation of so-called *rare events*, occurring through barriers that are higher than those associated to thermal fluctuations, often of the order of K_bT. These simulations do not only allow obtaining mechanistic insights but also to reconstruct the free-energy profile along the reaction coordinate to retrieve kinetic and thermodynamic data.

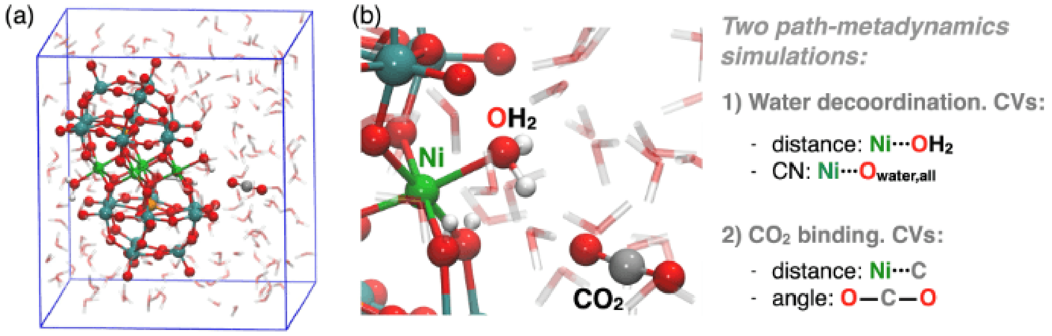


FIGURE 3.12: (a) Simulation cell used in DFT-MD simulations, containing a H_4**Ni**$_4$(2e) catalyst and a CO_2 molecule embedded in explicit water solvent, accounting for a number of atoms equal to 615. (b) Definition of collective variables (CV) combined in the path collective variable (σ) for each of the path-metadynamics simulations.

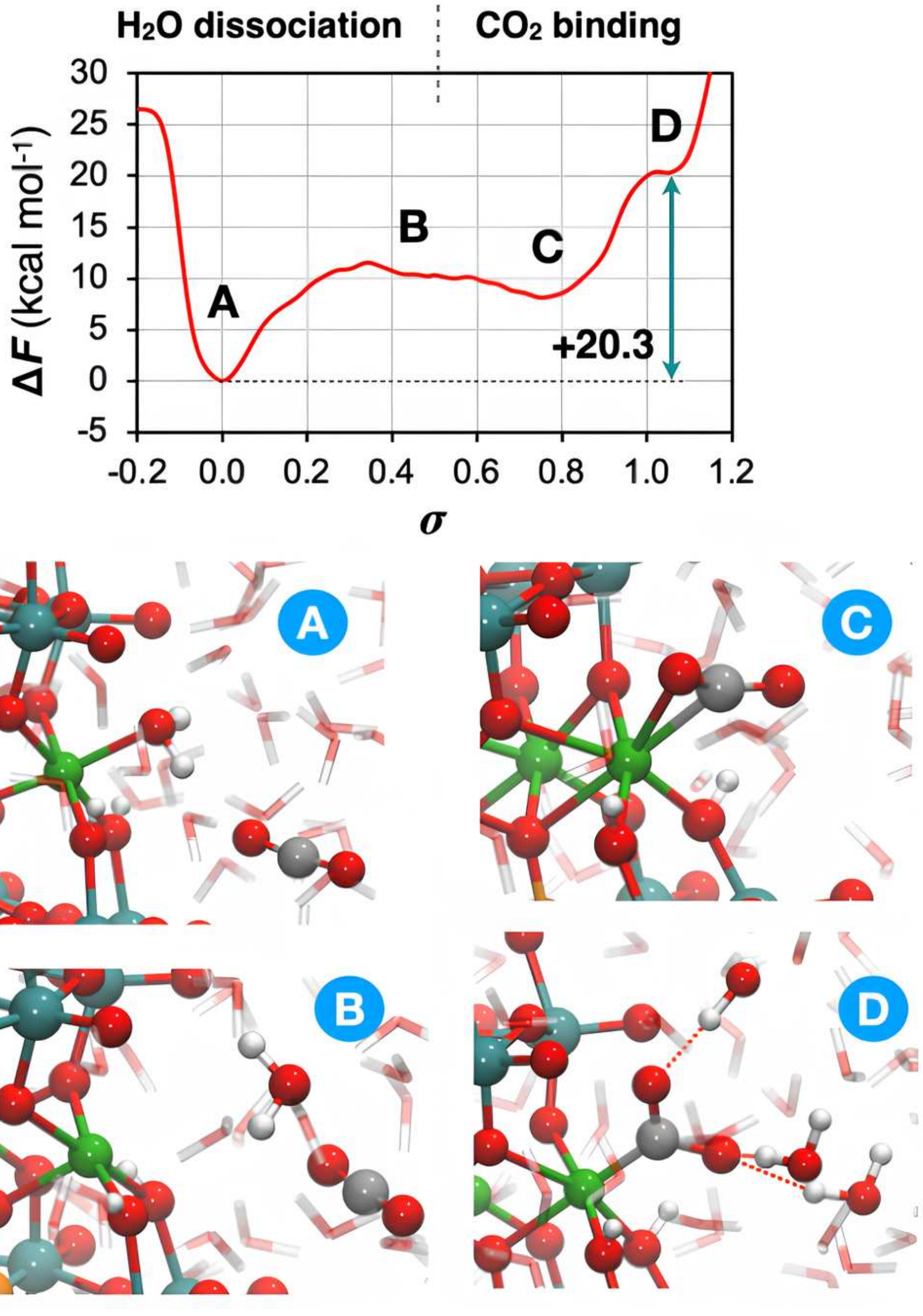


FIGURE 3.13: Reconstructed free-energy profile along the ligand exchange process of H_2O by CO_2 in by H_4**Ni**$_4$(2e). Representative snapshots of the most relevant species are illustrated as balls-and sticks.

Specifically, as illustrated in Figure 3.12b, two independent path-metadynamics simulations were conducted, aimed at investigating each of the elementary steps involved in the ligand-exchange process (see section 3.2.3 for details). The combination of the reconstructed free-energy landscapes leads to the profile shown in Figure 3.13, where σ represents the reaction coordinate, being 0 at the Ni–OH_2 state and 1 at the Ni–$COO^{\bullet -}$ intermediate. Interestingly, these simulations indicate that going from state A to B (aqua ligand release) entails a free-energy cost of only 10 kcal mol^{-1}, in sharp contrast with that of 26.5 kcal mol^{-1} determined via static calculations. Then, the approach of CO_2 leads to a transient η^2 physiosorbed intermediate C, which further evolves toward D, in which an electron from the POM framework has been effectively transferred to the CO_2 moiety. Importantly, the overall free-energy cost for the ligand-exchange process of +20.3 kcal mol^{-1} obtained from DFT-MD simulations support that the activation of CO_2 by $H_4\mathbf{Ni}_4$(2e) is indeed feasible at the experimental conditions, despite proceeding uphill. However, if no constraints in the shape of repulsive potential walls are included to prevent the protonation of the oxygen atoms of CO_2, we observe a rapid proton transfer from a neighboring Ni–O(H)–W site to the as formed Ni–$COO^{\bullet -}$ intermediate involving a solvent water molecule as a proton shuttle. Such a protonation is thus expected to stabilize the CO_2-bound intermediate, preventing reversibility in the CO_2 binding to some extent.

Of note, the cost of the individual CO_2 insertion step as obtained from DFT-MD simulations (ca. 10.3 kcal mol^{-1}) is much closer to that obtained from static calculations (7.6 kcal mol^{-1}), indicating that the overestimation of the free-energy cost of this process observed in static calculations with implicit solvent arises from the poor description or oversimplification of the variation of the water molecule environment in going from the coordination sphere of Ni to the solvent bulk.

Once shown that CO_2 can be activated on $H_4\mathbf{Ni}_4$(2e), we studied the following steps in the reaction mechanism to form the CO product relying on static DFT calculations. The calculated free-energy profile is represented in Figure 3.15. As anticipated by DFT-MD simulations, upon losing an electron from the W centres, the Ni^{II}–O–W^{VI} groups become less basic, making the protonation of the carboxylate radical anion (–$COO^{\bullet -}$) by a Ni–O(H)–W site spontaneous. This proton-transfer event was calculated to be exergonic by 11.2 kcal mol^{-1} and leads to the formation of **2′** (Figure 3.15 and Figure B.4). Besides stabilizing the CO_2-bound intermediate, such a proton transfer triggers a second SET from the reduced POM to the CO_2 moiety, forming a Ni–COOH intermediate in which both extra electrons have been already transferred to CO_2.

Again, as the charge density on the POM framework decreases in going from **2** to **2′**, its structural oxygen atoms become less basic. As such, the abstraction of a second proton by TEOA becomes favourable, being slightly exergonic, by 0.7 kcal mol^{-1}. Thus, the deprotonation of **2′** generates species **3** (Figure 3.15 and Figure B.4), from which the protonation of the OH group in the Ni–COOH intermediate by

$HTEOA^+$ induces the cleavage of the C–OH bond. This process leads to intermediate **4** (Figure 3.15 and Figure B.4), in which the CO product remains coordinated to the Ni centre, releasing a water molecule, and takes place through the **TS**, overcoming a low barrier of only 4.3 kcal mol^{-1}. The structure of the TS is represented in Figure 3.14. The formation of **4** from **3** was predicted to be endergonic by 1.5 kcal mol^{-1}. However, with the knowledge gained from path-metadynamics simulations, it is reasonable to think that this process would be exergonic if the release of the water molecule byproduct could be described within a more realistic explicit and dynamic environment. In fact, applying a free-energy correction equal to the difference in the cost of releasing a water molecule, as predicted by static DFT and DFT-MD (*vide supra*) renders this step exergonic by 15.0 kcal mol^{-1} (see Figure 3.15). Finally, the release of the CO product and regeneration of the active form of the catalyst driven by reduction and protonation events proceeds downhill and provides the thermodynamic driving force for the whole reaction (Figure 3.15).

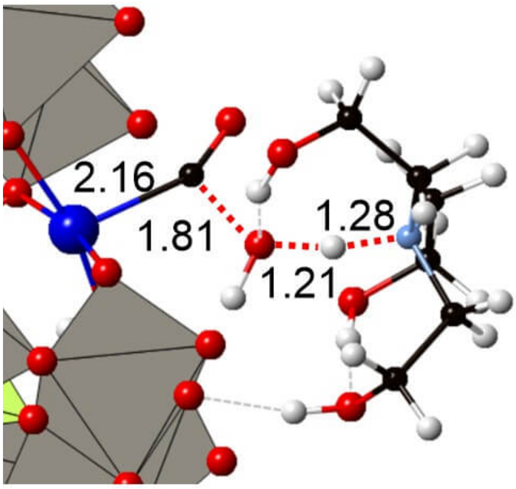


FIGURE 3.14: Transition state for the concerted OH protonation by $HTEOA^+$ and C–OH bond cleavage. Octahedral W is shown in gray, Ni in blue, C in black, O in red, N in cyan, and H in white.

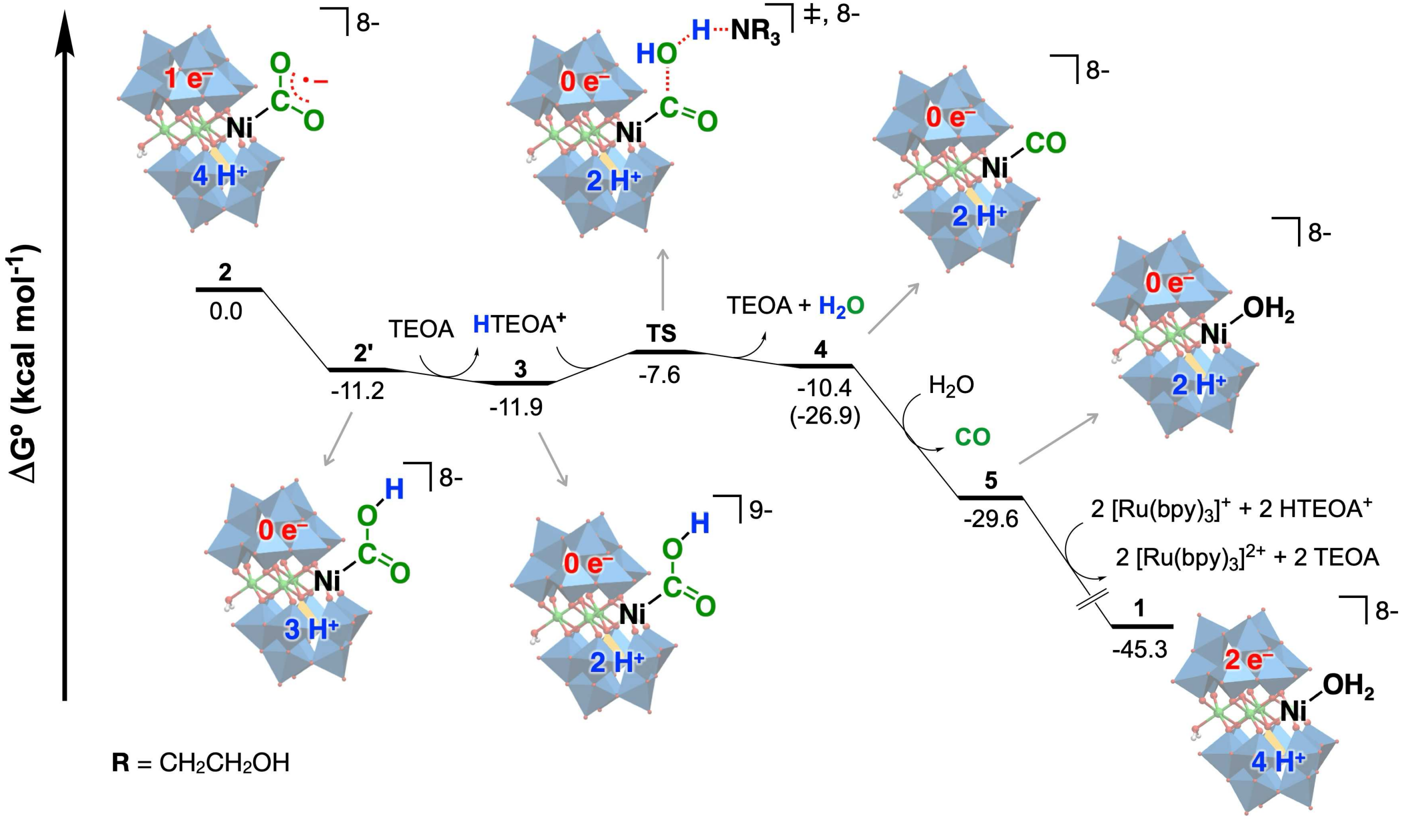


FIGURE 3.15: Calculated free energy profile for the formation of CO from the Ni-COO$^{\bullet-}$ intermediate (**2**) resulting from the activation of CO$_2$ on the H$_4$**Ni**$_4$(2e) catalyst. The value in parenthesis corresponds to the relative energy of species **4** upon applying a DFT-MD derived correction to the free-energy of a water molecule in solution.

3.2.5 Conclusion

With the gained knowledge about the **Ni_4**/**Ru^{2+}**/TEOA photocatalytic system and on the basis of previously reported reaction mechanisms for CO_2 reduction by transition-metal-based compounds bearing redox-active ligands, we propose the reaction mechanism illustrated in Figure 3.16. Under photocatalytic conditions, the catalyst was found to accept 2 electrons and 4 protons, from the **Ru^{2+}**/TEOA photocatalytic system via reductive quenching pathways, generating its active form. From this species, CO_2 coordinates through its carbon atom to a Ni^{II} ion, replacing an aqua ligand and generating a Ni^{II}–$COO^{\bullet -}$ intermediate. Two sequential proton shifts allows the migration of the second electron from the POM framework to the CO_2 moiety and the cleavage of the C–O bond, leading to the CO product in addition to a water molecule. Finally, the active form of the catalyst is regenerated by **Ru^{+}** and $HTEOA^{+}$, acting as electron and proton donor species, respectively, both formed via light-induced reductive quenching of the photosensitizer by TEOA, thereby closing the catalytic cycle. Our computational results suggest that while the redox activity of the POM ligand is key to enable efficient charge photoaccumulation, the ligand-exchange process whereby CO_2 binds to the catalytically active Ni centre may limit the efficiency of thermally activated chemical steps. Hence, there are strong indications that future efforts should focus on developing catalytic platforms that preserve (or outperform) the redox activity of **Ni_4**, while reducing the cost of aqua-by-CO_2 replacement in the coordination environment of the active metal, perhaps through the creation of more hydrophobic reaction cavities or secondary coordination spheres that further stabilize bound CO_2 species.

In a more specific sense, this first part of the chapter supported existing knowledge regarding the roles of all chemical components involved in the catalytic cycle while establishing new structure–function relationships under the specific experimental conditions and for the specific catalyst employed. The redox and acid-base properties of POMs previously reported in the literature were corroborated and contextualized, strengthening their connection to the experimental setup. The **Ni_4** system demonstrated the ability to store electrons in both trilacunary Keggin ligands, making it reactive enough to efficiently reduce CO_2, and interacting productively with the photocatalytic system via its reductive quenching pathway, in line with experimental findings. The role of TEOA as a dual electron and proton source by substituent of the amine group was confirmed. The combined electron relay function of the POMs and the intrinsic catalytic capability of Ni to activate CO_2 support the potential of Ni-substituted POMs as promising catalysts for CO_2 conversion. In particular, the ability of both Keggin ligands to accumulate electrons may account for the improved activity observed experimentally (Figure 3.3). In contrast, other Ni-containig systems studied in this work either possess only one Keggin ligand or have POM fragments that are spatially isolated, limiting the extent of electron storage per active Ni centre. This reduced capacity correlates with diminished reactivity. Overall, this subchapter presented a methodological framework to evaluate

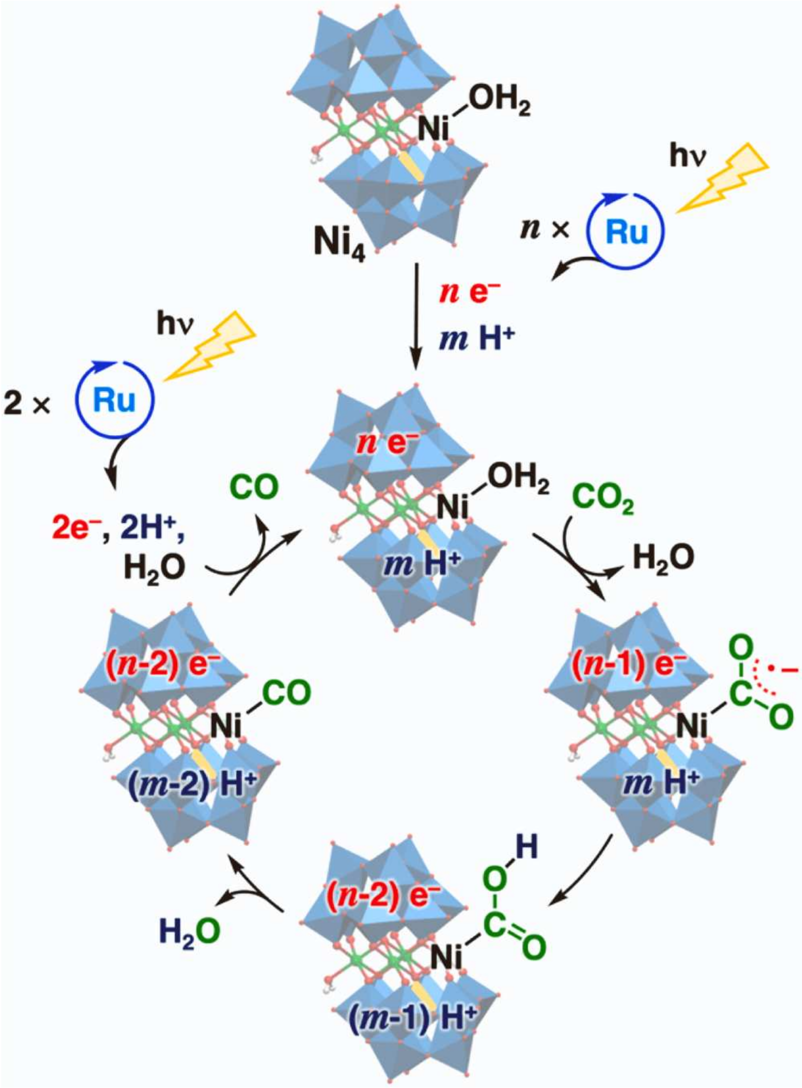


FIGURE 3.16: General scheme of the proposed mechanism, n = 2 and m = 4 for H_4**Ni$_4$**(2e) and n = 3 and m = 5 for H_5**Ni$_4$**(3e).

the reactivity of POM for CO_2 reduction into CO. The insights gained provide a basis for future catalyst design, with the potential to enhance the predictive power of computational modeling and guide the rational development of new photocatalytic systems for sustainable CO_2 conversion.

From a computational perspective, it is noteworthy that the ability of the active form of the catalyst to activate CO_2 could not be confirmed via static DFT calculations using implicit solvent models, due to a systematic overestimation of the free-energy cost associated to releasing an aqua ligand to the water solvent bulk. Instead, supporting the experimentally observed activity of this catalyst required the application of more sophisticated and far from routine work DFT-MD simulations on explicitly solvated systems and that of state-of-the-art enhanced-sampling techniques. We acknowledge that these simulations are way too expensive to be performed on a systematic fashion. However, one should be aware of the limitations and potential misinterpretations of computational results derived from conventional static DFT calculations and implicit solvent models when dealing with complex systems where the explicit role of the environment is poised to play such a determining role.

3.3 Porphyrin-POM photocatalytic dyads for CO_2 reduction

3.3.1 Introduction

Porphyrins are a fascinating class of heterocyclic macrocyclic compounds composed of four pyrrole subunits interconnected via methine bridges, forming a planar and extensively conjugated π-system. This structural arrangement underlies their intense coloration and strong absorption in the visible spectrum, typically with a pronounced Soret band around 400 nm and multiple Q-bands between 500 and 700 nm. These optical features are central to their biological roles in heme (oxygen transport in blood), chlorophyll (photosynthesis), and cytochromes (electron transport). Chemically, the porphyrin ring can coordinate metal ions in its central cavity to form metalloporphyrins (MPs), enabling functions like redox catalysis, light harvesting, and signal transduction. This metal coordination can significantly modulate their redox potentials, magnetic behaviour, and spectral properties, making them highly tunable for both biological and synthetic applications.[235]

From a synthetic standpoint, porphyrins can be produced via methods such as the Adler–Longo or Lindsey synthesis, which involve the condensation of pyrrole with aldehydes under acidic conditions. Transition metal-catalyzed approaches have also been explored to introduce novel functionalities and expand the porphyrin family.[236] Their chemical robustness and resistance to oxidation and degradation make them well-suited for use in harsh environments, and this same stability also explains their preservation in ancient fossils. Additionally, porphyrins can serve as building blocks for advanced materials like MOFs, where their regular arrangement and electronic properties enable applications in gas separation, electrocatalysis, photocatalysis and photonics.[237]

Structurally, porphyrins are mostly planar, but peripheral substitutions or macrocyclic distortions can affect their symmetry and reactivity. Substituted porphyrins, such as tetrakis(p-sulfonatophenyl)porphyrin ($TPPS_4$), have been employed to investigate metal-binding properties using UV-visible and NMR spectroscopy. These studies reveal that transition metals bind more strongly than alkali or alkaline earth metals due to better coordination geometry and ligand field stabilization.[238] Self-assembly of porphyrins into nanostructures via π–π stacking and hydrophobic interactions enables the construction of functional nanomaterials with enhanced light absorption and charge separation, attributes valuable in solar energy and photodynamic therapy.[239]

The versatility of porphyrins arises from their dual chemical reactivity, both as ligands and as photoactive units, making them indispensable in fields ranging from biochemistry and medicine to materials science and green catalysis. Advances in synthesis now allow for precise control over porphyrin core and peripheral modifications, enabling the tailored design of porphyrins for specific roles such as photosensitizers, enzyme mimics, molecular wires, or imaging agents. Their multifunctional nature, arising from a balance of rigid planarity, extended conjugation, and modifiable sites, underscores their centrality in both natural and engineered molecular systems.[240]

Metalloporphyrins as catalysts for CO_2 photo/electro-reduction

Metalloporphyrins (MPs) are synthetic analogs of naturally occurring porphyrin complexes, such as heme, and serve as promising catalysts for carbon dioxide reduction reactions (CO_2RR). Porphyrin-based molecular catalysts incorporating transition metals, typically Fe and Co, have proved effective in reducing CO_2. They have a high affinity for carbon dioxide thanks to the formation of a CO_2-metal bond, while in the excited state they can accept the electrons needed to reduce the substrate.[9, 10] The metal centre profoundly influences the CO_2 activation and reduction steps by modulating electron density and adsorption energies of key intermediates like –COOH and –CO.

For instance, first-principles studies and experimental work demonstrate that cobalt and iron porphyrins (CoTPP and FeTPP) exhibit superior CO_2RR electrocatalityc performance, as their d-band centres provide optimal binding strengths to intermediates, balancing activation and desorption processes.[11] This enables higher Faradaic efficiencies and lower overpotentials for CO generation. For example, Mn-TPP modified with sulfonate groups (Mn-TPPS), immobilized on glassy carbon electrode (GCE), shifts the C–C coupling pathway, facilitating formation of acetate, a two-carbon product rarely achieved with molecular catalysts.[241] The choice of metal centre not only governs product selectivity (e.g., CO vs HCOOH vs hydrocarbons) but also affects reaction energetics, with metals like Zn and Ni forming weaker interactions with intermediates, thereby displaying lower activity.[11] Moreover, structural modifications such as N-confusion, which is the substitution of one or more N of the porphyrin ligand by carbon, can raise the d-band centre, increasing electron donation to the CO_2 molecule and stabilizing the transition states, which further lowers the reduction potential.[242] Additionally, MPs can be functionalized or embedded within hybrid materials, such as covalent organic frameworks (COF) or MOFs, to enhance charge transfer and CO_2 adsorption capabilities.[12] These advances in MPs chemistry not only deepen the understanding of molecular electrocatalysis but also pave the way for designing efficient and selective systems for CO_2 valorization under ambient conditions.

Turning to photocatalysis a seminal example was reported in 2014 by Bonin *et al.* whereby an iron porphyrinic compound was able to perform CO_2RR in homogeneous conditions under visible-light irradiation in presence of an organic PS and a SED, with 100% selectivity for CO production, no deactivation of the catalyst after

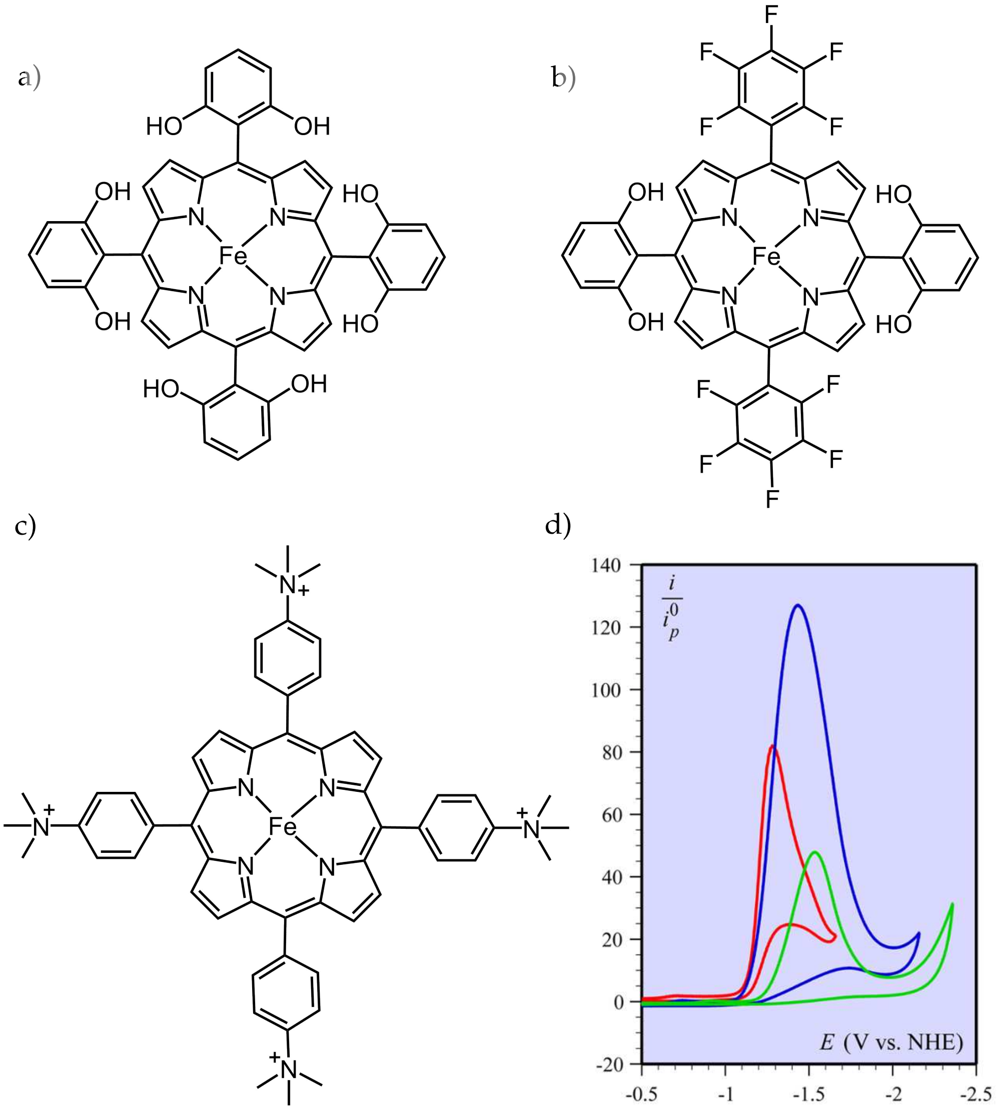


FIGURE 3.17: Representative examples of Fe-based porphyrin derivatives used for electrocatalytic reduction of CO_2 into CO. (d) CV of 1 mM CAT (green) (a), FCAT (blue) (b) and WSCAT (red) (c) in DMF + 0.1 M n-Bu_4NPF_6 at 0.1 V s^{-1} in the presence of 0.23 M CO_2 and of 1 M PhOH; i_p^0, the peak current of the reversible Fe^{II} /Fe^{I} wave at the same scan rate represents a one-electron stoichiometry. Reprinted from reference 246.

50 h and a turn-over number (TON) of 60.[4] But one of the fundamental advantages MPs is that they can be custom functionalised. For example, Sakai's group studied the effect of peripheral groups on the properties of such systems. By introducing sulfonatophenyl groups onto meso-positions of a cobalt porphyrin, a visible-light water-soluble catalyst has been obtained which exhibits a turn-over frequency (TOF) for CO production of 456 h^{-1} in presence of $\mathbf{Ru}^{2+}$ as PS.[243] Furthermore, replacing the sulphonatophenyl substituents with N-methylpyridinium groups and using a Cu^{I} complex as PS led to a TOF as high as 2680 h^{-1} (CO selectivity 77%).[244]

The photocatalytic enhancement observed was attributed to the 6-electron chargeable nature of the pyridinium derivative, enabling the rapid release of CO via the reduction of Co^{II} to Co^{I} by intramolecular transfer of electrons from the reducing equivalent stored to one of the N-methylpyridinium acceptors. Recently, Halime, Aukauloo and coll. reported that a urea-functionalized iron porphyrin produced CO with a TOF of 3720 h^{-1} (selectivity = 99%) with $\mathbf{Ru}^{2+}$ as PS, the urea functions acting as multipoint hydrogen bonding scaffold toward the CO_2 substrate. While previous reports proposed Fe^0 porphyrin species as the key intermediate, activation of CO_2 directly via the Fe^I state was suggested.[245]

Stabilising iron at low oxidation states should enhance the CO_2 reduction process. In this way, Costentin, Robert and Savéant evidenced via studies on the CO_2-to-CO electrochemical conversion that the presence of fluorine or quaternary amine substituents on the porphyrin ring has a favourable effect on the electrocatalytic performances, pushing the catalyst standard potential to positive values (see Figure 3.17).[246]

Two key Density Functional Theory (DFT) studies offer deep mechanistic insights into how transition metal porphyrins, specifically iron and cobalt, catalyze the electrochemical reduction of CO_2. Davethu and de Visser (2019)[13] focused on iron-porphyrin complexes, particularly TPP and a meso-substituted derivative. Their work revealed that the catalytic mechanism involves an initial reduction of the iron centre, enabling CO_2 binding and forming a reactive $[Fe^{III}(CO_2{}^{2-})(TPP^{\bullet-})]^{2-}$ intermediate. This species undergoes successive proton-coupled electron transfers (PCETs) to yield CO as the final product. Interestingly, the study showed that the redox activity of the ligand itself is essential in stabilizing the reactive intermediates, a feature that differentiates iron-porphyrin catalysis from simpler metal-ligand systems. Modifications to the second coordination sphere were found to alter CO_2 positioning slightly but did not significantly impact the electronic structure or thermodynamic profile of the reaction.[245, 246] In contrast, Shen *et al.* (2016) investigated the behaviour of cobalt-porphyrin catalysts.[14] Their DFT analysis detailed a mechanism where CO_2 binds to the Co^I centre, forming a $CO_2{}^-$ adduct that can follow multiple reaction pathways. The dominant route leads to CO formation through a decoupled proton and electron transfer, while minor pathways produce formic acid, through the key [Co(P)–(OCHO)] intermediate, and methane from subsequent CO reduction via further PCET steps. The results correlated well with experimental selectivity trends, where CO was the major product. The study emphasized the role of the cobalt centre in dictating both the binding strength of CO_2 and the energy barriers of subsequent reduction steps.

Together, these two studies illuminate how the nature of the central metal atom and the electronic properties of the porphyrin ligand influence the reactivity, selectivity, and efficiency of CO_2 electroreduction. Iron complexes tend to rely heavily on ligand-centred redox activity, whereas cobalt complexes utilize a more metal-centred approach with broader product distributions. Modifying either the metal or the surrounding ligand environment, including the second coordination sphere, can shift reaction energetics and intermediates, offering strategic levers for catalyst optimization. These findings provide a theoretical foundation for designing next-generation molecular catalysts tailored for sustainable carbon dioxide conversion.

3.3.2 Experimental background

As already stated above, POMs, which are composed of early-transition metals in high oxidation state such as W^{VI}, Mo^{VI} or V^{V}, are polynuclear molecular metal oxides recognized for their fast, reversible, multi-electron redox characteristics and their acido-basic properties.[33, 148, 149] Interestingly, it has been shown that the combination of MPs with ε-Zn_4PMo_{12} reduced Keggin POMs in a 3D material lead to high CO_2RR electrocatalytic performances, with the exertion of an electric field resulting in an oriented electron transport channel through the solid.[176] Two years later, Wang *et al.* studied by means of DFT a hypothetical 2D material associating via Mo≡N imido bonds MPs and $[Mo_6O_{19}]^{2-}$ (**Mo**$_6$) Lindqvist type-POMs, indicating that in this framework the in situ generated $2e^-/2H^+$ reduced **Mo**$_6$ POMs act as multi-electron regulators for the CO_2 reduction reaction performed by the MPs, the cobalt derivative representing a peculiarly promising electrocatalyst for CO_2RR.[247] On the basis of all these results, the group of P. Mialane and A. Dolbecq investigated the covalent association of POMs and MPs at the molecular level with three new **Mo**$_6$/MP dyads and their CO_2RR electro/photocatalytic activity.

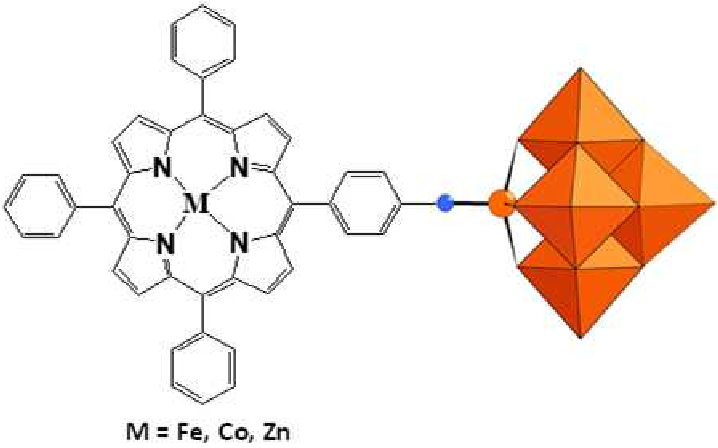


FIGURE 3.18: Representation of the compounds **Mo**$_6$-porph(M) (M = Fe, Co, Zn), blue sphere : N, orange sphere: O, orange octahedra: MoO_6.

These three dyads, **Mo**$_6$-porph(Zn), **Mo**$_6$-porph(Co) and **Mo**$_6$-porph(Fe), consist of MPs covalently bonded to a Lindqvist-type polymolybdate containing Zn, Co and

Fe and were synthesized by reacting the POM, $[Mo_6O_{19}]^{2-}$ (**Mo_6**), with metalated-TPP in presence of dicyclohexylcarbodiimide (DCC) in anhydrous dimethyl sulfoxide (DMSO) at 70°C.[138, 248, 249] The expected individual vibrations of **Mo_6** and MP have been confirmed recording the IR spectrum of the isolated product indicating the preservation of the nature of both species in the dyad. Still, a new and sharp vibration corresponding to Mo≡N stretching, confirmed the formation of the organoimido POM as represented in Figure 3.18. Additional characterizations were carried out, including ESI/MS, TGA, EDX, elemental analysis, and UV-Vis spectroscopy. No shift in the Soret band was observed compared to the MP alone; however, a notable increase in intensity was detected. Furthermore, ^1H NMR showed that the covalent grafting of the MP onto the POM platform induces a significant downfield shift of the MP protons, underlining an electron withdrawing effect of the POM but also the absence of free MP.

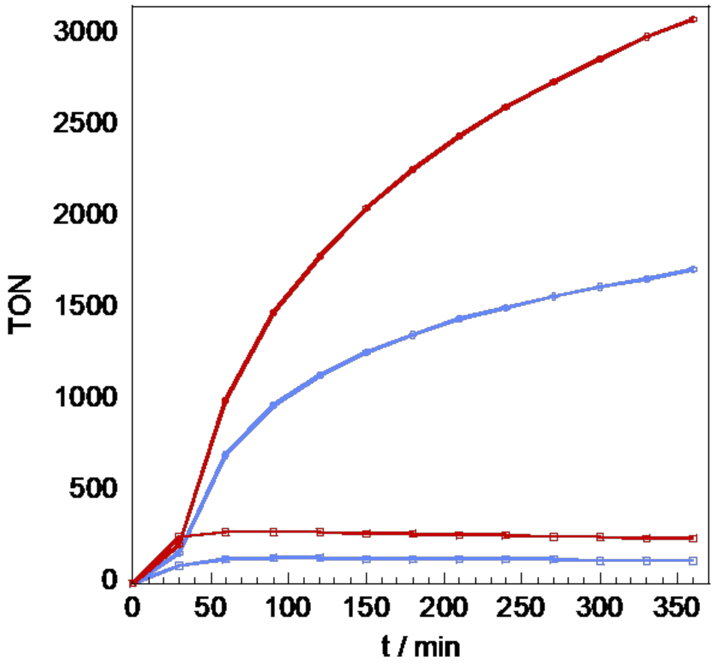


FIGURE 3.19: CO and H_2 TONs for **Mo_6**-porph(Fe) (4 μM) and FeTPP (4 μM) in the presence of $[Ru(dmbpy)_3](PF_6)_2$ (500 μM) and BIH (48 mM) in DMF:TEOA (5:1) solutions obtained under irradiation with a blue LED (10 mW cm^2). Red circles, CO production by **Mo_6**-porph(Fe); red squares, production of H_2 by **Mo_6**-porph(Fe); blue circles, CO production by FeTPP; blue squares, production of H_2 by FeTPP.

The photocatalytic performances of these dyads were explored in CO_2 saturated DMF:TEOA (5:1) solutions, with TEOA and dimethylphenylbenzimidazoline (BIH) as SEDs, and $[Ru(dmbpy)_3]^{2+}$ (dmbpy = 4,4'-dimethyl-2,2'-bipyridine) as PS. The solutions were irradiated with a blue LED and both gas and liquid phases were analyzed. Remarkably, as shown in Figure 3.19, the photocatalytic activity of the

Mo_6-porph(Fe) dyad was found to be significantly higher than that of the POM-free porph(Fe), indicating that the presence of the POM has indeed a positive impact on the catalysis. Among the series of dyads, the Fe-containing one was found to outperform its Zn- and Co-based analogs, and hence, subsequent studies focused on **Mo_6**-porph(Fe).

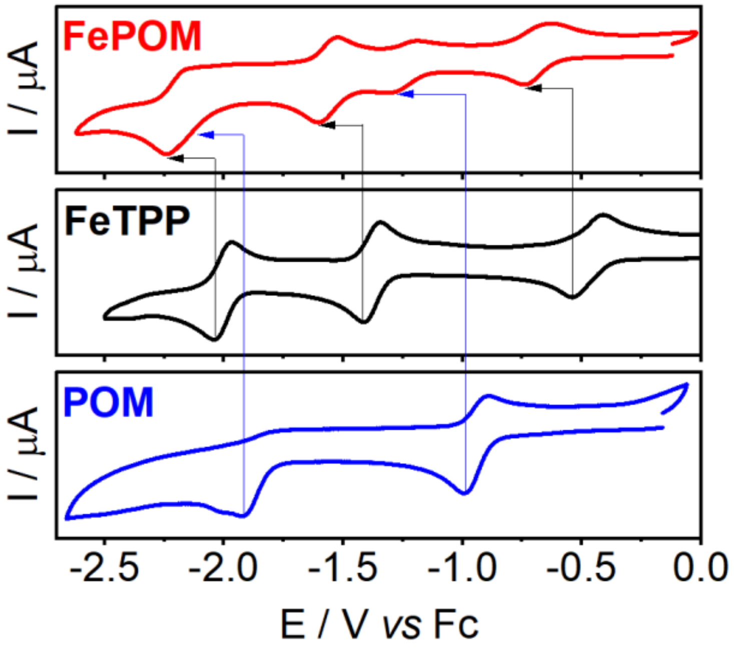


FIGURE 3.20: Cyclovoltammogram (CV) of FePOM in comparison with reference complexes FeTPP and POM [condition: 0.5 mM catalyst, 0.1 M $(TBA)PF_6$ in Ar-saturated DMF, scan rate 100 mV s^{-1}]. CVs of **Mo_6**-porph(Fe) (red line), Fe(TPP) (black line) and **Mo_6** (blue line).

Electrochemical studies were subsequently conducted to gain deeper insight into the catalytic properties of the dyads and the origin of their enhanced activity, with a focus on the **Mo_6**-porph(Fe) derivative. The cyclovoltammogram (CV) of the **Mo_6**-porph(Fe) dyad, represented in Figure 3.20, presents what it seems the combined reduction processes of both MP and **Mo_6** moieties. As previously reported,[250] the wave associated with the first reversible SET process of the Mo centre of **Mo_6**-porph(Fe) is cathodically shifted compared to the same process for **Mo_6** for the electron donation effect of the N-MP ligand compared to the oxo-ligand. The waves associated to the porphyrinic moiety of the dyad also appear to be shifted to negatives potentials when compared to Fe(TPP), ascribed to the fact that the presence of the POM ligand brings a doubly negative charge to the dyad. It is worth mentioning that no wave corresponding to the **Mo_6** and Fe(TPP) precursors were not observed, confirming the purity of the covalently linked product.

Despite extensive research on POMs and MPs as individual redox-active systems, the electronic behaviour of their hybrid assemblies upon reduction remains elusive and scarcely studied. This is particularly evident in their cyclic voltammetry profiles, where the overlapping redox events do not allow unambiguous interpretation. The underlying challenge lies in the limited understanding of how the electronic properties of one moiety may modulate or perturb those of the other. Specifically, the influence of the POM unit on the redox and catalytic behaviour of the MP core, especially in the context of CO_2 reduction, has not been fully elucidated.

To address this gap, this subchapter explores the electronic structure of the reduced intermediates in the POM–MP conjugate using DFT. By characterizing the spatial distribution of spin density, frontier molecular orbitals, and charge localization in each redox state, we aim to decode the mutual electronic effects exerted between the two components. This theoretical insight is essential to clearly identify the nature of the experimentally observed redox waves, as well as to understand how the POM moiety alters the catalytic landscape of the MP, potentially modifying its activity, selectivity, or stability toward CO_2 reduction. These findings may guide future design principles for modular hybrid catalysts with tunable redox and catalytic properties.

3.3.3 Computational methods

Density Functional Theory (DFT) calculations were performed using the B3LYP exchange correlation functional[81] on the Gaussian 16 (rev. C.02) quantum chemistry code.[119] The LANL2TZ basis[206] set and associated pseudopotentials [207] were used for Fe and Mo centres, supplementing them with a shell of f-type polarization functions.[208] The remaining atoms were described by the 6-311++G(d,p) basis set.[86, 209, 210] Solvent effects of DMF were included in the calculations through the IEF-PCM solvent model[120] as implemented in Gaussian 16. The nature of all minima on the potential energy surface was confirmed by frequency calculations. A free-energy variation of +1.9 kcal mol^{-1} was applied to all species to account for the change from the reference state at 1 atm in the gas phase used in Gaussian's thermochemistry equations to the standard state of 1 M at 25°C in solution. For DMF, this correction increases to $+3.4$ kcal mol^{-1} because, as the solvent, its standard state corresponds to 12.91 M rather than 1 M. The orbital occupancy was studied using the Natural Bond Orbital (NBO) analysis version 3.1 already implemented in Gaussian. Microkinetic model simulations were performed using the Achem 2.0 program.[251]

3.3.4 Results and discussion

The consecutive redox events of porph(Fe)-$\mathbf{Mo_6}$ were investigated by sequentially adding electrons to the system. For each redox state, the relative stabilities of different spin states were compared to ensure that the assignment and characterization of each redox wave was based on ground-state structures.

Characterization of the Fully Oxidized $\mathrm{Mo_6}$-porph(Fe)

We first examined the electronic structure of the fully oxidized $\mathbf{Mo_6}$-porph(Fe) complex using density functional theory (DFT). This system is derived from $[\mathrm{Fe^{III}TPP}]\mathrm{Cl}$, where the $\mathrm{Fe^{III}}$ centre is initially coordinated by a chloride ligand in addition to the porphyrin ring. This gives a five-coordinate iron centre with a square pyramidal geometry. Geometry optimization across different spin states shows that the square pyramidal structure is consistently maintained. In all cases, the $\mathrm{Fe^{III}}$ ion is slightly displaced out of the porphyrin plane. This distortion hints at a tendency towards off-plane-metal square based pyramid coordination, but the overall structure remains square pyramidal. Crystal field theory explains how the local geometry affects the energy levels of the Fe d-orbitals.[252] Figure 3.21 illustrates the typical d-orbital splitting for off-plane-metal square pyramid, in-plane-metal square based pyramid, and square planar ligand arrangements.

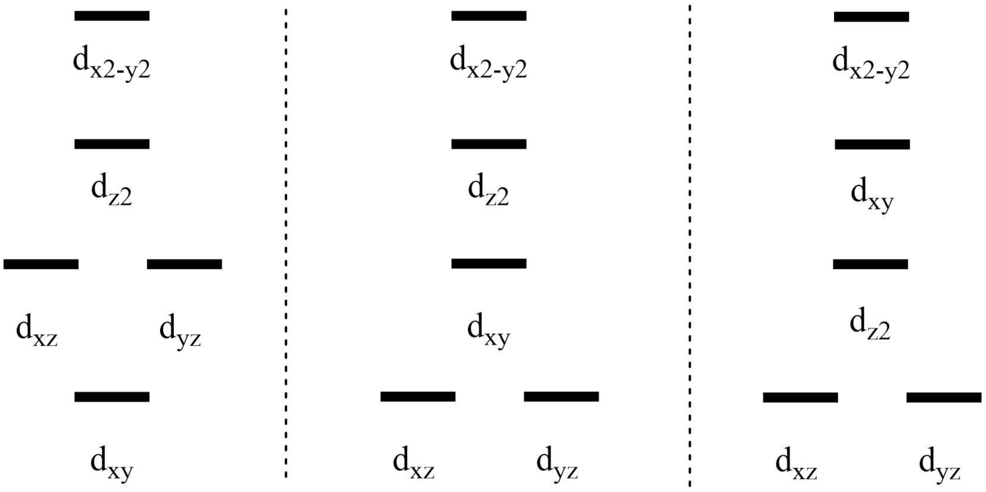
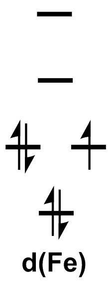
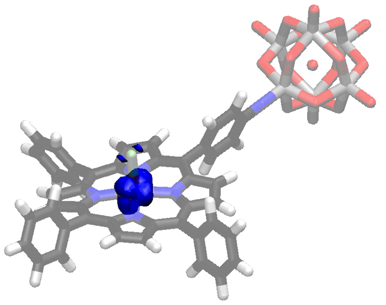
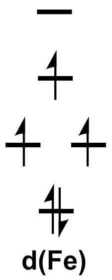
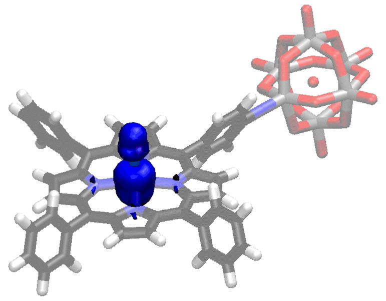
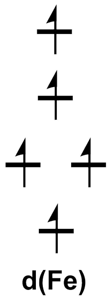
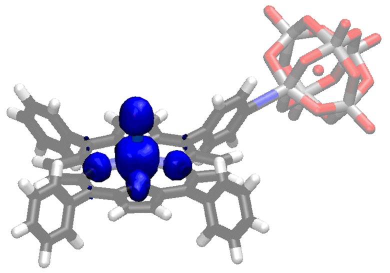


FIGURE 3.21: From left to right, Fe centre d-type orbital diagrams corresponding to the off-plane-metal square pyramid, in-plane-metal square based pyramid and square planar distribution of ligands according crystal field theory.

We carried out DFT calculations for the doublet, quartet, and sextet spin states. The quartet was found to be the most stable, followed closely by the sextet, which lies 1.5 kcal $\mathrm{mol^{-1}}$ higher in energy. In contrast, the doublet state was significantly less stable, with a relative energy increase of 14.6 kcal $\mathrm{mol^{-1}}$. To understand the origin of this instability, we performed a Natural Bond Orbital (NBO) analysis to examine the electronic configuration of the Fe centre in the doublet state. The analysis revealed a distorted d-orbital arrangement, in which the unoccupied, higher energy-lying d_{xy} and d_{z^2} are more poised to accept electron density from the ligands than in intermediate or high-spin configurations, thus shifting up the energy levels of the $\mathrm{Fe^{III}}$ centre.

The NBO analysis for the quartet state (Entry **2** of Table 3.2) shows that the d_{xz}

TABLE 3.2: Comparison of relative Gibbs free energies (kcal mol^{-1}) and electronic structures of the calculated spin states for the fully-oxidized **Mo_6**-porph(Fe) system.

Entry	2S+1	ΔG	Schematic molecular orbitals diagram[a]	Spin density distribution
1	2	+14.6		
2	4	+0.0		
3	6	+1.5		

and d_{yz} orbitals are nearly degenerate. This is consistent with both off and in-plane-metal square pyramid geometries. As in the doublet, the $d_{x^2-y^2}$ orbital is involved in back-donation from the porphyrin ring, albeit to a lesser extent. Compared to the doublet state, a weaker polarization of the electron density from the chloride to the Fe centre is observed in the quartet and sextet states. This is supported by the NBO results that show single occupancy of d_{xz} and by the spin density plot, which shows spin density on Cl^- ligand. The Fe–Cl bond length increases to 2.402 Å in the quartet state (compared to 2.274 Å in the doublet), indicating a weaker, more ionic interaction between Fe and Cl. In the sextet state, all five d-orbitals are nearly degenerate. Each orbital is singly occupied, matching the expected d^5 configuration of Fe^{III}. There is no significant charge transfer between Fe and the ligands. Accordingly, a greater amount of spin density is observed on the N atoms of the porphyrin ligand, leading to slightly longer Fe–N bond lengths in the sextet (2.080 Å vs. 2.010

Å in the quartet), consistent with increased electron repulsion and weaker bonding.

Even if NBO analysis seems to support the electronic interpretation of the studied spin configurations, it is worth bearing in mind its limitations when dealing with metal-based systems with complex electron configurations, or even with highly conjugated systems such as porphyrins, where π-electron are delocalized across the macrocyclic ring. However, NBO analysis is built around the idea of localized bonding, assigning electrons to well-defined σ- and π-bonds, lone pairs, and hybrid orbitals, which poorly reflects the delocalized reality of porphyrins. This can result in artificial fragmentation of the conjugated system and a misleading representation of resonance structures, especially when dealing with aromaticity or electron delocalization across the entire ring. Hence, although these analyses can provide a sound picture that allows for qualitative comparison between spin states, quantitative results such as orbital occupations and energies are to be interpreted with caution.

Moreover, many porphyrins, particularly metalloporphyrins, exhibit multireference character, meaning their electronic ground state cannot be represented by a single electronic configuration. Our DFT analyses relies on single-reference wavefunctions, which limits its ability to capture the electron correlation and configuration mixing necessary for accurately describing such systems.

In summary, the three spin states, explored for the non-reduced **Mo$_6$**-porph(Fe) system, can be classified as low-, intermediate-, and high-spin states, corresponding to the doublet, quartet and sextet, respectively. The quartet state is predicted to be most stable, followed closely by the sextet. The doublet is much less favourable. These results align with previous B3LYP-based studies. For example, Taylor *et al.*[15] also found nearly degenerate quartet and sextet states in similar Fe^{III} porphyrins. Experimental findings likewise support high-spin ground states under certain conditions.[253] It is important to note that Fe^{III} porphyrin spin states are sensitive to environmental factors such as axial ligands, solvent, and substituents.[254, 255] Thus, while the quartet state is favoured here, this may not hold for all systems.

Replacement in the dyad of the Cl^- ligand by DMF

As the as-synthesized product is chlorinated, a possible preliminary substitution of Cl^- by a DMF solvent molecule was explored computationally by means of DFT. This process is evaluated to be endergonic by a standard free-energy cost of 3.9 kcal mol^{-1}. As a solvent molecule, DMF is found in a much higher concentration (12.91M) than the **Mo$_6$**-porph(Fe) dyad (0.5 mM), it is expected to favour such ligand replacement process. In order to evaluate the mentioned concentration effect, a simple microkinetic model simulation was carried out starting from initial concentrations of the Cl coordinated system and DMF and establishing the direct and reverse rate constants of the process, to 5.44 10^7 and 3.93 10^{10} M^{-1} s^{-1}, respectively, as shown at the top of Figure 3.22. The total free energy barrier of the direct process, arise from assuming free-energy diffusion energy barriers of 3.0 kcal mol^{-1} and from the standard-state reaction free energy of 3.9 kcal mol^{-1}. The reverse process only

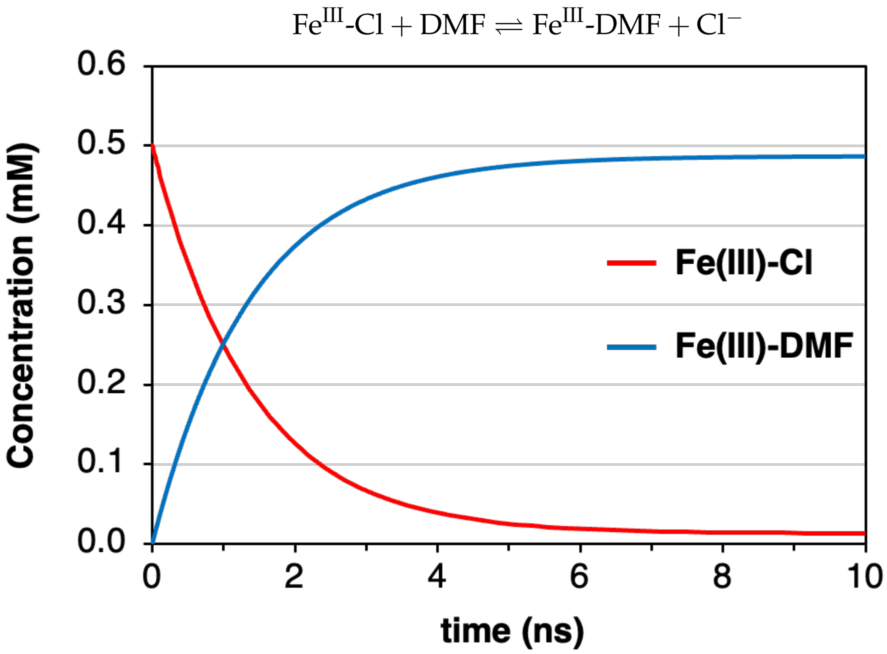


FIGURE 3.22: At the top the process evaluated by the microkinetic model. At the bottom, the evolution of the concentrations of the **Mo**$_6$-porph(Fe) species coordinated by Cl (red) or DMF (blue) ligands of the kinetic model build for the ligand replacement. The direct and reverse rate constants were set to 5.44 10^7 and 3.93 10^{10} M^{-1} s^{-1}, respectively. These arise from assuming diffusion free-energy barriers of 3.0 kcal mol^{-1} and from the standard-state reaction free energy (+3.9 kcal mol^{-1} for the direct process). Initial concentrations: [Cl^-] = [Fe(III)-DMF] = 0 M; [Fe(III)-Cl] = 0.5 mM; [DMF] = 12.91 M. Final concentrations (after reaching an equilibrium): [Cl^-] = [Fe(III)-DMF] = 0.49 mM; [Fe(III)-Cl] = 0.01 mM; [DMF] = 12.91 M.

contemplates the diffusion free-energy assumed of 3.0 kcal mol^{-1}. Both free-energy barriers were transformed into rate constants using the Arrhenius equation (Eq. 3.6). Overall, the simulation reveals that once the equilibrium is achieved after ca. 8 ns, 95% of the **Mo**$_6$-porph(Fe) dyad is coordinated to DMF, while only 5% remains coordinated with Cl^-.

$$k = \frac{K_b T}{h} \exp \frac{-\Delta G^{\ddagger}}{RT} \tag{3.6}$$

Analyzing the electronic nature of the Fe centre in the quartet state (Figure B.5), a similar pattern to that observed in the Cl^--coordinated system emerges. The coordination of DMF leads to notable electronic effects, as reflected in the spin density localized on the O atom analogously with the Cl^--coordinated system, shown in Entry **2** of Table 3.2.

First reduction of Mo$_6$-porph(Fe)

The system under study presents two main regions that are susceptible to reduction: the d-type orbitals of the Fe centre and the d-type orbitals of the Mo atoms in the Lindqvist-type POM cluster (**Mo**$_6$). The LUMO of the DMF-**Mo**$_6$-porph(Fe) species, depicted in Figure 3.23, is predominantly localized on the Fe^{III} ion. This electronic distribution indicates that the first reduction event primarily involves population of a d-type orbital on the iron centre. Upon reduction, the Fe^{III} ion is converted into an Fe^{II} species, which results in an increase in the electron density at the metal centre. This increased charge density weakens the coordination between the Fe centre and its axial ligand, thereby facilitating the release of the ligand into the bulk solvent. Overall, these observations suggest that the Fe site acts as the initial electron-accepting centre in the reduction process, with important consequences for the structural and electronic properties of the complex. The decoordination of the DMF ligand from the reduced Fe^{II} ion (top of Figure 3.24) has been evaluated to be exergonic by 5.2 kcal mol^{-1}. However, the high concentration of DMF in the solvent medium exerts an antagonistic effect, opposing the release of the coordinated molecule into the bulk. To assess whether the intrinsic exergonicity of the decoordination process is sufficient to overcome this thermodynamic penalty, a microkinetic model was constructed. In this model, the forward and reverse rate constants were set to 3.93×10^{10} s^{-1} and 6.06×10^6 M^{-1} s^{-1}, respectively. The forward rate constant corresponds to a diffusion-controlled process with an assumed free-energy barrier of 3.0 kcal mol^{-1}. In contrast, the reverse rate constant incorporates both the diffusion barrier and the standard-state reaction free energy of -5.2 kcal mol^{-1}. The respective free-energy barriers were converted into rate constants using the Arrhenius equation (Eq.3.6). These results suggest that, despite the high solvent concentration, the decoordination process remains favourable under the studied conditions. A simulation was

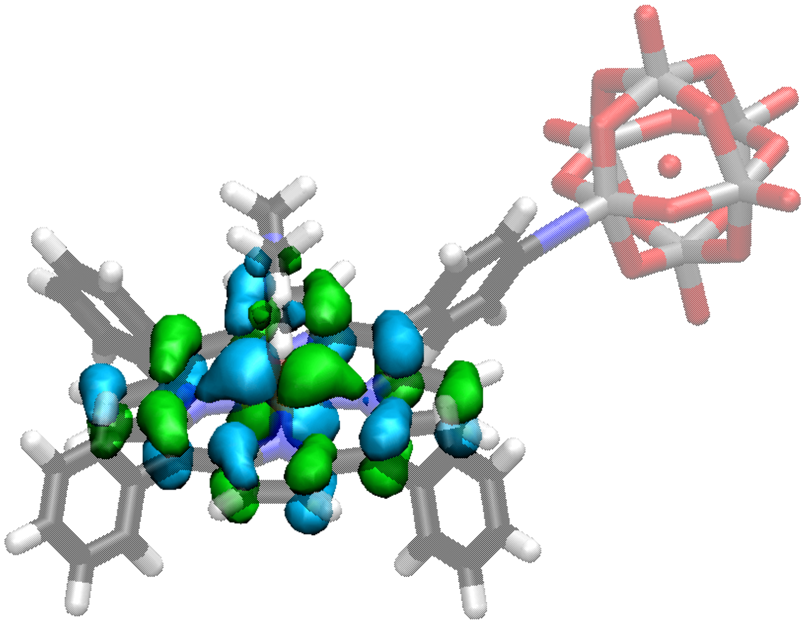


FIGURE 3.23: Representation of the LUMO of the quartet spin state of the DMF coordinated **Mo**$_6$-porph(Fe) species, coloring the positive phase in blue cyan and the negative in green, over the translucent structure of the same system coloring C in black, H in white, N in blue, Mo in gray and O in red.

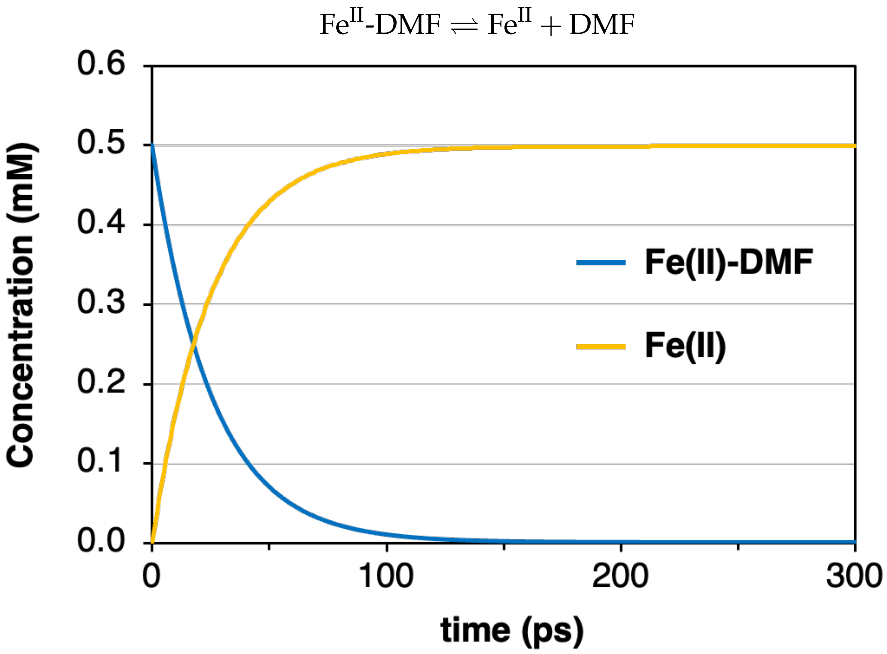


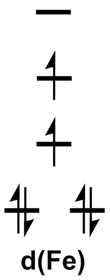
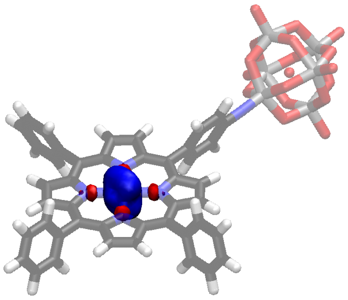
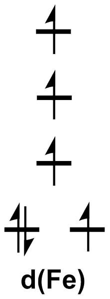
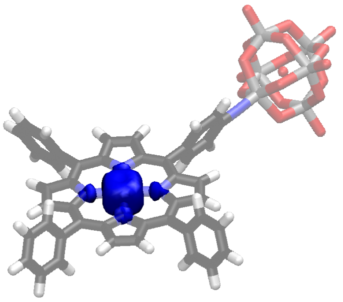
FIGURE 3.24: At the top the process evaluated by the microkinetic model. At the bottom, the evolution of the concentrations of the porph(Fe)-**Mo**$_6$ species coordinated by DMF (blue) and decoordinated (green) of the kinetic model build for the ligand decoordination process. The direct and reverse rate constants were set to 3.93 10^{10} s^{-1} and 6.06 10^6 M^{-1} s^{-1}, respectively. These arise from assuming diffusion free-energy barriers of 3.0 kcal mol^{-1} and from the standard-state reaction free energy (-5.2 kcal mol^{-1} for the direct process). Initial concentrations: [Fe(II)] = 0 M; [Fe(II)-DMF] = 0.5 mM; [DMF] = 12.91 M. Final concentrations (after reaching an equilibrium): [Fe(II)] = 0.5 mM; [Fe(II)-DMF] = 0 M; [DMF] = 12.91 M.

performed using the constructed microkinetic model, with initial concentrations set to 0.5 mM for the DMF-coordinated system and 12.91 M for the free DMF solvent molecules. The evolution of the concentrations of the coordinated and decoordinated species is shown in Figure 3.24. The simulation confirms that the exergonicity of the process is sufficient to overcome the energetic penalty associated with the high concentration of the solvent, reaching equilibrium within 150 ps, which is significantly faster than observed in the ligand-substitution simulation. These results demonstrate that the reduced Fe^{II} system facilitates the ligand decoordination process, which is a key step for enabling the second SET.

We next investigated in detail the electronic structure of the one electron-reduced **Mo**$_6$-porph(Fe) system, where the Fe centre is formally in the +2 oxidation state. Calculations across possible spin states reveal that the intermediate-spin triplet is the ground state, in line with previous studies on Fe^{II} porphyrins.[256] The high-spin quintet lies only 2.2 kcal mol^{-1} above the triplet, while the low-spin singlet is considerably destabilized, with a relative free energy of +29.2 kcal mol^{-1}. The low-spin state becomes a singlet which do not present spin density. In contrast, the triplet exhibit spin configuration exclusively on the Fe centre (Entry **1** in Table 3.3).

The high-spin quintet (Entry **2** in Table 3.3) presents a bulkier spin delocalization on the Fe centre as a result of presenting more singly occupied d-type orbitals.

TABLE 3.3: Comparison of relative Gibbs free energies (kcal mol^{-1}) and electronic structures of the calculated spin states for the one electron-reduced **Mo$_6$**-porph(Fe) system.

Entry	2S+1	ΔG	Schematic molecular orbitals diagram	Spin density distribution
1	3	+0.0		
2	5	+2.2		

In summary, the one electron-reduced state of the **Mo$_6$**-porph(Fe) system localizes the electron on the Fe^{III} centre, forming a Fe^{II} species stabilized in a triplet ground state. The increased electron density at the metal weakens the axial Fe–DMF interaction, promoting ligand dissociation. Despite the competitive effect of high DMF concentration in solution, kinetic modeling shows that this decoordination is both thermodynamically favourable and occurs on sub-nanosecond timescales. The observed spin-state preferences and electronic reorganization around Fe underscore its pivotal role in redox responsiveness and structural dynamics, establishing the basis for subsequent electron transfer processes.

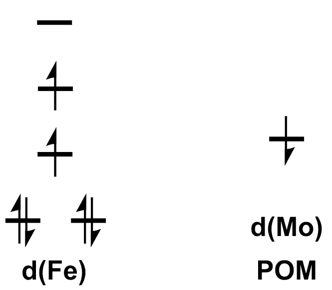
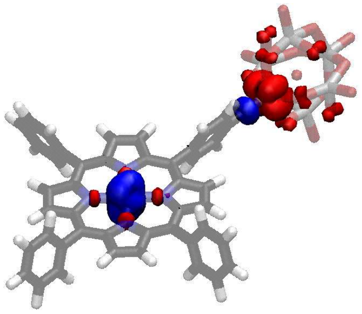
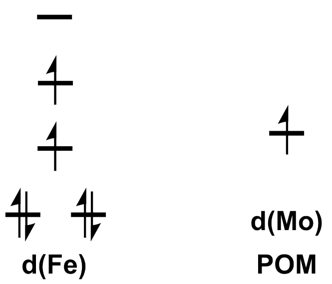
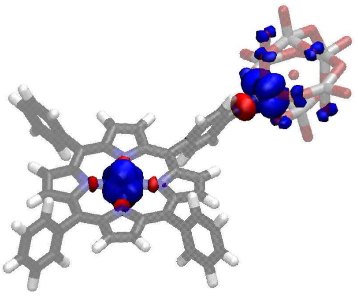
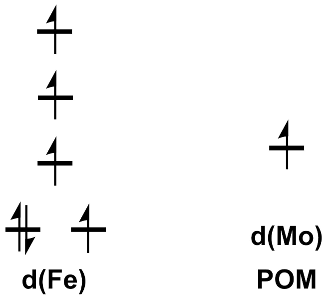
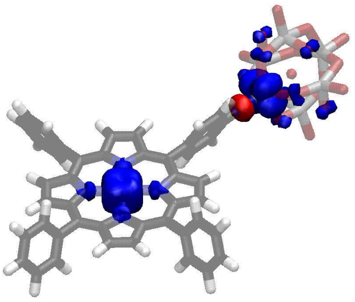
Second reduction of Mo$_6$-porph(Fe)

To explore the electronic structure of the two-electron reduced **Mo$_6$**-porph(Fe) complex, we investigated the nature and stability of its possible spin states, with particular focus on the location of the second added electron and the role of the POM fragment.

Starting from the one electron-reduced species, which exhibits a triplet ground state, analysis of the LUMO (Figure B.6, left) indicates that the second electron is likely accepted by the POM. The orbital is composed mainly of Mo-centred d-type orbitals, especially on the Mo bound to the imido ligand bridging the porphyrin and

POM. This prediction is confirmed by the highest beta SOMO of the two electron-reduced system (Figure B.6, right), which shows clear localization on the imido-coordinated Mo centre. Among the three calculated spin states, the doublet is the most stable (Entry **1** in Table 3.4). Its Fe-centred d-orbital configuration resembles that of the one electron-reduced triplet (Entry **1** in Table 3.3). The high-spin sextet (Entry **3** in Table 3.4) lies only 1.6 kcal mol^{-1} above the doublet. It retains the Fe local quintet configuration observed in the one electron-reduced high-spin state (Entry **2** in Table 3.3), with the additional unpaired electron on the POM leading to a sextet, overall. This modest energy difference reflects stabilization from the Fe high-spin configuration. The quartet (Entry **2** in Table 3.4) is significantly less favourable, lying 10.9 kcal mol^{-1} higher in energy. In this state, although the local spin configuration at the Fe centre resembles a triplet, the overall spin density adopts a different arrangement, which may be the primary cause of the state's destabilization. The

TABLE 3.4: Comparison of relative Gibbs free energies (kcal mol^{-1}) and electronic structures of the calculated spin states for the two electron-reduced **Mo_6**-porph(Fe) system.

Entry	2S+1	ΔG	Schematic molecular orbitals diagram	Spin density distribution
1	2	+0.0		
2	4	+10.9		
3	6	+1.6		

presence of an additional unpaired electron on the POM accounts for the quartet spin multiplicity of the system.

In summary, the two-electron reduced state of the **Mo$_6$**-porph(Fe) complex introduces an additional electron primarily localized on the imido-bound Mo site of the POM. The most stable configuration is a doublet, formed by coupling this added electron with a Fe-based triplet configuration. The close energy of the sextet reflects the stabilizing influence of the Fe high-spin state, while the quartet is destabilized by a higher Fe centre d-orbital rearrangement. These findings reinforce the importance of metal–ligand orbital interactions in dictating the redox and spin-state landscape of this hybrid system.

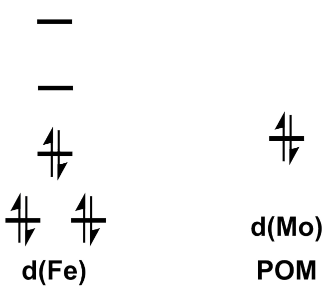
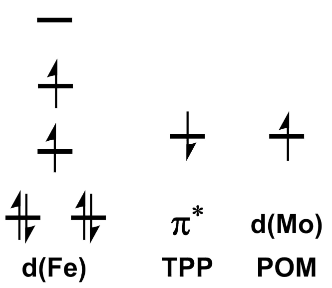
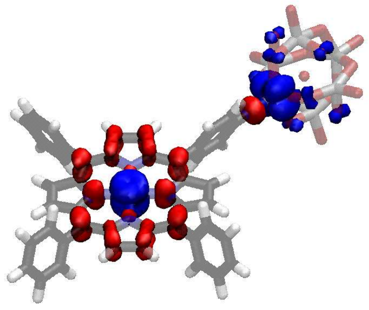
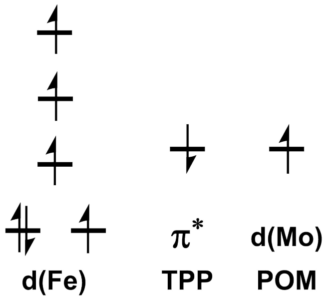
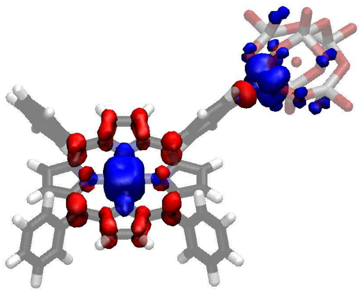
Third reduction of Mo$_6$-porph(Fe)

To investigate the three electron-reduced form of the **Mo$_6$**-porph(Fe) complex, we examined possible spin states and the preferred site for electron localization. The frontier unoccupied molecular orbitals of the two electron-reduced system show nearly degenerate energies between porphyrin- and POM-centred orbitals, suggesting multiple possible electron destinations.

Three spin states, singlet, triplet, and quintet, were optimized and analyzed (Table 3.5). The triplet was found to correspond to the ground state. The Fe centre adopts an intermediate-spin configuration as the spin density maps confirms a spin alpha unpaired electron on the POM (at the imido-bound Mo centre) and another spin-beta in a porphyrin-centred π^* orbital, so the Fe centre have to present a local triplet spin in order to achieve a global triplet spin configuration. The quintet, 2.9 kcal mol^{-1} above the triplet, retains the same porphyrin and POM electron distribution but involves a high-spin Fe^{II} centre. This spin configuration is energetically less favourable keeping parallel trends observed in less-reduced species. Of note, forcing a singlet spin state leads to a closed-shell solution with two electrons on the POM, although this lies 38.1 kcal mol^{-1} higher in energy than the triplet state.

Overall, the third reduction adds an electron to a porphyrin-based π^* orbital, forming a delocalized triplet state involving the porphyrin. This result reinforces the cooperative redox behaviour of the hybrid system, where the electronic structure flexibly adapts to added electrons by redistributing spin and charge across the Fe, porphyrin, and POM components.

TABLE 3.5: Comparison of relative Gibbs free energies (kcal mol^{-1}) and electronic structures of the calculated spin states for the three electron-reduced **Mo**$_6$-porph(Fe) system.

Entry	2S+1	ΔG	Schematic molecular orbitals diagram	Spin density distribution
1	1	+38.1		
2	3	+0.0		
3	5	+2.9		

Fourth reduction of Mo$_6$-porph(Fe)

The fourth reduction of the **Mo**$_6$-porph(Fe) system primarily involves the POM unit, as suggested by the LUMO, which features significant contributions from Mo d-type orbitals (Figure 3.25).

To understand the resulting electronic structure, we examined the doublet, quartet, and sextet spin states. Among these, the doublet is the most stable. The sextet lies 3.8 kcal mol^{-1} higher, while the quartet is the least favourable at +9.5 kcal mol^{-1} (Table 3.6). These results indicate that the added electron is best accommodated within the POM, favouring the doublet spin configuration. Considering that both electrons on the POM moiety couples without contributing to the overall spin configuration of the system, and one beta-spin electron is delocalyzed on the porphyrin in a π^* orbital, confirmed by the spin density distribution. The local spin configuration of the Fe centre have to retain a triplet, similar to the three electron-reduced triplet state.

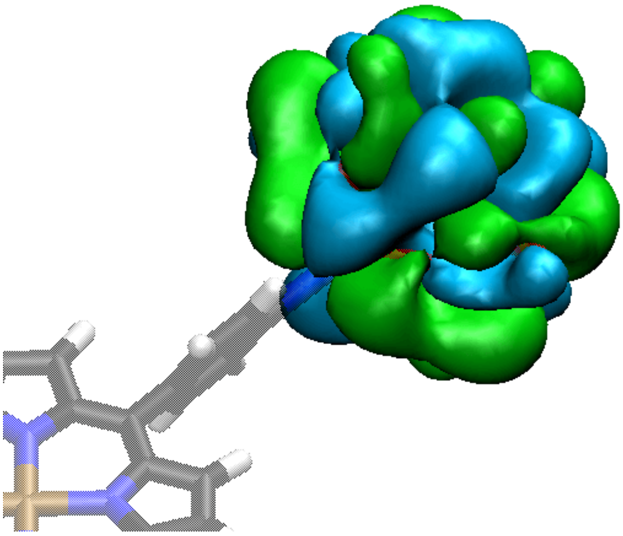


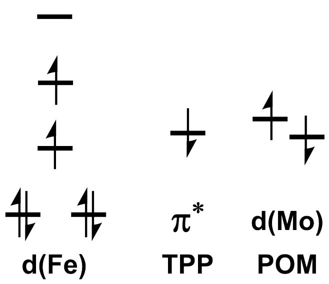
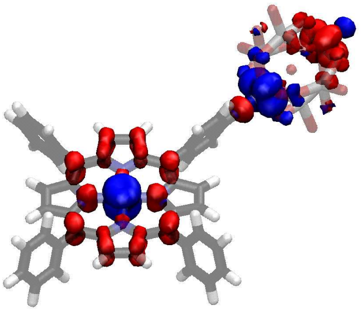
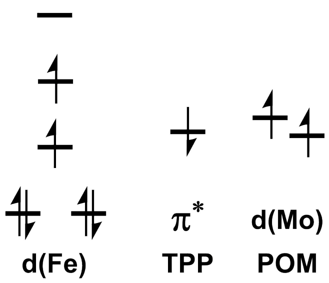
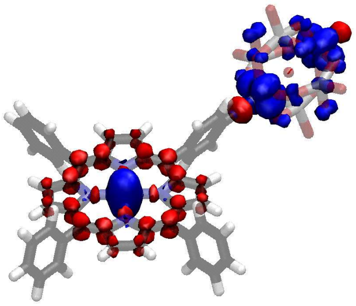
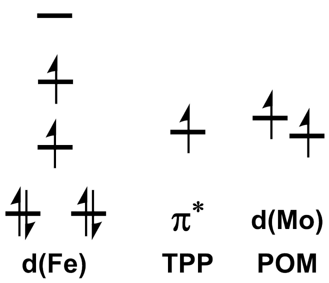
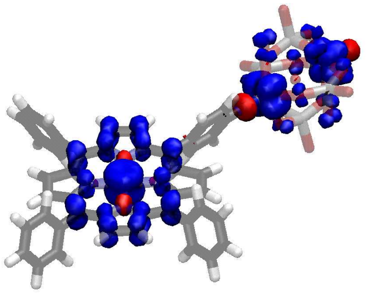
FIGURE 3.25: Representation of the LUMO of the doublet spin state of the three electron-reduced **Mo**$_6$-porph(Fe) species, the positive phase is colored in blue cyan and the negative in green, over the translucent structure of the system coloring C in black, H in white, N in blue, Mo in gray, Fe in ochre and O in red.

In the quartet configuration, the Fe–porphyrin unit adopts a local doublet state, exhibiting a distorted spin density for both the Fe centre and the porphyrin ligand (Entry **2** in Table 3.6). An additional two unpaired electron on the POM generates an overall quartet state. In the sextet state, the Fe–porphyrin unit remains a triplet, but the porphyrin-localized electron is uncoupled. Two additional unpaired electrons on the POM lead to an overall sextet spin configuration (Entry **3** in Table 3.6). Attempts to converge higher spin states, such as an overall octet or an Fe-centred quartet with a coupled two electrons on the POM moiety, were unsuccessful due to wave-function convergence issues.

Importantly, the HOMO in the most stable doublet state is porphyrin-centred, implying a potential role in electron transfer processes underlying the catalysis. This localized electron may participate in substrate activation, particularly relevant for applications like CO_2 reduction.

Overall, the fourth reduction of **Mo**$_6$-porph(Fe) clearly localizes the added electron on the POM, with spin density spread across the Mo centres, the Fe centre, and the porphyrin. The doublet state, combining a local Fe triplet coupled with a porphyrin π^* electron and a local open-shell singlet of the POM, is the ground state. Higher spin configurations reflect alternative coupling schemes but are less favourable.

TABLE 3.6: Comparison of relative Gibbs free energies (kcal mol^{-1}) and electronic structures of the calculated spin states for the fourth-reduced **Mo_6**-porph(Fe) system.

Entry	2S+1	ΔG	Schematic molecular orbitals diagram	Spin density distribution
1	2	+0.0		
2	4	+9.5		
3	6	+3.8		

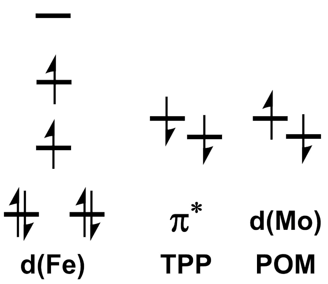
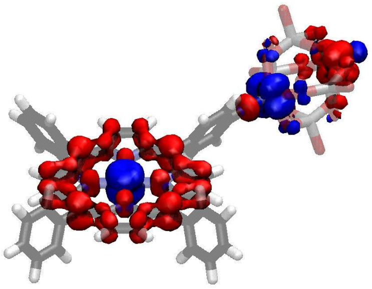
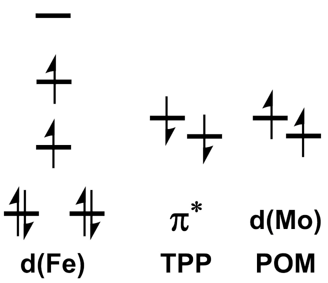
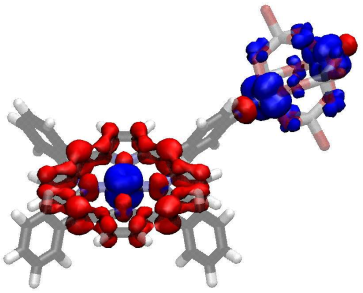
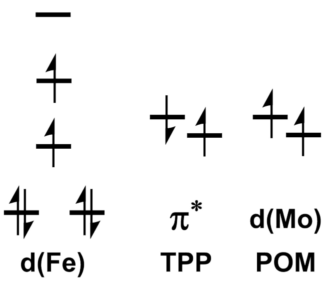
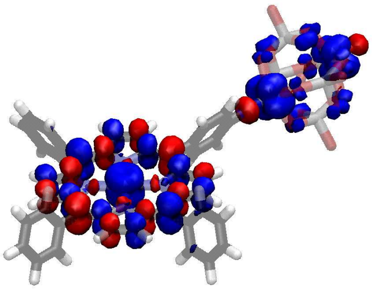
Fifth reduction of Mo_6-porph(Fe)

In the final fifth reduction step, the added electron is primarily localized on the porphyrin ligand, complementing the previous two porphyrin-centred and two POM-centred reductions.

Spin-state analysis identifies the quintet as the most stable configuration, followed closely by the triplet and an open-shell singlet, which lie only +1.3 and +2.1 kcal mol^{-1} higher in energy, respectively (Table 3.7). This narrow energy window indicates a delicate balance in spin coupling across the system. However, the septet state, in which one of the poprhyrin-centred electrons is decoupled from Fe electrons, is strongly disfavoured (+18.6 kcal mol^{-1}).

In the singlet state (Entry **1** in Table 3.7), the Fe centre maintains a local triplet

TABLE 3.7: Comparison of relative Gibbs free energies (kcal mol^{-1}) and electronic structures of the calculated spin states for the five electron-reduced **Mo_6**-porph(Fe) system.

Entry	2S+1	ΔG	Schematic molecular orbitals diagram	Spin density distribution
1	1	+2.1		
2	3	+1.3		
3	5	+0.0		

configuration, coupled with two unpaired electrons on the porphyrin. The two electrons on the POM are spin-paired, contributing no net spin. The triplet configuration (Entry **2** in Table 3.7) features a similar Fe–porphyrin arrangement, but with the two unpaired electrons on the POM now contributing to the total spin. This yields an overall triplet with delocalized spin density across all three components of the system. This configuration resembles those reported for formal Fe^0TPP complexes, where excess electrons are accommodated on the ligand.[256] In the quintet state (Entry **3** in Table 3.7), the porphyrin electrons adopt opposite spins, weakening their coupling with the Fe centre but yielding the lowest free energy overall. This subtle reorganization suggests that Fe's d-electron configuration is the dominant factor in stabilizing the system, while porphyrin and POM couplings serve a secondary role. Attempts to converge higher spin states, including a septet or nonet, were unsuccessful or led to significant destabilization. The septet, when achievable, disrupts the d-type orbital distribution of the Fe centre and leads to an energetically unfavourable arrangement (Table B.1).

For fifth reduction of $\mathbf{Mo_6}$-porph(Fe), the quintet state emerges as the ground state, but the small energy gaps to the singlet and triplet suggest accessible thermal population of multiple spin states. The Fe centre plays the primary role in dictating spin stability, while the porphyrin and POM flexibly accommodate electron and spin rearrangements.

Prediction of reduction potentials

Using the computed free energies of the most stable spin configurations for each reduced state of $\mathbf{Mo_6}$-porph(Fe), the corresponding reduction potentials for the associated redox processes were calculated. These values are reported in the second column of Table 3.8, and compared to the experimental ones, determined from the recorded CV traces shown in Figure 3.20. The reduction potentials were calculated using Eq. 3.7, where n represents the number of electrons transferred in the reduction process, equal to 1 for all cases considered in this study, and F denotes the Faraday constant.

$$E = \frac{\Delta \mathrm{G}}{-\mathrm{n}F} \tag{3.7}$$

To allow direct comparison with the experimental data, the calculated absolute reduction potentials were referenced to the ferrocene (Fc) electrode, which has an absolute potential of 4.98 V. This referencing enables a meaningful comparison between computed and experimentally measured redox potentials. The calculated potential for the first reduction process shows good agreement with the experimental value (Entry **1** of Table 3.8). The second reduction, which corresponds to the first reduction of the POM moiety (Entry **2** of Table 3.8), is predicted to occur at a more favourable potential than observed experimentally, at –0.95 V compared to the experimental value of –1.21 V. The potential of the third reduction, corresponding to the first reduction of the porphyrin moiety (Entry **3** of Table 3.8), is calculated at –1.83 V, being slightly more negative than the experimental value of –1.54 V, indicating an overestimation of reduction energy with respect to experiment. Beyond this point, making accurate quantitative predictions of redox potentials becomes significantly more

TABLE 3.8: Experimental vs. calculated reduction potentials for the five consecutive reduction processes of porph(Fe)-POM, referenced to Fc.

Entry	$\mathbf{E_{exp}(V)}$	$\mathbf{E_{calc}(V)}$
1	-0.67	-0.56
2	-1.21	-0.95
3	-1.54	-1.83
4	-1.98	-1.55
5	-2.18	-2.59

challenging, as the complexity of the electronic structure increases. As shown in Table 3.8 (Entries **4** and **5**), the calculated reduction potentials exhibit non-negligible discrepancies compared to experimental values, likely due to the increased multiconfigurational character of the electronic structures of four and five-electron reduced species, which cannot be accurately reproduced using mono-determinant methods such as DFT. Thus, while DFT can provide a reasonable qualitative picture of the redox sites involved in each reduction event, aiming to quantitatively reproduce experimental redox potentials for such complex electronic structures might require the application of more sophisticated multireference methods like complete active space self-consistent field (CASSCF).[16, 17] Nevertheless, these are highly computationally demanding and most likely unaffordable at this point in time for systems of the dimensions and nuances of $\mathbf{Mo_6}$-porph(Fe).

Moreover, the solubility and macroscopic behaviour of POMs in organic solvents might be affected by the presence of large conutercations like TBA.[202] Ion-pairing effects can significantly influence the redox behaviour of POMs upon reduction. Preliminary attempts to explicitly include ion-pairing effects by placing the countercation within the geometry, guided by Monte Carlo simulations, failed to reproduce the experimental reduction potentials. This suggests that accurately capturing the influence of TBA countercations requires more advanced computational approaches, likely involving explicit treatment of dynamic interactions through molecular dynamics or hybrid quantum/classical (QM/MM) methods.

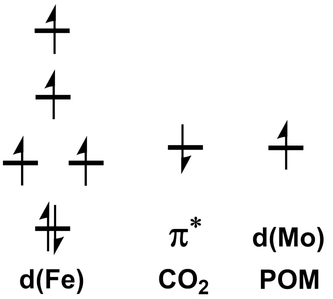
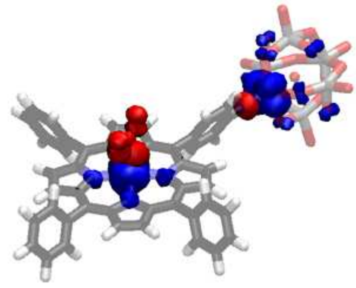
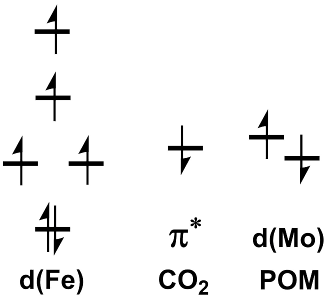
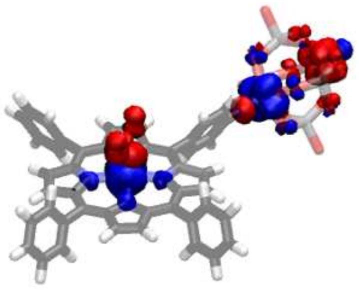
Experimental validation and mechanistic proposal

In parallel to the development of this theoretical work, our experimental colleagues at University of Paris-Saclay (Prof. Leibl and Dr. Gotico), carried out an in-depth experimental study of the redox properties of $\mathbf{Mo_6}$-porph(Fe) combining electrocatalytic approaches and spectroelectrochemical techniques. Importantly, these results fully supported our theoretical results regarding the identification of the redox sites involved in the sequence of redox events. Furthermore, they identified catalytic waves appearing at the third and fourth redox waves. This allowed proposing the complex reaction mechanism displayed in Figure 3.27, whereby CO_2 can be activated and reduced on the three-electron reduced species, but also and much faster on the four- and five-electron reduced ones (I_3, I_4 and I_5 in Figure 3.27).

To better understand these results and the origin of the enhanced catalytic activity, we investigated CO_2 activation on these species, processes I_3 and I_4 in Figure 3.27, and compared it with the corresponding process in the parent POM-free Fe-TPP system.

First of all a spin-configuration exploration was done for the products of I_3 and I_4 processes, as previously done. For the product of I_3, the most stable spin-configuration is the quintet (Entry 1 in Table 3.9). Compared with the non-coordinated system (Entry 2 of Table 3.5), the SET process from the porphyrin ligand to the CO_2 ligand was confirmed with the delocalization of the electron of the $CO_2^{\bullet-}$ radical

TABLE 3.9: Comparison of electronic structures of the calculated spin states of the CO_2 insertion products for the three and four electron-reduced **Mo**$_6$-porph(Fe) dyad.

Entry	Processes in Figure 3.27	Schematic molecular orbitals diagram	Spin density distribution
1	I_3		
2	I_4		

anion and its coupling with the unpaired electron of the Fe centre. Furthermore, the extra coordination of the CO_2 ligand changes the Fe coordination sphere from the square planar distribution to the off-plane-metal square pyramid (see Figure 3.21). For the product of I_4, the most stable spin-configuration is the quartet (Entry 2 of Table 3.9). The porphyrin–CO_2 SET is further confirmed by comparison with the non-coordinated system (Entry 1 in Table 3.6), showing that the electrons remain localized on the POM moiety and that the presence of the additional CO_2 ligand induces an orbital rearrangement at the Fe centre.

The evaluated ΔG for both processes is 19.6 kcal mol^{-1}, very similar to the compared Fe-TPP reference systems (18.6 kcal mol^{-1}) free energy. According to our calculations the CO_2 insertion process does not seem to be responsible for the experimentally observed differences in the CO generation rate for both systems. According to the reported mechanism by Visser *et. al.*, [13] the explored first SET is followed by a proton transfer process (P_3, P_4 and P_5 of Figure 3.27), which triggers a second SET, as observed in the first part of this chapter (intermediate **2′** in Figure 3.15). This second SET may determine the experimentally observed CO production rate, and is currently under investigation.

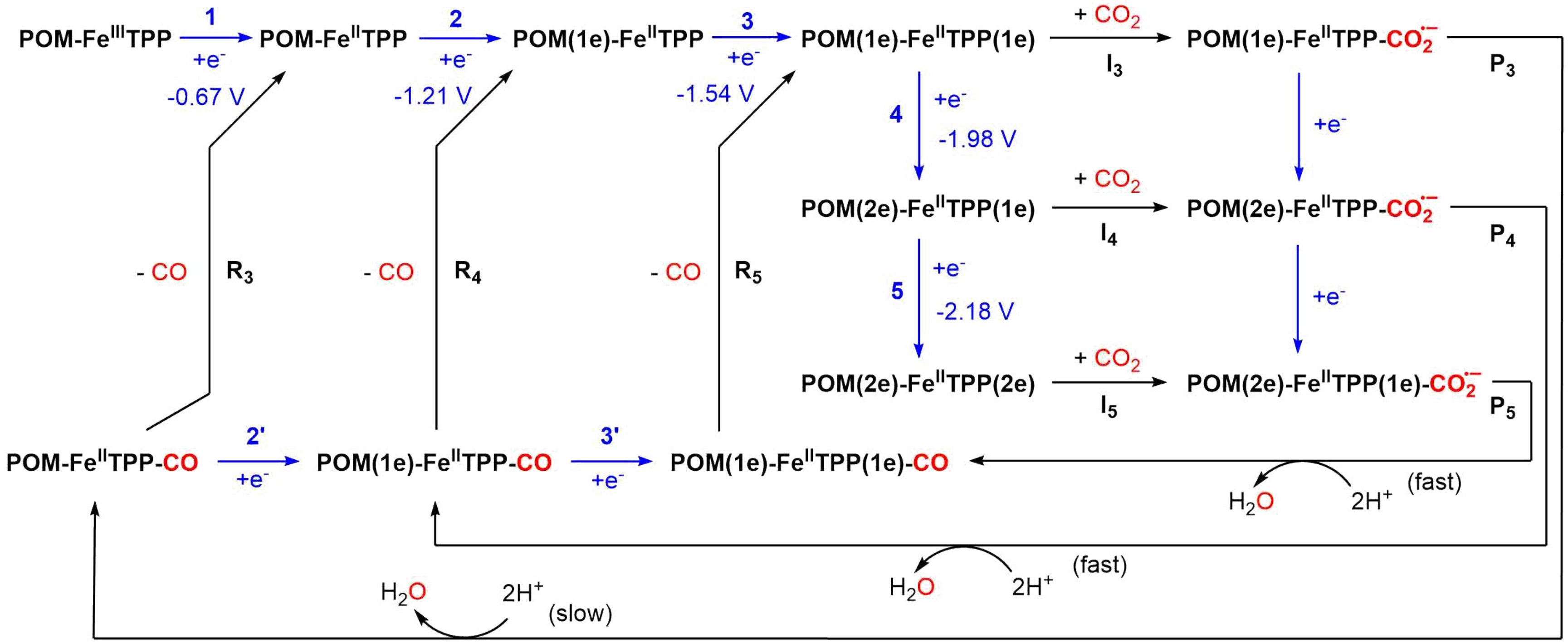


FIGURE 3.27: Proposed reaction mechanism of **Mo$_6$**-porph(Fe) (POMFeTPP) including all five reducing processes explored and possible other competing pathways. Reduction processes in blue. CO_2 atoms highlighted in red. Experimental observations in brackets.

3.3.5 Conclusion

This subsection provides an in-depth analysis of the electronic structure and redox properties of the hybrid catalytic dyad **Mo_6**-porph(Fe), consisting of a porphyrin-functionalized Lindqvist-type polyoxometalate (**Mo_6**). Through a combination of DFT calculations, NBO analyses, spin-density surface mapping, and microkinetic modelling, we reveal the interplay between the porphyrin and POM moieties across successive reduction events. Our results demonstrate that the system exhibits a sequence of redox processes localized initially on the Fe centre of the porphyrin, followed by electron uptake by the POM unit and the porphyrin ligand itself. The initial reduction converts Fe^{III} to Fe^{II}, facilitating ligand decoordination, while subsequent reductions localize electron density on the Mo centres and porphyrin π^* orbitals. Importantly, we show that the spin-state energetics are finely balanced, with intermediate-spin states consistently favoured. While we acknowledge the limitations of the employed DFT method in accurately describing complex electronic structures with multiple unpaired electrons, the region of the molecule that hosts the extra electron during each reduction event remains unchanged across all calculated spin states. This consistency supports the reliability of the qualitative characterization of the redox processes observed experimentally, which was further supported by spectroelectrochemical studies. Microkinetic simulations further confirm the thermodynamic and kinetic feasibility of key steps such as Cl-to-DMF ligand substitution and DMF decoordination, events that are essential for enabling subsequent electron transfer. The nature of the spin states and electron distribution depends strongly on the ligand environment and the order of reduction, highlighting the complex coupling between the porphyrin, POM, and axial ligands. Notably, the imido bridge links both moieties, making them more difficult to reduce and thereby increasing the overall reducing power compared to the free moieties.

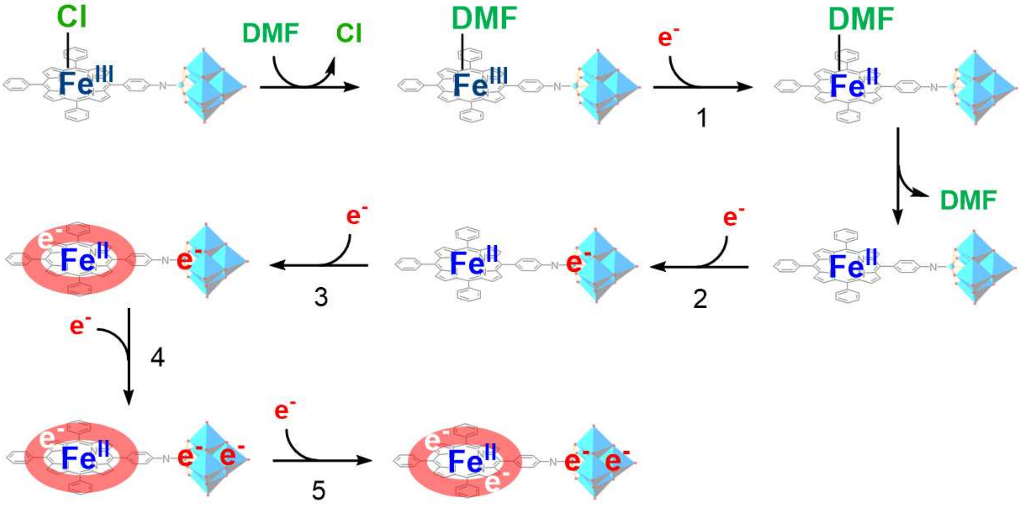


FIGURE 3.26: Schematic representation of the reducing steps of **Mo_6**-porph(Fe).

The observed sequence of Fe-, Mo-, and ligand-centred reductions, along with spin-coupling trends, offer mechanistic insight relevant to designing porphyrin–POM hybrids for redox catalysis, including CO_2 reduction. Overall, this work elucidates the rich redox landscape of these molecular assemblies and offers a theoretical foundation for the rational design of modular catalysts with tunable electronic properties.

3.4 Conclusion

The exploration of POMs as electron mediators in catalytic CO_2 reduction unveils a compelling framework for addressing the energetic and mechanistic challenges associated with transforming this chemically inert molecule. This work supports the idea that POMs are not mere redox-active agents; they also participate on cooperative interactions related to efficient electron management, charge stabilization, and molecular activation. These capabilities are indispensable for reactions involving stable molecules like CO_2, where overcoming kinetic barriers necessitates precise control over multi-electron and proton-coupled events. Unlike systems that rely solely on active metal centres for catalytic turnover, the studied Ni-substituted POM-based architecture exhibit catalytic properties, diversifying the conditions required to reduce CO_2. The Ni^{II} centres are preserved in fixed oxidation states, acting predominantly as anchoring sites for CO_2 coordination. The actual redox transformations occur within the POM structure itself, which temporarily accumulates the electrons necessary for subsequent reduction steps. This configuration underscores a novel strategy in catalyst design, leveraging redox-inert coordination motifs alongside highly delocalized electron storage domains. A key insight from the chapter is the ability of POMs to facilitate stepwise electron delivery. This property becomes particularly valuable in driving endergonic transformations, such as CO_2 reduction, where large energy inputs are required. POMs, due to their capacity for reversible redox cycling, can mitigate such demands by modulating the rate and energetics of individual electron transfers. This behaviour is not only a function of their composition but also influenced by the surrounding solvent, counterions, and potential ion-pairing interactions. Although not directly modeled in this study, such electrostatic effects are likely to enhance SET processes, particularly under conditions where proton-coupled electron transfer pathways are operative. Furthermore, the redox behaviour of POMs is intrinsically linked to their structural parameters, size, geometry, charge, and substituents, which influence their electronic landscapes. Structural variations such as lacunary formation or heteroatom substitution introduce new dimensions for tuning reactivity. This modularity allows the same POM framework to be optimized for different catalytic roles, from hydrogen evolution and nitrogen fixation to selective CO_2 conversion into fuels and chemicals. Importantly, this adaptability enables the reconfiguration of POMs into hybrid materials, photochemical assemblies, or surface-immobilized devices, bridging the gap between homogeneous and heterogeneous catalysis. This chapter also prompts a broader shift in how we conceptualize chemical processes involving CO_2. Rather than treating it as a terminal oxidation product, it is increasingly being incorporated into closed-loop systems where its reduction forms part of a synthetic cycle. In such contexts, POMs emerge not just as tools for overcoming reaction barriers but as enablers of circular carbon utilization. Their integration into electrocatalytic or light-driven platforms can thus serve as a blueprint for future energy storage and conversion technologies, where sustainability and reversibility are paramount.

In conclusion, the insights gained from this study position polyoxometalates as integral components in next-generation catalytic systems. Their multifunctionality, tunability, and redox flexibility provide a versatile platform for mediating complex transformations under mild conditions. More than just passive scaffolds or redox relays, POMs are shown to coordinate electron flow, stabilize intermediates, and interface with substrates in ways that expand the boundaries of molecular catalysis. This work not only advances the understanding of the field of CO_2 reduction but also contributes to a larger vision of harnessing molecular systems for sustainable chemical innovation.

Chapter 4

Conclusions and Perspectives

This PhD thesis has presented a computational investigation of molecular catalysts for the activation of chemically inert small molecules, such as H_2 and CO_2, with a particular emphasis on H_2 splitting and CO_2 reduction reactions. The catalytic systems considered in this thesis span molecular architectures ranging from discrete boron-nitride molecular Frustrated Lewis Pairs (FLPs) (Chapter 2) to more complex polyoxometalates (POMs) hybrid systems (Chapter 3). Our computational investigation used Density Functional Theory (DFT) calculations, data driven approaches, and advanced molecular dynamics simulations in order to i) establish structure-activity relationships in B/N molecular FLPs for H_2 splitting and further CO_2 hydrogenation, and ii) provide mechanistic insights of POM-based catalytic systems. Our primary objectives were both to advance our theoretical understanding of such systems and to possibly identify designing rules for efficient catalysts. We highlight below the key findings related to these two distinct families of catalysts and then propose possible perspectives for future work in these areas.

4.1 Frustrated Lewis Pairs for H_2 splitting

In chapter 2, we successfully developed multivariate models capable of predicting both the thermodynamics (ΔG) and kinetics ($\Delta G^{\ddagger}$) of H_2 splitting by intramolecular B/N-based Frustrated Lewis Pairs (FLPs) by combining DFT calculations and data-driven approaches.

A comprehensive automated data extraction pipeline was developed using in-house developed Python scripts to systematically process Gaussian output files and compute 38 molecular descriptors encompassing geometric, electronic, and electrostatic properties. The geometric descriptors include the $\mathrm{LA}\cdots\mathrm{LB}$ distance (d), the orbital directionality parameter (λ), a novel term we propose, and the dihedral angle between p-type orbital directions (γ). Electronic descriptors incorporate orbital energies of LA and LB centres, atomic charges from both Natural Population Analysis (NPA) and CHelpG methods, and the thermodynamically-derived free energies of hydride and proton attachment (FEHA and FEPA, respectively).

The statistical modelling framework employs both multivariate linear regression (MLR) and Partial Least Squares (PLS) regression with rigorous validation protocols. K-fold cross-validation was implemented to balance computational efficiency with statistical reliability, yielding predictive quality metrics (q^2) alongside fitting parameters (r^2) and root mean squared error (RMSE). A sophisticated data-driven variable selection protocol, inspired by Andries *et al.*, [7] was implemented to discriminate truly informative descriptors from noise by augmenting the descriptor set with intentionally non-informative random variables and iteratively filtering based on coefficient magnitudes across 10 independent PLS models.

We constructed and exploited a dataset of 112 intramolecular B/N FLPs, chosen to cover varied electronic and structural features through various skeletons and chemical substituents of both the Lewis acid and Lewis base centres. DFT calculations were carried out to obtain a set of chemically meaningful molecular descriptors, as well as associated Gibbs free energies (ΔG) and free-energy barriers ($\Delta G^{\ddagger}$) for H_2 splitting. By correlating these descriptors with reactivity metrics, we identified key features governing both thermodynamics and kinetics.

We first show that H_2 splitting reaction free energy is mainly governed by 3 descriptors, constituted of two electronic descriptors and one structural descriptor: the LA's hydride affinity (FEHA), followed by the LB's proton affinity (FEPA) and the $LA \cdots LB$ separation distance. Overall, for thermodynamic predictions (ΔG), our optimized three-descriptor MLR model achieved excellent performance (r^2 = 0.87, q^2 = 0.85, RMSE = 5.52 kcal mol^{-1}). The respective weights of FEHA and FEPA in the model are consistent with the idea that weaker hydride and proton affinities make the H_2 splitting less favourable, due to lower acidity and basicity of the FLP sites. More specifically, the dominant role of FEHA also confirms that the nature of the LA substituents, directly affecting its acidity, plays the dominant role in stabilizing the H_2 splitting product; that is, higher acidity correlates with more exergonic reactions. Shorter $LA \cdots LB$ distances and stronger LB basicity also favour H_2 splitting from a thermodynamic perspective, but to a lesser extent. The influence of d on ΔG may reflect the contribution of attractive electrostatic interactions between LA-H^- and $^+$H-LB groups in the zwitterionic product.

Importantly, our model provides to our knowledge the first quantitative model that explicitly assesses the individual contribution of key descriptors to the reaction free energy for H_2 splitting. Also, the MLR model we presented not only highlights the relative contributions of key FLP parameters, such as acidity, basicity, and the $LA \cdots LB$ distance, to their H_2 activation ability, but also functions as a straightforward and practical quantitative structure–activity relationship (QSAR) model, with a reasonable predictive capability with a RMSE below 5 kcal mol^{-1}.

We then developed models capable of reflecting and predicting the H_2 splitting reaction barrier ($\Delta G^{\ddagger}$). This required the incorporation of quadratic terms to capture optimal geometric configurations, resulting in a nine-descriptor model (r^2 = 0.85, q^2 = 0.81, RMSE = 5.13 kcal mol^{-1}). Interestingly, the model reveals that activation barriers are controlled primarily by geometric factors, such the LA$\cdots$LB distance (d and d^2 terms) and the angle that aligns the LB's lone pair with the LA's empty p orbital (λ and λ^2 terms), both displaying quadratic effects and optimal values near 3.0 Å and 70°, respectively. Notably, the electronic descriptors of the acid/base are found to play secondary roles on the H_2 activation barrier height, in sharp contrast with the reaction energy, which is mainly governed by electronic parameters.

External validation confirmed that the ΔG model generalizes well across diverse FLPs, including B/P and Al/N systems, ie not seen by our trained model, while the $\Delta G^{\ddagger}$ model remains reliable mainly for intramolecular B/N systems. Using the same descriptors than for H_2 splitting, we built a preliminary model for CO_2 hydrogenation towards formate production. Our preliminary model achieved good predictive quality, highlighting FEHA as the primary descriptor.

Overall, our work not only revealed the main parameters governing each of the reactivity metrics, but also give clear guidelines on how to tune FLP structures to target tailored reactivity toward H_2. Moreover, this study provides "ready-to-use" quantitative structure–reactivity guidelines and predictive tools to design optimal FLPs for H_2 splitting, which are expected to serve as efficient pre-screening tools to accelerate the rational discovery of new structures. Finally, preliminary results suggest that multivariate models built from the pool of descriptors considered in Chapter 2 hold the potential for capturing the intricacies of subsequent CO_2 hydrogenation promoted by the zwitterionic intermediate resulting from FLP-assisted H_2 splitting. However, current efforts are under way to expand the FLP dataset and broaden the pool of potential descriptors, hopefully allowing the development of more accurate and transferable models.

4.2 POM-based catalysts

Chapter 3 examines polyoxometalate (POM)-based catalysts for CO_2 reduction. We first present a mechanistic and computational investigation of a tetra-nickel containing sandwich Keggin-type POM, named $[Ni_4(H_2O)_2(PW_9O_{34})_2]^{10-}$ (**Ni_4**), tested for CO_2 reduction into CO in photocatalytic conditions using a Ru-based PS and TEOA as proton and electron donor and under visible light illumination. It revealed the intricate interplay between electron storage, proton management, and substrate activation that underlies efficient CO_2 conversion. Our computational results highlight the pivotal role of POMs as redox-active components. The proposed catalytic cycle begins with photoinduced electron and proton transfer from the **Ru^{2+}**/TEOA system, activating the **Ni_4** catalyst by enabling two-electron, four-proton accumulation. CO_2 then coordinates to a Ni^{II} centre via ligand exchange, followed by electron and proton transfer steps that facilitate C–O bond cleavage and CO formation. The cycle concludes with catalyst regeneration by **Ru^+** and $HTEOA^+$, completing a photo-driven, redox mediated process. Our computational findings indicate that the redox activity of the POM ligand plays a crucial role in enabling efficient charge photoaccumulation. However, the ligand exchange step, where CO_2 binds to the catalytically active Ni centre, may hinder the efficiency of the subsequent thermally activated chemical processes. These results strongly suggest that future research should aim to design catalytic systems that retain (or enhance) the redox performance of $\mathbf{Ni_4}$ while minimizing the energetic cost of replacing coordinated water with CO_2. This could potentially be achieved by introducing more hydrophobic reaction cavities or engineering secondary coordination spheres that better stabilize CO_2 once bound. Notably, a methodological breakthrough in this section lies in the recognition and resolution of limitations inherent in conventional static DFT approaches through advanced path-metadynamics simulations using CP2K with BLYP-D3 level theory. The systematic overestimation of solvation free energies for ligand dissociation processes (26.5 kcal mol^{-1} from static DFT vs. 10.0 kcal mol^{-1} from DFT-MD) was revealed and corrected through enhanced sampling techniques employing collective variables combining $\mathrm{Ni}\cdots\mathrm{O}$ distances and coordination numbers. This finding has broader implications for computational catalysis, suggesting that many previous theoretical studies may have underestimated the feasibility of ligand exchange processes in aqueous environments.

In parallel, a detailed electronic structure analysis of a hybrid porphyrin–POM covalent system ($\mathbf{Mo_6}$-porph(Fe)) elucidates the sequence and localization of reduction events through DFT calculations, NBO analysis, and microkinetic modelling to evaluate Cl^-/DMF ligand substitution dynamics. After the preliminary reduction of the Fe centre, electrons are initially stored at the Mo centres of the POM moiety before populating the porphyrin ligand. These findings reveal balanced spin-state energetics and underscore the importance of ligand environment and inter-fragment coupling in dictating redox behaviour. Key mechanistic steps, such as ligand substitution and decoordination, are shown to be both thermodynamically and kinetically feasible, providing a mechanistic rationale for the observed electrochemical properties and catalytic activity.

Overall, this chapter highlights the role of POMs in CO_2 reduction catalysis, positioning them as multifunctional components capable of mediating electron flow, stabilizing intermediates, and facilitating substrate activation. Both studies highlight the pivotal role of POMs as redox-active components, acting as electron reservoirs during the catalysis. Their ability to reversibly store electrons, coupled with structural tunability, allows them to overcome kinetic and thermodynamic barriers associated with CO_2 activation, either enabling or promoting this class of challenging reactivity. Hence, this work provides both a methodological and conceptual framework for the rational design of next-generation redox-active catalysts in photochemical and electrochemical platforms for CO_2 conversion.

Altogether, the findings encompassed in this thesis emphasize the value of combining data-driven models, electronic structure theory, and mechanistic analysis alongside experimental data to unravel complex catalytic processes. Whether through statistical modelling of FLPs or quantum chemical characterization of POM-based systems, this work provides generalizable strategies to rationally design and tune catalytic structures for small-molecule activation. Nevertheless, the inherent complexity of these systems demands further investigation to fully master the associated technologies. Computational tools are essential in this endeavour, as addressing such complexity requires the combined power of computational performance and human expertise.

4.3 Perspectives

4.3.1 Dealing with complex pathways and electronic structures

Across all the projects studied in this PhD thesis, mainly in the first part of chapter 3, several possible pathways have been considered and computationally evaluated in order to characterize the less energetic pathway. Such explorative study is computationally expensive, especially in large systems like POMs. Ideally, additional efforts should be devoted to explore all possible reaction pathways and side processes (also explaining H_2 formation) which would require a massive amount of computational resources. Moreover, POM-based systems involved in photo- and electrocatalytic reactions can adopt extremely intricate electronic structures with many unpaired electrons that can exhibit antiferromagnetic coupling. An accurate description of these systems most likely require sophisticated computational methodologies beyond DFT, such post-hartree fock methods, like CASSCF or its CASPT2 or NEVPT2 adaptations.[16] However, these are more computationally demanding than DFT, which makes their application to the large systems studied in Chapter 3 unaffordable. More efforts need to be done to enhance system description while balancing computational cost.

Regarding FLP chemistry, competitive pathways, like the poisoning pathways mentioned in section 2.1.1 or the formation of other products different than the targeted HCOOH may also exist and come into play depending on the FLP structure. The preliminary TS's $H\cdots H$ distance classification by families (section 2.3) suggest different TS-types named "early" and "late", presenting closest $H\cdots H$ distance to the H_2 molecule, the substrate, or to the products, respectively. A deeper study is required to further understand the features governing the exhibited TS, possibly predict the ocurrence of early or late TSs and their impact on the H_2 splitting process and subsequent CO_2 hydrogenation. Accounting for this complexity might require not only additional computations to characterize these pathways, but also more flexible models taking into account squared correlations, included in section 2.3.3, or even more sophisticated mathematical relationships. Adding complexity will be challenging in terms of methodology. This can be addressed using methods like random forest modelling, an ensemble learning technique that builds multiple decision trees during training and outputs the average prediction. By aggregating the results of many diverse trees, random forest increases model flexibility and can capture more complex, non-linear correlations in the data, beyond simple squared or linear relationships.

4.3.2 Encoding systems type and size variability

An extra factor which is key to ensure models' quality and reliability is the database size used for training the models. The chapter 2 data-driven study was based on 112 B/N FLP-containing molecules, a chemically restricted set of intramolecular FLPs. The generation of the data was time consuming, not by the computational cost of the DFT calculations required to compute each system (free energy profile, electronic descriptors, etc.), but the number of calculations submitted under human supervision, to ensure proper convergence of the desired geometries of FLP-systems and the corresponding intermediates. More variability requires expanding the chemical space under study providing more perspective in order to better understand the relevance of the discovered features and identifying new that emerge from the fact to include more FLP-types. The existing chemical range of the experimentally reported FLP-systems active for H_2 splitting and CO_2 activation (i.e. B/N, B/P, Al/N, AL/P, intermolecular, intramolecular, etc.) demands larger libraries of FLP molecules to include the mentioned variability in the data in order to generate more robust and reliable data-driven models. Still, the generation of this large human-supervised libraries becomes unaffordable upon their size. New active-learning techniques gained popularity to address this issue, making the human-supervision less required.[257] In combination with more compatible databases, the generation of larger datasets becomes accessible and will be addressed in the near future as continuation of the this PhD project.

4.3.3 Supramolecular effects

One of the main limitations in the computational study of complex systems like POMs is their relevant interactions with solvent molecules and counterions. For the **Ni$_4$** system, we observed that explicit description of the environment is key to model the step of CO_2 binding to catalytically active centres to form a TM-COO$^{\bullet-}$ intermediate, as static calculations tend to overestimate the cost of this process. For the **Mo$_6$**-porph(Fe) system, the CO_2 coordination to the Fe centre is evaluated to be endergonic for $\sim$20 kcal mol^{-1}, a similar situation than for the **Ni$_4$** system. Future efforts should be those devoted to revisit the activation of CO_2 on the Fe-based system. However, due to their large size of this systems, the explicit inclusion of solvent molecules at traditional DFT level of theory, becomes computationally expensive. For this reason, in terms of methodology, extra efforts need to be done to model these kinds of systems more accurately while keeping the calculations affordable to the actual available computational resources. Recently, the application of machine learning techniques has gained popularity and some machine learned potentials have been developed an included to new functionals, resulting in a promising strategy to overcome the issues mentioned above. Furthermore, machine-learned methodologies may be able to capture the complexity of CO_2 activation and ligand-exchange processes, bypassing the need of very expensive DFT-based potentials for enhanced-sampling simulations. Supramolecular effects are also observed in the FLP chemical behaviour, the chemically active dimerisation processes (Figure 2.8) are always competing with the single molecule reactivity, being an extra competitive pathway, mentioned in section 4.3.1. Hence, as part of future work involves the performance of classical Molecular Dynamics simulations in a systematic fashion, to evaluate the impact of the chemical environment of our systems on their electronic structure and properties. Recently, López and coworkers have published a python-based code called DESC, [18] to facilitate this procedure. Its application will thus be strongly considered.

References

(1) IPCC, *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: Cambridge, United Kingdom and New York, NY, USA, 2021; Vol. In Press.

(2) Welch, G. C.; Juan, R. R. S.; Masuda, J. D.; Stephan, D. W. *Science* **2006**, *314*, 1124–1126.

(3) Lang, Z.; Miao, J.; Lan, Y.; Cheng, J.; Xu, X.; Cheng, C. *APL Materials* **2020**, *8*.

(4) Bonin, J.; Robert, M.; Routier, M. *Journal of the American Chemical Society* **2014**, *136*, PMID: 25396278, 16768–16771.

(5) Riddhi, R. K.; Penas-Hidalgo, F.; Chen, H.; Quadrelli, E. A.; Canivet, J.; Mellot-Draznieks, C.; Solé-Daura, A. *Chem. Soc. Rev.* **2024**, -.

(6) Das, S.; Turnell-Ritson, R. C.; Dyson, P. J.; Corminboeuf, C. *Angewandte Chemie International Edition* **2022**, *61*, e202208987.

(7) Andries, J. P.; Vander Heyden, Y.; Buydens, L. M. *Analytica Chimica Acta* **2017**, *982*, 37–47.

(8) Wang, F.; Neumann, R.; de Graaf, C.; Poblet, J. M. *ACS Catalysis* **2021**, *11*, 1495–1504.

(9) Zou, L.; Sa, R.; Lv, H.; Zhong, H.; Wang, R. *ChemSusChem* **2020**, *13*, 6124–6140.

(10) Domingo-Tafalla, B.; Chatterjee, T.; Palomares, E. *Journal of Porphyrins and Phthalocyanines* **2023**, *27*, 23–46.

(11) Fang, J.; Zhu, Y.-N.; Long, X.; Li, X.-B.; Zhang, Q.; Yang, G.; Du, S.; Liu, Z.; Liu, Z.; Peng, F. *Journal of Colloid and Interface Science* **2024**, *668*, 366–374.

(12) Benseghir, Y.; Solé-Daura, A.; Cairnie, D. R.; Robinson, A. L.; Duguet, M.; Mialane, P.; Gairola, P.; Gomez-Mingot, M.; Fontecave, M.; Iovan, D.; Bonnett, B.; Morris, A. J.; Dolbecq, A.; Mellot-Draznieks, C. *J. Mater. Chem. A* **2022**, *10*, 18103–18115.

(13) Davethu, P. A.; de Visser, S. P. *The Journal of Physical Chemistry A* **2019**, *123*, PMID: 31283234, 6527–6535.

(14) Shen, J.; Kolb, M. J.; Göttle, A. J.; Koper, M. T. M. *The Journal of Physical Chemistry C* **2016**, *120*, 15714–15721.

(15) Ghosh, A.; Persson, B. J.; Taylor, P. R. *JBIC Journal of Biological Inorganic Chemistry* **2003**, *8*, 507–511.

(16) Eade, R. H.; Robb, M. A. *Chemical Physics Letters* **1981**, *83*, 362–368.

(17) Van Leest, N. P.; de Bruin, B. *Inorganic Chemistry* **2021**, *60*, PMID: 34096281, 8380–8387.

(18) Masip-Sánchez, A.; Poblet, J. M.; López, X. *Journal of Chemical Theory and Computation* **2025**, *21*, PMID: 40019021, 2472–2486.

(19) Kleidon, A. Understanding the Earth as a whole system: From the Gaia Hypothesis to Thermodynamic Optimality and Human Societies, 2020.

(20) Meadows, D. H.; Meadows, D. L.; Randers, J.; Behrens III, W. W., *The Limits to Growth*; Potomac Associates – Universe Books: New York, 1972, p 205.

(21) Tzimas, E.; Filiou, C.; Peteves, S.-D.; Veyret, J.-B., *Hydrogen storage: state-of-the-art and future perspective*; European Communities: 2003.

(22) Zhou, L. *Renewable and Sustainable Energy Reviews* **2005**, *9*, 395–408.

(23) Mazloomi, K.; Gomes, C. *Renewable and Sustainable Energy Reviews* **2012**, *16*, 3024–3033.

(24) D'Alessandro, D. M.; Smit, B.; Long, J. R. *Angewandte Chemie International Edition* **2010**, *49*, 6058–6082.

(25) Appel, A. M. et al. *Chemical Reviews* **2013**, *113*, PMID: 23767781, 6621–6658.

(26) Bard, A.; Parsons, R.; Jordan, J., *Standard Potentials in Aqueous Solution*; Monographs in Electroanalytical Chemistry and Electrochemistr; CRC Press: 2017.

(27) Topham, S.; Bazzanella, A.; Schiebahn, S.; Luhr, S.; Zhao, L.; Otto, A.; Stolten, D. In *Ullmann's Encyclopedia of Industrial Chemistry*; John Wiley Sons, Ltd: 2014, pp 1–43.

(28) Kuramochi, Y.; Ishitani, O.; Ishida, H. *Coordination Chemistry Reviews* **2018**, *373*, Coordination Chemistry for Energy, 333–356.

(29) Yamazaki, Y.; Takeda, H.; Ishitani, O. *Journal of Photochemistry and Photobiology C: Photochemistry Reviews* **2015**, *25*, 106–137.

(30) Takeda, H.; Cometto, C.; Ishitani, O.; Robert, M. *ACS Catalysis* **2017**, *7*, 70–88.

(31) Pope, M. T.; Jeannin, Y.; Fournier, M., *Heteropoly and isopoly oxometalates*; Springer: 1983; Vol. 33.

(32) Nyman, M. *Dalton Trans.* **2011**, *40*, 8049–8058.

(33) Pope, M. T.; Kortz, U. In *Encyclopedia of Inorganic and Bioinorganic Chemistry*; John Wiley Sons, Ltd: 2012.

(34) Wilkins, L. C.; Melen, R. L. In *Encyclopedia of Inorganic and Bioinorganic Chemistry*; John Wiley Sons, Ltd: 2017, pp 1–24.

(35) Pal, R.; Ghara, M.; Chattaraj, P. K. *Catalysts* **2022**, *12*, DOI: `10.3390/catal12020201`.

(36) Lam, J.; Szkop, K. M.; Mosaferi, E.; Stephan, D. W. *Chem. Soc. Rev.* **2019**, *48*, 3592–3612.

(37) Erker, G. *Comptes Rendus Chimie* **2011**, *14*, Future of sciences, sciences for the future: Chemistry and its interfaces with biology and physics, 831–841.

(38) Liu, L.; Lukose, B.; Jaque, P.; Ensing, B. *Green Energy Environment* **2019**, *4*, 20–28.

(39) Spies, P.; Kehr, G.; Bergander, K.; Wibbeling, B.; Fröhlich, R.; Erker, G. *Dalton Trans.* **2009**, 1534–1541.

(40) Rokob, T. A.; Hamza, A.; Pápai, I. *Journal of the American Chemical Society* **2009**, *131*, PMID: 19722636, 10701–10710.

(41) Barrales-Martínez, C.; Durán, R.; Jaque, P. *Chem. Sci.* **2023**, *14*, 11798–11808.

(42) Stephan, D. W. *Accounts of Chemical Research* **2015**, *48*, PMID: 25535796, 306–316.

(43) Fontaine, F.-G.; Rochette, É. *Accounts of Chemical Research* **2018**, *51*, PMID: 29308653, 454–464.

(44) Stephan, D. W.; Erker, G. *Chem. Sci.* **2014**, *5*, 2625–2641.

(45) Li, N.; Zhang, W.-X. *Chinese Journal of Chemistry* **2020**, *38*, 1360–1370.

(46) Ashley, A. E.; Thompson, A. L.; O'Hare, D. *Angewandte Chemie International Edition* **2009**, *48*, 9839–9843.

(47) Courtemanche, M.-A.; Pulis, A. P.; Rochette, É.; Légaré, M.-A.; Stephan, D. W.; Fontaine, F.-G. *Chem. Commun.* **2015**, *51*, 9797–9800.

(48) Fontaine, F.-G.; Stephan, D. W. *Current Opinion in Green and Sustainable Chemistry* **2017**, *3*, CO2 Capture and Chemistry 2017, 28–32.

(49) Fontaine, F.-G.; Courtemanche, M.-A.; Légaré, M.-A.; Rochette, É. *Coordination Chemistry Reviews* **2017**, *334*, Small Molecule Activation, 124–135.

(50) Wen, M.; Huang, F.; Lu, G.; Wang, Z.-X. *Inorganic Chemistry* **2013**, *52*, PMID: 24087841, 12098–12107.

(51) Jiang, B.; Zhang, Q.; Dang, L. *Org. Chem. Front.* **2018**, *5*, 1905–1915.

(52) Ghara, M.; Chattaraj, P. K. *Structural Chemistry* **2019**, *30*, 1067–1077.

(53) Ghara, M.; Pan, S.; Chattaraj, P. K. *Phys. Chem. Chem. Phys.* **2019**, *21*, 21267–21277.

(54) Liu, L.; Vankova, N.; Heine, T. *Phys. Chem. Chem. Phys.* **2016**, *18*, 3567–3574.

(55) Ye, J.; Yeh, B. Y.; Reynolds, R. A.; Johnson, J. K. *Molecular Simulation* **2017**, *43*, 821–827.

(56) Zhuang, D.; Li, Y.; Zhu, J. *Organometallics* **2020**, *39*, 2636–2641.

(57) Ghara, M.; Chattaraj, P. K. *Theoretical Chemistry Accounts* **2020**, *139*, 1–8.

(58) Guo, Y.; Li, S. *Inorganic Chemistry* **2008**, *47*, PMID: 18572910, 6212–6219.

(59) Spies, P.; Erker, G.; Kehr, G.; Bergander, K.; Fröhlich, R.; Grimme, S.; Stephan, D. W. *Chem. Commun.* **2007**, 5072–5074.

(60) Spies, P.; Schwendemann, S.; Lange, S.; Kehr, G.; Fröhlich, R.; Erker, G. *Angewandte Chemie International Edition* **2008**, *47*, 7543–7546.

(61) Grimme, S.; Kruse, H.; Goerigk, L.; Erker, G. *Angewandte Chemie International Edition* **2010**, *49*, 1402–1405.

(62) Rokob, T. A.; Bakó, I.; Stirling, A.; Hamza, A.; Pápai, I. *Journal of the American Chemical Society* **2013**, *135*, PMID: 23432375, 4425–4437.

(63) Liu Zeonjuk, L.; St. Petkov, P.; Heine, T.; Röschenthaler, G.-V.; Eicher, J.; Vankova, N. *Phys. Chem. Chem. Phys.* **2015**, *17*, 10687–10698.

(64) Yepes, D.; Jaque, P.; Fernández, I. *Chemistry – A European Journal* **2016**, *22*, 18801–18809.

(65) Pérez, P.; Yepes, D.; Jaque, P.; Chamorro, E.; Domingo, L. R.; Rojas, R. S.; Toro-Labbé, A. *Phys. Chem. Chem. Phys.* **2015**, *17*, 10715–10725.

(66) Cabrera-Trujillo, J. J.; Fernández, I. *Chem. Commun.* **2019**, *55*, 675–678.

(67) Rokob, T. A.; Hamza, A.; Stirling, A.; Soós, T.; Pápai, I. *Angewandte Chemie* **2008**, *120*, 2469–2472.

(68) Welch, G. C.; Stephan, D. W. *Journal of the American Chemical Society* **2007**, *129*, PMID: 17260994, 1880–1881.

(69) Heshmat, M. *The Journal of Physical Chemistry C* **2020**, *124*, PMID: 34122685, 10951–10960.

(70) Daru, J.; Bakó, I.; Stirling, A.; Pápai, I. *ACS Catalysis* **2019**, *9*, 6049–6057.

(71) Pu, M.; Privalov, T. *The Journal of Chemical Physics* **2013**, *138*, 154305.

(72) Pu, M.; Privalov, T. *ChemPhysChem* **2014**, *15*, 3714–3719.

(73) Pu, M.; Privalov, T. *ChemPhysChem* **2014**, *15*, 2936–2944.

(74) Pu, M.; Privalov, T. *Israel Journal of Chemistry* **2015**, *55*, 179–195.

(75) Liu, L.; Lukose, B.; Ensing, B. *The Journal of Physical Chemistry C* **2017**, *121*, 2046–2051.

(76) Chai, J.-D.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.

(77) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146401.

(78) Staroverov, V. N.; Scuseria, G. E.; Tao, J.; Perdew, J. P. *The Journal of Chemical Physics* **2003**, *119*, 12129–12137.

(79) Staroverov, V. N.; Scuseria, G. E.; Tao, J.; Perdew, J. P. *The Journal of Chemical Physics* **2004**, *121*, 11507–11507.

(80) Schirmer, B.; Grimme, S. *Chem. Commun.* **2010**, *46*, 7942–7944.

(81) Becke, A. D. *The Journal of Chemical Physics* **1993**, *98*, 5648–5652.

(82) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *Journal of Chemical Theory and Computation* **2006**, *2*, PMID: 26626525, 364–382.

(83) Zhao, Y.; Truhlar, D. G. *Theoretical chemistry accounts* **2008**, *120*, 215–241.

(84) Tomasi, J.; Mennucci, B.; Cammi, R. *Chemical Reviews* **2005**, *105*, PMID: 16092826, 2999–3094.

(85) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650–654.

(86) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. *The Journal of Chemical Physics* **1982**, *77*, 3654–3665.

(87) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. *J. Chem. Phys.* **1982**, *77*, 3654–3665.

(88) Spitznagel, G. W.; Clark, T.; Schleyer, P. v. R.; Hehre, W. J. *J. Comput. Chem.* **1987**, *8*, 1109–1116.

(89) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. R. *J. Comput. Chem.* **1983**, *4*, 294–301.

(90) Sharma, G.; Newman, P. D.; Platts, J. A. *Journal of Molecular Graphics and Modelling* **2021**, *105*, 107846.

(91) Erker, G.; Stephan, D. W., *Frustrated Lewis Pairs*; Springer: 2013; Vol. 1.

(92) Huang, F.; Jiang, J.; Wen, M.; Wang, Z.-X. *Journal of Theoretical and Computational Chemistry* **2014**, *13*, 1350074.

(93) Stephan, D. W.; Erker, G. *Angewandte Chemie International Edition* **2010**, *49*, 46–76.

(94) Kolychev, E. L.; Theuergarten, E.; Tamm, M. In *Frustrated Lewis Pairs II: Expanding the Scope*, Erker, G., Stephan, D. W., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2013, pp 121–155.

(95) Wass, D. F.; Chapman, A. M. In *Frustrated Lewis Pairs II: Expanding the Scope*, Erker, G., Stephan, D. W., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2013, pp 261–280.

(96) Wan, Q.; Lin, S.; Guo, H. *Molecules* **2022**, *27*, DOI: `10.3390/molecules27123734`.

(97) Ghuman, K. K.; Hoch, L. B.; Wood, T. E.; Mims, C.; Singh, C. V.; Ozin, G. A. *ACS Catalysis* **2016**, *6*, 5764–5770.

(98) Ghoussoub, M.; Yadav, S.; Ghuman, K. K.; Ozin, G. A.; Singh, C. V. *ACS Catalysis* **2016**, *6*, 7109–7117.

(99) Hamza, A.; Stirling, A.; András Rokob, T.; Pápai, I. *International Journal of Quantum Chemistry* **2009**, *109*, 2416–2425.

(100) Camaioni, D. M.; Ginovska-Pangovska, B.; Schenter, G. K.; Kathmann, S. M.; Autrey, T. *The Journal of Physical Chemistry A* **2012**, *116*, PMID: 22663774, 7228–7237.

(101) Skara, G.; De Vleeschouwer, F.; Geerlings, P.; De Proft, F.; Pinter, B. *Scientific Reports* **2017**, *7*, 16024.

(102) Zeonjuk, L. L.; Vankova, N.; Mavrandonakis, A.; Heine, T.; Röschenthaler, G.-V.; Eicher, J. *Chemistry – A European Journal* **2013**, *19*, 17413–17424.

(103) Das, S.; Laplaza, R.; Blaskovits, J. T.; Corminboeuf, C. *Angewandte Chemie International Edition* **2022**, *61*, e202202727.

(104) Bell, R. P.; Hinshelwood, C. N. *Proceedings of the Royal Society of London. Series A - Mathematical and Physical Sciences* **1936**, *154*, 414–429.

(105) Evans, M. G.; Polanyi, M. *Trans. Faraday Soc.* **1938**, *34*, 11–24.

(106) Ullrich, M.; Lough, A. J.; Stephan, D. W. *Journal of the American Chemical Society* **2009**, *131*, PMID: 19061321, 52–53.

(107) Neu, R. C.; Ouyang, E. Y.; Geier, S. J.; Zhao, X.; Ramos, A.; Stephan, D. W. *Dalton Trans.* **2010**, *39*, 4285–4294.

(108) Jiang, C.; Blacque, O.; Fox, T.; Berke, H. *Dalton Trans.* **2011**, *40*, 1091–1097.

(109) Ye, J.; Johnson, J. K. *ACS Catalysis* **2015**, *5*, 6219–6229.

(110) Mulliken, R. S. *The Journal of Chemical Physics* **1934**, *2*, 782–793.

(111) Parr, R. G.; Pearson, R. G. *Journal of the American Chemical Society* **1983**, *105*, 7512–7516.

(112) Zhang, J.; Shao, Y.; Li, Y.; Liu, Y.; Ke, Z. *Chinese Chemical Letters* **2018**, *29*, 1226–1232.

(113) Roesler, R.; Piers, W. E.; Parvez, M. *Journal of Organometallic Chemistry* **2003**, *680*, Frontiers in Boron Chemistry Dedicated to M. Frederick Hawthorne on his 75th Birthday, 218–222.

(114) Chernichenko, K.; Nieger, M.; Leskelä, M.; Repo, T. *Dalton Transactions* **2012**, *41*, 9029–9032.

(115) Chernichenko, K.; Kótai, B.; Pápai, I.; Zhivonitko, V.; Nieger, M.; Leskelä, M.; Repo, T. *Angewandte Chemie International Edition* **2015**, *54*, 1749–1753.

(116) Tussing, S.; Kaupmees, K.; Paradies, J. *Chemistry – A European Journal* **2016**, *22*, 7422–7426.

(117) Heshmat, M.; Ensing, B. *The Journal of Physical Chemistry A* **2020**, *124*, PMID: 32666803, 6399–6410.

(118) McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *72*, 5639–5648.

(119) Frisch, M. J. et al. Gaussian~16 Revision C.02, Gaussian Inc. Wallingford CT, 2016.

(120) Cancès, E.; Mennucci, B.; Tomasi, J. *The Journal of Chemical Physics* **1997**, *107*, 3032–3041.

(121) Pedregosa, F. et al. *Journal of Machine Learning Research* **2011**, *12*, 2825–2830.

(122) Tropsha, A. *Molecular Informatics* **2010**, *29*, 476–488.

(123) Abdi, H. *Wiley interdisciplinary reviews: computational statistics* **2010**, *2*, 97–106.

(124) Paradies, J. *Coordination Chemistry Reviews* **2019**, *380*, 170–183.

(125) Chernichenko, K.; Kótai, B.; Nieger, M.; Heikkinen, S.; Pápai, I.; Repo, T. *Dalton Transactions* **2017**, *46*, 2263–2269.

(126) Zhivonitko, V. V.; Sorochkina, K.; Chernichenko, K.; Kótai, B.; Földes, T.; Pápai, I.; Telkki, V.-V.; Repo, T.; Koptyug, I. *Physical Chemistry Chemical Physics* **2016**, *18*, 27784–27795.

(127) Son, J.-H.; Pudenz, M. A.; Hoefelmeyer, J. D. *Dalton Transactions* **2010**, *39*, 11081–11090.

(128) Bontemps, S.; Devillard, M.; Mallet-Ladeira, S.; Bouhadir, G.; Miqueu, K.; Bourissou, D. *Inorganic Chemistry* **2013**, *52*, PMID: 23544869, 4714–4720.

(129) Beckmann, J.; Hupf, E.; Lork, E.; Mebs, S. *Inorganic Chemistry* **2013**, *52*, PMID: 24080026, 11881–11888.

(130) Kutter, F.; Lork, E.; Beckmann, J. *Zeitschrift für anorganische und allgemeine Chemie* **2018**, *644*, 1234–1237.

(131) Mo, Z.; Kolychev, E. L.; Rit, A.; Campos, J.; Niu, H.; Aldridge, S. *Journal of the American Chemical Society* **2015**, *137*, 12227–12230.

(132) Mo, Z.; Rit, A.; Campos, J.; Kolychev, E. L.; Aldridge, S. *Journal of the American Chemical Society* **2016**, *138*, 3306–3309.

(133) Bhattacharjee, I.; Bhunya, S.; Paul, A. *Inorganic Chemistry* **2020**, *59*, 1046–1056.

(134) Delarmelina, M.; Carneiro, J. W. d. M.; Catlow, C. R. A.; Bühl, M. *Catalysis Communications* **2022**, *162*, 106385.

(135) Bronsted, J. *Chemical Reviews* **1928**, *5*, 231–338.

(136) Evans, M.; Polanyi, M. *Transactions of the Faraday Society* **1936**, *32*, 1333–1360.

(137) Mentel, Ł. mendeleev - A Python package with properties of chemical elements, ions, isotopes and methods to manipulate and visualize periodic table. Version v0.14.0, 2023.

(138) Xu, S.; Haeffner, F.; Li, B.; Zakharov, L. N.; Liu, S.-Y. *Angewandte Chemie International Edition* **2014**, *53*, 6795–6799.

(139) Bontemps, S.; Bouhadir, G.; Dyer, P. W.; Miqueu, K.; Bourissou, D. *Inorganic Chemistry* **2007**, *46*, PMID: 17523635, 5149–5151.

(140) Monakhov, K. Y.; Bensch, W.; Kögerler, P. *Chem. Soc. Rev.* **2015**, *44*, 8443–8483.

(141) Müller, A.; Gouzerh, P. *Chem. Soc. Rev.* **2012**, *41*, 7431–7463.

(142) Pope, M. T. In *Encyclopedia of Inorganic Chemistry*; John Wiley Sons, Ltd: 2006.

(143) Pratt III, H. D.; Anderson, T. M. *Dalton Trans.* **2013**, *42*, 15650–15655.

(144) Peake, C. L.; Kibler, A. J.; Newton, G. N.; Walsh, D. A. *ACS Applied Energy Materials* **2021**, *4*, 8765–8773.

(145) López, X.; Carbó, J. J.; Bo, C.; Poblet, J. M. *Chemical Society Reviews* **2012**, *41*, 7537–7571.

(146) Glass, E. N.; Fielden, J.; Huang, Z.; Xiang, X.; Musaev, D. G.; Lian, T.; Hill, C. L. *Inorganic Chemistry* **2016**, *55*, PMID: 27082443, 4308–4319.

(147) Proust, A.; Matt, B.; Villanneau, R.; Guillemot, G.; Gouzerh, P.; Izzet, G. *Chem. Soc. Rev.* **2012**, *41*, 7605–7622.

(148) Zhang, H.; Zhao, W.-L.; Li, H.; Zhuang, Q.; Sun, Z.; Cui, D.; Chen, X.; Guo, A.; Ji, X.; An, S.; Chen, W.; Song, Y.-F. *Polyoxometalates* **2022**, *1*, 9140011.

(149) Cameron, J. M.; Guillemot, G.; Galambos, T.; Amin, S. S.; Hampson, E.; Haidaraly, K. M.; Newton, G. N.; Izzet, G. *Chemical Society Reviews* **2022**, *51*, 293–328.

(150) Nikoloudakis, E.; Karikis, K.; Laurans, M.; Kokotidou, C.; Solé-Daura, A.; Carbó, J. J.; Charisiadis, A.; Charalambidis, G.; Izzet, G.; Mitraki, A.; Douvas, A. M.; Poblet, J. M.; Proust, A.; Coutsolelos, A. G. *Dalton Trans.* **2018**, *47*, 6304–6313.

(151) Dolbecq, A.; Dumas, E.; Mayer, C. R.; Mialane, P. *Chemical Reviews* **2010**, *110*, PMID: 20666374, 6009–6048.

(152) Mialane, P.; Mellot-Draznieks, C.; Gairola, P.; Duguet, M.; Benseghir, Y.; Oms, O.; Dolbecq, A. *Chemical Society Reviews* **2021**, *50*, 6152–6220.

(153) Samaniyan, M.; Mirzaei, M.; Khajavian, R.; Eshtiagh-Hosseini, H.; Streb, C. *ACS Catalysis* **2019**, *9*, 10174–10191.

(154) Buru, C. T.; Farha, O. K. *ACS Applied Materials & Interfaces* **2020**, *12*, PMID: 31961127, 5345–5360.

(155) Li, X.-X.; Liu, J.; Zhang, L.; Dong, L.-Z.; Xin, Z.-F.; Li, S.-L.; Huang-Fu, X.-Q.; Huang, K.; Lan, Y.-Q. *ACS Applied Materials & Interfaces* **2019**, *11*, PMID: 31240910, 25790–25795.

(156) Xie, S.-L.; Liu, J.; Dong, L.-Z.; Li, S.-L.; Lan, Y.-Q.; Su, Z.-M. *Chemical science* **2019**, *10*, 185–190.

(157) Xu, H.; You, S.; Lang, Z.; Sun, Y.; Sun, C.; Zhou, J.; Wang, X.; Kang, Z.; Su, Z. *Chemistry–A European Journal* **2020**, *26*, 2735–2740.

(158) Benseghir, Y.; Solé-Daura, A.; Mialane, P.; Marrot, J.; Dalecky, L.; Béchu, S.; Frégnaux, M.; Gomez-Mingot, M.; Fontecave, M.; Mellot-Draznieks, C., et al. *ACS Catalysis* **2021**, *12*, 453–464.

(159) Dolbecq, A.; Guirauden, A.; Fourmigué, M.; Boubekeur, K.; Batail, P.; Rohmer, M.-M.; Bénard, M.; Coulon, C.; Sallé, M.; Blanchard, P. *J. Chem. Soc., Dalton Trans.* **1999**, 1241–1248.

(160) Chen, J.-J.; Vilà-Nadal, L.; Solé-Daura, A.; Chisholm, G.; Minato, T.; Busche, C.; Zhao, T.; Kandasamy, B.; Ganin, A. Y.; Smith, R. M.; Colliard, I.; Carbó, J. J.; Poblet, J. M.; Nyman, M.; Cronin, L. *Journal of the American Chemical Society* **2022**, *144*, PMID: 35536652, 8951–8960.

(161) Cao, M.; Zi, J.; Sang, R.; Xu, L. *Dalton Trans.* **2023**, *52*, 13351–13357.

(162) Zhao, M.; Liu, Q.; Feng, Y.; Lv, H. *Dalton Trans.* **2024**, *53*, 18083–18088.

(163) Ren, H.; Wang, G.; Chen, T.; Wang, Q.; Ren, J.; Pang, H. *Advanced Sustainable Systems* **2025**, *9*, 2400752.

(164) Müller, W. E.; Wang, X.; Schröder, H. C., *Biomedical Inorganic Polymers;* Springer: 2013.

(165) Bijelic, A.; Rompel, A. *ChemTexts* **2018**, *4*, 10.

(166) Wang, S.-S.; Yang, G.-Y. *Chemical Reviews* **2015**, *115*, PMID: 25965251, 4893–4962.

(167) Mars, P.; van Krevelen, D. *Chemical Engineering Science* **1954**, *3*, The Proceedings of the Conference on Oxidation Processes, 41–59.

(168) Zhang, Y.; Li, Y.; Guo, H.; Guo, Y.; Song, R. *Materials Chemistry Frontiers* **2024**, *8*, 732–768.

(169) Remmers, M.; Mashtakov, B.; Repp, S.; Rein, A. S. J.; Wang, K.; Anjass, M.; Chen, Z.; Carrella, L. M.; Rentschler, E.; Streb, C. *Angewandte Chemie International Edition* **2025**, *64*, e202418864.

(170) Rossin, E.; Bonchio, M.; Natali, M.; Sartorel, A. *Sustainable Energy Fuels* **2024**, *8*, 1944–1952.

(171) Matt, B.; Fize, J.; Moussa, J.; Amouri, H.; Pereira, A.; Artero, V.; Izzet, G.; Proust, A. *Energy Environ. Sci.* **2013**, *6*, 1504–1508.

(172) Tzaguy, A.; Masip-Sánchez, A.; Avram, L.; Solé-Daura, A.; López, X.; Poblet, J. M.; Neumann, R. *Journal of the American Chemical Society* **2023**, *145*, PMID: 37642197, 19912–19924.

(173) Li, X.-H.; Chen, W.-L.; He, P.; Wang, T.; Liu, D.; Li, Y.-W.; Li, Y.-G.; Wang, E.-B. *Inorg. Chem. Front.* **2019**, *6*, 3315–3326.

(174) Li, G.; Gu, Y.; Ren, R.; Li, S.; Zhu, H.; Xue, D.; Kong, X.; Zheng, Z.; Liu, N.; Li, B.; Zhang, J. *Nanoscale* **2024**, *16*, 12550–12558.

(175) Kim, N.; Nam, J. S.; Jo, J.; Seong, J.; Kim, H.; Kwon, Y.; Lah, M. S.; Lee, J. H.; Kwon, T.-H.; Ryu, J. *Nanoscale Horizons* **2021**, *6*, 379–385.

(176) Wang, Y.-R.; Huang, Q.; He, C.-T.; Chen, Y.; Liu, J.; Shen, F.-C.; Lan, Y.-Q. *Nature communications* **2018**, *9*, 4466.

(177) Fabre, B.; Falaise, C.; Cadot, E. *ACS Catalysis* **2022**, *12*, 12055–12091.

(178) Gu, J.; Chen, W.; Shan, G.-G.; Li, G.; Sun, C.; Wang, X.-L.; Su, Z. *Materials Today Energy* **2021**, *21*, 100760.

(179) Cao, Y.; Chen, Q.; Shen, C.; He, L. *Molecules* **2019**, *24*, 2069.

(180) Song, Y.-F.; Tsunashima, R. *Chem. Soc. Rev.* **2012**, *41*, 7384–7402.

(181) Zang, D.; Wang, H. *Polyoxometalates* **2022**, *1*, 9140006.

(182) Talbi, K.; Penas-Hidalgo, F.; Robinson, A. L.; Gotico, P.; Leibl, W.; Mialane, P.; Gomez-Mingot, M.; Fontecave, M.; Solé-Daura, A.; Mellot-Draznieks, C., et al. *Applied Catalysis B: Environment and Energy* **2024**, *345*, 123681.

(183) Lv, H.; Guo, W.; Wu, K.; Chen, Z.; Bacsa, J.; Musaev, D. G.; Geletii, Y. V.; Lauinger, S. M.; Lian, T.; Hill, C. L. *Journal of the American Chemical Society* **2014**, *136*, PMID: 25243410, 14015–14018.

(184) Von Allmen, K.; Moré, R.; Müller, R.; Soriano-López, J.; Linden, A.; Patzke, G. R. *ChemPlusChem* **2015**, *80*, 1389–1398.

(185) Lv, H.; Chi, Y.; van Leusen, J.; Kögerler, P.; Chen, Z.; Bacsa, J.; Geletii, Y. V.; Guo, W.; Lian, T.; Hill, C. L. *Chemistry – A European Journal* **2015**, *21*, 17363–17370.

(186) Han, X.-B.; Qin, C.; Wang, X.-L.; Tan, Y.-Z.; Zhao, X.-J.; Wang, E.-B. *Applied Catalysis B: Environmental* **2017**, *211*, 349–356.

(187) Paille, G.; Boulmier, A.; Bensaid, A.; Ha-Thi, M.-H.; Tran, T.-T.; Pino, T.; Marrot, J.; Rivière, E.; Hendon, C. H.; Oms, O.; Gomez-Mingot, M.; Fontecave, M.; Mellot-Draznieks, C.; Dolbecq, A.; Mialane, P. *Chem. Commun.* **2019**, *55*, 4166–4169.

(188) Qin, L.; Wang, R.; Xin, X.; Zhang, M.; Liu, T.; Lv, H.; Yang, G.-Y. *Applied Catalysis B: Environmental* **2022**, *312*, 121386.

(189) Tanuhadi, E.; Cano, J.; Batool, S.; Cherevan, A.; Eder, D.; Rompel, A. *Journal of Materials Chemistry C* **2022**, *10*, 17048–17052.

(190) Chen, Y.; Guo, Z.-W.; Chen, Y.-P.; Zhuang, Z.-Y.; Wang, G.-Q.; Li, X.-X.; Zheng, S.-T.; Yang, G.-Y. *Inorganic Chemistry Frontiers* **2021**, *8*, 1303–1311.

(191) Zhang, G.; Wang, F.; Tubul, T.; Baranov, M.; Leffler, N.; Neyman, A.; Poblet, J. M.; Weinstock, I. A. *Angewandte Chemie* **2022**, *134*, e202213162.

(192) Chang, Q.; Meng, X.; Ruan, W.; Feng, Y.; Li, R.; Zhu, J.; Ding, Y.; Lv, H.; Wang, W.; Chen, G., et al. *Angewandte Chemie* **2022**, *134*, e202117637.

(193) Pichon, C.; Mialane, P.; Dolbecq, A.; Marrot, J.; Rivière, E.; Bassil, B. S.; Kortz, U.; Keita, B.; Nadjo, L.; Sécheresse, F. *Inorganic Chemistry* **2008**, *47*, 11120–11128.

(194) El Moll, H.; Rousseau, G.; Dolbecq, A.; Oms, O.; Marrot, J.; Haouas, M.; Taulelle, F.; Rivière, E.; Wernsdorfer, W.; Lachkar, D.; Lacôte, E.; Keita, B.; Mialane, P. *Chemistry – A European Journal* **2013**, *19*, 6753–6765.

(195) Rousseau, G.; Zhang, S.; Oms, O.; Dolbecq, A.; Marrot, J.; Liu, R.; Shang, X.; Zhang, G.; Keita, B.; Mialane, P. *Chemistry – A European Journal* **2015**, *21*, 12153–12160.

(196) Limburg, B.; Bouwman, E.; Bonnet, S. *ACS Catalysis* **2016**, *6*, 5273–5284.

(197) Gueret, R.; Poulard, L.; Oshinowo, M.; Chauvin, J.; Dahmane, M.; Dupeyre, G.; Lainé, P. P.; Fortage, J.; Collomb, M.-N. *Acs Catalysis* **2018**, *8*, 3792–3802.

(198) Pellegrin, Y.; Odobel, F. *Comptes Rendus Chimie* **2017**, *20*, Artificial photosynthesis / La photosynthèse artificielle, 283–295.

(199) Li, M.; Zhuang, B.; Lu, Y.; An, L.; Wang, Z.-G. *Journal of the American Chemical Society* **2021**, *143*, 773–784.

(200) Sampaio, R. N.; Grills, D. C.; Polyansky, D. E.; Szalda, D. J.; Fujita, E. *Journal of the American Chemical Society* **2020**, *142*, PMID: 31881154, 2413–2428.

(201) Jafari, S. A.; Entezari, M. H. *Journal of Molecular Liquids* **2020**, *312*, 113331.

(202) Misra, A.; Kozma, K.; Streb, C.; Nyman, M. *Angewandte Chemie International Edition* **2020**, *59*, 596–612.

(203) Song, J.; Luo, Z.; Zhu, H.; Huang, Z.; Lian, T.; Kaledin, A. L.; Musaev, D. G.; Lense, S.; Hardcastle, K. I.; Hill, C. L. *Inorganica Chimica Acta* **2010**, *363*, Special Volume: Protagonists in Chemistry: Dedication to Professor Achim Müller, 4381–4386.

(204) Chaumont, A.; Wipff, G. *The Journal of Physical Chemistry C* **2009**, *113*, 18233–18243.

(205) Álvarez-Moreno, M.; de Graaf, C.; López, N.; Maseras, F.; Poblet, J. M.; Bo, C. *Journal of Chemical Information and Modeling* **2015**, *55*, PMID: 25469626, 95–103.

(206) Roy, L. E.; Hay, P. J.; Martin, R. L. *J. Chem. Theory Comput.* **2008**, *4*, 1029–1031.

(207) Hay, P. J.; Wadt, W. R. *The Journal of Chemical Physics* **1985**, *82*, 270–283.

(208) Ehlers, A.; Böhme, M.; Dapprich, S.; Gobbi, A.; Höllwarth, A.; Jonas, V.; Köhler, K.; Stegmann, R.; Veldkamp, A.; Frenking, G. *Chemical Physics Letters* **1993**, *208*, 111–114.

(209) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *The Journal of Chemical Physics* **1972**, *56*, 2257–2261.

(210) Hariharan, P. C.; Pople, J. A. *Theoretica chimica acta* **1973**, *28*, 213–222.

(211) Clemente-Juan, J. M.; Andres, H.; Borrás-Almenar, J. J.; Coronado, E.; Güdel, H. U.; Aebersold, M.; Kearly, G.; Büttner, H.; Zolliker, M. *Journal of the American Chemical Society* **1999**, *121*, 10021–10027.

(212) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. *Computer Physics Communications* **2005**, *167*, 103–128.

(213) Kühne, T. D. et al. *The Journal of Chemical Physics* **2020**, *152*, 194103.

(214) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.

(215) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.

(216) LIPPERT, B. G.; HUTTER, J.; and, M. P. *Molecular Physics* **1997**, *92*, 477–488.

(217) VandeVondele, J.; Hutter, J. *The Journal of Chemical Physics* **2007**, *127*, 114105.

(218) Goedecker, S.; Teter, M.; Hutter, J. *Phys. Rev. B* **1996**, *54*, 1703–1710.

(219) Hartwigsen, C.; Goedecker, S.; Hutter, J. *Phys. Rev. B* **1998**, *58*, 3641–3662.

(220) Bussi, G.; Donadio, D.; Parrinello, M. *The Journal of Chemical Physics* **2007**, *126*, 014101.

(221) Rappé, A. K.; Casewit, C. J.; Colwell, K.; Goddard III, W. A.; Skiff, W. M. *Journal of the American chemical society* **1992**, *114*, 10024–10035.

(222) Díaz Leines, G.; Ensing, B. *Phys. Rev. Lett.* **2012**, *109*, 020601.

(223) Pérez de Alba Ortíz, A.; Tiwari, A.; Puthenkalathil, R. C.; Ensing, B. *The Journal of Chemical Physics* **2018**, *149*, 072320.

(224) Cohen, S. G.; Baumgarten, R. J. *Journal of the American Chemical Society* **1965**, *87*, 2996–2997.

(225) Cohen, S. G.; Parola, A.; Parsons, G. H. J. *Chemical Reviews* **1973**, *73*, 141–161.

(226) DeLaive, P. J.; Foreman, T. K.; Giannotti, C.; Whitten, D. G. *Journal of the American Chemical Society* **1980**, *102*, 5627–5631.

(227) Song, W.; Brennaman, M. K.; Concepcion, J. J.; Jurss, J. W.; Hoertz, P. G.; Luo, H.; Chen, C.; Hanson, K.; Meyer, T. J. *The Journal of Physical Chemistry C* **2011**, *115*, 7081–7091.

(228) Solé-Daura, A.; Benseghir, Y.; Ha-Thi, M.-H.; Fontecave, M.; Mialane, P.; Dolbecq, A.; Mellot-Draznieks, C. *ACS Catalysis* **2022**, *12*, 9244–9255.

(229) Prier, C. K.; Rankic, D. A.; MacMillan, D. W. C. *Chemical Reviews* **2013**, *113*, PMID: 23509883, 5322–5363.

(230) Wayner, D. D.; McPhee, D. J.; Griller, D. *Journal of the American Chemical Society* **1988**, *110*, 132–137.

(231) Stanbury, M.; Compain, J.-D.; Chardon-Noblat, S. *Coordination Chemistry Reviews* **2018**, *361*, 120–137.

(232) Isegawa, M.; Sharma, A. K. *Sustainable Energy Fuels* **2019**, *3*, 1730–1738.

(233) Petersen, H. A.; Myren, T. H.; Luca, O. R. *Inorganics* **2020**, *8*, 62.

(234) Zhang, A.; Yang, J.; Luo, Y.; Fan, S. *Scientific Reports* **2023**, *13*, 11305.

(235) Tahoun, M.; Gee, C. T.; McCoy, V. E.; Sander, P. M.; Müller, C. E. *RSC Adv.* **2021**, *11*, 7552–7563.

(236) Shinokubo, H.; Osuka, A. *Chem. Commun.* **2009**, 1011–1021.

(237) Rajasree, S. S.; Li, X.; Deria, P. *Communications Chemistry* **2021**, *4*, 47.

(238) Song, Z.; Adeyemo, A. O.; Baker, J.; Traylor, S. M.; Lightfoot, M. L. *Georgia journal of science: official publication of the Georgia Academy of Science* **2011**, *69*, 89.

(239) Liu, S.; Li, Z.; Tong, H.; Zhong, Y.; Bai, F. In *Self-Assembly of Materials and Their Applications*; IntechOpen: 2022.

(240) Chen, X.; Gao, J.; Jiang, D. *Chemistry Letters* **2015**, *44*, 1257–1259.

(241) Abdinejad, M.; Yuan, T.; Tang, K.; Duangdangchote, S.; Farzi, A.; Iglesias van Montfort, H.-P.; Li, M.; Middelkoop, J.; Wolff, M.; Seifitokaldani, A.; Voznyy, O.; Burdyny, T. *Chemistry – A European Journal* **2023**, *29*, e202203977.

(242) Ren, Z.; Zhao, B.; Xie, J. *Small* **2023**, *19*, 2301818.

(243) Call, A.; Cibian, M.; Yamamoto, K.; Nakazono, T.; Yamauchi, K.; Sakai, K. *ACS Catalysis* **2019**, *9*, 4867–4874.

(244) Zhang, X.; Cibian, M.; Call, A.; Yamauchi, K.; Sakai, K. *ACS Catalysis* **2019**, *9*, 11263–11273.

(245) Pugliese, E.; Gotico, P.; Wehrung, I.; Boitrel, B.; Quaranta, A.; Ha-Thi, M.-H.; Pino, T.; Sircoglou, M.; Leibl, W.; Halime, Z.; Aukauloo, A. *Angewandte Chemie International Edition* **2022**, *61*, e202117530.

(246) Costentin, C.; Robert, M.; Savéant, J.-M. *Accounts of Chemical Research* **2015**, *48*, PMID: 26559053, 2996–3006.

(247) Wang, C.; Zhu, C.-Y.; Zhang, M.; Geng, Y.; Li, Y.-G.; Su, Z.-M. *J. Mater. Chem. A* **2020**, *8*, 14807–14814.

(248) Araghi, M.; Mirkhani, V.; Moghadam, M.; Tangestaninejad, S.; Mohammdpoor-Baltork, I. *Dalton Trans.* **2012**, *41*, 3087–3094.

(249) Araghi, M.; Mirkhani, V.; Moghadam, M.; Tangestaninejad, S.; Mohammdpoor-Baltork, I. *Dalton Trans.* **2012**, *41*, 11745–11752.

(250) Strong, J. B.; Yap, G. P. A.; Ostrander, R.; Liable-Sands, L. M.; Rheingold, A. L.; Thouvenot, R.; Gouzerh, P.; Maatta, E. A. *Journal of the American Chemical Society* **2000**, *122*, 639–649.

(251) Braun, W.; Herron, J. T.; Kahaner, D. K. *International Journal of Chemical Kinetics* **1988**, *20*, 51–62.

(252) Burns, R. G., *Mineralogical applications of crystal field theory*; 5; Cambridge university press: 1993.

(253) Reed, C. A.; Guiset, F. *Journal of the American Chemical Society* **1996**, *118*, 3281–3282.

(254) Rydberg, P.; Olsen, L. *The Journal of Physical Chemistry A* **2009**, *113*, 11949–11953.

(255) Radoń, M. *Physical Chemistry Chemical Physics* **2024**, *26*, 18182–18195.

(256) Römelt, C.; Song, J.; Tarrago, M.; Rees, J. A.; van Gastel, M.; Weyhermüller, T.; DeBeer, S.; Bill, E.; Neese, F.; Ye, S. *Inorganic Chemistry* **2017**, *56*, PMID: 28379689, 4745–4750.

(257) Smith, J. S.; Nebgen, B.; Lubbers, N.; Isayev, O.; Roitberg, A. E. *The Journal of Chemical Physics* **2018**, *148*, 241733.

Appendix A

Chapter 2: Supporting Figures

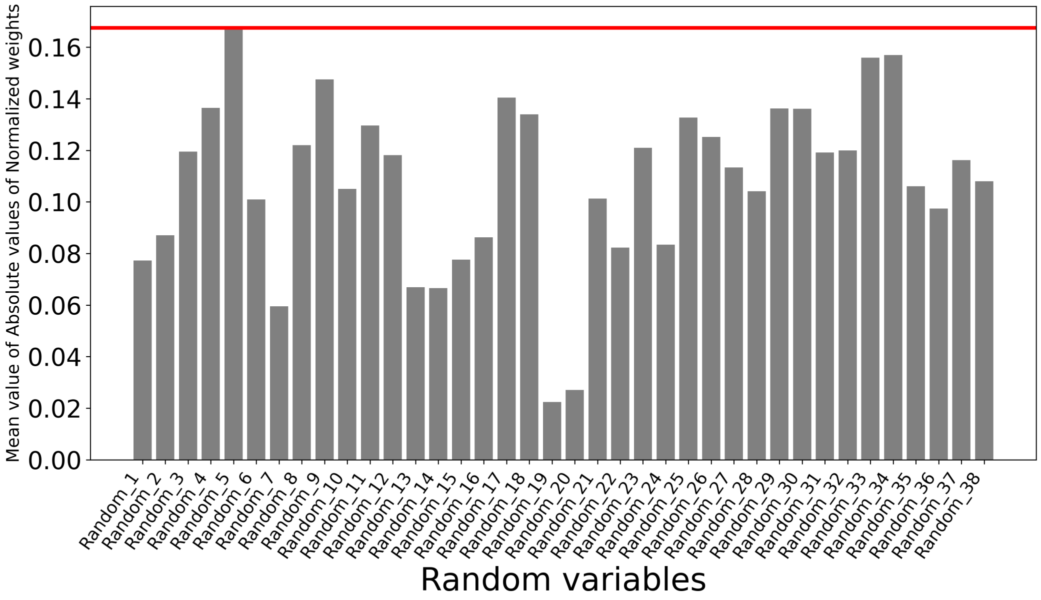


FIGURE A.1: Mean values of the absolute value of each coefficient of the generated random variables across the 10 repetitions models predicting $\Delta G^{\ddagger}$ at the first iteration of the data-driven method. Threshold established at the maximum weight among all random variables as a horizontal red line.

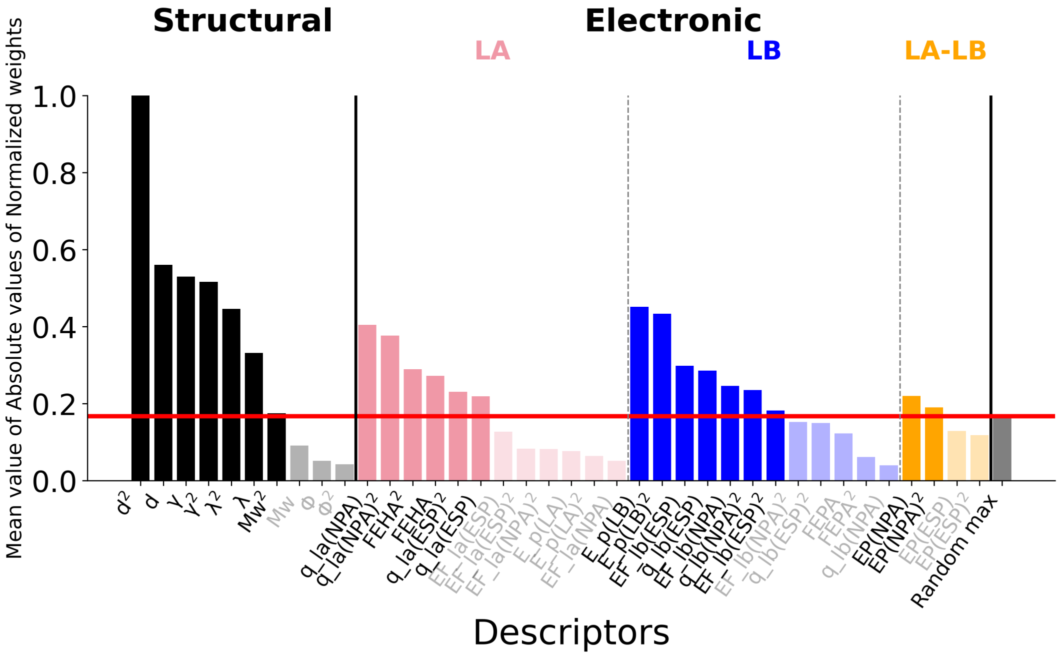


FIGURE A.2: Mean values of the absolute value of each coefficient of the original descriptors across the 10 repetitions models predicting $\Delta G^{\ddagger}$ at the first iteration of the data-driven method. The maximum weight of the generated random variables is represented in gray and establish the threshold represented as a horizontal red line. The highlighted descriptors over the established threshold are classified as informative.

Appendix B

Chapter 3: Supporting Figures

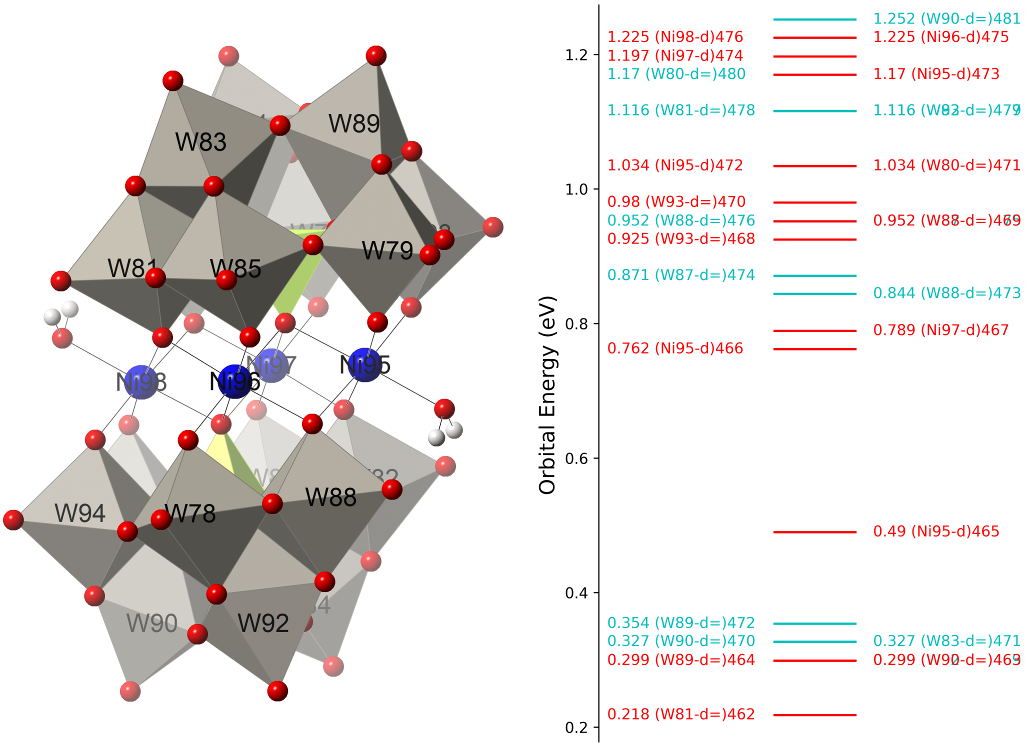


FIGURE B.1: Left: the polyhedral representation of **Ni_4** colouring the octahedric W in gray, the tetrahedral P in yellow, Ni atoms in blue, oxygen atoms in red and H atoms in white. Right: the MOs diagram showing the lowest-energy unoccupied spin-orbitals of the same system, coloured in blue the alpha spin-orbitals and red the beta ones.

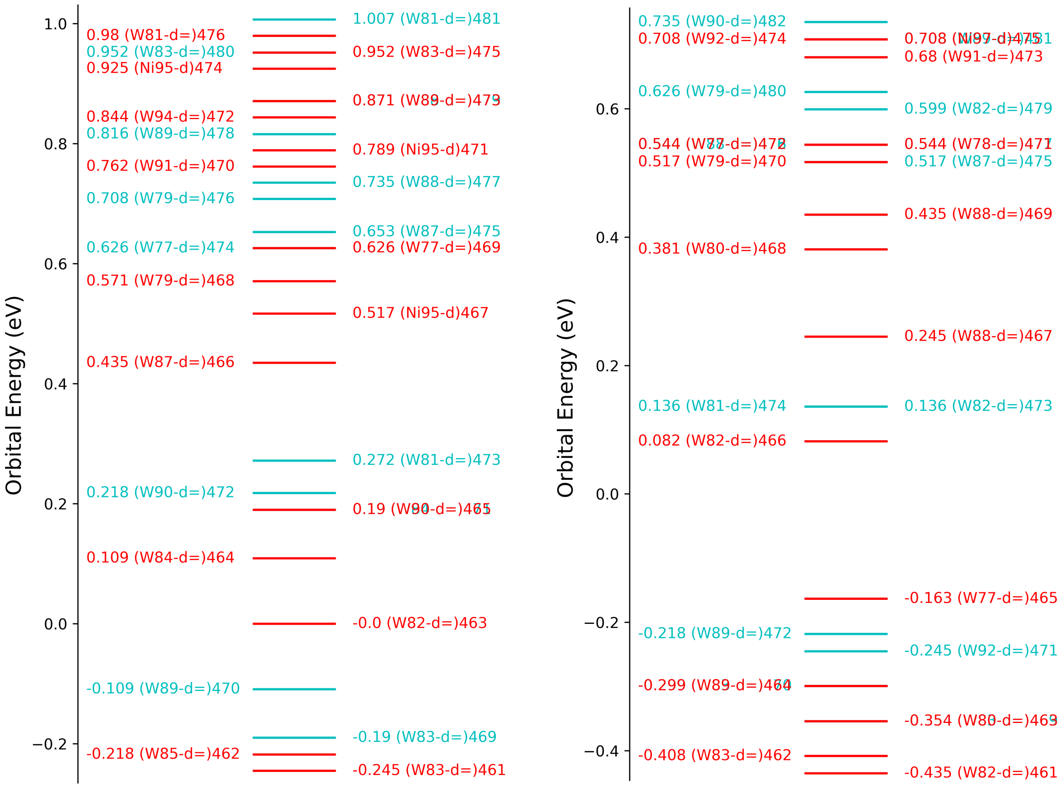


FIGURE B.2: Left: the MOs diagram of the unoccupied spin-orbitals of **Ni**$_4$(1e) system, coloured in blue the alpha spin-orbitals and red the beta ones. Right: the same for **Ni**$_4$(2e).

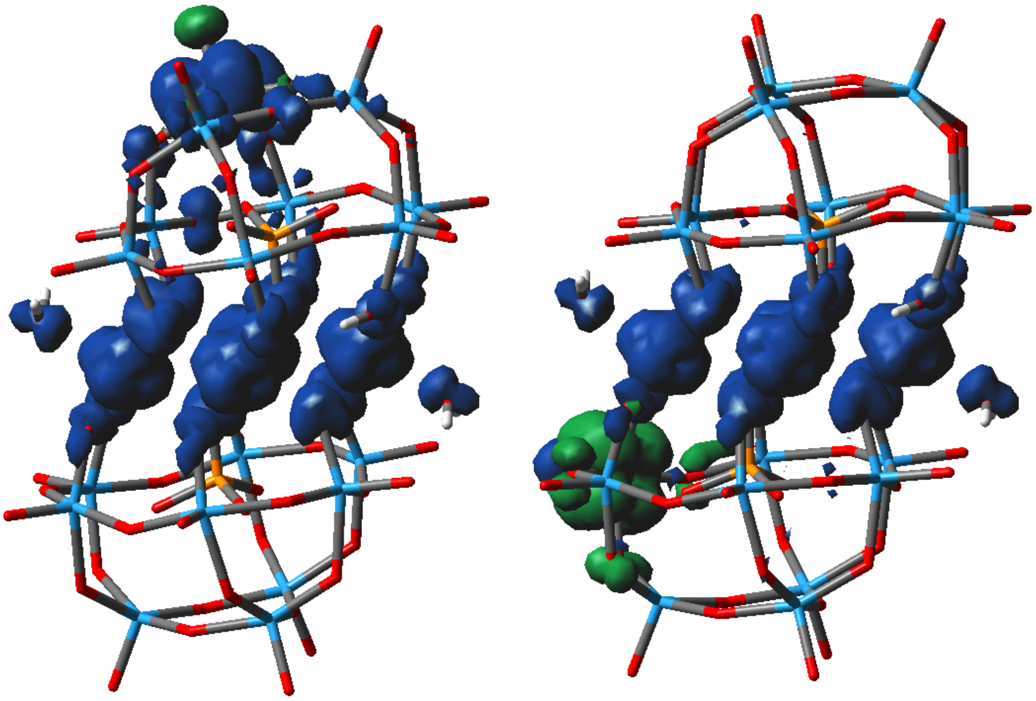
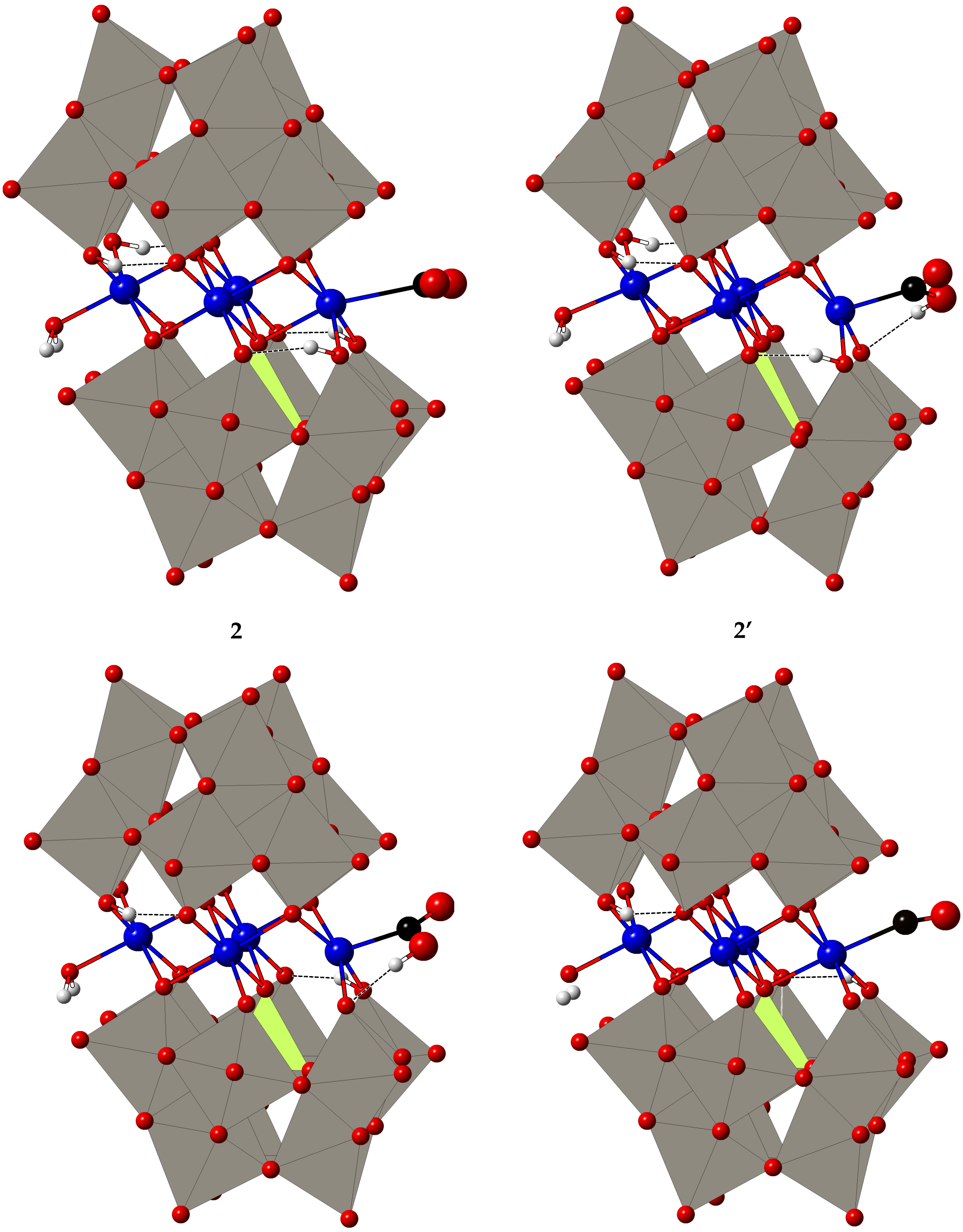


FIGURE B.3: Spin densities of **Ni**$_4$(1e) being at the left the electron in spin up configuration and at the right in spin down, projected on its optimized structure.



2 **2′**

3 **4**

FIGURE B.4: Hybrid Ball-and-Sticks-polyhedral representation of the intermediates of the calculated free-energy profile. Octahedral W in grey, the tetrahedral P in yellow, O atoms in red, Ni atoms in blue and H atoms in white. Hydrogen bonds are represented as dashed lines.

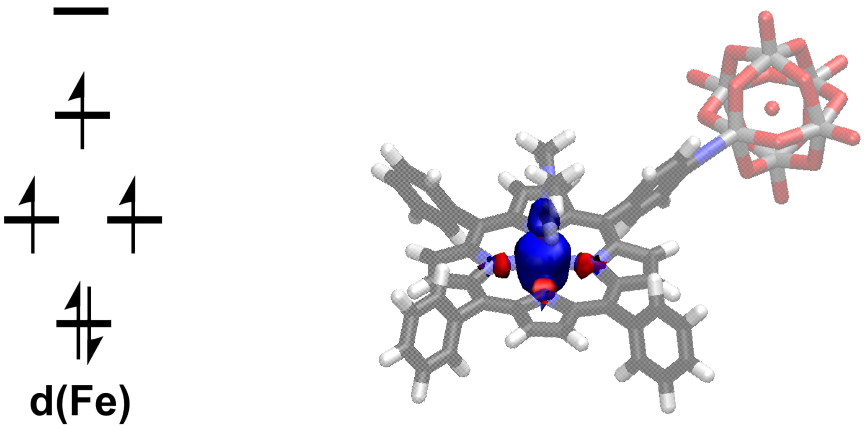


Figure B.5: At the left, the d-type orbital distribution of the Fe centre of the quartet spin state of the fully oxidized DMF coordinated **Mo**$_6$-porph(Fe) according to crystal field theory. At the right the correspondent spin density surface surface colouring the spin up density in blue and the spin down in red, over the translucent structure of the same system colouring C in black, H in white, N in blue, Mo in gray and O in red.

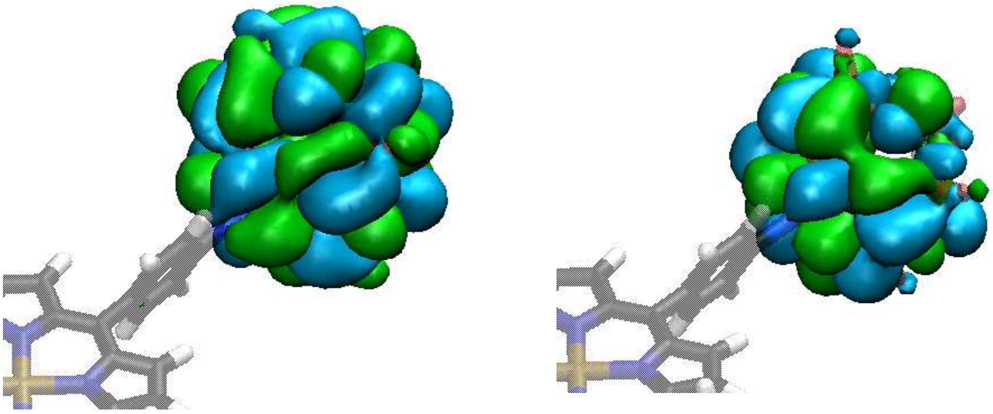


Figure B.6: Left: Representation of the LUMO of the triplet spin state of the one electron-reduced **Mo**$_6$-porph(Fe) species. Right: Representation of the highest energy beta SOMO of the doublet spin state for the two electron-reduced system. For both orbitals the positive phase is coloured in blue cyan and the negative in green, over the translucent structure of the system colouring C in black, H in white, N in blue, Mo in gray, Fe in ochre and O in red.

TABLE B.1: Comparison of relative Gibbs free energies (kcal mol^{-1}) and electronic structures of the calculated spin states for the five electron-reduced **Mo_6**-porph(Fe) system.

Entry	**2S+1**	**ΔG**	**Fe center d-type orbitals diagrams**	**Spin density distribution**
4	7	+18.6