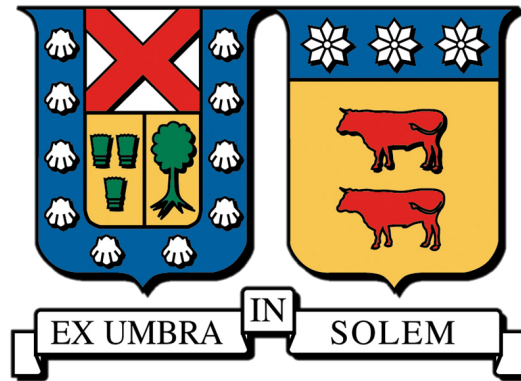


UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA

DEPARTAMENTO DE INGENIERÍA MECÁNICA



**Computational Electrostatics of Adsorbed Glucose and
Bilirubin Oxidases on Electrodes: Modeling and
Characterization of their Preferential Orientation**

CHRISTIAN ALFONSO HERRERA VERGARA

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF CIVIL MECHANICAL ENGINEER

Advisor: PhD. CHRISTOPHER COOPER VILLAGRÁN UTFSM
Co-advisor: PhD. GUIDO RAOS POLIMI

JULY-2026



CONSTANCIA DE VALIDACIÓN Y CONFIDENCIALIDAD DE MONOGRAFÍA A REPOSITORIO ACADÉMICO

1.- IDENTIFICACIÓN DEL TRABAJO ACADÉMICO

Tipo de monografía (marcar una opción): ☒ Memoria o trabajo de título ☐ Tesis de Postgrado

Título del trabajo: Computational Electrostatics of Adsorbed Glucose and Bilirubin Oxidases on Electrodes: Modeling and Characterization of their Preferential Orientation

Nombre del candidato(a): Christian Alfonso Herrera Vergara

Carrera / Grado: Ingeniería Civil Mecánica

Campus: Casa Central

Departamento: Departamento de Ingeniería Mecánica

2.- VALIDACIÓN DEL PROFESOR GUÍA/DIRECTOR DE TESIS

Yo, Christopher Cooper Villagrán, en mi calidad de profesor(a) guía/director(a) del trabajo académico mencionado anteriormente **DEJO CONSTANCIA** que:

- He revisado esta versión del documento y corresponde a la versión final aprobada del trabajo.
- El trabajo cumple con los requisitos académicos y de formato establecidos por la institución.

3.- EVALUACIÓN DE CONFIDENCIALIDAD POR PROPIEDAD INDUSTRIAL (marcar una opción)

☒ El trabajo **NO contiene** información que amerite confidencialidad y puede ser publicado de inmediato en repositorio con acceso abierto.

☐ El trabajo **CONTIENE** información con potenciales implicancias de propiedad industrial o intelectual y requiere un periodo de confidencialidad (**embargo**) por (**marcar una opción**):

☐ 6 meses ☐ 12 meses ☐ 2 años ☐ 3 años ☐ 5 años ☐ 10 años

Fundamentación de la necesidad de confidencialidad (obligatorio si se solicita embargo):

4.- FIRMAS

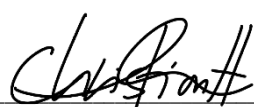
Profesor(a) guía o director(a) de memoria o tesis:

Fecha: 30 de Julio 2026

Firma: 

Estudiante o Candidato(a):

Fecha: 30 de julio 2026

Firma: 

Este formulario debe ser insertado como página 2 de la memoria o tesis, completado y firmado por estudiante y profesor(a) antes de la entrega en portal PRISMA de Biblioteca USM.

Acknowledgments

I am deeply grateful for the opportunities and companionship that have brought me here. I always had a greater affinity for numbers than for words, and science and mathematics were the path that naturally led me to study Civil Mechanical Engineering, given its strong connection to physics and mathematics. I wanted this thesis, which brings that stage to a close, to have an interesting numerical and physical-mathematical topic, and I tried to develop it in the best way possible, seeking to deliver a piece of work I could be proud of.

I am deeply grateful to my thesis advisors. To Professor Christopher Cooper, for his constant support throughout this thesis, for always responding quickly and effectively, and for his enthusiasm and passion for what he does, which is evident from a distance and has been a source of inspiration; he is a great professor and thesis advisor, and in my view one of the best at the entire university. I am also grateful to my co-advisor, Professor Guido Raos, who made this collaboration possible and showed great support throughout the development of this thesis, and who welcomed me from the very first moment in Milan during my exchange at the Politecnico di Milano, without hesitation, receiving me into his office as one more member of the group.

I am also grateful to Universidad Técnica Federico Santa María for its teaching and its learning-by-doing philosophy, as well as to the Department of Mechanical Engineering and its community, for being a department of students who are united and who help one another.

I am grateful to all the friends I made in Civil Mechanical Engineering throughout these years, especially to my closest group of friends from MEC and associated: Gustavo, Benjamín, Marcelo, Ignacio, Alonso, Diego, Juan, Adrián, Pablo, Sebastián, Hugo, Rolando, Enrique, Josué, Gaspar, and Esteban, for supporting one another along the way, for the conversations, the study groups, and the gatherings.

I am grateful to my high school friends who remain close to this day: to my friends from the “Cuarteto”, Diego, Sebastián, and Luis; and to my friends from the “Tomo Ltda.” group, Claudio, Cristóbal, Joaquín, Pedro, Sebastián, and Felipe.

Finally, I am deeply grateful to my parents, Christian and Georgina, for their unconditional and unrestricted support at every step; this thesis, and the completion of this degree, are also their achievement. As well as to my sister Valeria, with whom I have shared every stage of life since before the very first day.

Thank you to all of you, and to everyone else who has influenced my path along the way.

This research was partially supported by the computing infrastructure of the Scientific and Technological Center of Valparaiso, Federico Santa Maria Technical University (ANID PIA/APOYO AFB180002).

Abstract

Oxidoreductase enzymes play a central role in the development of bioelectrochemical devices, whose efficiency depends on how these macromolecules orient when adsorbed onto an electrode. This work models and characterizes the preferential orientation of glucose oxidase (GOx) and bilirubin oxidase (BOx) adsorbed on a flat electrode, evaluating their interaction energy against variations in pH (5, 7, and 9) and applied potential (± 25.7 mV and 0 mV), by solving the linearized Poisson-Boltzmann equation through the boundary element method (BEM) with the software PyGBe. To automate the systematic sampling of orientations and the sequential execution of the simulations required for each condition, an automation code was developed, extending previous work from this same research line. For each condition studied, the Boltzmann-optimal orientation was identified, and the orientational sensitivity to both variables was quantified.

Orientation consistently dominates over pH and, in most cases, over the applied potential, however, when comparing pH and potential directly, it is the potential that dominates in most conditions, a result opposite to that previously reported by similar studies, partly explained by methodological differences identified in those works. The macroscopic current does not necessarily follow the electrostatic preference of each enzyme, although both favor the positive potential electrostatically, BOx reaches its maximum current under negative potential, consistent with the redox role of each enzyme in an enzymatic fuel cell. The local charge distribution near the electrode reinforces, in most conditions, the macroscopic net-charge preference of the protein, in contrast with the stronger local dominance reported in previous work. Finally, electron transfer rates and the electrostatic fingerprints induced on the electrode were determined for both enzymes, extending the scope of previous studies focused solely on interaction energy.

Keywords: bioelectrochemistry, Poisson-Boltzmann equation, boundary element method, protein orientation, glucose oxidase, bilirubin oxidase, PyGBe.

Abstract in Spanish

Las enzimas oxidorreductasas cumplen un rol central en el desarrollo de dispositivos bioelectroquímicos, cuya eficiencia depende de la orientación con que estas macromoléculas se adsorben sobre un electrodo. Este trabajo modela y caracteriza la orientación preferencial de la glucosa oxidasa (GOx) y la bilirrubina oxidasa (BOx) adsorbidas sobre un electrodo plano, evaluando su energía de interacción frente a variaciones de pH (5, 7 y 9) y potencial aplicado (± 25.7 mV y 0 mV), mediante la resolución de la ecuación de Poisson-Boltzmann linealizada a través del método de elementos de frontera (BEM) con el software PyGBe. Para automatizar el muestreo sistemático de orientaciones y la ejecución secuencial de las simulaciones requeridas por cada condición, se desarrolló un código de automatización, extendiendo trabajo previo de esta misma línea de investigación. Para cada condición estudiada se identificó la orientación Boltzmann-óptima y se cuantificó la sensibilidad orientacional frente a ambas variables.

La orientación domina consistentemente sobre el pH y sobre el potencial aplicado en la mayoría de los casos, sin embargo, al comparar directamente pH y potencial, es el potencial quien domina en la mayoría de las condiciones, un resultado opuesto al reportado previamente por estudios similares, explicado en parte por asimetrías metodológicas identificadas en dichos trabajos. La corriente macroscópica no sigue necesariamente la preferencia electrostática de cada enzima, pese a que ambas favorecen electrostáticamente el potencial positivo, BOx alcanza su corriente máxima bajo potencial negativo, consistente con el rol redox de cada enzima en una celda de combustible enzimática. La distribución local de carga cerca del electrodo refuerza, en la mayoría de las condiciones, la preferencia macroscópica de carga neta de la proteína, en contraste con la mayor dominancia local reportada en trabajos anteriores. Finalmente, se determinaron las tasas de transferencia electrónica y las huellas electrostáticas inducidas sobre el electrodo para ambas enzimas, extendiendo el alcance de estudios previos centrados únicamente en la energía de interacción.

Palabras clave: bioelectroquímica, ecuación de Poisson-Boltzmann, método de elementos de frontera, orientación proteica, glucosa oxidasa, bilirrubina oxidasa, PyGBe.

Contents

Acknowledgments	ii
Abstract	iv
Abstract in Spanish	v
1 Introduction	1
1.1 Objectives	2
2 Theoretical Framework	3
2.1 Biocatalysis	3
2.1.1 Structure	4
2.1.2 Mechanism	5
2.1.3 Kinetics	8
2.1.4 Environmental Factors	10
2.1.5 Classification of Enzymes and Oxidoreductases	11
2.2 Bioelectrochemistry	13
2.2.1 Electron Transfer with Electrodes	13
2.2.2 Applications of Bioelectrochemical Systems	16
2.3 Implicit Solvent Model	18
2.3.1 Electrostatics in a Continuous Dielectric Medium	20
2.3.2 Poisson-Boltzmann Equation	21
2.4 Implicit Solvent Model Applied to the Protein-Surface System	23
2.4.1 Single-Surface System: Isolated Protein in the Solvent	24
2.4.2 Protein-Surface System	26
2.4.3 Internal Cavities and the Stern Layer	26
2.4.4 Energy Calculation in the Implicit Solvent Model	28
2.5 Boundary Element Method (BEM)	31
2.5.1 General Aspects of the Method	31

2.5.2	Mathematical Origin and Derivation of the Equation	32
2.5.3	Application of the BEM to the Implicit Solvent Model	36
2.5.4	Discretization of the Boundary Integral Equations	40
2.6	Statistical Thermodynamics and Electron Transfer Rate	43
2.6.1	Boltzmann Distribution over the Configuration Space	43
2.6.2	Marcus Theory for Electron Transfer	44
2.6.3	Expected Value of the Electron Transfer Rate	46
2.6.4	Expected Value of an Energetic Observable	46
2.6.5	Spearman Rank Correlation	47
2.7	Richardson Extrapolation	47
2.8	Software PyGBe	48
2.9	State of the Art: Prior Work Directly Preceding This Thesis	49
3	Methods	52
3.1	Definition of the Simulated System	52
3.1.1	Studied Enzymes	52
3.1.2	Electrode Geometry and Orientation Sampling	54
3.1.3	Evaluated Electrochemical Conditions	58
3.2	Mesh Convergence Study	58
3.3	Simulation Parameters and Input Configuration	59
3.4	Computational Environment	63
3.5	Computational Cost	64
3.6	Computational Code	66
3.6.1	Script Description	66
3.6.2	Computational Workflow	67
3.7	PyGBe Bug: Handling of Cavities in CPU Mode	76
4	Results and Discussion	78
4.1	Influence of the Applied Potential	78
4.2	Influence of pH	81
4.3	Influence of Orientation	85
4.3.1	Summary Table of Optimal Orientations and Ranking Consistency	86
4.3.2	Electron Transfer Kinetics	89
4.3.3	Orientational Landscape of Energy, Probability, and Weighted Current	91
4.3.4	Electrostatic Fingerprint Induced on the Electrode	101

5	Conclusions and Recommendations	108
5.1	Fulfillment of the Objectives	108
5.2	Main Findings	109
5.3	Recommendations	111
	Bibliography	113
	Appendix A Supplementary Tables	I
	Appendix B Supplementary Figures	V

List of Tables

3.1	Structural characteristics of the studied enzymes. The net charge and the electron transfer atom are explained in the text. Approximate dimensions were taken from the atomic coordinates of the PQR file [1].	54
3.2	Number of azimuthal rotation positions N_φ for each tilt value θ	57
3.3	Mesh convergence parameters and Richardson extrapolation results for the solvation energy.	59
3.4	Hardware specifications of the simulation environments (specifications per node in CCTVal).	63
3.5	Versions of PyGBe dependencies in each simulation environment.	64
3.6	Average computation time per orientation and average number of GMRES iterations, for each of the conditions studied. All runs executed on CCTVal.	65
3.7	Scalability of PyGBe with the number of CPU cores, compared to a single GPU run, for a representative case in CCTVal.	66
4.1	Range of variation of $\langle \Delta E_{\text{bind}} \rangle$ induced by the applied potential (at fixed pH) and by pH (at fixed potential), for both enzymes.	82

4.2	Thermodynamic, kinetic, and electrostatic results for Protein , bilirubin oxidase (6IQZ) and glucose oxidase (3QVP) across the eighteen $\text{pH} \times \Phi_e$ conditions studied. θ/φ_B and P_B are the angle (tilt/rotation) and the Boltzmann probability of the most probable orientation; $\Delta E_{\text{bind}}^{\text{min}}$ is its net adsorption energy. Dist_Z is the donor - acceptor distance evaluated at the Boltzmann optimal orientation: Cu T1 - electrode for BOx, N5(FAD) - electrode for GOx. $\mathbf{AA}_{\text{min}}$ is the closest residue to the electrode surface; Q10 is the summed net charge of the ten closest residues. ΔQ is the electrostatic charge perturbation induced by the protein on the electrode surface. $\langle \Delta E_{\text{bind}} \rangle$ is the expected value of the macroscopic behavior of the interaction energy. θ/φ_W and Contrib.% indicate the orientation and the percentage contribution of the dominant term in the Boltzmann-weighted sum of the electron transfer. Match indicates whether the weighted optimal orientation coincides with the Boltzmann optimal (“Boltzmann”) or with the purely geometric optimal of k_{ET} (“Raw k_{ET} ”). I_{total} is the total macroscopic current.	87
4.3	Orientation of highest and lowest electron transfer rate constant k_{ET} , for both enzymes.	91
4.4	Charge induced on the isolated electrode (without protein), $Q_{\text{electrodo}}$, for each applied potential.	101
A.1	Absolute free energy components of the isolated protein, for each protein and pH. .	I
A.2	Absolute free energy components of the isolated electrode, for each applied potential.	I
A.3	Range of variation in interaction energy ΔE_{bind} induced by orientation, potential, and pH, for each of the 18 conditions studied (all values in kcal/mol). The last column indicates which of the two variables (potential, pH, or both) the orientation range exceeds.	II
A.4	Three orientations with the highest and lowest electron transfer rate constant k_{ET} for both enzymes.	II
A.5	Optimal orientation, Boltzmann probability, and expected value $\langle \Delta E_{\text{solv}} \rangle^*$, calculated using a Boltzmann distribution defined exclusively by the solvation energy ΔE_{solv} (without the surface term), for the 18 conditions studied. The asterisk distinguishes this quantity from $\langle \Delta E_{\text{solv}} \rangle$ (Section 4.1), where the Boltzmann weights come from the complete binding energy; this table evaluates the robustness of the orientational ranking against the inclusion of the surface term. The last column reports the Spearman correlation by Equation (2.91), between the ranking of the 406 orientations induced by ΔE_{solv} and by ΔE_{bind} (Table 4.2).	III
A.6	Ten residues closest to the electrode, ranked by distance, for each of the nine conditions of BOx (6IQZ). Each cell shows residue name, number, and net charge in parentheses.	IV

A.7	Ten residues closest to the electrode, ranked by distance, for each of the nine conditions of GOx (3QVP). Each cell shows residue name, number, and net charge in parentheses.	IV
-----	--	----

List of Figures

2.1	Free enthalpy reaction coordinate diagram for an enzyme reaction [4].	9
2.2	Schematic illustration of how enzyme orientation (A) and immobilization linker length (B) affect the distance d between the electrode and the redox cofactor, and consequently the interfacial electron transfer rate. The enzyme is represented as a blue disk, the active site as a red triangle, and the primary electron acceptor as a orange circle [11].	14
2.3	Major applications of enzymatic electrochemistry: (a) biosensors, (b) renewable energy by enzymatic fuel cells, and (c) electrosynthesis [13].	16
2.4	Representation of a protein immersed in a solvent [16].	25
2.5	Surface definitions used to delimit the protein [16].	26
2.6	Representation of a protein with a Stern layer and two cavities filled with solvent. The gray regions contain salt, the white regions do not [16].	28
2.7	Auxiliary surface Γ' with hemisphere Γ'_{hem} to treat the singularity [16].	34
2.8	Discretization of a molecular surface into panels. The point charges q_k represent the interior charge distribution of the protein; Γ_i is the panel containing the collocation point and Γ_j the panel being integrated [16].	40
3.1	Triangulated mesh of the solvent-excluded surface (SES) of (a) Bilirubin Oxidase (6IQZ) and (b) Glucose Oxidase (3QVP) in their reference orientation ($\theta = 0^\circ$, $\varphi = 0^\circ$), 3 Å away from the electrode, modeled as a $100 \times 100 \times 10$ Å plate. The surface electrostatic potential is shown using a color scale (blue: negative; red: positive), calculated at pH 9.	55
3.2	Definition of the orientation angles of the protein relative to the electrode: tilt angle θ and azimuthal rotation angle φ . In this work, the reference axis for the initial orientation is the minimum moment of inertia axis of the protein, not the dipole moment. Adapted from Cooper [16].	56

3.3	Comparison of orientation sampling schemes on the configuration sphere. (A) Uniform geometric sampling in (θ, φ) ; the points cluster at the poles, overrepresenting those orientations in the statistical averaging. (B) Uniform solid angle sampling; the point density is approximately constant over the whole sphere, assigning equal statistical weight to each region of the orientation space.	57
3.4	Example of the hierarchical directory structure for a single simulation case. Directory names and file identifiers vary across scenarios, but the overall layout remains the same.	61
3.5	NanoShaper configuration file for the triangulation of the SES surface.	62
3.6	Global numerical parameters of the PyGBe solver (<code>.param</code>).	62
3.7	Dielectric topology and boundary conditions of the molecular system (<code>.config</code>). . .	63
3.8	Algorithmic flowchart of the computational pipeline developed for protein–electrode interaction energy simulations. The system is organized into operational contexts: manual pre-processing (Phase 0), cluster entry orchestration (Phase 1), automated GPU-accelerated BEM simulation pipeline (Phases 2–8), and local statistical post-processing (Phases 7–9), as detailed in Section 3.6.2.	75
4.1	Range of interaction energy ΔE_{bind} across the 406 sampled orientations, for each pH and applied potential Φ_e . Each floating bar spans from the Boltzmann-optimal orientation (the most favorable, lower end) to the least favorable one (upper end) found within the sampled configuration space.	79
4.2	Boltzmann-weighted average interaction energy $\langle \Delta E_{\text{bind}} \rangle$ as a function of the applied potential Φ_e , for each pH and both enzymes.	80
4.3	Boltzmann-weighted average interaction energy $\langle \Delta E_{\text{bind}} \rangle$ as a function of pH, for each applied potential Φ_e and both enzymes.	81
4.4	Electron transfer rate constant k_{ET} (logarithmic scale) (a.u.) as a function of orientation (θ, φ) , for the 406 sampled orientations. The red dot marks the orientation of maximum k_{ET} (minimum donor - acceptor distance). This map is identical for the nine pH and potential conditions of each protein.	90
4.5	Thermodynamic and kinetic maps for Bilirubin Oxidase (BOx) at pH 9, for the three electrode potentials studied (+25.7 mV, 0 mV, -25.7 mV). Columns: net interaction energy (ΔE_{bind}), Boltzmann probability, and weighted electron transfer rate. The axes represent the relative orientation of the protein on the electrode: tilt angle (θ) and rotation angle (φ) , with the protein electrode distance held fixed.	94

4.6	Thermodynamic and kinetic maps for Glucose Oxidase (GOx) at pH 9, for the three electrode potentials studied (+25.7 mV, 0 mV, -25.7 mV). Columns: net interaction energy (ΔE_{bind}), Boltzmann probability, and weighted electron transfer rate. The axes represent the relative orientation of the protein on the electrode: tilt angle (θ) and rotation angle (φ), with the protein electrode distance held fixed.	95
4.7	Total macroscopic current I_{total} (logarithmic scale) for the 18 conditions studied, grouped by protein and pH, with the three applied potentials shown in different colors.	98
4.8	Induced surface charge density perturbation ($\Delta\sigma$) for 6IQZ (left) and 3QVP (right) at pH 9 and the three electrode potentials studied (+25.7, 0, -25.7 mV), for the highest-probability orientation of each condition. The 10 closest residues to the sensor are labeled by name and residue number.	102
4.9	Induced surface charge density perturbation ($\Delta\sigma$) for 6IQZ (left) and 3QVP (right) at pH 7 and the three electrode potentials studied (+25.7, 0, -25.7 mV), for the highest-probability orientation of each condition. The 10 closest residues to the sensor are labeled by name and residue number.	103
4.10	Induced surface charge density perturbation ($\Delta\sigma$) for 6IQZ (left) and 3QVP (right) at pH 5 and the three electrode potentials studied (+25.7, 0, -25.7 mV), for the highest-probability orientation of each condition. The 10 closest residues to the sensor are labeled by name and residue number.	104
B.1	Solvation contribution to the expected value of the interaction energy, $\langle\Delta E_{\text{solv}}\rangle$, as a function of the applied potential Φ_e , for each pH and both enzymes.	V
B.2	Surface contribution to the expected value of the interaction energy, $\langle\Delta E_{\text{surf}}\rangle$, as a function of the applied potential Φ_e , for each pH and both enzymes.	V
B.3	Solvation contribution to the expected value of the interaction energy, $\langle\Delta E_{\text{solv}}\rangle$, as a function of pH, for each applied potential Φ_e and both enzymes.	VI
B.4	Surface contribution to the expected value of the interaction energy, $\langle\Delta E_{\text{surf}}\rangle$, as a function of pH, for each applied potential Φ_e and both enzymes.	VI
B.5	Thermodynamic and kinetic maps for Bilirubin Oxidase (BOx) at pH 5, for the three electrode potentials studied (+25.7 mV, 0 mV, -25.7 mV). Columns: net interaction energy (ΔE_{bind}), Boltzmann probability, and weighted electron transfer rate. The axes represent the relative orientation of the protein on the electrode: tilt angle (θ) and rotation angle (φ), with the protein electrode distance held fixed.	VII

B.6	Thermodynamic and kinetic maps for Bilirubin Oxidase (BOx) at pH 7, for the three electrode potentials studied (+25.7 mV, 0 mV, -25.7 mV). Columns: net interaction energy (ΔE_{bind}), Boltzmann probability, and weighted electron transfer rate. The axes represent the relative orientation of the protein on the electrode: tilt angle (θ) and rotation angle (φ), with the protein electrode distance held fixed.	VIII
B.7	Thermodynamic and kinetic maps for Glucose Oxidase (GOx) at pH 5, for the three electrode potentials studied (+25.7 mV, 0 mV, -25.7 mV). Columns: net interaction energy (ΔE_{bind}), Boltzmann probability, and weighted electron transfer rate. The axes represent the relative orientation of the protein on the electrode: tilt angle (θ) and rotation angle (φ), with the protein electrode distance held fixed.	IX
B.8	Thermodynamic and kinetic maps for Glucose Oxidase (GOx) at pH 7, for the three electrode potentials studied (+25.7 mV, 0 mV, -25.7 mV). Columns: net interaction energy (ΔE_{bind}), Boltzmann probability, and weighted electron transfer rate. The axes represent the relative orientation of the protein on the electrode: tilt angle (θ) and rotation angle (φ), with the protein electrode distance held fixed.	X
B.9	Distribution of ΔE_{bind} for the 406 sampled orientations across each of the eighteen studied conditions, colored according to the Boltzmann probability of each orientation (logarithmic scale). The label below each distribution indicates the maximum probability achieved under that condition, corresponding to the Boltzmann-optimal orientation reported in Table 4.2. The horizontal dashed line marks $\Delta E_{\text{bind}} = 0$; the vertical line separates BOx (left) from GOx (right).	XI

Chapter 1

Introduction

Oxidoreductase enzymes are central to the development of bioelectrochemical devices such as biosensors and enzymatic fuel cells, where their catalytic activity depends critically on how they adsorb and orient on an electrode surface. This orientation is governed by the electrostatic interaction between the protein and the electrode, which in turn depends on the pH of the solution and the potential applied to the electrode; the preferential, or optimal, orientation of the protein is understood here as the one with the highest Boltzmann probability, that is, the most electrostatically favorable orientation given the applied conditions. Predicting this preferential orientation computationally requires solving the Poisson-Boltzmann equation for the protein-electrode system across a large number of possible orientations and operating conditions, a high-dimensionality electrostatics problem for which explicitly modeling every solvent molecule is too computationally expensive. The implicit solvent model addresses this by replacing the surrounding water and ions with a continuous dielectric medium, while the boundary element method (BEM) solves the resulting equations by discretizing only the molecular surface, offering an efficient solution (Chapter 2). This work adopts this approach, solving the linearized Poisson-Boltzmann equation with the open-source software PyGBe.

Terrani [1] carried out the most complete study of this kind for two specific enzymes, glucose oxidase (GOx) and bilirubin oxidase (BOx), building orientational energy maps and Boltzmann probability distributions as a function of pH and electrode potential using the finite-difference software APBS, and concluding that pH is the dominant factor governing their preferential orientation. A related study by Ruiz et al. [2], applying the boundary element method with PyGBe to a structurally distinct [NiFe] hydrogenase, reached a similar conclusion regarding the dominance of pH. However, the specific comparison supporting this conclusion in Terrani’s work combines the solvation energy, averaged over the full orientational space, with the complete binding energy evaluated only at a single reference orientation; both studies are also obtained under a coarser

electrode model that includes an explicit Stern layer. This thesis extends Terrani’s line of research using PyGBe, reusing his molecular structure files to compare the resulting energies and orientations against those he reports, and evaluating the complete Boltzmann-averaged binding energy, rather than solvation alone, against both pH and potential under the same set of conditions, which, as shown in Chapter 4, leads to a markedly different conclusion regarding their relative influence. To carry out this systematic sampling across the combinations of pH, electrode potential, and spatial orientation studied in this work, an automation code was developed, extending previous work in this same research line [3]. Building on the electron transfer rate model developed by Ruiz et al. [2], this work further connects these orientational preferences to the resulting electron transfer kinetics and macroscopic current for GOx and BOx. It also examines the induced electrostatic fingerprint left by the adsorbed protein across the full range of applied potentials, extending Terrani’s own analysis of this effect, which was limited to the zero-potential condition.

1.1 Objectives

General Objective

The main objective of this work is to do the modeling and characterization of the preferential orientation of the adsorbed glucose oxidase and bilirubin oxidase enzymes. For this, the interaction energy of the system will be evaluated against variations in pH and electrode potential, solving the Poisson-Boltzmann equation through the boundary element method (BEM) using the software PyGBe.

Specific Objectives

- Establish the mathematical formulation of continuum electrostatics through the Poisson-Boltzmann equation for implicit solvent, and its corresponding computational discretization through the boundary element method (BEM), to represent the protein-surface interaction.
- Implement a Python code capable to automate the systematic sampling of spatial orientations and the sequential execution of PyGBe, to carry out the simulations under different conditions of pH and electrode potential.
- Post-process the numerical results from the simulations to determine the net interaction energy with respect to the isolated system, generating energy maps and Boltzmann probability maps.
- Analyze the adsorption stability and quantify the sensitivity of the preferential orientation of each enzyme against variations of pH and electrode potential.

Chapter 2

Theoretical Framework

This chapter contextualizes the background of the computational study conducted in this thesis, and develops the mathematical and physical tools used throughout. Section 2.1 introduces the fundamentals of enzymatic catalysis biochemistry, covering protein structure, catalytic mechanism, kinetics, and the environmental factors that govern enzyme activity. Section 2.2 describes how enzymes are coupled to electrodes in bioelectrochemical systems and the applications that motivate this work. Section 2.3 introduces the implicit solvent model, which replaces explicit water molecules with a continuous dielectric medium, and presents the Poisson-Boltzmann equation that governs electrostatics in the protein-solvent system. Section 2.4 extends this framework to the protein-electrode geometry, including internal cavities and the Stern layer. Section 2.5 then introduces the boundary element method as the numerical strategy to solve the resulting integral equations on the molecular surface. Section 2.6 covers the statistical thermodynamics needed to convert orientation-dependent energies into probability distributions, and the Marcus theory framework used to estimate electron transfer rates from those distributions. Section 2.7 presents Richardson extrapolation, used to estimate the mesh-independent solution from simulations run at different mesh densities. Section 2.8 introduces PyGBe, the numerical software used to solve the boundary element equations throughout this work. Finally, Section 2.9 reviews the computational antecedents that directly precede this thesis, situating the specific contribution of this work within that line of research.

2.1 Biocatalysis

Cellular metabolism includes hundreds of chemical reactions that must be carried out with kinetic precision, high selectivity and under conditions compatible with life. This capacity does not arise spontaneously from thermodynamic principles, but is mediated by highly specialized protein molecules called enzymes. Biocatalysis designates the scientific and technological field that takes

advantage of this molecular machinery to accelerate chemical reactions of industrial or scientific interest, without the catalyst being consumed or permanently altered in the process [1].

The term enzyme derives from the Greek *enzymos* (“leavened”), reflecting that the first recognized biocatalytic processes were those of fermentation. In its modern definition, enzymes are catalytic proteins, that is, macromolecules formed by chains of amino acids that specifically accelerate biochemical reactions by acting on specific substrates, and that operate in organized metabolic sequences responsible for the degradation of nutrients, the transformation of chemical energy and the biosynthesis of macromolecules [1]. Compared to conventional chemical catalysts, enzymatic catalysis presents substantial operational advantages, such as high specificity over the target substrate with minimal formation of byproducts, mild operating conditions of temperature and pressure, and intrinsic stereoselectivity that is particularly valuable in pharmaceutical synthesis [1, 4]. However, the operation of enzymes also brings challenges, their stability is limited under extreme conditions of temperature or pH, and the availability of suitable enzymes for certain reactions remains a restrictive factor [4].

2.1.1 Structure

Enzymes, as proteins in their totality, are built from amino acids, organic compounds that invariably present an amino group ($-\text{NH}_2$) and a carboxyl group ($-\text{COOH}$) attached to the same central carbon, to which a variable side chain called R-group is added. This side chain is what differentiates each of the twenty standard amino acids, determining their polarity, chemical reactivity and charge under physiological conditions [1, 5].

Amino acids are linked through peptide bonds, covalent bonds formed between the carbon of the carboxyl group of one residue and the nitrogen of the amino group of the next, with release of water at each condensation event . These peptide bonds have a partial double bond character that restricts their free rotation and that affects the secondary structure of the proteins . The resulting polypeptide chain is organized in four hierarchical levels of increasing complexity [1, 5]:

- **Primary structure:** The linear and unique sequence of amino acids along the polypeptide chain. From the classical perspective, the primary structure determines the higher hierarchical levels [1, 5].
- **Secondary structure:** Local and recurring folding patterns stabilized by hydrogen bonds between atoms of the peptide backbone. The most common ones are the α -helix, a right-handed spiral of 3.6 residues per turn, and the β -sheet, formed by extended segments arranged in a

parallel or antiparallel manner and stabilized by transverse hydrogen bonds between strands [1, 5].

- **Tertiary structure:** The overall three-dimensional conformation of a complete polypeptide chain, resulting from the interactions between side chains, disulfide bridges, hydrophobic contacts, ionic interactions, hydrogen bonds and van der Waals forces. Among all the forces that stabilize the protein structure, non-covalent interactions are the most important [5]. The polypeptide chains of large proteins always contain two or more autonomous three-dimensional units called subunits or domains. This level determines the geometry of the functional site of the enzyme, so it is the most closely linked to biological activity and proper biochemical functioning [1].
- **Quaternary structure:** A level exclusive to proteins with multiple subunits (oligomeric), it describes the spatial arrangement and interactions between multiple polypeptide chains that constitute the functional protein complex. The subunits are assembled through hydrophobic interactions, hydrogen bonds and van der Waals interactions [5].

2.1.2 Mechanism

The catalytic function of an enzyme resides in a three-dimensional region of its tertiary structure called the active site, a highly specialized region that, although it typically involves less than ten residues and occupies between 10 and 20% of the total volume of the protein, concentrates all the molecular specificity and catalytic power of the enzyme [1]. In this cavity, the substrate is positioned with precision through a network of non-covalent interactions that form the enzyme-substrate complex (ES).

To understand how enzymes accelerate biochemical reactions, it is necessary to start from the concept of activation energy, understood as the energetic barrier that the reactants must overcome to reach the transition state, the instance in which the original bonds break simultaneously with the formation of the new ones. This barrier represents a relevant kinetic bottleneck, since even in thermodynamically spontaneous reactions ($\Delta G < 0$), a high activation barrier can make the reaction practically unobservable. Enzymes act precisely on this point, stabilizing the transition state inside the active site, which reduces the energy barrier and increases the reaction rates by factors typically greater than 10^6 . It should be noted, however, that enzymatic catalysis does not modify the thermodynamic equilibrium of the reaction, but its effect is exclusively kinetic in nature [1].

This stabilization of the transition state is achieved through a combination of transient covalent interactions and, mainly, non-covalent interactions between the substrate and the residues of the active site. The free energy released when the ES complex is formed, called binding energy (ΔG_B), supports both the stabilization of the complex and the reduction of the activation energy. The relatively large size of enzymes responds precisely to this need to accommodate the diversity of functional groups required to accumulate multiple simultaneous interactions that together confer specificity and catalytic power [1]. The four families of non-covalent interactions involved are described below, following the classification used in the literature: electrostatic, van der Waals, hydrogen bonding, and hydrophobic interactions [1, 5, 6, 7]. Non-bonded interactions in force fields, however, are usually considered in only two groups, electrostatic and van der Waals, expressed through a Coulomb potential and a Lennard-Jones potential, respectively [8]:

$$E_{\text{non-bond}} = \sum_{i=1}^N \sum_{j=i+1}^N \left(4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right). \quad (2.1)$$

Hydrogen bonding is typically reproduced without an explicit term, using only this combination of electrostatic and van der Waals models [6, 8], and the hydrophobic effect is described not as an independent force but as a thermodynamic consequence of weaker dispersion interactions between packed nonpolar side chains, combined with an entropic contribution from solvent reorganization [8]. Hydrogen bonding and the hydrophobic effect are therefore physically real phenomena whose underlying physics is already captured by the electrostatic and van der Waals terms, without requiring an additional explicit force.

- **Electrostatic interactions:** Electrostatic interactions are established between the ionizable groups of the substrate, charged in aqueous solution, and the residues of opposite charge in the active site. Under physiological conditions, acidic amino acids acquire negative charge and basic ones acquire positive charge, and some atoms form dipoles by polarization; all these forces obey Coulomb’s law [1, 5].
- **Hydrogen bonds:** They are established when a hydrogen atom, bonded to an electronegative donor atom, interacts with another electronegative acceptor atom. This attraction arises from the partial positive charge that hydrogen acquires due to the difference in electronegativity, giving them a precise geometry that makes them central in the recognition of polar groups of the substrate and in the stabilization of the protein structure [1, 5].
- **Van der Waals interactions:** Short-range dispersion forces, with characteristic lengths of 0.3–0.5 nm and energies of 1–3 kcal/mol, that arise from direct contact between the atoms of the substrate and those of the active site. Although each individual interaction is weak, the

geometric complementarity between both structures allows the accumulation of hundreds of simultaneous contacts, resulting in a significant net energetic contribution [1, 5].

- **Hydrophobic interactions:** They are established between the nonpolar regions of the substrate and the hydrophobic zones of the active site. This behavior can be understood by the tendency of nonpolar molecules to aggregate with each other in aqueous solution, since such association is energetically favorable as it preserves the hydrogen bond network of water, which would be disturbed by the presence of nonpolar surfaces [1].

The geometric and chemical complementarity between the active site and the substrate, which is the basis of the high enzymatic selectivity, has been described through two conceptual models. The lock and key model postulates a rigid and preformed active site, structurally complementary to the substrate, so that only the correct molecule can fit into it. The induced fit model conceives the active site as a dynamic entity, the binding of the substrate induces conformational rearrangements in the enzyme that progressively optimize the specificity and catalytic efficiency [1].

Many enzymes require non-protein components called cofactors to perform their activity. These can play a structural role, stabilizing the conformation necessary for catalysis, or participate directly in the chemistry of the reaction by transferring functional groups, electrons or stabilizing charged intermediates [1]. Among the most common cofactors in oxidoreductases are nicotinamide compounds, NAD(P)(H), required by dehydrogenases; flavin compounds, FMN or FAD, required by oxidases; pyridoxal phosphate, PLP, in transaminases and decarboxylases; and thiamine pyrophosphate, TPP, in carbolicases [4]. According to their nature and behavior in the catalytic cycle, three categories are distinguished:

- **Cofactor:** It is incorporated into the active site and is quantitatively regenerated at the end of each cycle. Pyridoxal phosphate, present in various amino group transfer enzymes, is a representative example [1].
- **Coenzyme:** An organic molecule that acts as a co-catalyst by transferring chemical groups between substrates and unlike the cofactor, is transformed into a different product at the end of each cycle, and must be regenerated through independent secondary enzymatic reactions. It is frequently derived from vitamins [1].
- **Prosthetic group:** A non-protein component strongly bound to the apoenzyme, which remains structurally unchanged throughout the catalysis and forms an integral and permanent part of the holoenzyme. Characteristic examples are transition metal ions such as Fe^{2+} in oxidoreductases, and high-affinity organic cofactors such as FAD in flavoproteins [1].

The combination of the protein portion, called apoenzyme, with its cofactor constitutes the holoenzyme, the catalytically active form of the enzyme. In the absence of the appropriate cofactor, the apoenzyme lacks significant biological activity, which shows the indispensable character of this component [1].

2.1.3 Kinetics

The quantitative characterization of catalytic activity is carried out within the framework of Michaelis-Menten kinetics [1]. To understand its energetic basis, it is necessary to distinguish two independent aspects of a chemical reaction, on one hand thermodynamics, which determines whether the reaction is spontaneous through the variation of free enthalpy ΔG_r , that is, the energy difference between the initial state $E + S$ and the final state $E + P$, and on the other hand kinetics, which determines how fast it occurs. A reaction can be highly spontaneous ($\Delta G_r \ll 0$) and still occur at a negligible rate if there is a high energy barrier between the reactants and the products. The enzyme does not alter ΔG_r , but acts exclusively on that kinetic barrier [4].

This mechanism is illustrated in Figure 2.1. Starting from the ground state $E + S$, the enzyme and the substrate combine to form the ES complex, which constitutes a local minimum of free energy ($\Delta G_b < 0$), the formation of the complex is in itself energetically favorable. From that valley, the system must overcome the catalytic activation barrier ΔG_{kat}^* to reach the transition state and progress toward $E + P$. The key point is that this barrier is measured from the valley of the ES complex, not from the initial state $E + S$. Thus, although ΔG_{kat}^* may seem large in absolute value, the effective barrier that the system must overcome is $\Delta G_{kat}^* - |\Delta G_b|$, which is significantly smaller than the barrier of the uncatalyzed reaction ΔG_{unkat}^* . The net energy saving is [4]:

$$\Delta\Delta G^* = \Delta G_{unkat}^* + \Delta G_b - \Delta G_{kat}^*. \quad (2.2)$$

It is the step from ES to the transition state that limits the rate and determines k_{cat} .

The Michaelis-Menten model describes the reaction rate v as a function of the substrate concentration $[S]$ through two elementary steps, the reversible formation of the ES complex and its irreversible conversion into product with regeneration of the free enzyme [1, 4]:



where k_1 is the rate constant for the forward reaction forming the ES complex, k_{-1} is the rate constant for the reverse dissociation of ES back into free enzyme and substrate, and k_2 is the rate constant for the irreversible conversion of ES into product [1]. Under the steady-state hypothesis

for the ES complex ($d[\text{ES}]/dt \approx 0$) and under conditions of excess substrate relative to the enzyme, the Michaelis-Menten expression is obtained [1, 4]:

$$v = \frac{v_{\max} [\text{S}]}{K_m + [\text{S}]}, \quad (2.4)$$

where $v_{\max}$ is the maximum rate, reached when the enzyme is completely saturated with substrate, and K_m is the Michaelis constant, a measure of the enzyme-substrate affinity defined as the substrate concentration at which $v = v_{\max}/2$. A high value of K_m indicates loose binding between enzyme and substrate, while a low value indicates tight binding [1, 4].

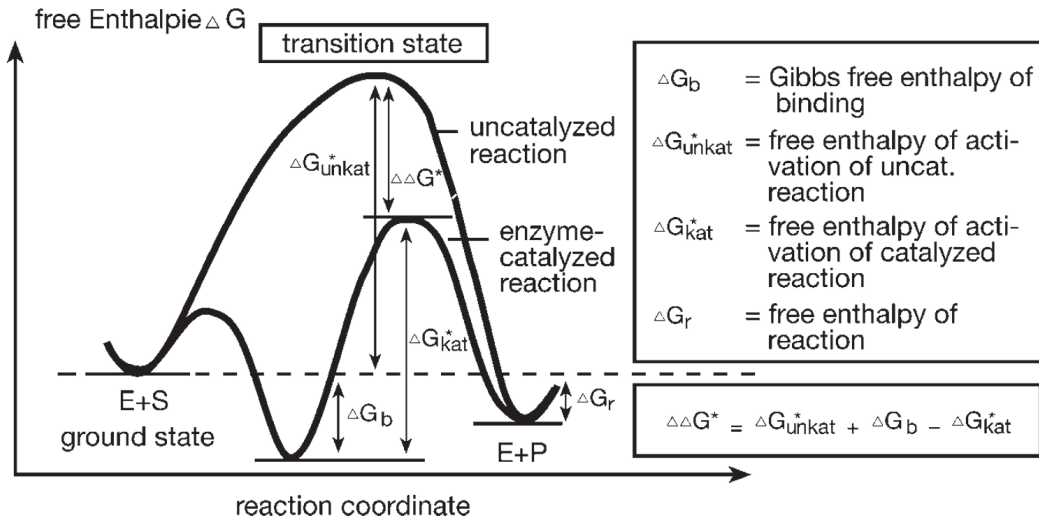


Figure 2.1: Free enthalpy reaction coordinate diagram for an enzyme reaction [4].

The curve v vs $[\text{S}]$ shows two well-differentiated kinetic regimes, at $[\text{S}] \ll K_m$ the rate is first order with respect to $[\text{S}]$, while at $[\text{S}] \gg K_m$ the enzyme operates under saturation and the rate is zero order, converging to $v_{\max}$ independently of the substrate concentration [1, 4]. The turnover number $k_{\text{cat}} = v_{\max}/[\text{E}]_{\text{total}}$ quantifies the number of substrate molecules converted per saturated enzyme molecule per unit time, while the ratio k_{cat}/K_m , called the specificity constant, integrates in a single parameter the affinity and the intrinsic catalytic efficiency, being the indicator of choice for comparing enzymes or evaluating different substrates [1, 4].

It should be noted that, although the parameters $v_{\max}$ and K_m can be determined experimentally for any enzyme, they provide little information about the rates or the chemical nature of the individual steps of the reaction, despite this, steady-state kinetics constitutes the standard framework for comparing and characterizing catalytic efficiency in biochemistry [1].

2.1.4 Environmental Factors

The catalytic activity measured under specific conditions does not reflect only the intrinsic properties of the enzyme, but the balance between multiple environmental variables. Every biocatalyst can be characterized in three fundamental performance dimensions: activity, selectivity and stability; for the great majority of enzymes, these dimensions remain favorable only within limited ranges of temperature (15 - 50 °C) and pH (5 - 10) [4]. The factors of greatest relevance are described below.

- **Temperature:** Increasing temperature raises the frequency of productive collisions between the enzyme and the substrate, increasing the catalytic rate up to an optimum located around 37 °C for most enzymes. However, when a critical threshold is exceeded, generally above 40 °C, the thermal energy begins to destabilize the non-covalent interactions that support the protein architecture, causing structural changes, bond breaking and loss of enzymatic activity, a phenomenon known as denaturation [1]. It should be noted that, if the conditions of the medium are restored, the denatured protein can recover its native conformation through renaturation [5].
- **pH:** The pH of the medium governs the protonation state of the ionizable residues of the protein. These functional groups transition between their charged and neutral forms when crossing their respective pK_a values [1]. These variations directly affect the geometry of the active site, the affinity for the substrate, and the catalytic efficiency [1].
- **Ionic strength:** The concentration of ions in the medium regulates the degree of screening of the electrostatic interactions on the protein surface [2]. Moreover, alterations in ionic strength can disturb the natural structure of the protein and induce its denaturation [5].
- **Inhibitors and activators:** Molecules other than the substrate can bind to the enzyme and modulate its activity in a reversible or irreversible manner. Reversible inhibition is classified into four categories according to the binding site of the inhibitor [1]:
 - *Competitive:* The inhibitor, typically a structural analog of the substrate, competes with it for access to the active site of the enzyme.
 - *Non-competitive:* The inhibitor binds both to the free enzyme and to the ES complex, decreasing the catalytic activity independently of the substrate concentration.
 - *Uncompetitive:* The inhibitor binds exclusively to the ES complex, the resulting conformational change in the ESI complex prevents the catalytic reaction from continuing.
 - *Mixed:* The inhibitor has affinity for both the free enzyme and the ES complex, although with different dissociation constants in each case, combining characteristics of the competitive and non-competitive mechanisms.

Activators, unlike inhibitors, act in the opposite direction, increasing the catalytic activity of the enzyme [1].

- **Substrate concentration:** The availability of substrate defines the kinetic regime under which the enzyme operates, this related to the Michaelis-Menten Equation 2.4, in the limit $[S] \ll K_m$, the reaction rate grows approximately linearly with the substrate concentration, while in the opposite extreme, $[S] \gg K_m$, the enzyme operates under saturation at maximum rate $v_{\max}$ [1].
- **Electrode potential:** In bioelectrochemical systems, the electric potential imposed on the electrode determines the surface charge density and, with it, the intensity and direction of the electrostatic field with which the adsorbed protein interacts. This external parameter exerts a significant influence on the orientational distribution of the enzymes on the electrode surface, favoring the configurations in which the electrostatic interaction is energetically more stable [2, 9].

2.1.5 Classification of Enzymes and Oxidoreductases

Since the first report of the Enzyme Commission in 1961, the International Union of Biochemistry and Molecular Biology (IUBMB) groups enzymes into six classes according to the type of reaction they catalyze. It is important to note that this scheme does not organize enzymes by their amino acid sequence or their three-dimensional structure, but by the nature of the chemical transformation they mediate [4]. Each enzyme is assigned an EC code of four digits, where the first indicates the class, the second the subclass, the third the sub-subclass and the fourth the serial number within it [1, 4]:

- **Oxidoreductases (EC 1):** They catalyze the transfer of electrons between a reduced donor substrate and an oxidized acceptor: $A_{\text{red}} + B_{\text{ox}} \rightleftharpoons A_{\text{ox}} + B_{\text{red}}$. Their subclasses include oxidases, dehydrogenases, peroxidases and reductases. Examples: glucose oxidase, cytochrome c oxidase, lactate dehydrogenase [1, 10].
- **Transferases (EC 2):** They catalyze the transfer of a functional group between molecules: $AB + C \rightarrow A + BC$. They include kinases, transaminases, methyltransferases and glycosyltransferases. Example: hexokinase in glycolysis [1, 10].
- **Hydrolases (EC 3):** They catalyze the breaking of covalent bonds through incorporation of water: $AB + \text{H}_2\text{O} \rightarrow A\text{-OH} + BH$. They include proteases, esterases, glucosidases and phosphatases. Example: trypsin [1, 10].

- **Lyases (EC 4):** They catalyze the breaking or formation of bonds through non-hydrolytic elimination or addition: $AB \rightleftharpoons A + B$. Example: fumarase [1, 10].
- **Isomerases (EC 5):** They catalyze intramolecular rearrangements between isomers: $AB \rightarrow BA$. Example: phosphohexose isomerase in glycolysis [1, 10].
- **Ligases (EC 6):** They catalyze the formation of a covalent bond between two substrates coupled to the consumption of ATP: $A + B + \text{ATP} \rightarrow AB + \text{ADP} + P_i$. Example: glutamine synthetase [1, 10].

Of the six classes, oxidoreductases (EC 1) are the most relevant for this work. With about 25 % of all known proteins, they constitute the most studied group in bioelectrocatalysis [10]. Their ability to mediate electron transfers allows the electrode to take the role of electron donor or acceptor in substitution of the natural substrate of the enzyme, a principle that underlies amperometric biosensors, enzymatic fuel cells and bioelectrosynthesis platforms [10].

Within the oxidoreductases, oxidases use molecular oxygen as the terminal electron acceptor, which makes them especially suitable for devices that operate in physiological environments [10]. A common structural feature in oxidoreductases of bioelectrochemical interest is the presence of a redox cofactor inside the protein matrix, which acts as the active node of the electronic exchange [10]. These cofactors are classified into two categories:

- **Metallic cofactors:** They comprise copper (Cu) centers, heme groups with iron (Fe), iron-sulfur clusters (Fe-S) and molybdenum centers. Multicopper oxidases (MCO) illustrate this category, their T1, T2 and T3 centers, differentiated by magnetic and spectroscopic properties, operate in a coordinated manner to catalyze the four-electron reduction of O_2 to H_2O without release of reactive oxygen species [1, 10].
- **Organic cofactors:** They include FAD (flavin adenine dinucleotide), FMN (flavin mononucleotide) and PQQ (pyrroloquinoline quinone). In flavoproteins, FAD alternates reversibly between its oxidized and reduced form (FADH_2) during the catalytic cycle [1, 10].

In both families, the cofactor is protected inside the protein, which preserves its electronic environment but places a barrier of several angstroms between the active site and the surface of the electrode. This tension between structural integrity and electronic accessibility is the central problem that motivates the development of the bioelectrochemical systems analyzed in this thesis [1, 10].

2.2 Bioelectrochemistry

2.2.1 Electron Transfer with Electrodes

Biological metabolic reactions are based on the catalysis of biomolecules. Enzyme active sites interact with substrates, converting them into specific products. Oxidoreductases transfer electrons through a series of single electron transfers. When an oxidoreductase is coupled to an electrode, create a simplified bioelectrochemical systems through which they can study interfacial mechanisms of electron transfer and develop technologies like biosensors, biofuel cells, and bioelectrosynthetic devices. This process of electron transfer can occur through two different mechanisms: Direct Electron Transfer (DET) and Mediated Electron Transfer (MET)[1, 10].

The characterization of bioelectrochemical systems also requires evaluating an enzyme’s catalytic efficiency when it is absorbed onto an electrode, the current transferred from the electrode to the protein must be considered. The orientation of the protein on the electrode surface is a critical factor in determining how accessible the catalytic center is. The way a protein is oriented on an electrode will define whether the electron and proton transfer necessary to complete the redox reaction will be facilitated or hindered, thereby affecting fundamental electron transfer efficiencies [2]. This electron transfer process is explained by Marcus theory for non-adiabatic systems, which will be developed in detail in Section 2.6.2.

Direct Electron Transfer (DET)

Direct Electron Transfer happens when electrons are exchanged directly between the electrode and the active site of the enzyme, without the need of chemical mediators. It is an ideal process for proteins with exposed redox centers. Its simplicity reduces costs and provides a direct way to study the relation between thermodynamics and enzyme kinetics [1, 10].

However, according to Marcus theory, DET is strictly limited by distance. The exponential decay of the electronic coupling (H_{DA}) demands that the redox cofactor is in extreme proximity to the electrode to allow the quantum tunneling effect. The literature does not converge on a single cutoff distance for this phenomenon, since it depends on the particular structure of each protein; reported values range from 14 Å, identified as the distance required for efficient DET [10], up to 26 Å, a commonly accepted threshold beyond which the transfer rate constant (k_{ET}) is considered negligible [11]. Beyond this threshold, the direct connection becomes unviable, forcing the use of Mediated Electron Transfer (MET).

Given this spatial sensitivity, the efficiency of DET depends drastically on the geometric orientation of the enzyme. As schematized in Figure 2.2, both the way the enzyme is oriented at the electrode surface and the length of the linker used for its immobilization determine how close the redox cofactor can get to the electrode, directly affecting the rate at which electrons are exchanged [11].

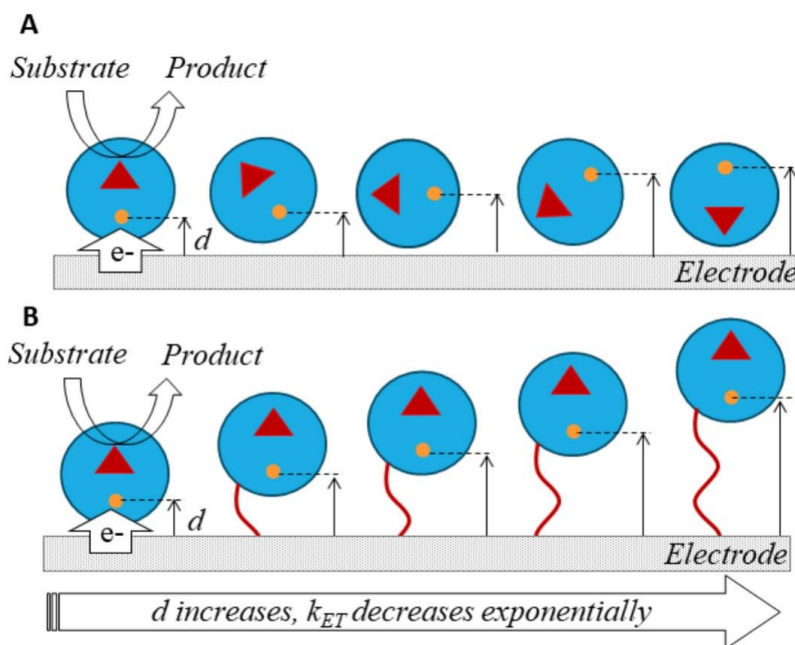


Figure 2.2: Schematic illustration of how enzyme orientation (A) and immobilization linker length (B) affect the distance d between the electrode and the redox cofactor, and consequently the interfacial electron transfer rate. The enzyme is represented as a blue disk, the active site as a red triangle, and the primary electron acceptor as an orange circle [11].

Also, these systems face two big operative limitations: only the first monolayer in direct contact with the electrode gets electrically connected, and the simple adsorption on bare electrodes usually generates slow kinetics that quickly denature the biocatalyst [1, 10, 11].

To overcome these drawbacks, the literature highlights three main immobilization strategies:

- Chemical modification of the electrode: Use of self-assembled monolayers or chemical treatments to mimic the enzyme's physiological partner or substrate, promoting interactions that force a thermodynamically favorable orientation [11].
- Molecular engineering of the protein: Structural modification, like addition of histidine tags, to create a specific anchor to make sure that the redox center points to the surface to minimize the tunneling distance [1, 10].

- Morphological modification (Nanomaterials): Use of tridimensional architectures like conducting hydrogels, nanoparticles, or nanotubes. These structures manage to wrap the proteins and connect them electrically regardless of their initial orientation by massively increasing the surface area [10, 11].

Mediated Electron Transfer (MET)

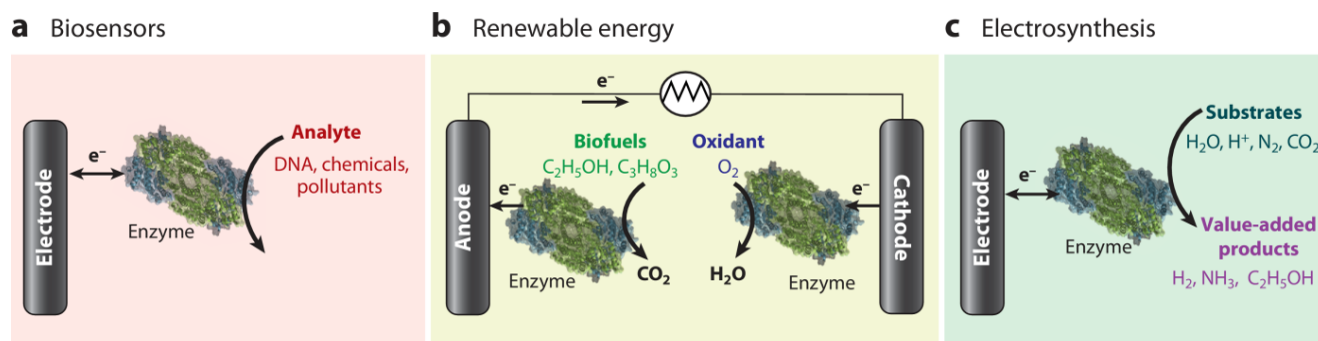
For those biocatalysts whose active center is deeply buried in the protein structure, the direct electrical connection becomes unviable due to the drastic attenuation of the tunneling effect; this is reported to occur at distances greater than 20 Å [1] or beyond 21 Å [12], depending on the source. In these scenarios, the system requires Mediated Electron Transfer (MET) [1]. This mechanism introduces an electroactive auxiliary species, or redox mediator, that acts as a transport bridge. Under this approach, Marcus theory governs the electron exchange between the enzyme and the mediator; therefore, the efficiency depends on this molecule being able to diffuse and penetrate to the active site, a process governed by its size and its hydrophobic or electrostatic interactions with the protein [11]. Once there, the mediator exchanges electrons, transitioning between its oxidized (M_{ox}) and reduced (M_{red}) states. Then, the reduced species diffuses towards the electrode to discharge the electrons, regenerating into its M_{ox} state to restart the cycle. So that this electrochemical recycling does not become the bottleneck of the catalysis, the mediator must possess a high interfacial electron transfer rate (k_{ET}) with the electrode [1, 11].

Unlike DET, the operation of this approach does not depend on the immobilization or orientation of the protein on the surface. By taking advantage of the diffusion of the mediator, enzymes can operate freely in solution, which allows for a massive increase in the amount of biocatalyst inside the system. For this cycle to be efficient, the chosen species must present high solubility, stability in both redox states, and ensure a favorable thermodynamic cascade. This means strictly positioning its potential at an intermediate point between the potential of the electrode and the enzyme, guaranteeing the appropriate driving force depending on whether it is an oxidation or reduction reaction [1, 11].

Despite its geometric flexibility, MET introduces important operational challenges. The dependence on diffusive molecules adds chemical impurities that require later separation stages. Also, at an analytical level, the electrochemical signals of the mediator often overlap and mask the real response of the enzyme [1, 10]. As a modern alternative to avoid these complications, the literature highlights the design of redox polymers. These three-dimensional matrices trap the enzyme on the electrode and conduct the electrons through their own structure, eliminating the need for free mediators in the electrolyte [10].

2.2.2 Applications of Bioelectrochemical Systems

The integration of biological catalysts with electrodes has positioned bioelectrochemical systems as a tool of great interest for the development of sustainable technologies and green chemistry. These systems stand out for allowing a direct and flexible conversion between chemical and electrical energy, operating under mild reaction conditions and with a high selectivity. Depending on their main purpose, the applications of bioelectrocatalysis are usually divided into two big branches: those oriented to analysis, like biosensors, and those focused on the generation of products or energy, where bioelectrosynthesis and enzymatic fuel cells stand out [1, 10]. As summarized in Figure 2.3.



Lee YS, et al. 2021
Annu. Rev. Phys. Chem. 72:467–88

Figure 2.3: Major applications of enzymatic electrochemistry: (a) biosensors, (b) renewable energy by enzymatic fuel cells, and (c) electrosynthesis [13].

Biosensors

Amperometric biosensors operate as analytical platforms that couple a biological recognition element to an electrochemical transducer. In these devices, the catalytic interaction between the enzyme and its specific substrate generates an electrical current that results directly proportional to the concentration of the objective analyte. The most representative commercial example is the glucose biosensor, which constitutes the base of the glucometers used at a global level for the control of diabetes. In them, one or several enzymes catalyze the oxidation of glucose to emit a quantifiable electrical signal. While this technology has already proven itself in medical diagnosis, its potential extends towards the quality control in the food and beverage industries, as well as the continuous monitoring of environmental parameters [1, 10].

Bioelectrosynthesis

Driven by the need to synthesize chemical products of high added value, bioelectrosynthesis uses electrical energy as a clean source of electrons. This supply provides the necessary reducing

equivalents to traction chemical reactions mediated by enzymes. The approach has demonstrated to be successful in applications like the production of hydrogen (H_2) in aqueous solutions through hydrogenases, or the reduction of carbon dioxide (CO_2) towards derived fuels, such as carbon monoxide (CO) or formate. By integrating these reductive pathways with fuel cells, the possibility to co-generate electricity and chemical compounds simultaneously arises. However, the practical design must deal with the limitation that these enzymes usually require to operate at relatively low reduction potentials [1].

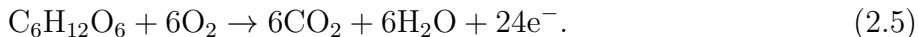
Enzymatic Fuel Cells (EFCs)

Enzymatic fuel cells represent one of the most promising alternatives for the generation of clean and sustainable energy. These electrochemical systems take advantage of the biological machinery to directly convert the chemical energy stored in a biofuel into electricity. Structurally, an EFC consists of an anode, where the enzymatic oxidation of the substrate takes place, and a cathode, in charge of reducing an oxidizing agent that, in general, is molecular oxygen (O_2). During its operation, the electrons released in the anodic compartment travel towards the cathode through an external electrical circuit to produce a useful current, while the charge balance is maintained through the internal migration of ions in the electrolyte [1].

Unlike inorganic cells catalyzed by precious metals, which are usually restricted to the oxidation of hydrogen or methanol, EFCs stand out for their capacity to operate with an extensive biodiversity of renewable sources, opening the spectrum to the use of sugars derived from biomass, ethanol or glycerol [1, 10].

Inside the development of this technology, the cells fed by glucose result particularly attractive for the biomedical sector. The high theoretical energy density of this monosaccharide (around 4125 [Wh/L]) and its constant presence in physiological fluids turn it into an ideal candidate to give autonomy to implantable devices, such as pacemakers and drug delivery systems. For the electrochemical exploitation of glucose, there are two oxidation routes with strong implications in the thermodynamics and the design of the device [1]:

1. **Total oxidation:** It emulates the natural process of cellular respiration, achieving the total extraction of 24 electrons per molecule to convert glucose into carbon dioxide and water:



Under ideal thermodynamic conditions (100% coulombic efficiency and without ohmic losses or overpotentials), this reaction generates a maximum reversible voltage of 1.24 V at room

temperature (298 K). Nonetheless, its implementation demands to orchestrate an extensive and complex cascade of multiple enzymes operating simultaneously, which supposes a critical obstacle for the practical design of the cell [1].

2. **Partial oxidation:** As a considerably more viable alternative, the partial oxidation of glucose to gluconolactone ($\text{C}_6\text{H}_{10}\text{O}_6$) is proposed:



Even though this pathway releases only 2 electrons per molecule and limits the maximum voltage to 1.18 V, it presents determinant operational advantages. Its architecture is much more simple by depending on a single, highly stable and efficient enzyme. In this area, Glucose Oxidase (GOx) is the predominant option due to its extreme specificity. The catalytic activity of this flavoprotein is mediated by the FAD (flavin adenine dinucleotide) cofactor, which is found embedded in its structure [1, 10].

On the other hand, to consolidate a totally enzymatic architecture, the cathodic reaction is usually coupled to multi-copper oxidases (MCOs), like laccase or Bilirubin Oxidase (BOx). These specialized enzymes use their metallic copper centers to catalyze the direct reduction of molecular oxygen to water, consuming four electrons without releasing undesired reactive oxygen species (ROS):



The assembly of a system with GOx in the anode and BOx in the cathode perfectly illustrates the synergy of bioelectrochemical devices under mild operative conditions. However, for this technology to scale towards commercially competitive power densities, the main scientific challenge lies in optimizing the efficient transfer of electrons. The active centers responsible for the catalysis (the FAD in GOx and the copper ions in BOx) are found deeply buried in the insulation of the protein matrix, making difficult the direct electrical contact with the surface of the electrode [1, 10].

2.3 Implicit Solvent Model

In biomolecular simulations, the aqueous solvent plays a determining role in electrostatics and, consequently, in protein functionality. Explicitly modeling each water molecule, as done in molecular dynamics simulations, becomes computationally expensive for large systems, which are frequently required to describe bulk solution. The implicit solvent model proposes as alternative to treat the solvent as a dielectric continuum characterized by its relative permittivity, eliminating the need

to track each water molecule individually. This drastically reduces the degrees of freedom of the problem, replacing explicit molecular trajectories with the solution of partial differential equations over the solute and solvent domains.

Integration of Solvent Degrees of Freedom

The formal basis of this approximation can be derived from statistical mechanics [1, 14]. Consider a solute with coordinates $\mathbf{X} \equiv \{x_1, x_2, \dots\}$ immersed in a solvent with coordinates $\mathbf{Y} \equiv \{y_1, y_2, \dots\}$, in contact with a thermal bath at temperature T . The state of the system is described by the joint probability distribution $P(\mathbf{X}, \mathbf{Y})$. For any generic quantity $Q(\mathbf{X})$ that depends only on the solute configuration, its expected value is:

$$\langle Q(\mathbf{X}) \rangle = \int Q(\mathbf{X}) P(\mathbf{X}, \mathbf{Y}) d\mathbf{X} d\mathbf{Y}. \quad (2.8)$$

Since Q does not depend on $\mathbf{Y}$, it is possible to integrate out the solvent coordinates and define a reduced distribution that depends exclusively on the solute:

$$\bar{P}(\mathbf{X}) = \int P(\mathbf{X}, \mathbf{Y}) d\mathbf{Y}. \quad (2.9)$$

Substituting Equation (2.9) into Equation (2.8), the expected value simplifies to:

$$\langle Q(\mathbf{X}) \rangle = \int Q(\mathbf{X}) \bar{P}(\mathbf{X}) d\mathbf{X}. \quad (2.10)$$

For a system in thermal equilibrium at temperature T , the reduced distribution takes the Boltzmann form:

$$\bar{P}(\mathbf{X}) \propto e^{\frac{-W(\mathbf{X})}{k_B T}}, \quad (2.11)$$

where $W(\mathbf{X})$ corresponds to the potential of mean force, which represents the reversible work done by the average forces and should not be confused with the average potential energy. This effective potential implicitly contains all the solvent effects on the solute, so that the generic quantity is expressed solely in terms of the solute coordinates:

$$\langle Q(\mathbf{X}) \rangle = \int Q(\mathbf{X}) e^{\frac{-W(\mathbf{X})}{k_B T}} d\mathbf{X}. \quad (2.12)$$

In this way, by marginalizing over the solvent degrees of freedom, all relevant information about its influence on the equilibrium properties of the system is retained, considerably reducing the computational cost compared to explicit approaches.

2.3.1 Electrostatics in a Continuous Dielectric Medium

The starting point of macroscopic electrostatic theory is Gauss's law, which relates the electric field $\mathbf{E}$ to a volumetric charge distribution ρ in vacuum:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0}, \quad (2.13)$$

where ε_0 is the permittivity of vacuum. Since in electrostatics magnetic effects are not considered, the field is irrotational, which guarantees the existence of a scalar potential ϕ such that $\mathbf{E} = -\nabla\phi$. Substituting into Equation (2.13) gives the Poisson equation in vacuum:

$$\nabla^2\phi = -\frac{\rho}{\varepsilon_0}. \quad (2.14)$$

The general solution of this equation for an arbitrary charge distribution can be written in integral form as:

$$\phi(\mathbf{r}) = \frac{1}{4\pi\varepsilon_0} \int_{\Omega} \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\Omega'. \quad (2.15)$$

This formulation, however, is valid only in vacuum. When the electric field acts on a dielectric material, the molecules of the medium respond to the applied field and become polarized. The macroscopic effect of this molecular response is represented by the electric polarization $\mathbf{P}$ in the medium, defined as the density of dipole moments per unit volume, whose associated potential is:

$$\phi(\mathbf{r}) = \frac{1}{4\pi\varepsilon_0} \int_{\Omega} \mathbf{P}(\mathbf{r}') \cdot \nabla' \left[\frac{1}{|\mathbf{r} - \mathbf{r}'|} \right] d\Omega'. \quad (2.16)$$

Overlapping this potential with that of the free charge distribution (Equation (2.15)) and applying integration by parts, it arrives at an expression equivalent to Equation (2.15) but with an effective charge density $\rho - \nabla \cdot \mathbf{P}$. The quantity $\rho_b = -\nabla \cdot \mathbf{P}$ corresponds to the polarization charge, so Gauss's law generalizes to:

$$\nabla \cdot \mathbf{E} = \frac{1}{\varepsilon_0}(\rho - \nabla \cdot \mathbf{P}). \quad (2.17)$$

This expression motivates the definition of the electric displacement vector $\mathbf{D}$:

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P}, \quad (2.18)$$

with which the macroscopic Gauss's law takes the compact form:

$$\nabla \cdot \mathbf{D} = \rho, \quad (2.19)$$

where ρ represents only the free charge density. For a linear and isotropic dielectric, the polarization is proportional to the applied field, $\mathbf{P} = \varepsilon_0 \chi_e \mathbf{E}$, with χ_e the electric susceptibility of the medium. This implies that $\mathbf{D} = \varepsilon \mathbf{E}$, where:

$$\varepsilon = \varepsilon_0(1 + \chi_e) \quad (2.20)$$

is the permittivity of the medium. Using $\mathbf{E} = -\nabla\phi$ in Equation (2.19), the macroscopic version of the Poisson equation is obtained:

$$\nabla^2 \phi = -\frac{\rho}{\varepsilon}. \quad (2.21)$$

Jump Conditions at Dielectric Interfaces

In systems with more than one dielectric region, it is necessary to establish the conditions that the solutions of Equation (2.21) must satisfy at the interface Γ between two media with permittivities ε_1 and ε_2 . These conditions are derived from the integral forms of Gauss's law and the irrotationality of the field, and result in:

$$(\mathbf{D}_2 - \mathbf{D}_1) \cdot \mathbf{n} = \sigma_f, \quad (2.22)$$

$$(\mathbf{E}_2 - \mathbf{E}_1) \cdot \mathbf{t} = 0, \quad (2.23)$$

where $\mathbf{n}$ is the unit normal vector to Γ , $\mathbf{t}$ is the tangent vector, and σ_f is the surface density of free charge. In the absence of free charges at the interface ($\sigma_f = 0$), the normal condition reduces to $\varepsilon_1 \partial\phi_1/\partial n = \varepsilon_2 \partial\phi_2/\partial n$, while the tangential condition implies the continuity of the potential: $\phi_1|_{\Gamma} = \phi_2|_{\Gamma}$. It is worth noting that σ_f does not include the polarization surface charge, which is already implicitly contained in the definition of $\mathbf{D}$.

2.3.2 Poisson-Boltzmann Equation

The physiological environment in which biomolecules are found consists essentially of water with dissolved salt. When dissolved, the salt generates ions that are free to move in the solvent and that, in the presence of an electric field, redistribute spatially in response to it. This causes the charge density of the solvent to depend on the electrostatic potential ϕ itself, coupling the ionic distribution with the field that generates it.

To describe this system, a solution with N_s ionic species is considered, where species i has charge q_i and mean density c_{0i} . Assuming that the force on the ions is purely electrostatic and modeling the ions as point charges in a continuous solvent, the equilibrium distribution of each

species follows Boltzmann statistics through the radial distribution function:

$$g_i(\mathbf{r}) = e^{-q_i\phi(\mathbf{r})/k_B T}, \quad (2.24)$$

so that the local concentration of species i is $c_i(\mathbf{r}) = c_{0i} g_i(\mathbf{r})$. This approximation becomes exact in the limit of very dilute solutions, where van der Waals interactions, solvation effects and ion-ion correlation effects are negligible compared to the electrostatic interaction. Substituting the ionic charge density $\rho = \sum_i q_i c_{0i} g_i(\mathbf{r})$ into the Poisson Equation (2.21), the nonlinear Poisson-Boltzmann equation (PBE) is obtained:

$$\nabla^2 \phi = -\frac{1}{\varepsilon} \sum_{i=1}^{N_s} q_i c_{0i} e^{-q_i\phi/k_B T}. \quad (2.25)$$

Linearization of the PBE

The nonlinear PBE can be simplified when the electrostatic potential is small compared to the thermal energy, a condition that is written as $|q_i\phi| \ll k_B T$. In this regime, the exponential admits a first-order Taylor expansion:

$$e^{-q_i\phi/k_B T} \approx 1 - \frac{q_i\phi}{k_B T}. \quad (2.26)$$

Substituting Equation (2.26) into Equation (2.25), the sum separates into two terms:

$$-\frac{1}{\varepsilon} \sum_{i=1}^{N_s} q_i c_{0i} \left(1 - \frac{q_i\phi}{k_B T} \right) = -\frac{1}{\varepsilon} \underbrace{\sum_{i=1}^{N_s} q_i c_{0i}}_{=0} + \frac{\phi}{\varepsilon k_B T} \underbrace{\sum_{i=1}^{N_s} q_i^2 c_{0i}}_{>0}. \quad (2.27)$$

The first term of Equation (2.27) vanishes due to the electroneutrality condition of the bulk solvent ($\sum_i q_i c_{0i} = 0$). The second one survives because $q_i^2 > 0$ regardless of the sign of the charge, so that the PBE linearizes as:

$$\nabla^2 \phi = \kappa^2 \phi, \quad (2.28)$$

where κ is the Debye-Hückel screening parameter:

$$\kappa^2 = \frac{1}{\varepsilon k_B T} \sum_{i=1}^{N_s} c_{0i} q_i^2. \quad (2.29)$$

The inverse κ^{-1} is the Debye length, which represents the characteristic scale over which the electrostatic interactions are screened by the ions in solution.

For the particular case of a symmetric monovalent salt such as sodium chloride, where the sodium ion has charge $q_+ = e^+$ and the chlorine ion has charge $q_- = -e^+$, with equal mean density c_0 , the nonlinear PBE (2.25) takes the form:

$$\begin{aligned}\nabla^2\phi &= -\frac{c_0}{\varepsilon} \left(e^+ e^{-e^+\phi/k_BT} + e^- e^{+e^+\phi/k_BT} \right) \\ &= \frac{2c_0e^+}{\varepsilon} \sinh\left(\frac{e^+\phi}{k_BT}\right),\end{aligned}\tag{2.30}$$

where it was used that $q_- = -e^+$ and the identity $\sinh(x) = (e^x - e^{-x})/2$. Linearizing each exponential of the sinh to first order,

$$e^{+e^+\phi/k_BT} \approx 1 + \frac{e^+\phi}{k_BT}, \quad e^{-e^+\phi/k_BT} \approx 1 - \frac{e^+\phi}{k_BT},\tag{2.31}$$

and subtracting both expressions, the constant terms cancel out exactly:

$$e^{+e^+\phi/k_BT} - e^{-e^+\phi/k_BT} \approx \frac{2e^+\phi}{k_BT}.\tag{2.32}$$

Substituting Equation (2.32) into Equation (2.30), one obtains

$$\nabla^2\phi \approx \frac{2c_0e^+}{\varepsilon} \cdot \frac{1}{2} \cdot \frac{2e^+\phi}{k_BT} = \frac{2c_0(e^+)^2}{\varepsilon k_BT} \phi,\tag{2.33}$$

so that for NaCl the screening parameter takes the value

$$\kappa^2 = \frac{2c_0(e^+)^2}{\varepsilon k_BT},\tag{2.34}$$

consistent with the general expression given by Equation (2.29). For a physiological concentration of 150 mM NaCl, Equation (2.34) gives $\kappa \approx 0.125 \text{ \AA}^{-1}$. The validity condition of the linearization corresponds to $|\phi| \ll k_BT/e^+ \approx 25.7 \text{ mV}$ at room temperature [1, 15]. The LPBE given by Equation (2.28) is the one that governs the solvent domain in this work.

2.4 Implicit Solvent Model Applied to the Protein-Surface System

The implicit solvent model translates the electrostatic problem of a biomolecule immersed in its aqueous environment into a set of dielectric regions, each one governed by the differential equation that corresponds to its physical composition, the Poisson equation in the ion-free regions, and the linearized Poisson-Boltzmann equation in those that contain dissolved salt. The different regions

are coupled through the jump conditions Equations (2.22) and (2.23). In this case, none of the internal interfaces of the system, protein-solvent, protein-cavity, protein-Stern, has any charge physically deposited on it; the only charges present are the point charges q_k inside the protein. Because of this, in all these interfaces the particular case of Equation (2.22) with $\sigma_f = 0$ is used, which reduces the condition to the simple continuity of the electric displacement. This section presents the three systems of increasing complexity used in this work and then extends the boundary integral formulation Equation (2.63), obtained for the Laplace operator, to the Poisson and Yukawa operators that govern, respectively, the solute and the solvent.

2.4.1 Single-Surface System: Isolated Protein in the Solvent

The most elementary case consists of a protein, represented by the domain Ω_1 with permittivity $\varepsilon_1 \approx 2 - 4$, immersed in a saline aqueous solvent that occupies the domain Ω_2 with permittivity $\varepsilon_2 \approx 80$, as illustrated in Figure 2.4. The atomic positions $\mathbf{r}_k$ and the point charges q_k that describe the internal charge distribution of the protein are obtained from its crystallographic structure, available in the Protein Data Bank (PDB), combined with a force field that also assigns the van der Waals radius of each atom. Under these conditions, the electrostatic potential of the system satisfies:

$$\begin{aligned}
\nabla^2 \phi_1(\mathbf{r}) &= - \sum_k \frac{q_k}{\varepsilon_1} \delta(\mathbf{r}, \mathbf{r}_k) && \text{in } \Omega_1, \\
\nabla^2 \phi_2(\mathbf{r}) &= \kappa^2 \phi_2(\mathbf{r}) && \text{in } \Omega_2, \\
\phi_1 &= \phi_2 && \text{on } \Gamma, \\
\varepsilon_1 \frac{\partial \phi_1}{\partial \mathbf{n}} &= \varepsilon_2 \frac{\partial \phi_2}{\partial \mathbf{n}} && \text{on } \Gamma.
\end{aligned} \tag{2.35}$$

The first equation corresponds to the Poisson equation for a distribution of point charges inside the protein, the second one is the linearized PBE (2.28) in the solvent. The two interface conditions are derived directly from Equations (2.23) and (2.22), the continuity of the potential reflects that the electric field must remain bounded on Γ , while the continuity of the electric displacement is a consequence of the fact that there is no free charge deposited exactly on the interface ($\sigma_f = 0$).

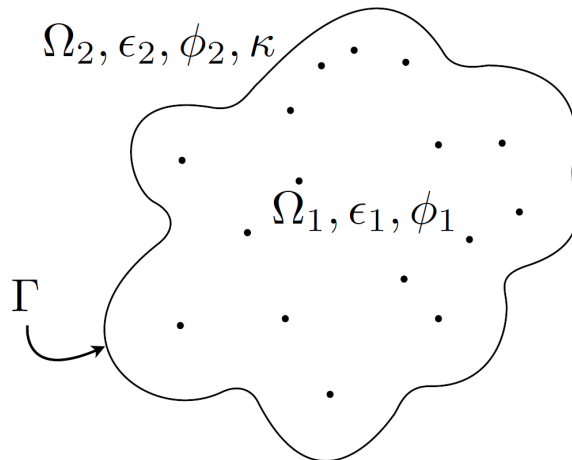


Figure 2.4: Representation of a protein immersed in a solvent [16].

Definition of the Protein-Solvent Interface

The specific choice of the surface Γ_1 is not unique. There are at least three commonly used definitions, all built from the atomic van der Waals radius r_{vdW} , assigned to each atom by the force field used to describe the protein. The simplest one is the van der Waals surface (Figure 2.5a), which delimits the protein as the union of the individual atomic spheres, this is the definition used in generalized Born methods, where the protein is approximated as a set of spheres.

The other two definitions are built by rolling a probe sphere around the protein, which represents the physical size of a solvent water molecule ($r_W \approx 1.4 \text{ \AA}$). The difference between them lies in which part of that sphere is used to trace the resulting surface. The Solvent Accessible Surface (SAS, Figure 2.5b) is obtained by following the center of the probe sphere while it rolls in contact with the atoms, without penetrating them. The Solvent Excluded Surface (SES, Figure 2.5c), on the other hand, is obtained by following the outer edge of that same sphere and corresponds to the closest distance a water molecule can physically reach with respect to the protein. In most regions, where the probe sphere rolls freely over the convex surface of the protein, the SES coincides with the van der Waals surface itself. The difference appears in small spaces and channels, there the sphere does not fit completely, and while its center is necessarily kept away from the deepest wall, its outer edge does manage to touch it, capturing that detail that the SAS does not resolve. Because of this, The SES represents the protein surface and cavities more accurately, which is the reason it is adopted in this work, computed using the **NanoShaper** software.

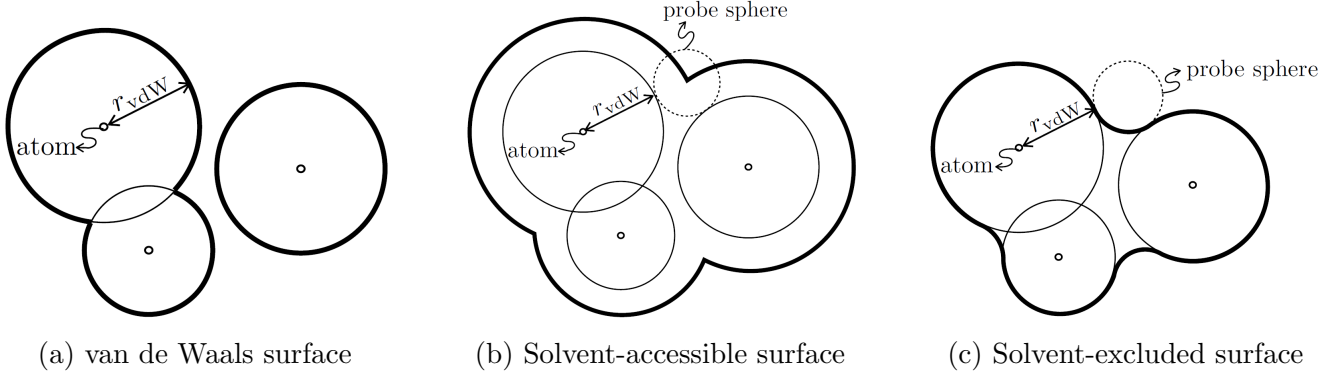


Figure 2.5: Surface definitions used to delimit the protein [16].

2.4.2 Protein-Surface System

When there is another surface in the solvent besides the protein, such as a flat electrode for example, a second boundary domain Γ_2 is introduced, on which a boundary condition is prescribed. This condition can fix directly the electric potential (Dirichlet type) or the surface charge density (Neumann type), depending on which of the two quantities is known experimentally. The system extends to:

$$\begin{aligned}
 \nabla^2 \phi_1(\mathbf{r}) &= - \sum_k \frac{q_k}{\varepsilon_1} \delta(\mathbf{r}, \mathbf{r}_k) && \text{in } \Omega_1, \\
 \nabla^2 \phi_2(\mathbf{r}) &= \kappa^2 \phi_2(\mathbf{r}) && \text{in } \Omega_2, \\
 \phi_1 = \phi_2, \quad \varepsilon_1 \frac{\partial \phi_1}{\partial \mathbf{n}} &= \varepsilon_2 \frac{\partial \phi_2}{\partial \mathbf{n}} && \text{on } \Gamma_1, \\
 \phi_2 = \phi_0 \quad \text{or} \quad -\varepsilon_2 \frac{\partial \phi_2}{\partial \mathbf{n}} &= \sigma_0 && \text{on } \Gamma_2,
 \end{aligned} \tag{2.36}$$

where ϕ_0 is the prescribed potential and σ_0 the prescribed surface charge density on the electrode.

2.4.3 Internal Cavities and the Stern Layer

Larger proteins can have pockets of solvent trapped inside them, as well as a narrow region adjacent to the SES where Poisson-Boltzmann theory overestimates the ionic concentration. Both effects require extending the two region model with additional domains and surfaces, as illustrated in Figure 2.6.

Internal Cavities

An internal cavity is a domain completely surrounded by the protein, with the same permittivity as the bulk solvent ($\varepsilon \approx 80$), since it consists of water trapped inside it. The potential inside it is treated as governed by the linearized PBE (2.28), in the same way as in the bulk solvent, although since the cavity is isolated from the exterior solvent, the limiting case $\kappa \rightarrow 0$, which reduces the equation to the Laplace equation, is also a reasonable approximation. At the interface separating the protein from the cavity, the same continuity conditions from Equation (2.35) are imposed.

Stern Layer

Poisson-Boltzmann theory treats ions as point charges with no volume, which leads to an overestimation of the ionic density in the immediate vicinity of charged surfaces. The Stern layer corrects this effect by introducing a thin region (thickness ~ 2 Å) adjacent to the SES, free of ions, with a permittivity similar to the solvent. Since there are no free ions in this region, the potential satisfies the Laplace equation:

$$\nabla^2 \phi_S(\mathbf{r}) = 0 \quad \text{in the Stern layer.} \quad (2.37)$$

Combined Five-Region System

Considering a protein with its Stern layer and two internal cavities, the complete system involves five domains: exterior solvent (Ω_1), Stern layer (Ω_2), protein interior (Ω_3) and two cavities (Ω_4 , Ω_5), connected by four interfaces:

$$\begin{aligned}
\nabla^2 \phi_1(\mathbf{r}) &= \kappa^2 \phi_1(\mathbf{r}) && \text{in } \Omega_1 \text{ (solvent),} \\
\nabla^2 \phi_2(\mathbf{r}) &= 0 && \text{in } \Omega_2 \text{ (Stern),} \\
\nabla^2 \phi_3(\mathbf{r}) &= - \sum_k \frac{q_k}{\varepsilon_3} \delta(\mathbf{r}, \mathbf{r}_k) && \text{in } \Omega_3 \text{ (protein),} \\
\nabla^2 \phi_4(\mathbf{r}) &= \kappa^2 \phi_4(\mathbf{r}) && \text{in } \Omega_4 \text{ (cavity 1),} \\
\nabla^2 \phi_5(\mathbf{r}) &= \kappa^2 \phi_5(\mathbf{r}) && \text{in } \Omega_5 \text{ (cavity 2),} \\
\phi_1 = \phi_2, \quad \varepsilon_1 \frac{\partial \phi_1}{\partial \mathbf{n}} &= \varepsilon_2 \frac{\partial \phi_2}{\partial \mathbf{n}} && \text{on } \Gamma_1 \text{ (solvent-Stern),} \\
\phi_2 = \phi_3, \quad \varepsilon_2 \frac{\partial \phi_2}{\partial \mathbf{n}} &= \varepsilon_3 \frac{\partial \phi_3}{\partial \mathbf{n}} && \text{on } \Gamma_2 \text{ (Stern-protein),} \\
\phi_3 = \phi_4, \quad \varepsilon_3 \frac{\partial \phi_3}{\partial \mathbf{n}} &= \varepsilon_4 \frac{\partial \phi_4}{\partial \mathbf{n}} && \text{on } \Gamma_3 \text{ (protein-cavity 1),} \\
\phi_3 = \phi_5, \quad \varepsilon_3 \frac{\partial \phi_3}{\partial \mathbf{n}} &= \varepsilon_5 \frac{\partial \phi_5}{\partial \mathbf{n}} && \text{on } \Gamma_4 \text{ (protein-cavity 2).}
\end{aligned} \quad (2.38)$$

This system requires four surfaces for its description, compared to the single dielectric interface Γ of the base system, or the two surfaces, one dielectric interface and one prescribed external boundary, of the protein-surface system.

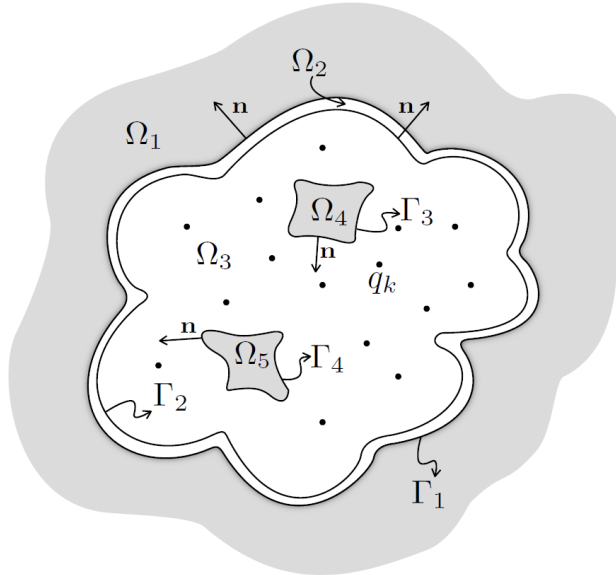


Figure 2.6: Representation of a protein with a Stern layer and two cavities filled with solvent. The gray regions contain salt, the white regions do not [16].

2.4.4 Energy Calculation in the Implicit Solvent Model

For a system of proteins dissolved in solvent, the total free energy can be separated into a polar (electrostatic) contribution and a non-polar one (van der Waals and cavity entropy). In this work it is assumed that all the free energy is electrostatic, which is a good approximation for proteins with highly charged regions. The polar component can be divided into three terms according to their physical origin, the Coulomb energy between the point charges of the solute, the solvation energy generated by the polarization of the surrounding medium, and the surface energy associated with a surface with prescribed charge or potential.

Coulomb Energy

The Coulomb energy comes from the mutual interaction between the point charges q_k located at the atoms of the protein, with no effect of the surrounding dielectric medium involved:

$$F_{\text{Coulomb}} = \frac{1}{2} \sum_i \sum_{j \neq i} \frac{q_i q_j}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad (2.39)$$

The factor $1/2$ avoids counting each pair of charges twice. This energy corresponds to the one the protein would have isolated in vacuum, and serves as a reference to define the solvation energy below.

Solvation Energy

When the protein is dissolved, the solvent polarizes in response to the charges of the solute, which is known as the reaction potential ϕ_{reac} :

$$\phi_{\text{reac}}(\mathbf{r}) = \phi_{\text{total}}(\mathbf{r}) - \phi_{\text{Coulomb}}(\mathbf{r}), \quad (2.40)$$

where ϕ_{Coulomb} is the potential that the point charges would generate in vacuum and ϕ_{total} the potential obtained by solving the complete system of the implicit solvent model. The solvation energy is calculated by evaluating this reaction potential at each point charge:

$$F_{\text{solv}} = \frac{1}{2} \int_{\Omega} \rho \phi_{\text{reac}} d\Omega = \frac{1}{2} \sum_{k=0}^{N_q} q_k \phi_{\text{reac}}(\mathbf{r}_k). \quad (2.41)$$

This quantity represents the energy needed to take the protein from vacuum into the solvent. When there are other elements in the medium, such as an electrode, the reaction potential also incorporates its contribution, so that F_{solv} generalizes the classical notion of solvation of an isolated molecule.

Surface Energy

The presence of a surface with prescribed charge or potential introduces an additional energy. Within the linearized framework of the PBE, this energy is expressed as:

$$\begin{aligned} F_{\text{surface}} &= \frac{1}{2} \int_{\Gamma} G_c \sigma_0^2 d\Gamma && \text{(prescribed surface charge } \sigma_0), \\ F_{\text{surface}} &= \frac{1}{2} \int_{\Gamma} G_p \phi_0^2 d\Gamma && \text{(prescribed potential } \phi_0), \end{aligned} \quad (2.42)$$

where G_c and G_p are the Green's kernels corresponding to the Neumann and Dirichlet problems respectively, defined such that $\phi(\sigma, R, \mathbf{x}) = G_c(R, \mathbf{x})\sigma$ and $\sigma(\phi, R, \mathbf{x}) = G_p(R, \mathbf{x})\phi$ within the linearization.

Induced Surface Charge Density

The presence of the protein induces a surface charge density σ on the electrode, which is obtained directly from the discontinuity of the electric displacement Equation (2.22) on Γ_2 :

$$\sigma(\mathbf{r}) = -\varepsilon_2 \nabla \phi_2(\mathbf{r}) \cdot \mathbf{n} \Big|_{\Gamma_2}. \quad (2.43)$$

This distribution captures the electrostatic footprint of the protein on the electrode.

Interaction Energy

When more than one body coexists in the solvent, such as a protein and an electrode, both interact electrostatically. The total free energy of the coupled system is the sum of the three previous contributions:

$$F_{\text{total}} = F_{\text{Coulomb}} + F_{\text{solv}} + F_{\text{surface}}. \quad (2.44)$$

The interaction energy is defined as the difference between the energy of the coupled system and the sum of the energies of each component evaluated in isolation:

$$F_{\text{interaction}} = F_{\text{total}}^{\text{system}} - \sum_{i=1}^{N_c} F_{\text{total}}^{(i)}, \quad (2.45)$$

where N_c is the number of components of the system and $F_{\text{total}}^{(i)}$ is calculated for component i in isolation. When the interacting system corresponds to two bound molecules, this interaction energy coincides with the binding energy.

In the particular case of a rigid protein that only changes orientation with respect to the electrode, F_{Coulomb} is invariant, the distances between the internal point charges of the protein are not modified by a rotation of the whole body. Therefore, this term appears with the same value both in $F_{\text{total}}^{\text{system}}$ and in the energy of the isolated protein, and cancels out exactly when calculating Equation (2.45). The interaction energy then reduces to the sum of the solvation and surface variations between the coupled system and each isolated component:

$$F_{\text{interaction}} = \Delta F_{\text{solv}} + \Delta F_{\text{surface}}. \quad (2.46)$$

Induced Surface Charge Perturbation

Analogous to the interaction energy defined in Equation (2.45), it is useful to isolate the contribution of the protein to the induced surface charge density from the baseline charge that the electrode

already carries due solely to an applied potential ϕ_e . In the absence of any solute, a charged electrode immersed in an electrolytic solution still accumulates a surface charge density $\sigma_{\text{electrode}}(\mathbf{r})$ purely as a capacitive response to ϕ_e , obtained from Equation (2.43) evaluated for the isolated electrode system. When the protein is adsorbed, the total induced density $\sigma_{\text{total}}(\mathbf{r})$ reflects both this baseline and the additional polarization caused by the protein's own charge distribution. The charge density perturbation is therefore defined as:

$$\Delta\sigma(\mathbf{r}) = \sigma_{\text{total}}(\mathbf{r}) - \sigma_{\text{electrode}}(\mathbf{r}), \quad (2.47)$$

which isolates the electrostatic footprint specifically attributable to the protein, independent of the reference potential applied to the electrode. This quantity is directly comparable across different values of ϕ_e or pH, since each condition is referenced to its own electrode-only baseline.

2.5 Boundary Element Method (BEM)

2.5.1 General Aspects of the Method

Boundary methods are those in which the approximate solution satisfies the governing equation (also referred to as the field equation), such that the problem's unknowns are restricted solely to the boundary. The governing equation may be satisfied in its homogeneous form or in a specific form featuring a right hand side term. Boundary methods can be derived by selecting a weighting function w in two ways:

- Selecting a function w that satisfies the governing equation in its homogeneous form.
- Utilizing functions that satisfy these equations such that the problem is constrained strictly to the boundary. The most common functions applied to the right hand side are Dirac Delta functions, which yield a value at a specific point when integrated over the domain.

The Boundary Element Method (BEM) is a numerical technique dedicated to solving Boundary Integral Equations (BIE). This is particularly useful for certain Partial Differential Equations (PDE) where the differential operator possesses a fundamental solution, as these can be represented as a boundary [17].

This principle offers distinct advantages over volumetric methods, such as finite differences or finite elements. BEM only requires the discretization of the domain's surface or boundary, necessitating fewer nodes and degrees of freedom to achieve comparable accuracy, which results in smaller systems of equations. Another benefit is that modifying the mesh geometry is simpler than

in volumetric methods, which are prone to distorted elements or poor aspect ratios. Furthermore, volumetric methods often require “staircase” approximations to represent interface surfaces, whereas BEM provides a faithful and precise representation of boundaries, capturing fine details without significantly increasing the computational cost.

Conversely, a drawback of BEM is that its discretization generates dense matrices, in contrast to the typical sparse matrices found in volumetric methods. This increases the computational complexity and cost of the solution, especially when using unaccelerated techniques like Gaussian elimination. Consequently, the implementation of acceleration techniques such as multipole methods, treecode, or Fast Fourier Transforms (FFT) is often required.

2.5.2 Mathematical Origin and Derivation of the Equation

In most literature, the Laplace equation serves as the starting point for explaining the mathematical foundations of BEM, and this work is no exception. This second order partial differential equation, which has a well known fundamental solution, describes various phenomena in physics and engineering, including heat transfer, potential flow, and electrostatics. Moreover, it represents a simplified case of the Poisson equation where the source term is zero, allowing for a more straightforward application of BEM and its subsequent extension to the Poisson and Poisson-Boltzmann equations.

Consider a scalar quantity ϕ within a domain Ω with a boundary Γ , subject to Dirichlet and Neumann boundary conditions, ϕ_0 and $\partial\phi_0/\partial\mathbf{n}$ respectively, expressed as:

$$\begin{aligned} \nabla^2\phi &= 0 \text{ in } \Omega, \\ \phi &= \phi_0 \text{ or } \frac{\partial\phi}{\partial\mathbf{n}} = \frac{\partial\phi_0}{\partial\mathbf{n}} \text{ on } \Gamma. \end{aligned} \tag{2.48}$$

By introducing a test function w , Equation (2.48) can be rewritten in its weak integral form as:

$$\int_{\Omega} \nabla^2\phi(\mathbf{r}'_{\Omega})w(\mathbf{r}'_{\Omega}) d\Omega' = 0, \tag{2.49}$$

where the evaluation point $\mathbf{r}_{\Omega}$ is located within the domain Ω , and the integration is performed over the primed variables.

When a linear differential operator (such as the Laplacian) is applied to its free-space Green’s

function, the result is the negative Dirac delta function ($-\delta(\mathbf{r})$), expressed as:

$$\nabla^2 G_L(\mathbf{r}, \mathbf{r}') = -\delta(\mathbf{r}). \quad (2.50)$$

This implies that applying the differential operator to the Green's function yields zero throughout the domain, except at the point $\mathbf{r} = \mathbf{r}'$, where an infinite spike occurs with an integral equal to unity. This behavior describes the system's response to a point excitation, such as an impulse or a point charge.

This relationship defines the free-space Green's function for the Laplace operator, given by:

$$G_L(\mathbf{r}, \mathbf{r}') = \frac{1}{4\pi|\mathbf{r} - \mathbf{r}'|}. \quad (2.51)$$

Substituting the Laplace free-space Green's function as the test function in Equation (2.49) yields:

$$\int_{\Omega} \nabla^2 \phi(\mathbf{r}'_{\Omega}) G_L(\mathbf{r}_{\Omega}, \mathbf{r}'_{\Omega}) d\Omega' = 0. \quad (2.52)$$

Recalling the vector analogue of the product rule for derivatives:

$$\nabla \cdot (\nabla \phi \cdot G_L) = \nabla^2 \phi \cdot G_L + \nabla \phi \cdot \nabla G_L, \quad (2.53)$$

$$\nabla \cdot (\phi \nabla G_L) = \nabla \phi \cdot \nabla G_L + \phi \nabla^2 G_L. \quad (2.54)$$

Accordingly, Equation (2.52) can be rewritten by modifying its integrand using Equation (2.53) as:

$$\int_{\Omega} [\nabla \cdot (\nabla \phi(\mathbf{r}'_{\Omega}) G_L(\mathbf{r}_{\Omega}, \mathbf{r}'_{\Omega})) - \nabla \phi(\mathbf{r}'_{\Omega}) \nabla G_L(\mathbf{r}_{\Omega}, \mathbf{r}'_{\Omega})] d\Omega' = 0. \quad (2.55)$$

The second term in Equation (2.55) can then be rewritten using the property in Equation (2.54):

$$\int_{\Omega} [\nabla \cdot (\nabla \phi(\mathbf{r}'_{\Omega}) G_L(\mathbf{r}_{\Omega}, \mathbf{r}'_{\Omega})) - \nabla \cdot (\phi(\mathbf{r}'_{\Omega}) \nabla G_L(\mathbf{r}_{\Omega}, \mathbf{r}'_{\Omega})) + \phi(\mathbf{r}'_{\Omega}) \nabla^2 G_L(\mathbf{r}_{\Omega}, \mathbf{r}'_{\Omega})] d\Omega' = 0. \quad (2.56)$$

Since the first two terms of the integral fall under the divergence operator, applying the divergence theorem to Equation (2.56) and incorporating the definition from Equation (2.50) results in:

$$\int_{\Gamma} \nabla \phi(\mathbf{r}'_{\Gamma}) \cdot \mathbf{n} G_L(\mathbf{r}_{\Omega}, \mathbf{r}'_{\Gamma}) d\Gamma' - \int_{\Gamma} \phi(\mathbf{r}'_{\Gamma}) \nabla G_L(\mathbf{r}_{\Omega}, \mathbf{r}'_{\Gamma}) \cdot \mathbf{n} d\Gamma' - \int_{\Omega} \phi(\mathbf{r}'_{\Omega}) \delta(\mathbf{r}_{\Omega}) d\Omega' = 0. \quad (2.57)$$

In this context, $\mathbf{n}$ represents the outward-pointing normal vector to Γ relative to the domain Ω ,

while $\mathbf{r}_\Gamma$ denotes a point on the surface Γ . Applying the sampling property of the Dirac delta, Equation (2.57) becomes:

$$\phi(\mathbf{r}_\Omega) = \int_{\Gamma} G_L(\mathbf{r}_\Omega, \mathbf{r}'_\Gamma) \frac{\partial}{\partial \mathbf{n}} \phi(\mathbf{r}'_\Gamma) d\Gamma' - \int_{\Gamma} \phi(\mathbf{r}'_\Gamma) \frac{\partial}{\partial \mathbf{n}} G_L(\mathbf{r}_\Omega, \mathbf{r}'_\Gamma) d\Gamma'. \quad (2.58)$$

Equation (2.58) is the integral formulation of the Laplace Equation (2.48). Here, $\mathbf{r}$ can be evaluated anywhere within the domain Ω , while $\mathbf{r}'$ is restricted to the boundary Γ .

This implies that $\phi(\mathbf{r}_\Omega)$ can be determined at any point in the domain Ω solely from surface integrals evaluated on Γ . Thus, knowing the values of ϕ and its normal derivative on Γ (ϕ_0 and $\partial\phi_0/\partial\mathbf{n}$) is sufficient. This aligns with the definition of boundary methods, where the field equation is satisfied throughout the domain, leaving the unknowns exclusively on the boundary.

A challenge arises, however, as $\mathbf{r}$ approaches the limit $\mathbf{r} \rightarrow \Gamma$, where the free-space Green's function G_L exhibits a singularity at $\mathbf{r}_\Gamma = \mathbf{r}'_\Gamma$. To address this, an auxiliary surface Γ' is introduced, identical to the original surface Γ except in a small region replaced by a hemisphere Γ'_{hem} of radius ε centered at $\mathbf{r}$. The integral over Γ' is split into two parts: the region where Γ and Γ' coincide, and the region over the hemisphere Γ'_{hem} . This procedure allows Γ' to bypass the singularity via Γ'_{hem} ; as the radius vanishes ($\varepsilon \rightarrow 0$), the surface Γ' reverts to Γ , recovering the original geometry.

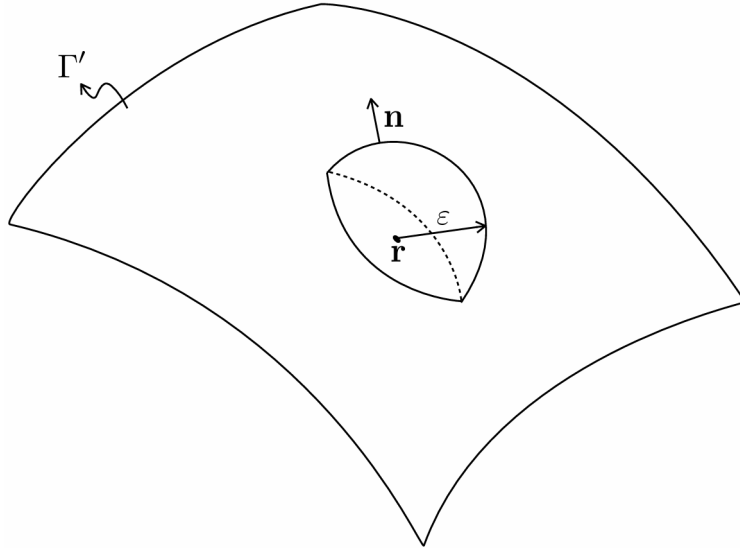


Figure 2.7: Auxiliary surface Γ' with hemisphere Γ'_{hem} to treat the singularity [16].

The first term of Equation (2.58) involving G_L over the hemisphere Γ'_{hem} is:

$$\int_{\Gamma'_{\text{hem}}} \frac{\partial}{\partial \mathbf{n}} \phi(\mathbf{r}') G_L(\mathbf{r}, \mathbf{r}') d\Gamma' = \frac{1}{4\pi} \int_{\Gamma'_{\text{hem}}} \frac{\partial}{\partial \mathbf{n}} \phi(\mathbf{r}') \frac{1}{|\mathbf{r} - \mathbf{r}'|} d\Gamma' = \frac{1}{4\pi\varepsilon} \int_{\Gamma'_{\text{hem}}} \frac{\partial}{\partial \mathbf{n}} \phi(\mathbf{r}') d\Gamma'. \quad (2.59)$$

On the hemisphere Γ'_{hem} , the distance $|\mathbf{r} - \mathbf{r}'| = \varepsilon$ is constant. Taking the limit $\varepsilon \rightarrow 0$ to recover the original geometry Γ :

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{4\pi\varepsilon} \int_{\Gamma'_{\text{hem}}} \frac{\partial}{\partial \mathbf{n}} \phi(\mathbf{r}') d\Gamma' = \lim_{\varepsilon \rightarrow 0} \frac{1}{4\pi\varepsilon} \frac{\partial}{\partial \mathbf{n}} \phi(\mathbf{r}) \int_{\Gamma'_{\text{hem}}} d\Gamma' = \lim_{\varepsilon \rightarrow 0} \frac{1}{4\pi\varepsilon} \frac{\partial}{\partial \mathbf{n}} \phi(\mathbf{r}) 2\pi\varepsilon^2 = 0. \quad (2.60)$$

As $\varepsilon \rightarrow 0$ implies $\mathbf{r}' \rightarrow \mathbf{r}$, the approximation of the area around the singularity as a hemisphere remains valid. For small ε and a smooth surface Γ , the region near Γ'_{hem} can be considered approximately flat. Consequently, this term (the second integral on the right-hand side of Equation (2.58)) vanishes and contributes nothing when evaluated on the boundary.

For the second term of Equation (2.58) involving $\partial G_L / \partial \mathbf{n}$ over the surface Γ'_{hem} :

$$\int_{\Gamma'_{\text{hem}}} \phi(\mathbf{r}') \frac{\partial}{\partial \mathbf{n}} G_L(\mathbf{r}, \mathbf{r}') d\Gamma' = -\frac{1}{4\pi} \int_{\Gamma'_{\text{hem}}} \phi(\mathbf{r}') \frac{(\mathbf{r} - \mathbf{r}') \cdot \mathbf{n}}{|\mathbf{r} - \mathbf{r}'|^3} d\Gamma' = -\frac{1}{4\pi} \int_{\Gamma'_{\text{hem}}} \phi(\mathbf{r}') \frac{\varepsilon}{\varepsilon^3} d\Gamma'. \quad (2.61)$$

Since $(\mathbf{r} - \mathbf{r}')$ is always parallel to the normal vector $\mathbf{n}$ with a constant magnitude ε , the limit $\varepsilon \rightarrow 0$ leads to:

$$\lim_{\varepsilon \rightarrow 0} -\frac{1}{4\pi} \int_{\Gamma'_{\text{hem}}} \phi(\mathbf{r}') \frac{\varepsilon}{\varepsilon^3} d\Gamma' = -\frac{\phi(\mathbf{r})}{4\pi} \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^2} \int_{\Gamma'_{\text{hem}}} d\Gamma' = -\frac{\phi(\mathbf{r})}{4\pi} \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^2} 2\pi\varepsilon^2 = -\frac{\phi(\mathbf{r})}{2}. \quad (2.62)$$

This integral yields a finite free term. Based on the results from Equations (2.60) and (2.62), the limit of Equation (2.58) as $\mathbf{r}_\Omega \rightarrow \mathbf{r}_\Gamma$ can be expressed as:

$$\frac{\phi(\mathbf{r}_\Gamma)}{2} + \int_{\Gamma} \phi(\mathbf{r}'_\Gamma) \frac{\partial}{\partial \mathbf{n}} G_L(\mathbf{r}_\Gamma, \mathbf{r}'_\Gamma) d\Gamma' = \int_{\Gamma} G_L(\mathbf{r}_\Gamma, \mathbf{r}'_\Gamma) \frac{\partial}{\partial \mathbf{n}} \phi(\mathbf{r}'_\Gamma) d\Gamma'. \quad (2.63)$$

These integrals are interpreted in the sense of the Cauchy Principal Value to handle the singularities when the integration and evaluation points coincide. Thus, Equation (2.63) represents the integral formulation of the Laplace equation on the domain boundary and is used for the implementation of the Boundary Element Method [16].

2.5.3 Application of the BEM to the Implicit Solvent Model

Equation (2.63) solves the homogeneous case governed by the Laplace operator. Its extension to the implicit solvent model requires incorporating two additional elements into that same derivation, a source term for the point charges of the solute and a new differential operator, the Yukawa operator, to describe the ionic screening of the solvent.

Extension to the Poisson Equation: The Source Term

The Poisson equation in the protein, $\nabla^2 \phi_1 = -\rho/\varepsilon_1$ with $\rho = \sum_k q_k \delta(\mathbf{r} - \mathbf{r}_k)$, can be separated into its homogeneous (Laplacian) part and the point charge source term. The homogeneous part follows exactly the derivation of section 2.5.2, since the differential operator is the same. The source term, instead, is incorporated by multiplying the complete Poisson equation by the test function G_L and integrating over Ω_1 :

$$\int_{\Omega_1} \nabla^2 \phi_1(\mathbf{r}') G_L(\mathbf{r}, \mathbf{r}') d\Omega' = -\frac{1}{\varepsilon_1} \int_{\Omega_1} G_L(\mathbf{r}, \mathbf{r}') \sum_k q_k \delta(\mathbf{r}' - \mathbf{r}_k) d\Omega'. \quad (2.64)$$

The left-hand side is identical to the one from the Laplace derivation and leads to the same result as in Equation (2.63), the right-hand side collapses through the sampling property of the Dirac delta, giving the fundamental Coulomb solution, Equation (2.51), evaluated at each charge:

$$-\frac{1}{\varepsilon_1} \sum_k q_k G_L(\mathbf{r}, \mathbf{r}_k) = -\frac{1}{\varepsilon_1} \sum_k \frac{q_k}{4\pi|\mathbf{r} - \mathbf{r}_k|}. \quad (2.65)$$

Adding both results and taking the limit $\mathbf{r}_{\Omega_1} \rightarrow \mathbf{r}_{\Gamma_1}$, with the same hemisphere argument from Section 2.5.2, the boundary integral equation for the Poisson equation evaluated on the SES is obtained:

$$\frac{1}{2} \phi_1(\mathbf{r}_{\Gamma_1}) + \int_{\Gamma_1} \frac{\partial G_L}{\partial \mathbf{n}}(\mathbf{r}_{\Gamma_1}, \mathbf{r}') \phi_1(\mathbf{r}') d\Gamma' - \int_{\Gamma_1} G_L(\mathbf{r}_{\Gamma_1}, \mathbf{r}') \frac{\partial \phi_1}{\partial \mathbf{n}}(\mathbf{r}') d\Gamma' = \frac{1}{\varepsilon_1} \sum_k \frac{q_k}{4\pi|\mathbf{r}_{\Gamma_1} - \mathbf{r}_k|}. \quad (2.66)$$

Extension to the Linearized PBE: The Yukawa Operator

In the solvent domain, the differential operator is no longer the Laplacian but the Yukawa operator, $(\nabla^2 - \kappa^2)$, which appears in the linearized PBE (2.28). Its Green's function is obtained by solving:

$$(\nabla^2 - \kappa^2) G_Y(\mathbf{r}, \mathbf{r}') = -\delta(\mathbf{r} - \mathbf{r}'), \quad (2.67)$$

whose solution in three dimensions is:

$$G_Y(\mathbf{r}, \mathbf{r}') = \frac{e^{-\kappa|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|}. \quad (2.68)$$

Multiplying the linearized PBE by G_Y and integrating over Ω_2 , the same algebraic procedure from Section 2.5.2 leads to an expression analogous to Equation (2.63), but for ϕ_2 . There is, however, an important difference, Ω_2 is the region surrounding the protein, so the normal vector $\mathbf{n}$, defined pointing from Ω_1 toward Ω_2 , is the inward normal with respect to Ω_2 and not the outward one required by Green's identity. Applying the identity with the outward convention proper to Ω_2 is equivalent to change the sign of $\mathbf{n}$ in the derivation, which changes the sign of the two layer terms. Taking the limit toward the interface, the boundary integral equation for the linearized PBE evaluated on Γ_1 results in:

$$\frac{1}{2}\phi_2(\mathbf{r}_{\Gamma_1}) - \int_{\Gamma_1} \frac{\partial G_Y}{\partial \mathbf{n}}(\mathbf{r}_{\Gamma_1}, \mathbf{r}') \phi_2(\mathbf{r}') d\Gamma' + \int_{\Gamma_1} G_Y(\mathbf{r}_{\Gamma_1}, \mathbf{r}') \frac{\partial \phi_2}{\partial \mathbf{n}}(\mathbf{r}') d\Gamma' = 0. \quad (2.69)$$

Unlike the Coulomb potential, G_Y decays exponentially with distance, which physically reflects the screening of the electrostatic interactions by the dissolved ions.

Compact Notation: Layer Operators

Equations (2.66) and (2.69) share the same structure, in both there appear two surface integrals identical in form, one with the Green's function and another with its normal derivative, evaluated against ϕ and $\partial\phi/\partial\mathbf{n}$ respectively. This structure repeats for any differential operator with a known fundamental solution, so it is convenient to introduce a common notation. For a Green's function G_X associated to an operator X (L for Laplace/Poisson, Y for Yukawa), the single layer operator V_X^Γ and the double layer operator K_X^Γ are defined as:

$$\begin{aligned} V_X^\Gamma(\psi)(\mathbf{r}) &= \int_{\Gamma} G_X(\mathbf{r}, \mathbf{r}') \psi(\mathbf{r}') d\Gamma', \\ K_X^\Gamma(\psi)(\mathbf{r}) &= \int_{\Gamma} \frac{\partial G_X}{\partial \mathbf{n}}(\mathbf{r}, \mathbf{r}') \psi(\mathbf{r}') d\Gamma'. \end{aligned} \quad (2.70)$$

With this notation, Equations (2.66) and (2.69) can be rewritten in compact form as the interior equation and the exterior equation:

$$\frac{\phi_1(\mathbf{r}_{\Gamma_1})}{2} + K_L^{\Gamma_1}(\phi_1|_{\Gamma_1}) - V_L^{\Gamma_1} \left(\frac{\partial \phi_1}{\partial \mathbf{n}} \Big|_{\Gamma_1} \right) = \frac{1}{\varepsilon_1} \sum_k \frac{q_k}{4\pi|\mathbf{r}_{\Gamma_1} - \mathbf{r}_k|}, \quad (2.71)$$

$$\frac{\phi_2(\mathbf{r}_{\Gamma_1})}{2} - K_Y^{\Gamma_1}(\phi_2|_{\Gamma_1}) + V_Y^{\Gamma_1} \left(\frac{\partial \phi_2}{\partial \mathbf{n}} \Big|_{\Gamma_1} \right) = 0. \quad (2.72)$$

Single-Surface System

Substituting the interface conditions $\phi_1|_{\Gamma_1} = \phi_2|_{\Gamma_1} \equiv \phi|_{\Gamma_1}$ and $\varepsilon_2 \partial\phi_2/\partial\mathbf{n} = \varepsilon_1 \partial\phi_1/\partial\mathbf{n}$ into Equation (2.72), the two unknowns of the problem are reduced to $\phi|_{\Gamma_1}$ and $\partial\phi_1/\partial\mathbf{n}|_{\Gamma_1}$, and the system formed by Equations (2.71) and (2.72) can be written in matrix form as:

$$\begin{bmatrix} \frac{1}{2} + K_L^{\Gamma_1} & -V_L^{\Gamma_1} \\ \frac{1}{2} - K_Y^{\Gamma_1} & \frac{\varepsilon_1}{\varepsilon_2} V_Y^{\Gamma_1} \end{bmatrix} \begin{bmatrix} \phi|_{\Gamma_1} \\ \frac{\partial\phi_1}{\partial\mathbf{n}}|_{\Gamma_1} \end{bmatrix} = \begin{bmatrix} \frac{1}{\varepsilon_1} \sum_k \frac{q_k}{4\pi|\mathbf{r}_{\Gamma_1} - \mathbf{r}_k|} \\ 0 \end{bmatrix}. \quad (2.73)$$

Extension to the Protein-Surface System

When the solvent domain is also bounded by an additional surface Γ_2 , the exterior Equation (2.72) must be evaluated on both surfaces. Since Green's identity integrates over the entire boundary of the domain, and that boundary is now $\Gamma_1 \cup \Gamma_2$, both surfaces contribute and the two-surface exterior BIE for a general interior point $\mathbf{r}$ reads:

$$\frac{\phi_2(\mathbf{r})}{2} - K_Y^{\Gamma_1}(\phi_2|_{\Gamma_1})(\mathbf{r}) + V_Y^{\Gamma_1} \left(\frac{\partial\phi_2}{\partial\mathbf{n}} \Big|_{\Gamma_1} \right) (\mathbf{r}) - K_Y^{\Gamma_2}(\phi_2|_{\Gamma_2})(\mathbf{r}) + V_Y^{\Gamma_2} \left(\frac{\partial\phi_2}{\partial\mathbf{n}} \Big|_{\Gamma_2} \right) (\mathbf{r}) = 0. \quad (2.74)$$

Evaluated on Γ_1 , it receives an additional contribution from the prescribed condition $\phi_2|_{\Gamma_2} = \phi_0$:

$$\frac{\phi|_{\Gamma_1}}{2} - K_Y^{\Gamma_1}(\phi|_{\Gamma_1}) + \frac{\varepsilon_1}{\varepsilon_2} V_Y^{\Gamma_1} \left(\frac{\partial\phi_1}{\partial\mathbf{n}} \Big|_{\Gamma_1} \right) + V_Y^{\Gamma_2} \left(\frac{\partial\phi_2}{\partial\mathbf{n}} \Big|_{\Gamma_2} \right) = K_Y^{\Gamma_2}(\phi_0), \quad (2.75)$$

while evaluated on Γ_2 , it incorporates the contributions from Γ_1 :

$$-K_Y^{\Gamma_1}(\phi|_{\Gamma_1}) + \frac{\varepsilon_1}{\varepsilon_2} V_Y^{\Gamma_1} \left(\frac{\partial\phi_1}{\partial\mathbf{n}} \Big|_{\Gamma_1} \right) + V_Y^{\Gamma_2} \left(\frac{\partial\phi_2}{\partial\mathbf{n}} \Big|_{\Gamma_2} \right) = - \left(\frac{1}{2} - K_Y^{\Gamma_2} \right) (\phi_0). \quad (2.76)$$

Together with the interior Equation (2.71), these three equations determine the three unknowns $\phi|_{\Gamma_1}$, $\partial\phi_1/\partial\mathbf{n}|_{\Gamma_1}$ and $\partial\phi_2/\partial\mathbf{n}|_{\Gamma_2}$, giving the system:

$$\begin{bmatrix} \frac{1}{2} + K_L^{\Gamma_1} & -V_L^{\Gamma_1} & 0 \\ \frac{1}{2} - K_Y^{\Gamma_1} & \frac{\varepsilon_1}{\varepsilon_2} V_Y^{\Gamma_1} & V_Y^{\Gamma_2} \\ -K_Y^{\Gamma_1} & \frac{\varepsilon_1}{\varepsilon_2} V_Y^{\Gamma_1} & V_Y^{\Gamma_2} \end{bmatrix} \begin{bmatrix} \phi|_{\Gamma_1} \\ \frac{\partial\phi_1}{\partial\mathbf{n}}|_{\Gamma_1} \\ \frac{\partial\phi_2}{\partial\mathbf{n}}|_{\Gamma_2} \end{bmatrix} = \begin{bmatrix} \frac{1}{\varepsilon_1} \sum_k \frac{q_k}{4\pi|\mathbf{r}_{\Gamma_1} - \mathbf{r}_k|} \\ K_Y^{\Gamma_2}(\phi_0) \\ - \left(\frac{1}{2} - K_Y^{\Gamma_2} \right) (\phi_0) \end{bmatrix}. \quad (2.77)$$

The first row corresponds to the interior equation of the protein evaluated on the SES; the second one, to the exterior equation of the solvent evaluated on the SES, with the influence of the

external surface incorporated in the independent term; and the third one, to the exterior equation evaluated directly on the outer boundary.

The system of Equations (2.77) assumes a Dirichlet condition on the surface. If instead, a Neumann condition with known charge density is prescribed ($-\varepsilon_2 \partial\phi_2/\partial\mathbf{n} = \sigma_0$), the roles on Γ_2 are reversed, the potential $\phi_2|_{\Gamma_2}$ moves to the vector of unknowns and the normal derivative, now known, is absorbed into the right-hand side through the single-layer operator, giving [9]:

$$\begin{bmatrix} \frac{1}{2} + K_L^{\Gamma_1} & -V_L^{\Gamma_1} & 0 \\ \frac{1}{2} - K_Y^{\Gamma_1} & \frac{\varepsilon_1}{\varepsilon_2} V_Y^{\Gamma_1} & -K_Y^{\Gamma_2} \\ -K_Y^{\Gamma_1} & \frac{\varepsilon_1}{\varepsilon_2} V_Y^{\Gamma_1} & \frac{1}{2} - K_Y^{\Gamma_2} \end{bmatrix} \begin{bmatrix} \phi|_{\Gamma_1} \\ \frac{\partial\phi_1}{\partial\mathbf{n}}|_{\Gamma_1} \\ \phi_2|_{\Gamma_2} \end{bmatrix} = \begin{bmatrix} \frac{1}{\varepsilon_1} \sum_k \frac{q_k}{4\pi|\mathbf{r}_{\Gamma_1} - \mathbf{r}_k|} \\ \frac{\sigma_0}{\varepsilon_2} V_Y^{\Gamma_2} \\ \frac{\sigma_0}{\varepsilon_2} V_Y^{\Gamma_2} \end{bmatrix}. \quad (2.78)$$

Generalization to Multi-Surface Systems

The five region system, Equation (2.38), follows exactly the same logic as the two previous cases, only now there are four surfaces instead of one. Each surface still contributes an interior equation and an exterior equation, constructed in the same way as in Equations (2.71) and (2.72). The difference is that, since there are several surfaces, a given equation can receive contributions coming from surfaces other than the one being evaluated, as happened with Γ_2 in the previous section. In this case, the cross contributions are more numerous and make the resulting system more complex. The system presents a 2×2 diagonal block for each surface Γ_j , formed by the single and double layer operators of ϕ_{i,Γ_j} and $\partial\phi_{i,\Gamma_j}/\partial\mathbf{n}$ evaluated on that same surface, while the off-diagonal blocks correspond to cross evaluations, where region i is evaluated on a surface Γ_m different from the one where the source charges or potentials are, Γ_j .

These cross contributions are classified according to the relative position between the evaluation surface and the source surface, whether they coincide (same surface), or if they are different, whether one lies inside or outside the region enclosed by the other. This classification only determines in which equation of the system each contribution is added, the form of the operators V_X^Γ and K_X^Γ involved does not change with respect to what was already defined. Assembling all the contributions for the four interfaces of the system, Equation (2.38), gives a linear system of eight unknowns, two per surface, whose block structure reflects which pairs of regions share a physical interface and which do not.

2.5.4 Discretization of the Boundary Integral Equations

To obtain a numerical solution, the continuous surface Γ is approximated by a finite set of N_p flat triangular panels Γ_j , $j = 1, \dots, N_p$, as illustrated in Figure 2.8. The unknown boundary functions, ϕ and $\partial\phi/\partial\mathbf{n}$, are assumed to be constant inside each panel, taking the value evaluated at its centroid.

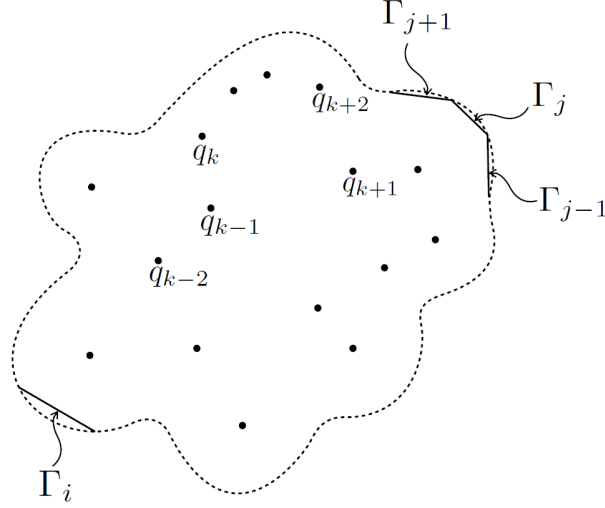


Figure 2.8: Discretization of a molecular surface into panels. The point charges q_k represent the interior charge distribution of the protein; Γ_i is the panel containing the collocation point and Γ_j the panel being integrated [16].

Under this piecewise constant approximation, the continuous layer operators defined in Equation (2.70) take a discretized form. The integral equations are evaluated placing the observation point at the centroid $\mathbf{r}_i$ of every receiving panel Γ_i . This turns the integral over the whole surface Γ into a sum of integrals over each individual panel Γ_j , weighted by the constant value the field takes on it:

$$\begin{aligned} V_{X,\text{disc}}^{\mathbf{r}_i}(\psi|_{\Gamma}) &= \sum_{j=1}^{N_p} \psi(\mathbf{r}_j) \int_{\Gamma_j} G_X(\mathbf{r}_i, \mathbf{r}') d\Gamma', \\ K_{X,\text{disc}}^{\mathbf{r}_i}(\psi|_{\Gamma}) &= \sum_{j=1}^{N_p} \psi(\mathbf{r}_j) \int_{\Gamma_j} \frac{\partial G_X}{\partial \mathbf{n}}(\mathbf{r}_i, \mathbf{r}') d\Gamma', \end{aligned} \quad (2.79)$$

where $\mathbf{r}_j$ is the centroid of panel Γ_j , used to take the constant field value out of the integral. The primed coordinate $\mathbf{r}'$ moves continuously over the surface of the panel Γ_j being integrated.

The remaining isolated integral in each term becomes a single scalar value for every pair of panels, so the discretized operators turn into the dense matrices:

$$\begin{aligned}
V_{X,ij} &= \int_{\Gamma_j} G_X(\mathbf{r}_i, \mathbf{r}') d\Gamma', \\
K_{X,ij} &= \int_{\Gamma_j} \frac{\partial G_X}{\partial \mathbf{n}}(\mathbf{r}_i, \mathbf{r}') d\Gamma'.
\end{aligned} \tag{2.80}$$

Here the Green's function depends on the distance between the fixed collocation point $\mathbf{r}_i$ and the integration variable $\mathbf{r}'$, which still moves over the entire surface of panel Γ_j and is not evaluated only at its centroid, that is, only the unknown fields ϕ and $\partial\phi/\partial\mathbf{n}$ are approximated as constant per panel, not the Green's function itself.

This way, the matrix systems derived in Section 2.5.3 become linear systems for the discrete values of ϕ and $\partial\phi/\partial\mathbf{n}$. The discretized matrix looks exactly like Equation (2.73), or any other matrix equation from that section, but now the elements of the matrix are $N_p \times N_p$ sub-matrices rather than integral operators, and the right hand side and unknown vectors are N_p -dimensional, with one entry per panel.

This procedure converts the original boundary integral equations into a finite linear system $A\mathbf{x} = \mathbf{b}$, whose dimension scales with the total number of panels, $2N_p$ for the single-surface case of Equation (2.73), and proportionally larger for systems with additional surfaces. Unlike domain discretization methods, where the resulting matrix is usually sparse, the matrix A obtained here is dense, since every panel interacts with every other panel through the long-range Green's function, no matter how far apart they are. Because of this density, direct matrix factorization becomes very expensive, which is why $A\mathbf{x} = \mathbf{b}$ is solved in practice using iterative solvers combined with acceleration techniques that approximate the dense matrix vector product at a lower computational cost.

Numerical Integration of the Panels

Each entry of Equation (2.80) is itself an integral over a single panel, usually evaluated with Gauss quadrature rules. The free-space Green's functions G_L and G_Y , however, contain a $1/r$ singularity, so the accuracy of this evaluation depends strongly on the distance between the collocation point and the panel being integrated. For this reason, three integration regimes are usually identified:

- **Far-away integrals:** when the collocation point is far enough from the panel compared to its size, the Green's function and its normal derivative change smoothly across the panel surface, and the integral can be well approximated with a low-order Gauss quadrature rule.

This distance is usually compared against a characteristic panel length derived from its area, which sets a threshold beyond which a panel is treated as far away.

- **Near-singular integrals:** when two panels are spatially close but different ($i \neq j$), the Green’s function shows a higher variation across the panel as the collocation point gets closer to it. To capture this behavior correctly, a higher order quadrature scheme is needed, with a denser distribution of integration points than in the far away case.
- **Singular integrals:** when the collocation point coincides with the panel being integrated ($i = j$), the $1/r$ kernel diverges inside the integration domain itself, and no quadrature rule, no matter how refined, can resolve this divergence correctly. These self-interactions are instead handled by isolating the singularity and applying a dedicated semi-analytical integration scheme over the panel geometry.

Discretized Free Energy Calculation

Once the linear system is solved for $\phi|_\Gamma$ and $\partial\phi/\partial\mathbf{n}|_\Gamma$ on every surface, these discrete unknowns can be substituted directly into the free energy expressions of Section 2.4.4, giving forms suitable for numerical evaluation.

The Coulomb energy, Equation (2.39), depends only on the point charges of the protein and is evaluated directly from their tabulated positions and magnitudes, requiring no further discretization.

The solvation energy, Equation (2.41), requires the reaction potential ϕ_{reac} evaluated at every charge location $\mathbf{r}_k$. Since each $\mathbf{r}_k$ lies inside the protein domain Ω_1 and not on its boundary, the explicit Coulomb singularity that appears when representing ϕ_1 at an interior point coincides exactly with ϕ_{Coulomb} in the definition, Equation (2.40), and the two cancel out, leaving only the contribution of the layer potentials:

$$\phi_{\text{reac}}(\mathbf{r}_k) = -K_L^{\mathbf{r}_k}(\phi_1|_{\Gamma_1}) + V_L^{\mathbf{r}_k}\left(\frac{\partial\phi_1}{\partial\mathbf{n}}\Big|_{\Gamma_1}\right). \quad (2.81)$$

This expression is discretized over the same N_p panels used to build the linear system, following the same procedure as Equation (2.80) but fixing the observation point at $\mathbf{r}_k$ instead of at a receiving panel’s centroid, and using the already solved values of $\phi_1|_{\Gamma_1}$ and $\partial\phi_1/\partial\mathbf{n}|_{\Gamma_1}$ from Equation (2.73) (or its multi-surface generalizations) as input. Substituting the resulting discretized reaction potential into Equation (2.41) for each charge location gives the discretized solvation energy.

The surface energy, Equation (2.42), follows from the same piecewise-constant approximation, assuming ϕ and $\partial\phi/\partial\mathbf{n}$ constant inside each panel reduces the surface integral to a finite sum weighted by the panel areas A_j :

$$\begin{aligned} F_{\text{surface}} &= \frac{1}{2} \sum_{j=1}^{N_p} \phi(\mathbf{r}_j) \sigma_{0j} A_j && \text{(prescribed charge),} \\ F_{\text{surface}} &= \frac{1}{2} \sum_{j=1}^{N_p} \phi_{0j} \sigma(\mathbf{r}_j) A_j && \text{(prescribed potential),} \end{aligned} \quad (2.82)$$

where $\sigma = \varepsilon \partial\phi/\partial\mathbf{n}$ is the induced surface charge density obtained from the discretized solution, consistent with Equation (2.43).

Discretized Induced Charge

Integrating σ over the discretized electrode surface (panel areas A_j , following the same piecewise-constant approximation as in Equation (2.82)) gives the total induced charge:

$$Q = \sum_{j=1}^{N_p} \sigma(\mathbf{r}_j) A_j, \quad (2.83)$$

which can likewise be decomposed as $Q_{\text{total}} = Q_{\text{electrode}} + \Delta Q$, where $\Delta Q = Q_{\text{total}} - Q_{\text{electrode}}$ quantifies the net charge redistribution induced specifically by the adsorbed protein, consistent with the perturbation defined in Equation (2.47).

2.6 Statistical Thermodynamics and Electron Transfer Rate

Once the free energy of the system has been characterized over the whole configuration space, the goal is to determine which is the preferential microstate under thermodynamic equilibrium conditions, and what the expected macroscopic electron transfer rate derived from it is. This requires two elements, the Boltzmann probability distribution over that configuration space, and the expression of the electron transfer rate as a function of the microstate of the system.

2.6.1 Boltzmann Distribution over the Configuration Space

Let γ be a microstate of a thermodynamic system within a configuration space Λ . According to the Boltzmann distribution, the probability $P(\gamma)$ of finding the system in that microstate at equilibrium at a temperature T depends on its total free energy:

$$P(\gamma) = \frac{\int_{\gamma} \exp\left(-\frac{F_{\text{total}}}{k_B T}\right) d\lambda}{\int_{\Lambda} \exp\left(-\frac{F_{\text{total}}}{k_B T}\right) d\Lambda}, \quad (2.84)$$

where F_{total} is the total free energy and k_B is the Boltzmann constant. The integral in the denominator corresponds to the partition function, normalizing the distribution over the whole space Λ , while the integral in the numerator considers the accumulated Boltzmann density over the range of configurations that defines γ [16].

The specific nature of γ depends on the problem under study. In this work, γ defines the relative orientation of the protein-electrode system, whose concrete geometric formulation is presented in the Methods chapter (Section 3.1.2).

In practice, the space Λ is discretized into a finite set of N sampled microstates. Approximating the integral by a pointwise evaluation of F_{total} at each microstate, the discrete probability distribution results in:

$$P_i = \frac{\exp\left(-\frac{F_{\text{total},i}}{k_B T}\right)}{\sum_{j=1}^N \exp\left(-\frac{F_{\text{total},j}}{k_B T}\right)}, \quad (2.85)$$

where the index $i = 1, \dots, N$ enumerates all the evaluated microstates. The one with the maximum probability P_i corresponds to the preferential configuration of the system at equilibrium.

2.6.2 Marcus Theory for Electron Transfer

The electron transfer process is explained by Marcus theory for non-adiabatic systems. In this type of reaction, the electronic coupling between the reactant pair is quite weak. This theoretical model expresses the electron transfer rate constant, k_{ET} , as follows:

$$k_{ET} = \frac{2\pi}{\hbar} \frac{H_{DA}^2}{\sqrt{4\pi\lambda RT}} e^{-\frac{(\Delta G_{\text{reac}} + \lambda)^2}{4\lambda RT}}. \quad (2.86)$$

This constant is described by the following factors: absolute temperature (T), the thermodynamic driving force of the reaction or standard free energy (ΔG_{reac}), the nuclear reorganization energy (λ), and the electronic coupling matrix element (H_{DA}).

The reorganization energy (λ) includes two effects, the fluctuations in the internal vibrational coordinates of the reactants and changes in the orientational coordinates of the surrounding solvent molecules following charge, while the coupling matrix element (H_{DA}), determines the dependence of k_{ET} on the electron pathway and decays exponentially with the separation distance ($\mathbf{r}$) between the donor and the acceptor.

$$H_{DA}^2 = (H_{DA}^0)^2 e^{-\beta \mathbf{r}}. \quad (2.87)$$

Here, β is an attenuation parameter identifying what the medium that surrounds the electrochemical system (the enzyme in this case) is doing. The H_{DA}^0 represents the maximum electronic coupling when the donor and acceptor are in contact. The biological environments of enzymes adsorbed onto an electrode may be complex and not necessarily reflected accurately by the Marcus theory; however, the rate at which the electron transfers is greatly influenced by the electron tunneling path and the critical distance between the electrode and active site [1, 11, 18].

Simplified Electron Transfer Rate Model

The analytical modeling of all the dynamic factors involved in electron transfer at a protein-surface interface, described by Marcus theory in Equation (2.86), can be very complex to model. For this reason, a simplification is adopted based on the nature of the simulations in the PyGBe software. Since these simulations use a PQR file that treats enzymes as rigid bodies with static dielectric boundaries in all their orientations, the vibrational reorganization energy (λ) is considered constant. Under this scenario, for the calculation of direct electron transfer, it is possible to group the thermodynamic and vibrational components of the Marcus equation into a single constant (C). Thus, the model assumes a classical hopping behavior, where the electron transfer rate decays exponentially with distance:

$$k_{ET} = C e^{-\beta \mathbf{r}}. \quad (2.88)$$

In this expression, β determines the decay rate through the protein medium. Although this parameter can vary depending on the secondary structure of the enzyme, in the absence of specific empirical information for each variant, a widely validated consensus value in the literature of around $\beta = 1.4 \text{ \AA}^{-1}$ is adopted [2, 19], which imposes a strong exponential penalty as the distance increases.

2.6.3 Expected Value of the Electron Transfer Rate

The macroscopic current produced by a protein-electrode system does not correspond to the rate of a single microstate, but to the weighted average of all the microstates accessible under thermodynamic equilibrium. This expected value is calculated as:

$$\langle k_{\text{ET}} \rangle = \sum_{i=1}^N P_i \cdot k_{\text{ET}}(\mathbf{r}_i), \quad (2.89)$$

where P_i is the Boltzmann probability of microstate i (Equation (2.85)), and $k_{\text{ET}}(\mathbf{r}_i)$ is the Marcus rate (Equation (2.88)) evaluated at the donor-acceptor distance r_i corresponding to that microstate [2]. As is the case with the configuration space Λ of Section 2.6.1, the explicit relation between each microstate i and its distance $\mathbf{r}_i$ depends on the concrete geometric parametrization adopted for the system, detailed in the Methods chapter.

Equation (2.89) constitutes the physical quantity that integrates the thermodynamic analysis, represented by the Boltzmann probability, with the kinetic analysis, represented by the Marcus rate, into a single macroscopic observable proportional to the current density of the electrode. A high value of $\langle k_{\text{ET}} \rangle$ indicates that the statistically most probable microstates are also kinetically favorable.

2.6.4 Expected Value of an Energetic Observable

The same weighted-average reasoning applied to $\langle k_{\text{ET}} \rangle$ in Section 2.6.3 extends naturally to any orientation-dependent energetic quantity. Let E_i denote the value of such a quantity for microstate i , for instance, the solvation energy difference $\Delta E_{\text{solv},i}$, the surface energy difference $\Delta E_{\text{surf},i}$, or their sum, the reduced interaction (binding) energy $\Delta E_{\text{bind},i}$ (Equation (2.46)). Its Boltzmann-weighted average over the configuration space Λ is defined as:

$$\langle E \rangle = \sum_{i=1}^N P_i E_i, \quad (2.90)$$

where P_i is the Boltzmann probability of microstate i (Equation (2.85)). While E_i describes the value of the observable at a single, specific microstate, $\langle E \rangle$ describes its average over the system as a whole, accounting for the relative weight of every accessible microstate under thermodynamic equilibrium.

2.6.5 Spearman Rank Correlation

Beyond the expected value of a single energetic observable (Section 2.6.4), it is also useful to compare how two different observables order the same set of sampled microstates, for instance whether the orientations favored by the solvation term $\Delta E_{\text{solv},i}$ remain favored once the surface term is included in $\Delta E_{\text{bind},i} = \Delta E_{\text{solv},i} + \Delta E_{\text{surf},i}$. The choice between Pearson's and Spearman's coefficients depends on what kind of association is of interest, Pearson's quantifies linear proportionality, whereas Spearman's quantifies monotonic association, regardless of the specific functional form relating the two variables. Since the question here is whether two observables induce the same ordering of the sampled microstates and not whether they are related by a straight line, Spearman's coefficient is the conceptually appropriate choice, rather than a nonparametric stand-in for Pearson's coefficient [20].

Given two observables E_i and E'_i evaluated at the same N microstates, each microstate is assigned a rank $r(E_i)$ and $r(E'_i)$ according to its position when the N values are sorted independently, from lowest to highest. The Spearman coefficient is then defined as:

$$\rho = 1 - \frac{6 \sum_{i=1}^N d_i^2}{N(N^2 - 1)}, \quad d_i = r(E_i) - r(E'_i), \quad (2.91)$$

where d_i is the rank difference for microstate i . A value of $\rho = 1$ indicates that both observables induce an identical ordering of the N microstates (a perfect monotonic relationship, not necessarily a linear one), $\rho = 0$ indicates no rank association, and $\rho = -1$ indicates a fully inverted ordering.

Unlike the Pearson coefficient, which requires the two observables to be proportional to one another to approach its extreme values, the Spearman coefficient is invariant under any monotonic transformation of either variable [20]. This makes it the appropriate choice when comparing the orientation ranking induced by ΔE_{solv} alone against that induced by the complete binding energy ΔE_{bind} (Section 4.3.1), since a high ρ value only certifies that the ranking of orientations is largely preserved, without implying that the surface term is a linear, proportional perturbation of the solvation term.

2.7 Richardson Extrapolation

Richardson extrapolation is a numerical technique [16] to estimate the exact solution of a system f_{exact} by using a series of calculations with different resolutions, from coarse to fine meshes. For this to work, the results must be within the asymptotic convergence range, which means they converge at a constant rate as the mesh is refined. To do this analysis, we need simulations with at least three different mesh densities. We define these results as f_i for $i = \{1, 2, 3\}$, where index 1 is the

finest mesh, 2 is the intermediate, and 3 is the coarsest. Similarly, the mesh spacing is identified as h_i following that same order. With these, we define a constant refinement ratio r , which shows how much the mesh is refined in each step $r = h_2/h_1 = h_3/h_2$. Here, the parameter h can represent spacing, area, or volume. With this, we can check the stability of the method by calculating the observed order of convergence p :

$$p = \frac{\log\left(\frac{f_3 - f_2}{f_2 - f_1}\right)}{\log(r)}. \quad (2.92)$$

If the value of p matches the theoretical order of the solver, it confirms that the simulations are in the asymptotic range. Under this condition, the estimated exact solution is calculated with the following equation:

$$f_{exact} \approx f_1 + \frac{f_1 - f_2}{r^p - 1}. \quad (2.93)$$

This f_{exact} value represents the continuum limit, which is the theoretical result we would get with an infinitely dense mesh. This ensures that the reported results are independent of the mesh geometry and truly represent the physical behavior of the system. Finally, this procedure allows us to quantify the approximation error of the mesh being used relative to the extrapolated exact value [16].

2.8 Software PyGBe

The boundary integral equations and the discretized energy expressions derived in the previous sections were solved in this work using PyGBe¹ (Python GPU Boundary Elements), an open-source software distributed under the MIT license and developed by Cooper and Barba [21].

Hybrid Architecture

PyGBe combines three layers. Python for the high-level control flow and the management of input and output files; C++ for the numerically intensive operations, mainly the evaluation of the single and double layer operators; and CUDA, accessed from Python through PyCUDA, to move the most expensive evaluations to graphics processing units (GPU). This hybrid architecture allows combining the flexibility of a high-level language with the performance needed to assemble and solve the dense linear systems described in the previous section.

¹<https://github.com/pygbe/pygbe>

Acceleration of the Matrix-Vector Product

Since the matrices resulting from the discretization are dense, evaluating the product $A\mathbf{x}$ directly has a cost of $\mathcal{O}(N^2)$. PyGBe reduces this cost through a treecode algorithm, which hierarchically groups the panels into an octree and approximates the joint contribution of distant groups of panels through a multipole expansion, instead of explicitly evaluating every panel to panel interaction. This reduces the complexity of the system to $\mathcal{O}(N \log N)$, which makes the mesh required to solve protein scale geometries computationally feasible.

Input Files

Running a calculation in PyGBe requires three types of input files, which separate the geometric, numerical, and physical aspects of the problem. The mesh files describe the triangulated discretization of each surface, generated externally with **NanoShaper** from a PQR file, which extends the standard PDB format by adding partial charges and Van der Waals radii for each atom (see Section 3.1.1), representing the atomic structure of the protein. A parameter file `.param` controls the numerical behavior of the solver, and a configuration file `.config` specifies the dielectric topology of the modeled system, that is the number and nature of the regions and surfaces involved, corresponding for example to one of the systems introduced in Section 2.4, together with the boundary condition prescribed on each surface.

2.9 State of the Art: Prior Work Directly Preceding This Thesis

Characterizing the interaction between an enzyme and a charged electrode as a function of spatial orientation is a high-dimensionality problem, the protein has thousands of atoms with a heterogeneous charge distribution, surrounded by an ionic solution whose behavior depends on pH and ionic strength, and must be evaluated in hundreds of different orientations. Although molecular dynamics offers the highest theoretical resolution, its computational cost is prohibitive for the protein sizes and orientational space required in this work [9, 16]. The approach adopted, described in detail in Chapter 2, is the implicit solvent model solved through the boundary element method (BEM) via PyGBe, which captures the dominant electrostatic physics at a fraction of the computational cost.

Rather than reviewing the broader field of protein-electrode simulation, this section focuses specifically on the small set of studies that form the direct research line leading to this thesis:

- **Cooper [16]:** Developed and validated PyGBe, a BEM software to solve the coupled Poisson and Poisson-Boltzmann equations in the context of the implicit solvent model. Unlike volumetric methods, the BEM discretizes only the molecular surface, reducing the degrees of freedom and representing the protein geometry with greater fidelity. Cooper validated the code for solvation and binding energies, and demonstrated its ability to predict preferential orientations of proteins on charged surfaces with results consistent with molecular dynamics and experimental observations.
- **Urzúa et al [9]:** Extended PyGBe to incorporate a homogeneous external electric field, allowing the study of the orientation of adsorbed proteins under applied potentials. The sampling of the orientational space was performed by defining the tilt angle and rotation angle with respect to the dipole moment vector of the protein. The model was validated against experimental measurements of enzymatic activity of trypsin adsorbed on carbon electrodes at different potentials, finding that a negative field favors orientations with greater exposure of the active sites, in agreement with the observations. The code associated with this work [3] constitutes the direct basis of the automation code implemented in the present thesis.
- **Ruiz et al [2]:** Applied the BEM approach with Poisson-Boltzmann to the study of the orientation and catalytic activity of a [NiFe] hydrogenase on a flat electrode, building energy maps and Boltzmann probability distributions as a function of pH, ionic strength and electrode potential. Connecting these distributions with an electron transfer rate model based on Marcus theory, the macroscopic current density was estimated under different operating conditions. The main finding was that pH is the dominant parameter in the orientational distribution and adsorption energy, followed by ionic strength, while the electrode potential produces a secondary but observable effect on the preferential orientations.
- **Terrani [1]:** Carried out the most complete study of this type applied to the enzymes that are the subject of the present thesis, examining the interaction of glucose oxidase (GOx) and bilirubin oxidase (BOx) with a capacitor electrode model using finite differences with the APBS software. Orientational energy landscapes and Boltzmann probability maps were built for pH = 5, 7 and 9, and electrode potentials of up to $\pm 10 k_B T / e$, finding a dominant influence of pH with significant qualitative changes in the preferential orientation of GOx when transitioning between pH 5 and pH 7, attributable to residues with pK_a values in that interval. In addition, the molecular input files for both enzymes were built and validated, including the force field parametrization for the non-standard cofactors, these files are reused in their entirety in the present work.

The present work extends this line of research using PyGBe on the same two enzymes studied

by Terrani [1], with the purpose of validating his findings through the BEM method and comparing both numerical approaches. To this end, a Python automation code is developed that extends the work of Urzúa et al. [3], incorporating improvements in the orientational sampling and in the consistency of the reference system between different pH values. Additionally, the calculation of the electron transfer rate and the macroscopic current based on the Marcus model of Ruiz et al. [2] is integrated, weighted by the Boltzmann probability distribution.

Chapter 3

Methods

The electrostatic simulations were carried out using the PyGBe software [21], which solves the linearized Poisson-Boltzmann equation at the protein-electrode interface using the boundary element method (BEM). The following sections describe, in order, the definition of the simulated system (Section 3.1), the mesh convergence study (Section 3.2), the simulation parameters (Section 3.3), the computational environment (Section 3.4), the computational cost of the simulations (Section 3.5), the automation scripts (Section 3.6) and the identification and resolution of a software bug regarding internal cavities in CPU mode (Section 3.7).

3.1 Definition of the Simulated System

3.1.1 Studied Enzymes

Two oxidoreductases used in glucose-based enzymatic fuel cells were studied: glucose oxidase (GOx, EC 1.1.3.4), responsible for the anodic glucose oxidation reaction, and bilirubin oxidase (BOx, EC 1.3.3.5), used at the cathode for the reduction of molecular oxygen. Their main structural characteristics are summarized in Table 3.1. Both enzymes are oxidases, so they use molecular O_2 as the final electron acceptor, and together they form the catalytic core of a glucose-based enzymatic fuel cell. Only the structural aspects relevant to the simulation setup are reported here.

Preparation of Molecular Structures

The molecular structures used as input for the electrostatic simulation were prepared by Terrani [1]. His procedure is summarized below, since the resulting PQR files were fully reused in this work.

Although GOx is biologically a homodimer, only chain A of the 3QVP file [22] (a monomer) was used. This decision was originally made by Terrani for two reasons: each subunit contains its

own FAD cofactor and is functionally independent, and studying a monomer substantially reduces the computational cost without losing relevant information about the interaction with the electrode. In this work the same choice was kept for the same reasons, and also because the high-resolution meshes used here would make doubling the protein volume computationally too expensive.

Solving the Poisson-Boltzmann equation numerically requires PQR files, which extend the PDB format by adding partial charges and Van der Waals radii. Terrani generated these files using PDB2PQR [23], which:

1. Completes the missing heavy atoms in the crystallographic structure.
2. Determines the protonation state of each ionizable residue at each pH using PROPKA [24].
3. Assigns charges and radii according to the AMBER99 force field [25].

The ligands (FAD, NAG, GOL, MAN/BMA in GOx; Cu and analogous ions in BOx) are not parametrized by AMBER99, so they received charges calculated with the AM1-BCC method [26, 27] (Antechamber/AmberTools) [25, 28]. The Cu atoms were assigned an oxidation state of +1, following the literature [29, 30], and the charges of the implicit hydrogens were transferred to the corresponding heavy atom.

Starting from these PQR files, this work generated the .xyzr files required by NanoShaper [31] to build the triangulated solvent-excluded surface (SES), which PyGBe uses as the calculation mesh. This conversion keeps the atomic coordinates and radii from the PQR file unchanged, and it was done with a custom preprocessing script.

When NanoShaper was run with cavity detection enabled, around 10 internal cavities per protein were identified. However, not all of them were modeled: each additional cavity requires an independent mesh in PyGBe and increases video memory (VRAM) usage considerably, which limits how many regions can be modeled at the same time. For this reason, only the largest cavities were included, since these are the ones most likely to affect the solvation energy, and in some cases they are also located near the active catalytic centers. Following this criterion, the 2 largest cavities were modeled for GOx (3QVP) [22], and the 1 relevant cavity was modeled for BOx (6IQZ) [32].

Reference Atom for Electron Transfer

Calculating the electron transfer rate, Equation (2.88), requires identifying a single representative atom at the active center of each enzyme. The distance from this atom to the electrode, $r(\theta, \varphi)$, is measured for each simulated orientation [2]:

- **Bilirubin oxidase (6IQZ):** The type 1 copper atom (T1). This metal center, located at the periphery of the protein, works as the entry point for electrons and can accept them directly from the electrode surface [33].
- **Glucose oxidase (3QVP):** The N5 nitrogen of the isoalloxazine ring in the FAD cofactor. This atom is the reactive core where catalysis takes place [34], that is, the site where electrons are transferred to the cofactor during glucose oxidation.

For both enzymes, the attenuation parameter was fixed at the consensus value from the literature, $\beta = 1.4 \text{ \AA}^{-1}$, which is discussed in the theoretical framework (Section 2.6.2).

Table 3.1: Structural characteristics of the studied enzymes. The net charge and the electron transfer atom are explained in the text. Approximate dimensions were taken from the atomic coordinates of the PQR file [1].

Property	Glucose oxidase (GOx)	Bilirubin oxidase (BOx)
PDB Code	3QVP	6IQZ
EC Number	1.1.3.4	1.3.3.5
Biological source	<i>Aspergillus niger</i>	<i>Myrothecium verrucaria</i>
Crystallographic resolution ($\text{\AA}$)	1.2	1.5
Molecular weight (kDa)	65.92	61.79
Modeled residues	581	535
Heavy atoms (homoenzyme)	4 452	4 390
Approx. dimensions ($x \times y \times z$, $\text{\AA}$)	$67 \times 70 \times 47$	$73 \times 60 \times 59$
Redox cofactor	FAD	Cu T1 + TNC
DET reference atom	N5 of FAD	Type 1 Cu (T1)
Modeled internal cavities	2	1
Net charge pH 5 (e)	+1.0	-2.0
Net charge pH 7 (e)	-25.0	-13.0
Net charge pH 9 (e)	-29.0	-17.0

3.1.2 Electrode Geometry and Orientation Sampling

The electrode was modeled as a flat rectangular dielectric interface with dimensions $100 \times 100 \times 10 \text{ \AA}$, with the active plane at $z = 0$. Terrani [1] showed, through a convergence study in APBS [23], that the electrostatic energies become stable for domains larger than $75 \text{ \AA}$. This means that the $100 \times 100 \text{ \AA}$ geometry leaves a free margin of 10 to $15 \text{ \AA}$ between the protein and the edge of the domain. Since PyGBe uses a different BEM scheme, this conclusion was checked independently by comparing it with a larger electrode of $250 \times 250 \times 10 \text{ \AA}$. The differences in solvation energy were below 0.009 kcal/mol , and the maximum difference in total binding energy was 0.033 kcal/mol , without changing the relative order between orientations. Based on this result, the $100 \times 100 \times 10 \text{ \AA}$

geometry was adopted.

The minimum distance between the electrode plane ($z = 0$) and the lowest point of the protein SES mesh was fixed at $H = 3 \text{ \AA}$ for all orientations. No explicit Stern layers were modeled around the electrode or the protein, so this distance is enough to make sure that the protein mesh does not overlap with the electrode, while still representing the adsorption condition. In the XY plane, the protein was centered at $(0, 0)$ relative to the electrode geometry, which extends from -50 to $+50 \text{ \AA}$ in both X and Y . Figure 3.1 shows this setup for BOx and GOx in the reference orientation.

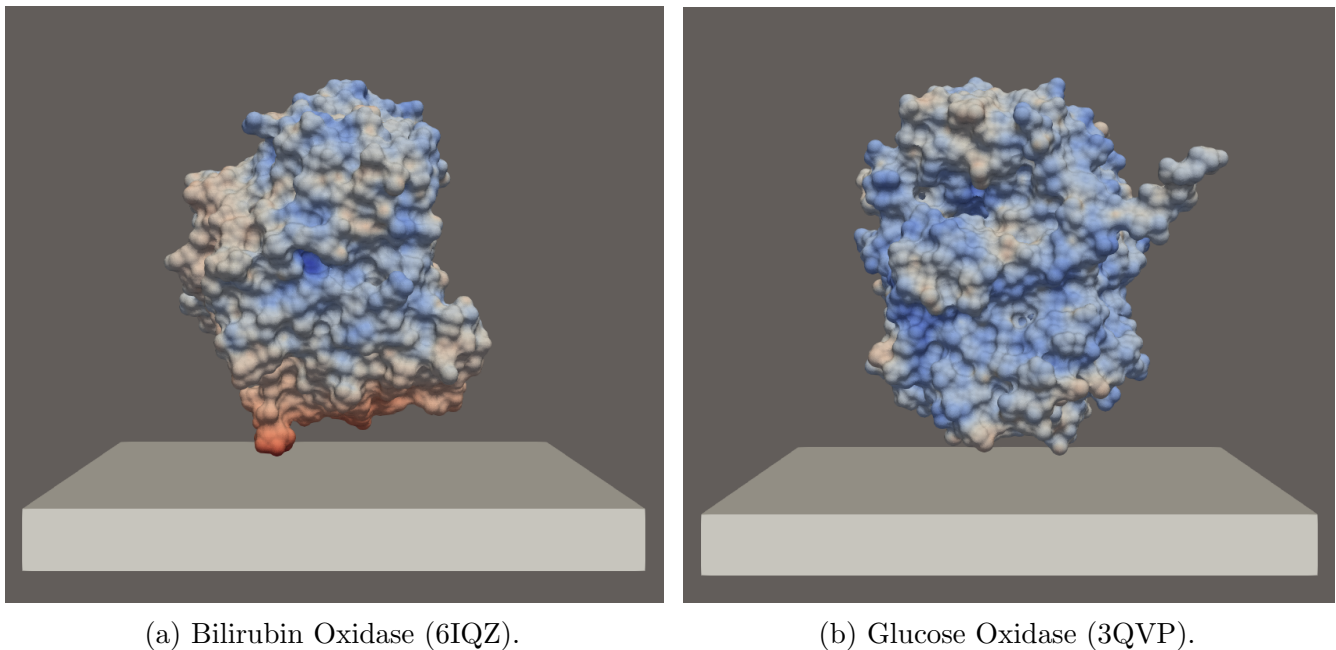


Figure 3.1: Triangulated mesh of the solvent-excluded surface (SES) of (a) Bilirubin Oxidase (6IQZ) and (b) Glucose Oxidase (3QVP) in their reference orientation ($\theta = 0^\circ$, $\varphi = 0^\circ$), $3 \text{ \AA}$ away from the electrode, modeled as a $100 \times 100 \times 10 \text{ \AA}$ plate. The surface electrostatic potential is shown using a color scale (blue: negative; red: positive), calculated at pH 9.

Reference System and Initial Orientation

The orientation of the protein relative to the electrode is described using two angles (Figure 3.2): the tilt angle $\theta \in [0, 180]$, measured from the electrode normal $\hat{\mathbf{z}}$, and the azimuthal rotation angle $\varphi \in [0, 360[$, measured around the minimum-inertia axis of the protein.

The initial orientation of the protein, that is, its reference position before any rotation, is set by aligning the minimum moment of inertia axis with the electrode normal $\hat{\mathbf{z}}$. This is different from the more common approach based on the dipole moment [9, 16]. This choice was made because the inertia tensor does not change with pH, it is calculated only from the heavy atoms (excluding

hydrogens), and their positions stay the same under different protonation states. The dipole moment, on the other hand, depends on the charge distribution, which does change with pH, so it would give a different starting orientation for each condition tested. Using the inertia tensor guarantees the same sampling space for every scenario, which makes it easier to compare results across different pH values.

After diagonalizing the inertia tensor over the heavy atoms, the three resulting eigenvectors define the principal axes of the protein. In the code, $\hat{\mathbf{x}}$ is assigned to the eigenvector with the largest moment of inertia, $\hat{\mathbf{y}}$ to the intermediate one, and $\hat{\mathbf{z}}$ to the smallest one, which matches the longest axis of the protein. This full three-axis assignment, and not just the alignment of $\hat{\mathbf{z}}$, is what makes the orientations fully reproducible. If only the tilt axis were fixed, the rotation angle φ would be defined relative to an arbitrary starting point, making it impossible to directly compare different conditions, or to compare with future work.

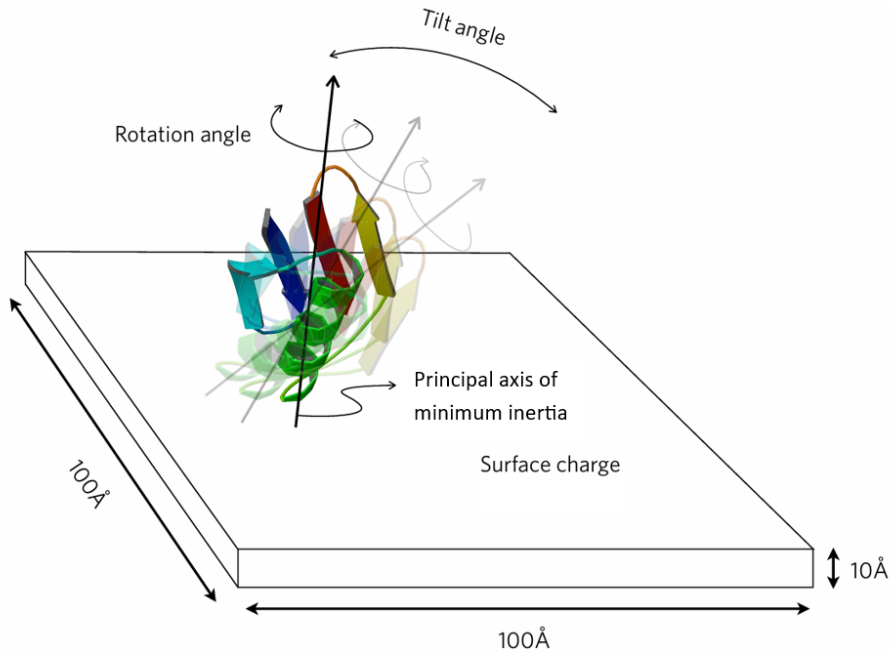


Figure 3.2: Definition of the orientation angles of the protein relative to the electrode: tilt angle θ and azimuthal rotation angle φ . In this work, the reference axis for the initial orientation is the minimum moment of inertia axis of the protein, not the dipole moment. Adapted from Cooper [16].

The rotations are applied as follows: the tilt angle θ rotates the protein counterclockwise around the $\hat{\mathbf{x}}$ axis, and the azimuthal angle φ rotates it counterclockwise around the minimum-inertia axis of the protein (initially aligned with $\hat{\mathbf{z}}$).

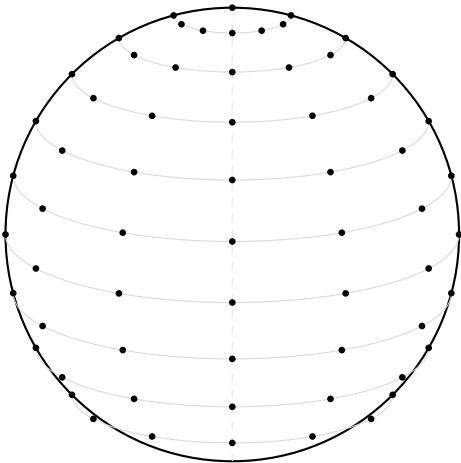
Solid Angle Sampling

The orientation space, which corresponds to the microstate space Λ in Equation (2.85), was sampled using 19 tilt values of θ , evenly spaced over $[0, 180]$ in steps of 10° . For each value of θ , the number of azimuthal positions N_φ scales with the sine of the angle, so the equator ($\theta = 90$) has 36 positions, while each pole has only 1. The full set of values is listed in Table 3.2, giving a total of $N = 406$ orientations.

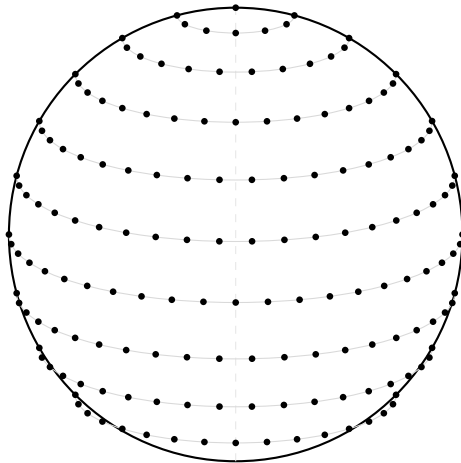
Table 3.2: Number of azimuthal rotation positions N_φ for each tilt value θ .

θ ($^\circ$)	0	10	20	30	40	50	60	70	80	90
N_φ	1	6	12	17	23	27	31	33	35	36
θ ($^\circ$)	100	110	120	130	140	150	160	170	180	Total
N_φ	35	33	31	27	23	17	12	6	1	406

This scheme makes sure that each orientation covers roughly the same area on the configuration sphere. It follows the solid angle principle and avoids the bias of a uniform (θ, φ) grid, where points cluster near the poles and give those orientations more weight than they should have in the Boltzmann averaging Equation (2.85). Figure 3.3 shows this contrast.



A. Geometric Sampling



B. Solid Angle Sampling

Figure 3.3: Comparison of orientation sampling schemes on the configuration sphere. **(A)** Uniform geometric sampling in (θ, φ) ; the points cluster at the poles, overrepresenting those orientations in the statistical averaging. **(B)** Uniform solid angle sampling; the point density is approximately constant over the whole sphere, assigning equal statistical weight to each region of the orientation space.

Furthermore, the Coulomb energy reported by PyGBe varied by less than 10^{-5} kcal/mol across the full set of sampled orientations for each protein and condition studied in this work, confirming

the cancellation predicted by Equation (2.46). Consequently, these negligible fluctuations are attributed to numerical error, specifically floating-point precision limits arising from coordinate transformations when rotating the PQR files.

3.1.3 Evaluated Electrochemical Conditions

Nine environmental scenarios per protein were evaluated, combining three pH conditions with three electrode potentials. Each scenario was computed over the 406 orientations, which gives $2 \text{ proteins} \times 9 \text{ scenarios} \times 406 \text{ orientations} = 7\,308$ independent PyGBe simulations.

pH 5, 7, and 9 were chosen because they cover the physiological range where both enzymes stay catalytically active without risk of protein denaturation [1]. The protonation state of each residue was determined by PROPKA during the preparation of the PQR files [1].

Three electrode potentials were simulated: $\phi_0 = -25.7, 0$, and $+25.7$ mV, which correspond to $-1, 0$, and $+1 k_B T/e$ at 298 K. The magnitude $|\phi_0| = 25.7 \text{ mV} = k_B T/e$ is the theoretical validity limit of the LPBE, since its derivation requires $|e\phi/k_B T| \ll 1$. Ruiz et al. [2], using the same PyGBe code, worked with potentials of up to $|\phi_0| = 50 \text{ mV}$ and found that the results stayed consistent even beyond this strict limit. However, this work used a more conservative criterion, staying within the regime where the linearization is rigorously valid.

3.2 Mesh Convergence Study

Applying this methodology to our system, the Richardson extrapolation was performed using the solvation energies (E_{solv}) obtained from three distinct mesh densities. The grid refinement was controlled by the Grid scale parameter with values of 1.0 (f_3 , coarsest), 2.0 (f_2 , intermediate), and 4.0 (f_1 , finest). Due to the non-linear nature of the triangulated mesh generation, the actual grid refinement ratios were $r_1 = h_3/h_2 = 3.6801$ and $r_2 = h_2/h_1 = 3.9468$, yielding an average ratio of $r_{\text{avg}} \approx 3.8135$.

By substituting the computed energies into Equation 2.92, the observed order of convergence was determined to be $p \approx 1.5098$. Consequently, the estimated continuum limit for the solvation energy on Equation 2.93 resulted in $f_{\text{exact}} = -1658.9185 \text{ kcal/mol}$. The detailed mesh parameters, calculated energies, and their relative errors with respect to this theoretical exact value are summarized in Table 3.3.

As shown in Table 3.3, the relative error drops sharply as the mesh is refined. For the coarsest configuration the error stands at 4.099%. Doubling the density parameter reduces the error to 0.543%. Finally, under the production condition selected for this work (Grid scale = 4.0), the relative error asymptotically decays to a marginal 0.072%. This corresponds to a minimal absolute difference of only -1.1953 kcal/mol away from the continuum limit, rigorously validating the choice of this mesh density to ensure highly accurate physical results with an optimized computational cost.

Table 3.3: Mesh convergence parameters and Richardson extrapolation results for the solvation energy.

Mesh Level	Grid Scale	Avg. Elem. Size ($\text{\AA}^2$)	E_{solv} (kcal/mol)	Relative Error (%)
Coarse (f_3)	1.0	0.30079	-1726.9143	4.099
Intermediate (f_2)	2.0	0.08173	-1667.9307	0.543
Fine (f_1)	4.0	0.02071	-1660.1130	0.072
<i>Extrapolation Outputs:</i>				
Average ratio $r_{\text{avg}} = 3.8135$		Observed convergence order $p = 1.5098$		
Extrapolated value $f_{\text{exact}} = -1658.9185$ kcal/mol		Fine mesh absolute error = -1.1953 kcal/mol		

3.3 Simulation Parameters and Input Configuration

For the preparation and execution of the simulations, two main tools are coupled: NanoShaper, responsible for generating the molecular meshes, and PyGBe, the solver based on the boundary element method (BEM).

The geometric definition of the solute is carried out through the NanoShaper configuration file, presented in Figure 3.5. The key parameters used are: `Grid_scale = 4.0`, which determines the density of the triangulated mesh and whose choice was validated by Richardson extrapolation (see Table 3.3); `Probe_Radius = 1.4 \AA`, corresponding to the standard radius of a water molecule used in the definition of the solvent excluded surface (SES); and `Cavity_Detection_Filling = true`, which enables the detection of internal cavities isolated from the solvent.

The execution of the PyGBe solver is governed by two complementary types of files that separate the numerical aspects from the topological ones. The general parameters file (`.param`) remains identical for all study cases considered in this work, which guarantees numerical consistency in the comparisons made. On the other hand, the configuration file (`.config`) is specific for each individual case, adapting to the particular geometries and physical conditions of each simulation.

Both files are presented in Figures 3.6 and 3.7, respectively.

Figure 3.6 defines the global numerical parameters of the solver: the floating point precision (`double`), the Gauss quadrature points per element ($K = 1$, $Nk = 5$, $K_fine = 19$), the residual convergence tolerance for the GMRES method (10^{-4}), the maximum number of iterations (1000), the multipolar expansion order of the *treecode* ($P = 4$), and the indicator `GPU = 1`, which activates hardware acceleration through CUDA.

Figure 3.7 defines the dielectric topology and boundary conditions of the molecular problem. The `FILE` blocks declare the system meshes and their interface type, while the `FIELD` blocks specify the physical properties of each contiguous dielectric region. In the present work, four regions are considered:

1. **The outer solvent:** The largest volume region, described by the linearized Poisson-Boltzmann equation (`LorY = 2`) with a dielectric constant $\varepsilon_2 = 80$ and a Debye screening parameter $\kappa = 0.125 \text{ \AA}^{-1}$.
2. **The protein interior:** Modeled by the Laplace equation (`LorY = 1`) with a dielectric constant $\varepsilon_1 = 3.23$ and no ionic penetration ($\kappa \approx 0$), hosting the point partial charges specified in the PQR file.
3. **The internal cavity interior:** Also described by the Laplace equation (`LorY = 1`), but with the dielectric properties of bulk water ($\varepsilon = 80$) and no associated charges.
4. **The electrode surface:** An interface with a Dirichlet boundary condition on which a constant surface electric potential (ϕ_0) is imposed.

The physical conditions represented by these parameters correspond to enzymes immersed in an implicit solvent that emulates water with salt ions at room temperature ($T = 298 \text{ K}$). A physiological monovalent salt solution ($z_j = \pm 1 e$) at a concentration of 0.15 mol/L is considered, equivalent to an ionic strength $I = 0.15 \text{ mol/L}$ [1]. Under these conditions, the Debye coupling parameter takes the value $\kappa \approx 0.125 \text{ \AA}^{-1}$, used consistently across all pH scenarios evaluated. The internal dielectric constant of the protein was set to $\varepsilon_1 = 3.23$ [35].

To manage the large number of study cases, together with their respective input files (`.config`, `.pqr`, meshes in `geometry/`) and output files (`output/`), the workspace is organized in a modular way. The hierarchical structure of the project directory, which hosts both the Python and Bash automation scripts and the structural data, is detailed in Figure 3.4. It is important to note that this figure shows the directory tree for a single representative case; in practice, the folder names

and file identifiers change for each simulation scenario, but the overall structure remains the same. Once these structural and configuration components are established, the operational workflow and the code execution logic are described in detail in the following section.

```

6iqz_pH_9_sensor_final_25.7mV/
├── 6iqz_pH_9/
│   ├── 6iqz_pH_9.param
│   ├── 6iqz_pH_9.pqr
│   ├── 6iqz_pH_9_EF1.0.config
│   ├── 6iqz_pH_9_reserva.pqr
│   ├── geometry/
│   │   ├── 6iqz_pH_9_gs40.face
│   │   ├── 6iqz_pH_9_gs40.vert
│   │   ├── 6iqz_pH_9_gs40_cavity_1.face
│   │   ├── 6iqz_pH_9_gs40_cavity_1.vert
│   │   ├── sensor_100x100x10_d02.face
│   │   └── sensor_100x100x10_d02.vert
│   ├── output/
│   │   ├── 2026-06-01-130419-output.log
│   │   ├── 2026-06-01-130419-resultspickle
│   │   ├── 2026-06-01-132517-phi.txt
│   │   ├── 2026-06-01-132527-phi_reac.txt
│   │   └── sensor_d02_100x100x10_1.4203e-4.phio
│   └── 6iqz_pH_9_ET.CU_3_0_180
├── 6iqz_pH_9_sensor_0_0_3_25.7mV
├── README.txt
├── __pycache__/
│   └── read_data.cpython-38.pyc
├── codebash_EF1.0/
│   ├── code.conformation.6iqz_pH9_sensor.cctval
│   ├── code.conformation.6iqz_pH9_sensor.tulio.sh
│   ├── logs_6iqz.6361533.err
│   ├── logs_6iqz.6361533.out
│   └── registro_EF1.0.log
├── conformation.py
├── electron_transfer.py
├── generador_config_file.py
├── move_protein_z-axis.py
├── nohup.out
├── plot_surface_fingerprint.py
├── read_data.py
├── run_log.txt
├── total_run.sh
└── vtk_generator_gr_linux.py

```

Figure 3.4: Example of the hierarchical directory structure for a single simulation case. Directory names and file identifiers vary across scenarios, but the overall layout remains the same.

```

Grid_scale           = 4.0
Grid_perfil          = 90.0
XYZR_FileName        = 6iqz_pH_9.xyzr
Build_epsilon_maps   = false
Build_status_map     = true
Save_Mesh_MSMS_Format = true
Compute_Vertex_Normals = true
Surface              = ses
Smooth_Mesh          = true
Skin_Surface_Parameter = 0.45
Cavity_Detection_Filling = true
Conditional_Volume_Filling_Value = 80.0
Keep_Water_Shaped_Cavities = false
Probe_Radius         = 1.4
Accurate_Triangulation = true
Triangulation         = true
Check_duplicated_vertices = true
Save_Status_map       = false
Save_PovRay           = false

```

Figure 3.5: NanoShaper configuration file for the triangulation of the SES surface.

```

Precision    double
K            1
Nk           5
K_fine       19
threshold    0.8
BSZ          128
restart      200
tolerance    1e-4
max_iter     1000
P            4
eps          1e-12
NCRIT        1000
theta        0.5
GPU          1

```

Figure 3.6: Global numerical parameters of the PyGBe solver (.param).

FILE	geometry/6iqz_pH_9_gs40				dielectric_interface		
FILE	geometry/6iqz_pH_9_gs40_cavity_1				internal_cavity		
FILE	geometry/sensor_100x100x10_d02				dirichlet_surface		
	sensor_d02_100x100x10_1.4203e-4.phi0						

PARAM	LorY	E?	Dielec	kappa	charges?	coulomb?	
	charge_file			Nparent	parent	Nchild	children
FIELD	2	0	80	0.125	0	0	
	NA			0	NA	2	0 2
FIELD	1	1	3.23	1e-12	1	1	
	6iqz_pH_9.pqr			1	0	1	1
FIELD	1	0	80	1e-12	0	0	
	NA			1	1	0	NA
FIELD	NA	1	NA	NA	0	0	
	NA			1	2	0	NA

Figure 3.7: Dielectric topology and boundary conditions of the molecular system (.config).

3.4 Computational Environment

The simulations were run in two different environments. The first one, called Tulio, is a shared computing server of the Numerical Molecular Modeling Group at UTFSM ¹, located in the Department of Mechanical Engineering at UTFSM, with no queue system. The second one is the high-performance cluster of the Scientific and Technological Center of Valparaiso (CCTVal), managed with SLURM, which was accessed through two partitions: batch (CPU nodes, wn35 and wn36) and gpuk (GPU nodes, gp08 and gp09). Jobs were launched in parallel on different nodes, but each job ran on a single node. The hardware specifications per node are summarized in Table 3.4.

Table 3.4: Hardware specifications of the simulation environments (specifications per node in CCTVal).

Environment	CPU	GPU	RAM
Tulio	2× Intel Xeon E5-2680 v3 @ 2.50 GHz (48 logical threads)	NVIDIA Tesla K40c (12 GB VRAM)	96 GB
CCTVal (batch)	2× Intel Xeon Gold 6136 @ 3.00 GHz (24 physical cores, 48 logical threads)	-	256 GB
CCTVal (gpuk)	Intel Xeon E5-2643 v2 @ 3.50 GHz (12 physical cores)	2× NVIDIA Tesla K20m (5 GB VRAM each)	64 GB

In Tulio, PyGBe was pre-installed by the research group. In CCTVal, it was installed by the

¹<https://bem4solvation.github.io>

author following the official repository guide ², under a conda environment managed with Anaconda3 2020.11. During the installation, an incompatibility was detected when compiling PyGBe with NumPy 1.24 or higher, caused by the removal of the numpy.bool alias in that version; for this reason NumPy 1.23.5 was used. The versions of all relevant dependencies in both environments are detailed in Table 3.5.

Table 3.5: Versions of PyGBe dependencies in each simulation environment.

Dependency	Tulio	CCTVal
PyGBe	0.3.2+69.gb857d66	0.3.2+72.gb268acd ^b
Python	3.8	3.8.20
NumPy	1.24.4	1.23.5
SciPy	1.10.1	1.10.1
PyCUDA	2024.1.1	2019.1.2
CUDA	11.2	10.1
SWIG	- ^a	3.0.12
matplotlib	3.7.5	3.7.2

^a Not required, PyGBe was already pre-compiled in Tulio.

^b Environment updated from version 0.3.2+70.g35648d5 after resolving the CPU cavity bug described in Section 3.7.

3.5 Computational Cost

Table 3.6 summarizes, for each of the conditions studied, the average wall-clock time required by PyGBe to solve a single orientation, together with the average number of GMRES iterations required for convergence. All large-scale runs (the full set of 406 orientations per condition) were executed on CCTVal (Section 3.4), Tulio was used only for code development, testing, and isolated single-orientation runs, such as those used to generate the electrostatic fingerprints presented in Section 4.3.4.

Multiplying the average time per orientation by the 406 sampled orientations gives the total computational cost of each condition, summing over all conditions yields a combined sequential cost of approximately 115 days. This total was substantially reduced in practice to roughly 25 days by distributing the 18 conditions across multiple nodes in parallel

Most conditions were run on a single GPU node. The exception is GOx at pH 5, which due to GPU availability at the time, was run on CPU nodes instead (40–48 physical cores), resulting in per-orientation times roughly double those of comparable GPU runs, and, for $\Phi_e = -25.7$ mV

²<https://github.com/pygbe/pygbe/blob/master/docs/installation.rst>

specifically, the highest per-orientation cost of the entire dataset. It is important to note that the apparent gap between GPU and CPU performance in Table 3.6 should not be overstated for the CCTVal environment specifically. A dedicated scalability test, run on CCTVal for a single representative case, shows that the actual GPU-to-CPU speedup there is more modest than it might first appear. Table 3.7 reports the wall-clock time and parallel efficiency for this test as a function of the number of CPU cores, alongside the equivalent GPU run, both on CCTVal. Increasing the CPU core count from 3 to 48 yields a speedup of $10.39\times$, with efficiency dropping from near-ideal (0.986 at 6 cores) to 0.649 at 48 cores, consistent with the expected loss of parallel efficiency at higher core counts due to communication overhead. The GPU run, in turn, is only about $1.4\times$ faster than the 48-core CPU run (629.5 s versus 873.7 s), a considerably smaller margin than the raw single-GPU versus single-CPU-node comparison in Table 3.6 might suggest, since that comparison does not account for the number of CPU cores actually available per node. Accordingly, running 3QPV at pH 5 on GPU instead of CPU on CCTVal would likely not have produced a substantially faster result. This comparability between devices is specific to the CCTVal hardware, however on Tulio (Table 3.4), CPU runs were observed to be roughly 4–6 times slower than GPU runs.

Table 3.6: Average computation time per orientation and average number of GMRES iterations, for each of the conditions studied. All runs executed on CCTVal.

Protein	pH	Φ_e (mV)	Device	Average time/orient (s)	GMRES iter
6IQZ	5	+25.7	GPU	1490.6	74
6IQZ	5	0	GPU	1487.6	74
6IQZ	5	−25.7	GPU	1489.2	74
6IQZ	7	+25.7	GPU	1212.9	59
6IQZ	7	0	GPU	1212.6	59
6IQZ	7	−25.7	GPU	1212.5	59
6IQZ	9	+25.7	GPU	1068.9	51
6IQZ	9	0	GPU	1068.6	51
6IQZ	9	−25.7	GPU	1066.8	51
3QVP	5	+25.7	CPU (48)	2148.3	75
3QVP	5	0	CPU (48)	2148.6	75
3QVP	5	−25.7	CPU (40)	2579.2	75
3QVP	7	+25.7	GPU	1074.7	43
3QVP	7	0	GPU	1074.8	43
3QVP	7	−25.7	GPU	1074.2	43
3QVP	9	+25.7	GPU	1032.9	41
3QVP	9	0	GPU	1032.3	41
3QVP	9	−25.7	GPU	1032.4	41

Interestingly, the number of GMRES iterations decreases monotonically with pH within each protein (BOx: $74 \rightarrow 59 \rightarrow 51$; GOx: $75 \rightarrow 43 \rightarrow 41$), loosely tracking the increase in the protein’s

net charge magnitude (Table 3.1). This is only a preliminary observation, not a firm conclusion, confirming a causal link would require examining the conditioning of the underlying linear system. It is nonetheless consistent with a broader pattern seen throughout the Results chapter, where near-neutral charge conditions (pH 5 for both enzymes) repeatedly show atypical behavior relative to the other pH values studied.

Table 3.7: Scalability of PyGBe with the number of CPU cores, compared to a single GPU run, for a representative case in CCTVal.

Device	N_{CPU}	Time (s)	Speedup (Efficiency)
CPU	3	9075.8	1.00 (1.000)
CPU	6	4604.0	1.97 (0.986)
CPU	12	2453.8	3.70 (0.925)
CPU	24	1403.4	6.47 (0.808)
CPU	48	873.7	10.39 (0.649)
GPU	—	629.5	14.42 (—)

3.6 Computational Code

3.6.1 Script Description

The main components of the computational system developed in this work are briefly described below, before detailing its execution flow.

- **split_cavities.py:** From the unified triangulated mesh generated by NanoShaper, it identifies and separates the disconnected geometric components: the outer shell of the molecule (`split_0`) and each internal cavity (`split_1`, `split_2`, ...).
- **generador_config_file.py:** Generates a copy of the PyGBe `.config` file with the paths of the triangulated meshes and the PQR file. These are updated using a unique suffix per run, preventing successive iterations from overwriting the base files.
- **move_protein_z-axis.py:** Performs the rigid geometric transformation of the molecule, aligns it with its principal axis of inertia, applies the tilt (θ) and azimuth (ϕ) rotations, and translates it to a distance H Å above the XY plane at the origin $(0,0)$.
- **conformation.py:** The main engine of the orientation loop. It generates the solid-angle sampling grid, coordinates the calls to `move_protein_z-axis.py` and the PyGBe solver for each orientation, and consolidates the energetic results into a single output file.

- **read_data.py:** An input/output utility library shared between modules. It implements functions to read `.pqr`, `.vert`, `.face`, `.config`, and `.param` files in the formats required by PyGBe.
- **electron_transfer.py:** Executed independently after the simulation. It replicates the orientation grid, determines the distance between the selected active atom and the electrode plane in each orientation, and calculates the electron transfer rate k_{ET} using the Marcus/Moser-Dutton model.
- **plot_surface_fingerprint.py:** Generates the electrostatic fingerprint of the biosensor for a specific orientation: a 2D heatmap of the induced surface charge density σ on the electrode, overlaid with the position of the protein residues closest to this surface.
- **vtk_generator_gr_linux.py:** Converts the meshes and the electrostatic potential field calculated by PyGBe into `.vtk` files for 3D visualization in ParaView.
- **Codigo_Post_Final_Christian.py:** Performs the final statistical analysis over the complete set of simulation cases. It calculates Boltzmann probabilities, weighted electron transfer rates, and their expected value, producing four heatmaps in the orientation space (θ, ϕ) and a report of the optimal configurations.
- **code_conformation_tulio.sh:** Bash script for execution on the local server. It sets up the computational environment and coordinates the main flow by consecutively calling `generador_config_file.py` and `conformation.py`.
- **code_conformation_cctval:** Homologous script to the previous one, adapted for execution on the high-performance cluster using the SLURM manager. It implements the same calling flow, adding the hardware reservation directives (GPU, CPU, and memory), the log redirection (`.out/.err`) and the loading of cluster modules (CUDA and Anaconda).

3.6.2 Computational Workflow

In order to detail the procedure in a concrete and reproducible way, the workflow is described using the simulation of the 6iqz enzyme at pH 9 facing the electrode charged at +25.7 mV as a reference case. Its directory structure is illustrated in Figure 3.4. The same procedure was systematically applied to the 18 simulation cases considered in this work (2 proteins $\times$ 3 pH values $\times$ 3 electrode potentials). Additionally, a complete flowchart of the execution pipeline is provided in Figure 3.8. It is important to emphasize that this flowchart represents the general algorithmic architecture of the code, applicable to any molecular system, despite being exemplified here through the reference

case. This diagram maps out the sequential phases and computational routines detailed in the following subsections.

Phase 0: Manual Preparation of the Simulation environment

This phase is a manual prerequisite that must be completed before launching any automated script. Its objective is to build the PyGBe working folder with all the necessary files in the structure and nomenclature required by the solver.

0.1 Obtaining and standardizing the PQR file: The `.pqr` file contains the atomic positions, partial charges, and van der Waals radio of the protein. In this work, the file was obtained from previous work [1]; in general terms, it can be generated using the PDB2PQR server or tools like APBS. A common issue is that non-standard residues and catalytic metals, in this case the copper centers of Bilirubin Oxidase, appear in the original file under the `HETATM` tag. Since PyGBe and the developed post-processing modules exclusively expect the `ATOM` tag, the file was standardized using the following `awk` command executed in the cluster terminal:

```
LC_NUMERIC=C awk '
  $1=="ATOM" || $1=="HETATM" {
    printf "ATOM %6s %-4s %-4s %4s %12.5f %12.5f %12.5f %10s %10s\n",
      $2,$3,$4,$5,$6,$7,$8,$9,$10
  }
  $1=="TER" { print "TER" }
  $1=="END" { print "END" }
' proteina.pqr > proteina_std.pqr
```

It should be noted that the resulting file features fixed-width columns based on precise character positions. This structural characteristic is directly exploited by the script `move_protein_z-axis.py`, where the instruction `atom_name = line[12:17].strip()` extracts the atom name using absolute slicing rather than whitespace separation. Consequently, this implicit format dependency must be carefully considered if the PQR file is generated using a different external tool.

0.2 Generation of the XYZR file: NanoShaper requires an input file in `.xyzr` format, with columns for position (x, y, z) and atomic radius r (van der Waals radius). This file was generated directly from the standardized PQR file using:

```
awk '{print $6, $7, $8, $10}' protein_std.pqr > protein.xyzr
```

0.3 Configuration and execution of NanoShaper: NanoShaper is the software used for the triangulation of the molecular surface. From the `.xyzr` file, it generates a triangulated mesh of the

solvent excluded surface (SES). The configuration parameters used in this work are presented in Figure 3.5. The execution was carried out with:

```
./NanoShaper config
```

NanoShaper produced two output files: `triangulatedSurf.face` (triangles defined by vertex indices in base 1) and `triangulatedSurf.vert` (Cartesian coordinates of the vertices). This unified mesh contains both the outer shell of the molecule and all internal cavities as a single geometric entity.

0.4 Mesh splitting: PyGBe distinguishes the outer shell from the internal cavities through different surface types in the `.config` file (`dielectric_interface` and `internal_cavity`, respectively), so it is necessary to separate both components into independent files. The script `split_cavities.py` uses the `trimesh` library to identify the disconnected components of the unified mesh through `mesh.split()`, storing each resulting submesh as `triangulatedSurf_split_i.face` and `triangulatedSurf_split_i.vert`. By convention, index 0 corresponds to the main outer shell and the subsequent indices to the internal cavities. The number of triangles of each component is reported to the console for verification.

0.5 PyGBe folder organization: Once the separated meshes were obtained, all files were renamed and organized according to the canonical directory structure of PyGBe, as illustrated for the reference case in Figure 3.4. The main mesh file was named `6iqz_pH_9_gs40.{face,vert}`, where the suffix `gs40` indicates the grid scale used in NanoShaper, and the internal cavity mesh was named `6iqz_pH_9_gs40_cavity_1.{face,vert}`. The standardized PQR file was copied as `6iqz_pH_9.pqr`. The electrode geometry was stored in `sensor_100x100x10_d02.{face,vert}` [3], and the Dirichlet boundary condition was specified in the file `sensor_d02_100x100x10_1.4203e-4.phi0`, which imposes a potential of +25.7 mV on the sensor surface. Additionally, to evaluate a zero potential on the electrode, a sensor file ending in `1.0000e-12.phi0` was used; this minimal value simulates a potential of 1.81×10^{-7} mV, which is practically zero, preventing division by zero errors in the code execution. The electrode consists of a flat surface of $100 \times 100 \times 10$ Å, with its upper face aligned with the *XY* plane at the origin of coordinates.

0.6 PyGBe configuration files: The execution of the solver is governed by two complementary files: `.param` and `.config`, which are presented in Figures 3.6 and 3.7, respectively. While the former establishes the global numerical parameters of the solver, the latter (`.config`) must be strictly adapted to each simulation scenario. In this way, the configuration file defines the access

paths to the meshes, the names of the structural files, and the particular physical conditions that characterize each independent study case.

0.7 Calculation of reference energies (isolated system): `Codigo.Post.Final.Christian.py` requires as reference the solvation energies $E_{\text{solv}}^{\text{iso}}$ and surface energies $E_{\text{surf}}^{\text{iso}}$ of the molecule and the electrode independently, that is, without mutual interaction. To obtain them, two additional PyGBe directories were built: one containing only the protein and its cavities (without the sensor in the `.config`), and another containing only the electrode (without the protein). In each case, PyGBe was executed once and the resulting energy values were recorded manually in the `casos_simulacion` dictionary of the post-processing script. As an example, for enzyme 6iqz at pH 9, the values obtained were $E_{\text{solv}} = -1660,11$ kcal/mol and $E_{\text{surf}} \approx 0$ kcal/mol for the sensor without applied potential, or $E_{\text{surf}} = -11,82$ kcal/mol for the electrode charged at +25,7 mV.

Phase 1: Orchestration of the Computing Cluster

The simulations were launched using Bash scripts that differ depending on the available computing environment, although both execute the same sequence of calls to the Python modules.

On the local server Tulio, the script `total_run.sh` records time stamps in `run_log.txt`, accesses the subdirectory `code.bash.EF1.0/` and launches the driver `code_conformation_6iqz_ph9_sensor_tulio.sh`, redirecting all stdout and stderr to the log file `registro_EF1.0.log`. This log contains the informative banner, the complete angular grid, and the diagnostic output for each orientation.

On the high performance cluster CCTval, the equivalent script `code_conformation_6iqz_ph9_sensor_cctval` was submitted to the SLURM queue manager using `sbatch`. The `#SBATCH` headers define the job resources such as the number of nodes, GPUs, memory, and maximum wall time. In addition, the script configures the Conda and CUDA environments and creates `$TMPDIR` for the CUDA compilation cache, which is necessary to avoid permission conflicts on shared filesystems. Once the job is completed, SLURM automatically generates the log files `logs_6iqz-{JOB}.out` and `logs_6iqz-{JOB}.err` and sends an email notification to the user.

It is important to mention that, before launching, both scripts require a manual configuration that depends on the specific simulation case. In particular, the following must be updated: the names of the working directories and geometries (`.face/.vert`); the tilt angle range (minimum, maximum and number of subdivisions, e.g., 0° , 180° , $N = 19$); and the minimum distance H between the outer mesh of the molecule and the electrode plane.

Phase 2: Generation of the Dynamic Configuration File

The script `generador_config_file.py` is the first one called by the Bash driver. It receives as arguments the case folder, the base names of the PQR and the mesh, the tilt range, the height H , and the electric field EF . The script builds the suffix `name1 = _3_0-180` and generates a modified copy of the base `.config` file, updating three types of paths: the main protein mesh, the internal cavities, and the `.pqr` file, adding the suffix to each one. The sensor lines and the electric field are copied without any modification. This `.config` file with the suffix works as a template for all the orientations, its mesh paths will be overwritten iteratively by `move_protein_z-axis.py` in the following phase.

Phase 3: Orientation Loop Engine

The script `conformation.py` is the central component of the pipeline. Before entering the loop, it reads the generated `.config` file to extract the sensor paths and the potential file, which will be included in the final report. Then, it builds the orientation grid using solid angle sampling: for each of the `tilt_N = 19` values of θ in $[0, 180]$, it calculates $N_\phi = \lfloor 36 \cdot \sin(\theta) \rfloor$ uniform rotations in $[0, 360[$. The exact poles ($\theta = 0$ and $\theta = 180$) produce only 1 point each, while the equator ($\theta = 90$) produces 36 points. The total number of orientations is 406. For each pair (θ_i, ϕ_i) , the loop executes three sub-steps in sequence and accumulates the results.

Phase 4: Geometric Transformation

The geometric transformation is executed by `move_protein_z-axis.py`, which is called from `conformation.py` using `os.system`. This script performs the rigid movement of the protein following the workflow described below.

First, the `read_data.readpqr` function reads the atomic positions and charges from the initial PQR file. Only heavy atoms are considered (excluding hydrogens) to calculate a realistic inertia tensor. This decision ensures that the principal axes are consistent in different protonation states, because the hydrogen amount changes with the pH, but the heavy atom structure remains the same. Then, the eigenvalues and eigenvectors of the inertia tensor are calculated using `numpy.linalg.eigh` and ordered from the highest to the lowest eigenvalue.

At this point, a chirality check is applied: if $\det(\mathbf{V}) < 0$, where $\mathbf{V}$ is the matrix of eigenvectors, one of the eigenvectors is inverted to guarantee a proper rotation. This correction is not only numerical; it is a biological necessity. Without this step, the diagonalization can produce a transformation with determinant -1 , which represents a combination of a rotation and a mirror

reflection. This ambiguity is not a limitation of the numerical implementation, but a direct consequence of the mathematics of the eigenvalue problem itself, if $\mathbf{v}$ is a solution of $(A - \lambda I)\mathbf{x} = 0$, then $-\mathbf{v}$ is equally valid, since the mathematical problem only defines the alignment of the orthogonal basis but is indifferent to its handedness, so the sign of each eigenvector is undefined by construction. Any correct implementation, including `numpy.linalg.eigh`, faithfully respects this mathematical property. Therefore, the chirality correction must always be applied explicitly. Without it, the stereochemistry of the protein would be inverted, invalidating any subsequent biological analysis.

Once the rotation matrix is defined, the main mesh is detected with `glob` using the base name of the files. Therefore, it is a requirement that the external surface and all the internal cavity meshes use the same name format, with only a different number at the end. For example, `6iqz_pH_9_gs40.vert` and `6iqz_pH_9_gs40_cavity_1.vert` must share the base name `6iqz_pH_9_gs40` that is written in the `.config` file. This permits the script to automatically include all the cavities in the transformation without specifying them one by one.

After that, the tilt (θ) and azimuth (ϕ) rotations are applied, and the molecule is translated to place its lowest point exactly at H Å above the XY plane at the origin $(0, 0)$. The new coordinates are saved in a new `.pqr` file and in new `.vert/.face` files, adding the suffix `_3.0-180`.

Finally, all the reading and writing operations in this phase are done by `read_data.py`. This script contains functions like `readpqr` (for atoms and charges), `readvertex` (for meshes), and `read_fields` and `read_surface` (to read the `.config` file of PyGBe), among others.

Phase 5: BEM Solver (PyGBe)

PyGBe receives the rotated geometry and the updated `.config` file. It solves the Poisson-Boltzmann equation using the Boundary Element Method (BEM). The solver produces an output log (`output_aux...`) that contains the solvation energy (E_{solv}), the surface energy (E_{surf}), the Coulomb energy (E_{coul}), and the total computation time.

Phase 6: Results Extraction and Cleanup

The function `scanOutput()`, defined inside `conformation.py`, parses the output log searching for E_{solv} , E_{surf} , E_{coul} , and the computation time. The values are accumulated in lists during the loop. Once the loop is finished, the final results file is written with the `.param` header, file metadata, and the complete table of energies for each orientation. After that, all temporary files are deleted: the

auxiliary log, the `.config` with suffix, the PQR and meshes with suffix, and the `.txt` and `.pickle` files inside the `output/` directory.

Phase 7: Electron Transfer Analysis

The script `electron_transfer.py` is executed independently and manually after the complete loop is finished. It reads the original protein mesh (`.vert`) and the original PQR file, and reconstructs the same grid of 406 orientations using the same solid angle algorithm. For each orientation, it calculates the position of the selected active atom after applying the same geometric transformation described in Phase 4. The vertical distance r from that atom to the sensor plane is then used in the simplified Marcus/Moser-Dutton model of the electron transfer rate Equation (2.88). The final output is a file containing the orientations, the distance r , and the calculated electron transfer rate constant for each case.

Phase 8: Optional Visualization Analyses

Once the orientation of interest is identified (typically the one with the highest Boltzmann probability), two optional analyses can be executed independently. Both require the `phi.txt` file generated by PyGBe for that specific orientation.

The script `vtk_generator_gr_linux.py` converts the meshes and the electrostatic potential field calculated by PyGBe into `.vtk` files for three-dimensional visualization in ParaView. The script `plot_surface_fingerprint.py` reads the sensor mesh and extracts the $\partial\phi/\partial n$ values from the second half of the `phi` vector, converts them to induced charge density σ using Equation (2.43), and identifies the 10 closest residues to the sensor using the alpha carbon atom as the representative position of each amino acid. This total density is then referenced against an additional PyGBe run of the isolated electrode under the same applied potential, isolating the electrostatic footprint specifically attributable to the adsorbed protein according to Equation (2.47), while the total induced charge is computed according to Equation (2.83). The output is a 2D heat map of $\Delta\sigma$ in a top view of the sensor, with the closest amino acids labeled by name and residue number and exported as a PDF.

Phase 9: Integrated Post-Processing

The script `Codigo_Post_Final_Christian.py` is executed independently, outside the main loop, using the output files collected from all clusters and organized manually into folders by case. It operates on the 18 simulation cases ($2 \text{ proteins} \times 3 \text{ pH values} \times 3 \text{ potentials}$). For each case, it reads the energy and electron transfer files, calculates the Boltzmann probabilities, and combines

the thermodynamic and kinetic information to obtain the electron transfer rate constant weighted by orientation.

The script produces four heat maps in the (θ, ϕ) space, exported as PDF files:

1. `heatmap_1_energy.pdf`: distribution of the net interaction energy ΔE_{bind} .
2. `heatmap_2_probability.pdf`: Boltzmann probability distribution P_i .
3. `heatmap_3_raw_k_et.pdf`: electron transfer rate constant k_{ET} in logarithmic scale, with the kinetic maximum indicated.
4. `heatmap_4_weighted_k_et_log.pdf`: product $k_{\text{ET}}(r_i) \cdot P_i$ in logarithmic scale, marking the orientation that maximizes the total macroscopic current.

In addition, a text report `Report_{case}.txt` is generated, listing the top three and bottom three orientations for each criterion (thermodynamic, kinetic, and macroscopic), together with their percentage contribution to the total current.

Note: Storage Management During the Loop

PyGBe normally generates four files per orientation inside the `output/` directory: a text log (`*-output.log`), a results pickle (`*-resultspickle`), and two potential files (`*-phi.txt` and `*-phi_reac.txt`). The latter two can be relatively large; with 406 orientations, the `phi` files can easily exceed 15 GB of disk space (for the proteins analyzed in this study and the mesh density used). For this reason, `conformation.py` deletes the `.txt` and `.pickle` files from the `output/` directory at the end of each iteration, keeping only the log file:

```
os.system(f'rm -f {Nombre_carpeta_principal}/output/*.txt
{Nombre_carpeta_principal}/output/*pickle')
```

At the end of the complete loop, the auxiliary log, the suffixed `.config`, the suffixed `.pqr`, and all suffixed `.vert/.face` meshes are also deleted. As a consequence, to run the optional analyses described in Phase 8, the procedure above must be followed: execute the pipeline for a single orientation with the cleanup lines commented out.

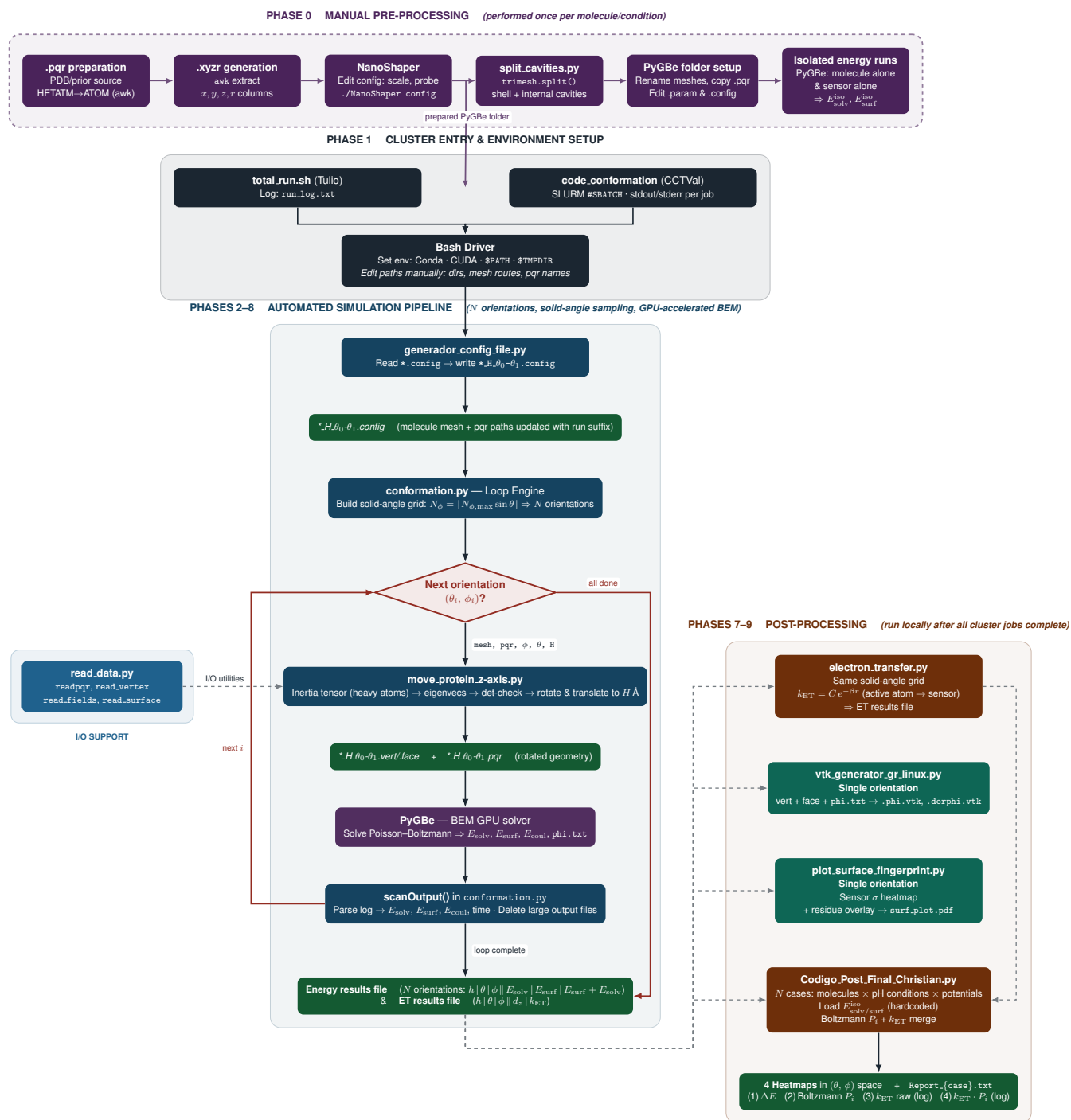


Figure 3.8: Algorithmic flowchart of the computational pipeline developed for protein–electrode interaction energy simulations. The system is organized into operational contexts: manual pre-processing (Phase 0), cluster entry orchestration (Phase 1), automated GPU-accelerated BEM simulation pipeline (Phases 2–8), and local statistical post-processing (Phases 7–9), as detailed in Section 3.6.2.

3.7 PyGBe Bug: Handling of Cavities in CPU Mode

During the simulations of this work, an inconsistency was detected between the results obtained in the local GPU server (Tulio) and those obtained in the CPU nodes of the CCTVal cluster, using identical input files. It was verified that the PyGBe outputs matched in geometry, number of elements, equations, regions, parameters, and Coulomb and surface energies, except for the expected CPU execution warning. The only real differences were the solvation energy and a noticeable increase in the GMRES iterations, which ruled out a configuration or data problem and focused the diagnosis on the linear system solving process.

For geometries with a single external dielectric interface (no internal cavities), the results were consistent between GPU and CPU. When internal cavities were incorporated, however, the solvation energy in CPU deviated significantly from the GPU reference. This was first observed for the 3QVP protein at pH 7 in front of an electrode charged at 25.7 mV:

- Tulio (GPU): -2417 kcal/mol
- CCTVal (CPU): -2852 kcal/mol

This resulted in a discrepancy greater than 10%, well above the expected numerical tolerance. To make sure the problem was not just with this geometry or cluster, the test was repeated with the lys example, officially distributed with PyGBe (one dielectric interface and three internal cavities), run without modifications on Tulio:

- Tulio (GPU): -581 kcal/mol
- Tulio (CPU): -555 kcal/mol

The deviation turned out to be reproducible in an independent geometry and persisted even on the same machine, which ruled out hardware or cluster configuration causes and pointed to an error in the solver itself, present only when combining internal cavities with CPU execution. Version differences between environments were also considered (see Table 3.5) where both presented the error.

Diagnosis and Correction

The problem was traced to the function `generateRHS` in `matrixfree.py`. This function assembles the right-hand side (RHS) of the linear system, adding the contribution of the point charges to the interior or exterior half of the RHS depending on whether the affected surface is a parent or a child of the region with charge. For child surfaces, such as internal cavities, the contribution should have been added to the exterior half, but the code was adding it by mistake to the interior half:

```
F[s_start:s_start+s_size] += aux      # incorrect
F[s_start+s_size:s_start+2*s_size] += aux  # correct
```

This block only runs when the region with charge has child surfaces, that is, with internal cavities, which explains why the error did not show up with a single dielectric interface. The equivalent GPU routine, `generateRHS_gpu`, did distinguish correctly between the surface types, which explains the absence of discrepancy on GPU. The correction was incorporated into the official PyGBe ³ repository [21] (commit b268acd, pull request #94, “bugfix RHS CPU for cavities”). After applying it, the 3QVP and lys cases were re-run on CPU, obtaining results consistent with their GPU references, with differences attributable only to floating point precision. PyGBe was subsequently reinstalled on CCTVal under version 0.3.2+72.gb268acd, and all CPU simulations that are part of the final results of this work were run on this corrected version.

Contribution: The author’s role was to systematically identify and isolate the problem, verify that the outputs matched except for solvation energy, GMRES iterations and simulation time; detect the discrepancy in his own case (3QVP); reproduce it in an unmodified PyGBe example (lys) to confirm that it was a general problem of the solver; and narrow down the conditions to the combination of internal cavities with CPU execution. This isolation of the root cause, achieved through this reproducibility testing, was what allowed the specific line of code responsible for the bug to be identified and corrected.

³<https://github.com/pygbe/pygbe>

Chapter 4

Results and Discussion

4.1 Influence of the Applied Potential

As defined in Equations (2.45) and (2.46), the interaction energy ΔE_{bind} reported throughout this chapter, and its solvation and surface components, ΔE_{solv} and ΔE_{surf} , correspond to differences with respect to the isolated protein and isolated electrode under the same pH and applied potential. The absolute energies of these isolated components are not individually reported in this chapter, since no part of the analysis depends on them directly, but are provided in the Appendix, in Tables A.1 and A.2 for reference.

Figure 4.1 shows, for each pH and each potential applied to the electrode, the range spanned by the interaction (binding) energy ΔE_{bind} across the 406 sampled orientations, where the lower bound of each bar corresponds to the Boltzmann-optimal orientation, i.e., the most favorable one, and the upper bound to the least favorable one. For both enzymes, the most favorable value within a given pH is always found when the potential is positive, and this energy well becomes progressively deeper as the pH increases. This is consistent with an electrostatically favorable interaction between the net negative charge of the proteins (Table 3.1) and a positively charged electrode, whose magnitude grows with the available charge of the protein. Analogously, the least favorable value is always found at the negative potential. When the applied potential is close to zero, on the other hand, the range of variation is much smaller for both enzymes, which makes physical sense since the protein no longer experiences a significant perturbation from the electrode.

Regarding the dependence on pH, the transition from pH 5 to pH 7 shows the most pronounced jump in the vast majority of the analyzed series (seven out of eight combinations of enzyme, potential, and range endpoint), while the transition from pH 7 to pH 9 is comparatively smoother; this trend is more marked in glucose oxidase (GOx). This pattern of transitions is consistent

with the discontinuity in the protonation state of the protein described in Section 4.2. The only exception is the favorable endpoint of BOx at +25.7 mV, where the jump from pH 7 to pH 9 (-1.28 kcal/mol) is larger than that from pH 5 to pH 7 (-0.62 kcal/mol).

It is important to note that this first analysis considers only the extreme values (minimum and maximum) and the range they define, that is, the best and worst possible orientations under each condition, without yet incorporating the statistical weight of each orientation or the overall behavior of the system.

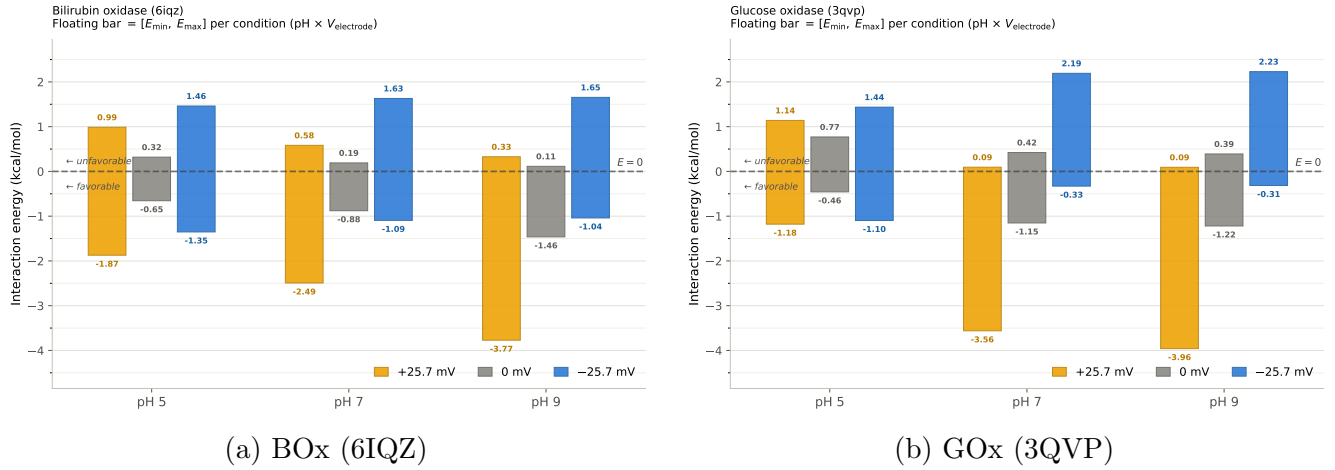


Figure 4.1: Range of interaction energy ΔE_{bind} across the 406 sampled orientations, for each pH and applied potential Φ_e . Each floating bar spans from the Boltzmann-optimal orientation (the most favorable, lower end) to the least favorable one (upper end) found within the sampled configuration space.

Applying the Boltzmann probability distribution to the interaction energy yields the expected value, or macroscopic behavior, of this quantity, $\langle \Delta E_{\text{bind}} \rangle$ described by the Equation (2.90), calculated over the complete set of 406 orientations, as a function of the electrode potential for each pH, and plotted in Figure 4.2. There it can be seen that, for both enzymes, $\langle \Delta E_{\text{bind}} \rangle$ decreases strictly monotonically with Φ_e at pH 7 and 9, whereas at pH 5 both show a local minimum of favorability, a small maximum, precisely at $\Phi_e = 0$ mV.

Interpreting $\Phi_e = 0$ mV as an internal reference of the system, even though $\langle \Delta E_{\text{bind}} \rangle$ is already negative in the absence of an applied potential (reflecting an intrinsically favorable interaction between the protein and the neutral electrode), it is observed that applying a negative potential systematically reduces this favorability at pH 7 and 9 for both enzymes. In BOx, the values at these pH levels become more positive under negative potential, but remain within the favorable threshold very close to zero ($\Delta = +0.31$ and $+0.48$ kcal/mol relative to $\Phi_e = 0$). This effect is

notably more pronounced in GOx, which even crosses into positive (unfavorable) values at both pH levels ($\Delta = +0.80$ and $+0.90$ kcal/mol, for pH 7 and 9 respectively). These observations are, in general, consistent with the expected behavior in front of a charged electrode as a function of the net charge of the protein (Table 3.1), although not strictly proportional to it, both in BOx ($-13e$ at pH 7 vs. $-17e$ at pH 9) and in GOx ($-25e$ vs. $-29e$). The resulting values at $\Phi_e = -25.7$ mV remain very close to each other despite the difference in charge. The crossing of GOx into positive values, on the other hand, does seem to relate to its greater charge magnitude compared to BOx, reflected in a larger displacement relative to the zero potential.

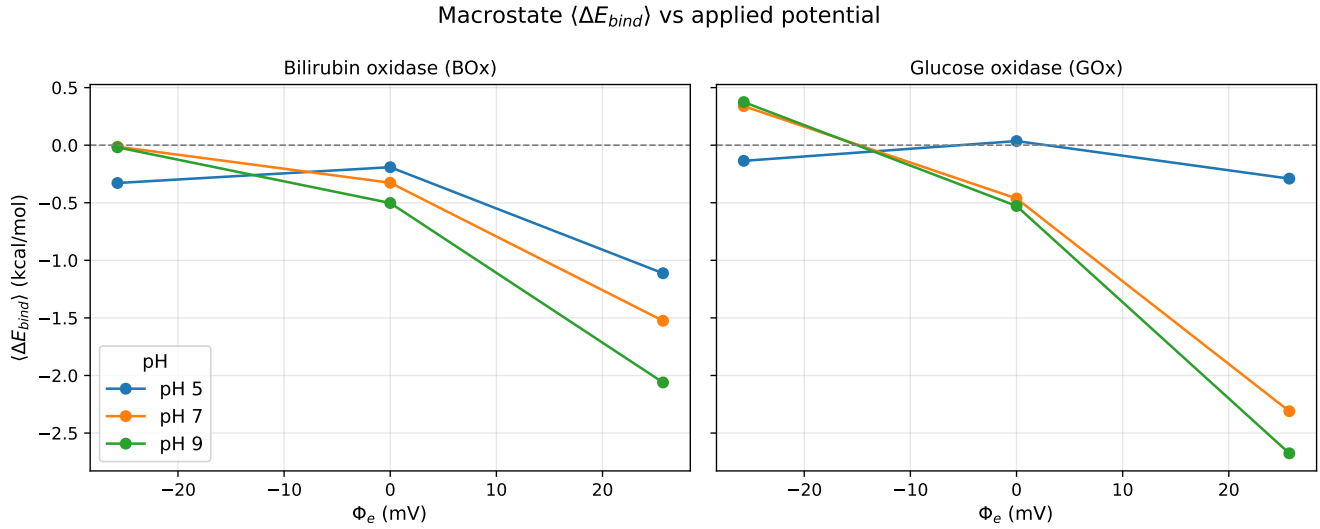


Figure 4.2: Boltzmann-weighted average interaction energy $\langle \Delta E_{bind} \rangle$ as a function of the applied potential Φ_e , for each pH and both enzymes.

These net charge values also help explain the particular behavior observed at pH 5, where both enzymes exhibit the smallest net charge magnitude (BOx: $-2e$; GOx: $+1e$, the only positive one in the set), placing them in a sort of threshold close to neutrality. For BOx, the behavior at pH 5 follows a trend relatively similar to that of the other pH levels; for GOx, on the other hand, the trend departs more clearly from the rest and turns out to be the flattest of all cases, possibly due to being the only condition with a positive net charge. Even so in GOx at pH 5, the potential of $+25.7$ mV remains the most favorable, although by a much smaller margin than at the other pH levels. This is qualitatively analogous to what was reported by Ruiz et al. [2] for his hydrogenase, where the capacitive net-charge model likewise did not hold at one of the extreme pH values of the study (pH 9, in his case, in contrast to pH 5 in the present work); in both cases, the exception occurs at one end of the pH range studied, although not necessarily the same one. Taken together, these results indicate that, for the interaction energies studied, the electrostatic response coherently follows the sign of the net charge of the protein at pH 7 and 9, with the partial exception of pH 5.

It is interesting to note that, when decomposing $\langle \Delta E_{\text{bind}} \rangle$ into its solvation and surface contributions (Appendix Figs. B.1 and B.2), each component generally follows the same pattern just described for the total interaction energy. However, it is observed that $\langle \Delta E_{\text{solv}} \rangle$ remains on the favorable (more negative) side under all conditions, with the sole exception of GOx at pH 5 and $\Phi_e = 0$ mV; it is the surface term $\langle \Delta E_{\text{surf}} \rangle$ that drives the crossing into positive values observed in GOx at pH 7 and 9 under $\Phi_e = -25.7$ mV, as well as the approach to values close to zero exhibited by BOx under that same condition.

4.2 Influence of pH

Figure 4.3 shows the same Boltzmann-weighted average interaction energy $\langle \Delta E_{\text{bind}} \rangle$ presented in the previous section, now as a function of pH for each applied potential (the separate solvation and surface contributions as a function of pH are shown in Appendices B.3 and B.4). For both enzymes, at $\Phi_e = 0$ and $+25.7$ mV, $\langle \Delta E_{\text{bind}} \rangle$ becomes more favorable as pH increases, consistent with the growth in magnitude of the net negative charge of both proteins (Table 3.1). At $\Phi_e = -25.7$ mV, however, this trend is reversed, $\langle \Delta E_{\text{bind}} \rangle$ becomes less favorable when going from pH 5 to pH 7 (in GOx, even crossing into positive, i.e., unfavorable, values)

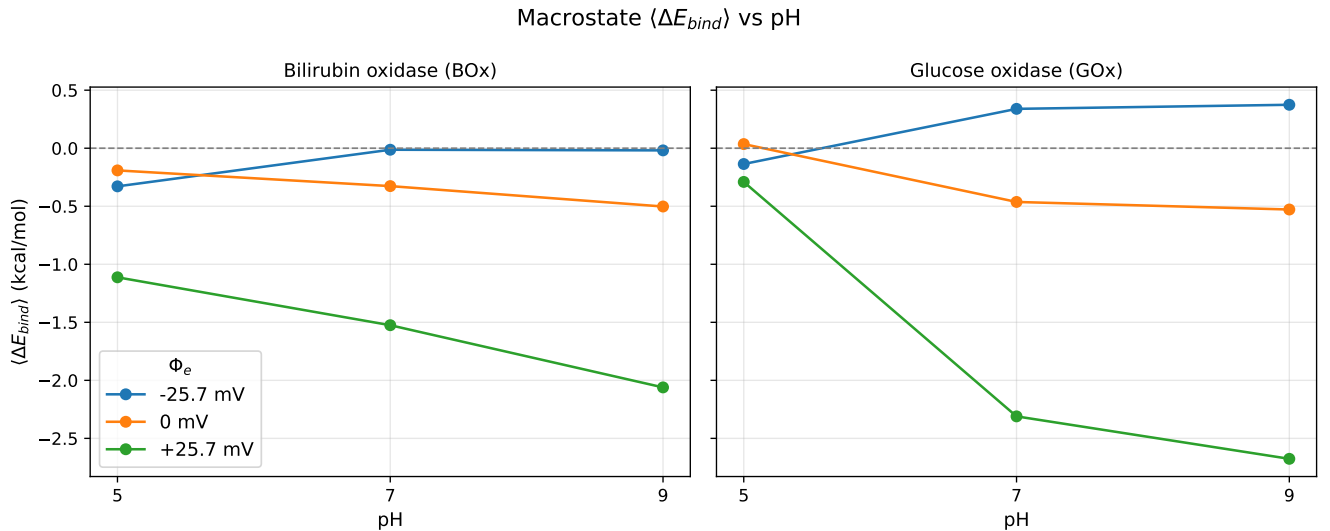


Figure 4.3: Boltzmann-weighted average interaction energy $\langle \Delta E_{\text{bind}} \rangle$ as a function of pH, for each applied potential Φ_e and both enzymes.

Unlike GOx, whose net charge changes drastically between pH 5 ($+1e$) and pH 7 ($-25e$) before stabilizing toward pH 9 ($-29e$), BOx exhibits a more gradual charge transition ($-2e$, $-13e$, $-17e$). This discontinuity, particularly marked in GOx, is consistent with the pKa range of several ionizable

residues, in particular aspartic acid (pKa 3.8), glutamic acid (pKa 4.5), and histidine (pKa 6.5), which fall precisely within the pH 5 - 7 interval, according to the PROPKA predictions used in the preparation of the structures [1]. Since the PQR files used in this work are the same ones originally prepared for that study, this explanation applies directly to the behavior observed here, and not merely by analogy between similar systems. Consequently, both enzymes show a more pronounced jump in $\langle \Delta E_{\text{bind}} \rangle$ between pH 5 and pH 7 than between pH 7 and pH 9, with the notable exception of BOx at $\Phi_e = +25.7$ mV, discussed in the previous section.

To quantitatively assess which of the two variables, potential or pH, exerts a greater influence on the interaction energy, the range of variation of $\langle \Delta E_{\text{bind}} \rangle$ induced by each was compared, holding the other fixed (Table 4.1). For BOx, the range induced by the potential goes from 0.92 to 2.04 kcal/mol, whereas the range induced by pH goes from 0.31 to 0.95 kcal/mol, so the potential clearly dominates, except for an almost exact tie between its lowest value (0.92 kcal/mol, at pH 5) and the highest value induced by pH (0.95 kcal/mol, at $\Phi_e = +25.7$ mV). For GOx, the range by potential goes from 0.33 to 3.05 kcal/mol, and the range by pH goes from 0.51 to 2.39 kcal/mol, here the potential dominates broadly at pH 7 and 9, but its effect at pH 5 (0.33 kcal/mol) is smaller than the effect of pH under any of the three potentials studied, coinciding with the near-zero net charge condition already discussed above. In any case, these results indicate that, for both enzymes, the applied potential exerts a greater influence than pH on the interaction energy, in apparent contrast with the conclusion of weak potential influence reported by Terrani [1].

Table 4.1: Range of variation of $\langle \Delta E_{\text{bind}} \rangle$ induced by the applied potential (at fixed pH) and by pH (at fixed potential), for both enzymes.

	Range by potential (kcal/mol)		
	pH 5	pH 7	pH 9
BOx	0.92	1.51	2.04
GOx	0.33	2.65	3.05
	Range by pH (kcal/mol)		
	-25.7 mV	0 mV	+25.7 mV
BOx	0.32	0.31	0.95
GOx	0.51	0.56	2.39

It is important to point out several methodological differences between the present work and that of Terrani [1] that could explain the apparent discrepancy regarding the relative influence of the potential. First, both studies employ different Poisson-Boltzmann formulations, whereas this work uses the linearized equation (LPBE) via PyGBe, Terrani uses the full nonlinear Poisson-Boltzmann equation via APBS, validated through a pointwise comparison against the linearized version from

a previous work (obtaining $R \approx 1$ for GOx at pH 5, 7, and 9 under $\Phi_e = 10 k_B T/e$, his Fig. 4.3). Given that this validation shows a strong correspondence between both formulations, it is unlikely that this difference alone explains the observed discrepancy; nonetheless, it cannot be entirely ruled out, since the validation was carried out for a single case and at a considerably higher potential ($10 k_B T/e$) than the low-potential regime ($\pm 1 k_B T/e$) on which the analysis of the present work is focused.

Second, while this work restricts the applied potential to the strict range of validity of the LPBE ($\pm 1 k_B T/e$, equivalent to ± 25.7 mV), Terrani additionally evaluates considerably larger potentials ($\pm 10 k_B T/e$); in that regime, the base charge of the electrode is so large ($1139.79 e$, according to his Table 4.3) that any perturbation induced by the protein turns out to be comparatively negligible, which could dilute the apparent influence of the potential when averaging over the entire range studied [1]. In fact, the author himself points out that the non-negligible growth observed specifically at $\Phi_e = 1 k_B T/e$ is due to the fact that the base charge of the electrode is there much smaller ($15.00 e$, notably close to the value of $15.64 e$ obtained in this work for the same condition, Section 4.3.4), so that the perturbation from the protein is proportionally more relevant in that regime for both studies. However, this coincidence in the mechanism (the relative magnitude of the protein's perturbation compared to the base charge of the electrode) does not translate into the same qualitative result, whereas in the present work $\langle \Delta E_{\text{surf}} \rangle$ becomes consistently more favorable (negative) at a fixed positive potential and increasing pH (see Figure B.4), Terrani [1] reports the opposite.

Third, Terrani's conclusions regarding the relative influence of potential and pH compare quantities of a different nature. His Boltzmann-averaged figures for each variable (Figs. 4.4 and 4.5) are restricted exclusively to ΔE_{solv} , with a probability distribution calculated only from that term. When visually inspecting his Figures 4.4 and 4.5 independently, pH still appears to dominate over potential even within this comparison restricted to solvation-only and macroscopically averaged. This contrasts with what is found in this work, where, when comparing the complete Boltzmann-averaged binding energy for both variables, the applied potential dominates over pH (Table 4.1); and, as shown later (Section 4.3.1), even when restricting the comparison to solvation only (matching Terrani's more limited quantity, with the Boltzmann probability calculated solely from that same energy), the potential still dominates in this work. This suggests that the discrepancy is not explained solely by the use of macroscopic solvation instead of macroscopic binding in Terrani's comparison, and could instead reflect a more fundamental difference between the two systems or methods.

Fourth, both works obtain the solvation energy through structurally different routes. In this work, E_{solv} is a native output of the PyGBe boundary element formulation, which additionally reports the Coulombic E_{Coulomb} and surface E_{surf} contributions as three separate quantities. Terrani [1], using APBS, follows a different route, APBS returns only a single total electrostatic free energy that combines the classical Coulombic term together with entropic and osmotic contributions into one number, without separating out the solvation contribution. To isolate it, he defines a reference energy calculated over the same geometry, but replacing the solvent dielectric constant ($\epsilon = 80$) with that of the protein ($\epsilon = 3.23$) throughout the entire domain, including the region that should physically correspond to water; the solvation energy is then the difference between the total and this reference. Terrani himself points out that this solvation energy difference, taken on its own, does not correspond to a physically meaningful quantity (the physically relevant quantity being the complete binding energy, obtained only once the surface term is added), although this does not imply any error in the calculation itself. In this work, by contrast, no such intermediate reference is required, since the isolated protein and electrode retain the physically correct dielectric constants at all times. This constitutes a methodological difference arising from the different native capabilities of APBS and PyGBe, not an error on either side, and it could partly contribute to the discrepancies discussed above.

This referencing scheme in APBS may serve an additional purpose beyond isolating the solvation contribution, finite-difference methods require point charges to be interpolated onto a fixed background grid, and the resulting potential is sensitive both to the grid resolution and to the exact position of the protein relative to that grid. Subtracting a reference calculation without the dielectric interface is intended to cancel this grid-dependent error, but this correction is only strictly valid when the protein occupies a fixed position relative to the grid; when the protein is reoriented, as required by the orientational sampling central to this line of research, the interpolation error changes with each new position and is no longer reliably canceled by the reference subtraction (C. Cooper and G. Raos, personal communication). The boundary element method adopted in this work is not subject to this limitation, the surface mesh is generated directly on the molecular geometry and moves rigidly with the protein under rotation, without relying on a fixed background grid, and is therefore not expected to introduce this type of position-dependent error.

Fifth, Terrani’s reference configuration, although effective, does not correspond to a complete isolation, the nominal center coordinate of 70 Å is not the actual protein-electrode separation once the protein’s own size is discounted. The resulting effective separation is approximately 35–47 Å (4.4–6 Debye lengths), enough to keep the residual field below $\sim 1\%$, a small, though not strictly zero, influence of the electrode on the protein in this reference configuration. Additionally, cavity

treatment differs between the studies discussed, APBS based calculations implicitly include all internal solvent-filled cavities, since the finite-difference grid spans the entire domain without requiring separate meshing, whereas PyGBe-based calculations require each cavity to be meshed explicitly, this work includes only the largest cavity (or two, for GOx) near the catalytic center, omitting the smaller ones due to computational cost. Although this is expected to have a smaller effect than the differences discussed above, it remains an additional, unquantified source of methodological divergence.

It is worth noting that this discrepancy is not limited to Terrani’s work. Ruiz et al. [2], studying a [NiFe] hydrogenase also with PyGBe, similarly conclude that pH is the dominant factor governing adsorption, followed by ionic strength, and only then by the applied potential. Since Ruiz et al. also uses PyGBe, this rules out the choice of solver (PyGBe versus APBS) as the sole explanation for the discrepancy with this work; the methodological difference most plausibly shared between Terrani and Ruiz et al. and absent in this work, is the explicit Stern layer, corresponding to a 2 Å region around both the protein and the electrode where the local ionic concentration is set to zero, keeping the dielectric constant of water ($\epsilon \approx 80$) but removing Debye-Hückel screening in that zone, where the potential there satisfies the Laplace equation instead of the linearized PBE, consistent with the Stern layer formulation presented in Section 2.4.3 (Equation 2.38). This results in a total protein-electrode separation of 4.1 Å in their scheme, compared to the 3 Å used in this work, where the entire separation is free water subject to full exponential screening. Although the absence of screening within the Stern layer would, on its own, tend to better preserve the potential in that zone, the larger total separation distance imposed by that layer appears to be the dominant factor, so that the resulting net effect depends on the balance between both mechanisms, and cannot be anticipated solely from a qualitative screening argument. The author considers this difference in distance and in the use of Stern layers, together with the grid-dependent interpolation error discussed above, to be the most plausible explanation among those discussed in this section, although, like the previous ones, it remains an independently unverified hypothesis; identifying the precise mechanism behind the observed discrepancy would require a dedicated comparison that exceeds the scope of this work.

4.3 Influence of Orientation

Having examined how the applied potential and pH influence the average energetic behavior of the system (Sections 4.1 and 4.2), this section delves further into the role of the specific orientation of the protein on the electrode surface.

Figure 4.1 illustrates this influence directly, the height of each bar is the energy range from $\Delta E_{\text{bind}}^{\text{min}}$ to $\Delta E_{\text{bind}}^{\text{max}}$ spanned by the 406 sampled orientations within a given pH and potential condition. Comparing this range in each of the 18 conditions against the ranges induced by potential and by pH (Appendix, Table A.3), orientation dominates over pH in 17 of the 18 conditions (the only exception being GOx at pH 5 and $\Phi_e = +25.7$ mV), and over potential in 12 of the 18, dominating in all cases with positive potential (+25.7 mV), in 4 of the 6 with negative potential (−25.7 mV), and only at pH 5 when the potential is zero. In this last case ($\Phi_e = 0$ mV), the range by orientation turns out to be smaller than that induced by potential at pH 7 and 9 for both enzymes, whereas at pH 5 exceeds it. The range by orientation reaches its maximum in BOx at pH 9 and $\Phi_e = +25.7$ mV (4.10 kcal/mol), and its minimum in BOx at pH 5 and $\Phi_e = 0$ mV (0.97 kcal/mol).

Terrani [1] reports a similar trend, namely that variations by orientation frequently exceed those induced by potential. The analysis of the 18 conditions presented here confirms that general trend, but shows that it is not universal, it consistently dominates over pH, and over potential only when the latter is nonzero.

4.3.1 Summary Table of Optimal Orientations and Ranking Consistency

Table 4.2 brings together, for each of the 18 conditions studied, the thermodynamic (orientation, adsorption energy and Boltzmann probability), kinetic (donor–acceptor distance, orientation, and dominant contribution to the weighted current), and electrostatic (nearest residue, local charge and induced charge perturbation) results presented throughout this chapter, condensing into a single view the behavior of the system under each combination of protein, pH, and potential. In addition to the energy $\Delta E_{\text{bind}}^{\text{min}}$ and the probability P_B of the Boltzmann-optimal orientation, the expected value $\langle \Delta E_{\text{bind}} \rangle$ is included, corresponding to the same quantity discussed in Sections 4.1 and 4.2 and plotted in Figures 4.2 and 4.3, which integrates the information from the 406 sampled orientations.

The remainder of this section develops in detail the different aspects summarized in this table, Section 4.3.2 presents the electron transfer geometry between each catalytic center of the protein and the electrode; Section 4.3.3 presents the complete thermodynamic-kinetic landscape for the 406 orientations; and Section 4.3.4 presents the electrostatic fingerprints associated with the charge perturbation ΔQ that the protein leaves on the electrode in the orientation of highest probability.

Table 4.2: Thermodynamic, kinetic, and electrostatic results for **Protein**, bilirubin oxidase (6IQZ) and glucose oxidase (3QVP) across the eighteen $\mathbf{pH} \times \Phi_e$ conditions studied. θ/φ_B and P_B are the angle (tilt/rotation) and the Boltzmann probability of the most probable orientation; $\Delta E_{\text{bind}}^{\text{min}}$ is its net adsorption energy. Dist_Z is the donor - acceptor distance evaluated at the Boltzmann optimal orientation: Cu T1 - electrode for BOx, N5(FAD) - electrode for GOx. $\mathbf{AA}_{\text{min}}$ is the closest residue to the electrode surface; **Q10** is the summed net charge of the ten closest residues. ΔQ is the electrostatic charge perturbation induced by the protein on the electrode surface. $\langle \Delta E_{\text{bind}} \rangle$ is the expected value of the macroscopic behavior of the interaction energy. θ/φ_W and **Contrib.%** indicate the orientation and the percentage contribution of the dominant term in the Boltzmann-weighted sum of the electron transfer. **Match** indicates whether the weighted optimal orientation coincides with the Boltzmann optimal (“Boltzmann”) or with the purely geometric optimal of k_{ET} (“Raw k_{ET} ”). I_{total} is the total macroscopic current.

Protein	pH	Φ_e (mV)	θ/φ_B (deg)	$\Delta E_{\text{bind}}^{\text{min}}$ (kcal/mol)	P_B %	Dist_Z (Å)	$\mathbf{AA}_{\text{min}}$	Q10 (e)	ΔQ (e)	$\langle \Delta E_{\text{bind}} \rangle$ (kcal/mol)	θ/φ_W (deg)	Match	Contrib. %	I_{total} (a.u.)
6IQZ	5	+25.7	140/93.91	-1.8721	1.43	42.12	ASP 285	-4	2.305	-1.1122	110/316.36	Raw k_{ET}	17.65	2.364×10^{-14}
6IQZ	5	0	150/63.53	-0.6546	0.57	35.88	ALA 243	-3	1.923	-0.1911	110/316.36	Raw k_{ET}	24.48	2.487×10^{-13}
6IQZ	5	-25.7	100/339.43	-1.3520	2.53	19.54	GLY 358	+1	-1.424	-0.3286	110/316.36	Raw k_{ET}	26.45	1.149×10^{-12}
6IQZ	7	+25.7	70/272.73	-2.4931	1.91	20.06	GLY 194	-4	2.486	-1.5249	70/272.73	Boltzmann	32.43	3.718×10^{-14}
6IQZ	7	0	70/272.73	-0.8777	0.66	20.06	GLY 194	-4	2.970	-0.3264	110/316.36	Raw k_{ET}	20.53	1.903×10^{-13}
6IQZ	7	-25.7	100/339.43	-1.0937	2.54	19.54	GLY 358	+1	-0.893	-0.0131	110/316.36	Raw k_{ET}	21.84	8.094×10^{-13}
6IQZ	9	+25.7	70/272.73	-3.7709	9.04	20.06	GLY 194	-4	3.638	-2.0608	70/272.73	Boltzmann	62.93	9.096×10^{-14}
6IQZ	9	0	70/272.73	-1.4620	1.36	20.06	GLY 194	-4	4.121	-0.5022	110/316.36	Raw k_{ET}	16.59	1.675×10^{-13}
6IQZ	9	-25.7	100/339.43	-1.0394	2.29	19.54	GLY 358	+1	-0.685	-0.0187	110/316.36	Raw k_{ET}	18.41	6.033×10^{-13}
3QVP	5	+25.7	160/120.00	-1.1784	1.51	36.15	ASP 195	-3	1.445	-0.2904	90/280.00	Raw k_{ET}	35.21	2.122×10^{-16}
3QVP	5	0	60/255.48	-0.4568	0.60	35.86	GLY 274	+3	-1.169	+0.0357	90/280.00	Raw k_{ET}	41.90	4.016×10^{-16}
3QVP	5	-25.7	60/255.48	-1.0966	1.70	35.86	GLY 274	+3	-0.990	-0.1361	90/280.00	Raw k_{ET}	43.50	3.787×10^{-16}
3QVP	7	+25.7	90/110.00	-3.5616	5.39	34.43	THR 124	-4	4.034	-2.3104	90/280.00	Raw k_{ET}	69.24	2.498×10^{-15}
3QVP	7	0	90/140.00	-1.1514	0.87	39.48	THR 390	-2	3.550	-0.4630	90/280.00	Raw k_{ET}	49.01	4.803×10^{-16}
3QVP	7	-25.7	20/300.00	-0.3283	0.94	43.57	LYS 13	0	0.600	+0.3391	90/280.00	Raw k_{ET}	23.33	3.801×10^{-17}
3QVP	9	+25.7	90/110.00	-3.9591	6.62	34.43	THR 124	-4	4.489	-2.6762	90/280.00	Raw k_{ET}	68.94	2.188×10^{-15}
3QVP	9	0	90/140.00	-1.2186	0.89	39.48	THR 390	-2	3.765	-0.5281	90/280.00	Raw k_{ET}	49.43	4.930×10^{-16}
3QVP	9	-25.7	20/300.00	-0.3142	1.00	43.57	LYS 13	0	0.620	+0.3743	90/280.00	Raw k_{ET}	24.83	3.656×10^{-17}

A first look at Table 4.2 anticipates two patterns that are developed in the rest of the section. First, the Match column shows that the weighted orientation coincides with the geometric optimum of k_{ET} (*Raw* k_{ET}) in the vast majority of conditions, and with the Boltzmann optimum in only two cases, both for BOx at $\Phi_e = +25.7$ mV (pH 7 and 9), a pattern examined in detail in Section 4.3.3. Second, the Boltzmann-optimal orientation remains identical between pH 7 and pH 9 for a given potential, in both enzymes, suggesting an energetic surface comparatively insensitive to this variable within that range, a behavior revisited in Section 4.3.3.

As an additional verification, Table A.5 (Appendix) reports, for the 18 conditions, the same optimal orientation and macroscopic energy analysis presented in Table 4.2, but considering only ΔE_{solv} and with the Boltzmann probability likewise calculated solely from that term, instead of the complete binding energy, together with the Spearman rank correlation between both orderings (Section 2.6.5). This emulates Terrani’s conditions, who processes the solvation energy macroscopically and calculates its Boltzmann probability likewise only from that same term. It also provides a quantitative answer to a qualitative observation repeatedly inferred during the development of this work, namely that the ranking of orientations by solvation energy appears, from visual inspection of the simulations’ `*-output.log` files, to be a reasonable approximation to the ranking by complete binding energy, even though the latter is the physically correct quantity. Although intuitive, this is the first time this relationship has been verified quantitatively.

It is worth distinguishing the quantity reported in Table A.5, denoted $\langle \Delta E_{\text{solv}} \rangle^*$, from the quantity $\langle \Delta E_{\text{solv}} \rangle$, the solvation component of $\langle \Delta E_{\text{bind}} \rangle$ plotted in Figures B.1 and B.3 (Appendix). Both quantities average the same solvation energy $\Delta E_{\text{solv},i}$ over the 406 orientations, but differ in the Boltzmann weights used: $\langle \Delta E_{\text{solv}} \rangle$, shown in those figures, uses the Boltzmann weights calculated over the complete binding energy, whereas $\langle \Delta E_{\text{solv}} \rangle^*$, reported only in Table A.5, uses weights calculated solely from the solvation energy itself, as if this were the total available energy of the system, completely ignoring the surface term. This second quantity is not plotted separately in this work; it is tabulated only for the purpose of the robustness check described above, evaluating both the orientational ranking and the macroscopic preference trend.

The results confirm this intuition, since the optimal orientation coincides between both criteria in 15 of the 18 conditions, with a consistently high Spearman correlation ($\rho = 0.83$ – 1.00). Across the 18 conditions, both the depth of the energy well $\Delta E_{\text{solv}}^{\text{min}}$ and the probability $P_{B,\text{solv}}$ turn out to be systematically smaller than their complete binding counterparts in Table 4.2, consistent with the surface term, under the most favorable conditions, contributing negatively and further deepening the energy well, although without substantially altering the relative ranking among orientations.

The cases at zero potential show a perfect Spearman correlation ($\rho = 1.0000$), an expected result given that the surface term is practically null (on the order of 10^{-9}) under $\Phi_e = 0$ mV, making ΔE_{bind} and ΔE_{solv} identical under that condition. Of the three conditions where the optimal orientation does not coincide exactly, in BOx at pH 5 and $\Phi_e = +25.7$ mV the optimal orientation by solvation is the 4th best in the complete binding ranking; in GOx at pH 9 and $\Phi_e = -25.7$ mV, the 15th; and in GOx at pH 7 and $\Phi_e = -25.7$ mV, the 24th.

The same macroscopic trend described for $\langle \Delta E_{\text{bind}} \rangle$, namely that the protein electrostatically prefers the positive potential electrode (Sections 4.1 and 4.2), is preserved when using exclusively $\langle \Delta E_{\text{solv}} \rangle^*$, with the sole exception of GOx at pH 5, already identified earlier as the condition of near zero net charge. Likewise, the dominance of the applied potential over pH is preserved, since comparing $\langle \Delta E_{\text{solv}} \rangle^*$ instead of $\langle \Delta E_{\text{bind}} \rangle$, the potential still dominates over pH in this work, as noted in Section 4.2 when discussing the comparison with Terrani; this suggests that the discrepancy with his conclusion is not explained solely by the use of macroscopic solvation instead of macroscopic binding, and could instead reflect a more fundamental difference between the two systems or methods.

Taken together, this indicates that using only ΔE_{solv} as an approximation, both for the orientational ranking and for the macroscopic trend, is reasonable in most conditions, although not exact, and quantitatively confirms what until now had only been observed qualitatively.

4.3.2 Electron Transfer Kinetics

The electron transfer rate constant k_{ET} , calculated using the simplified Marcus/Moser–Dutton model, Equation (2.88), depends exclusively on the distance $r(\theta, \varphi)$ between the reference atom of the active center (the T1 copper in BOx, the N5 nitrogen of FAD in GOx) and the plane of the electrode. Being a purely geometric quantity, determined by the rigid orientation of the protein and not by the electrostatic conditions of the system, k_{ET} is invariant with respect to pH and applied potential; a single map per protein describes the 406 sampled orientations under any of the 9 combinations of pH and potential studied, and constitutes the common kinetic reference to which those 9 conditions refer in the following sections.

Figure 4.4 shows this map for both enzymes, on a logarithmic scale given the strong exponential dependence of k_{ET} on distance. The maximum is located at $(\theta, \varphi) = (110.00, 316.36)$ for BOx and at $(90.00, 280.00)$ for GOx; since this is a fixed value repeated identically across the nine conditions of each enzyme, it is not included as an additional column in Table 4.2. Comparing the best case of each protein, BOx transfers electrons several orders of magnitude faster than GOx, consistent

with its smaller minimum donor–acceptor distance. k_{ET} is expressed in arbitrary units, a product of the constant C that groups together the thermodynamic and vibrational factors of the complete Marcus equation, Equation (2.88); its absolute magnitude is not comparable to an experimental rate, although the relative differences between orientations, such as the one observed here between proteins, are indeed physically meaningful.

These distances can be contrasted with the direct electron transfer (DET) thresholds discussed in the theoretical framework (Section 2.2). The optimal orientation of BOx (17.83 Å) falls within the 14–26 Å range reported for DET, although above the 14 Å value associated with efficient transfer; that of GOx (21.67 Å) already lies at the lower limit of the 20–21 Å threshold, beyond which direct transfer becomes markedly less favorable and typically requires redox mediators (MET). The least favorable orientations of both enzymes (53–55 Å) considerably exceed these thresholds, falling outside any viable DET regime. It should be noted that these distances include the fixed 3 Å separation imposed between the protein and the electrode in all orientations (Section 3.1.2), so that the minimum attainable distance, and with it the maximum possible k_{ET} , is limited by this design choice, and not solely by the geometry of the protein.

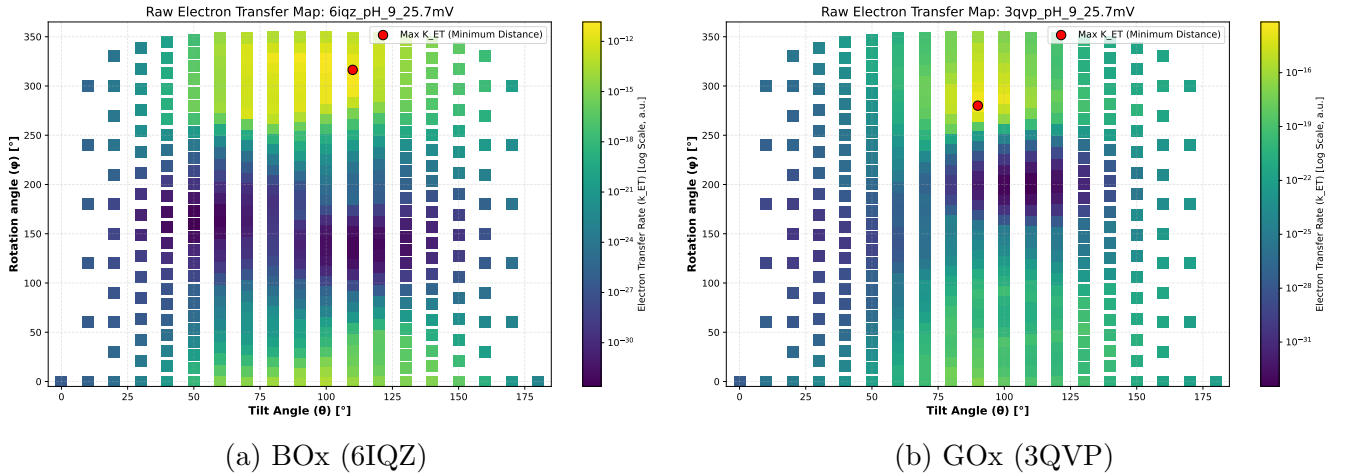


Figure 4.4: Electron transfer rate constant k_{ET} (logarithmic scale) (a.u.) as a function of orientation (θ, φ) , for the 406 sampled orientations. The red dot marks the orientation of maximum k_{ET} (minimum donor - acceptor distance). This map is identical for the nine pH and potential conditions of each protein.

The regions of high and low k_{ET} are not randomly distributed across the (θ, φ) space, they form contiguous, extended patches. In BOx (Figure 4.4a), a yellow band occupies the upper part of the map and a dark blue band the lower-middle part, with a relatively smooth transition between them, consistent with the donor–acceptor distance varying continuously with (θ, φ) , such that angularly

close orientations share similar k_{ET} values. In GOx (Figure 4.4b) the same two patches can be identified, but the transition between them is notably more abrupt. One possible explanation is geometric, the MAN-MAN-BMA-NAG-NAG oligosaccharide parameterized by Terrani [1] projects outward from the protein like an arm (visible in the solvent-excluded surface shown in Figure 3.1b), which in certain orientations could interpose itself toward the face facing the electrode, abruptly pushing the FAD’s N5 away from the measurement plane. Although this hypothesis is not verified quantitatively in this work, it is consistent with the known structure of the oligosaccharide. As an additional hypothesis, likewise not explored in depth here, this greater angular sensitivity of GOx could be related to its overall shape, the principal moments of inertia calculated for both proteins show that the relative dispersion among the three principal axes is approximately twice as large in GOx as in BOx, suggesting a comparatively less spherical, more elongated geometry, which would be consistent with a greater sensitivity of the donor–acceptor distance to changes in orientation. Neither explanation is conclusively confirmed in this work, and both remain as possible directions for future work. This concept of patches is revisited in the following section, where it is contrasted with the Boltzmann probability distribution, whose structure, as will be seen later, does not always exhibit the same spatial continuity.

Table 4.3 summarizes, for each enzyme, the orientation of highest and lowest k_{ET} among the 406 sampled (details of the three best and three worst in the Appendix, Table A.4). In BOx, the range goes from 17.83 to 53.64 Å ($\Delta r = 35.80$ Å), a drop of $\sim 3.1 \times 10^{21}$ -fold in k_{ET} ; in GOx, from 21.67 to 55.05 Å ($\Delta r = 33.38$ Å), a drop of $\sim 9.9 \times 10^{19}$ -fold. In both cases, a variation of tens of Å produces drops of several orders of magnitude, illustrating the already well-known strong exponential dependence of Equation (2.88).

Table 4.3: Orientation of highest and lowest electron transfer rate constant k_{ET} , for both enzymes.

Protein		θ (°)	φ (°)	r (Å)	k_{ET} (a.u.)
BOx (6IQZ)	Best	110.00	316.36	17.8317	1.4391×10^{-11}
	Worst	50.00	160.00	53.6356	2.4485×10^{-33}
GOx (3QVP)	Best	90.00	280.00	21.6732	6.6441×10^{-14}
	Worst	110.00	196.36	55.0482	3.3887×10^{-34}

4.3.3 Orientational Landscape of Energy, Probability, and Weighted Current

Figures 4.5 and 4.6 present, for BOx and GOx at pH 9, the complete picture of the 406 sampled orientations under the three potentials studied, organized into three columns: the net interaction

energy ΔE_{bind} , the Boltzmann probability P_i , and the weighted current $k_{\text{ET}}(r_i) \cdot P_i$. The panels corresponding to the remaining combinations of protein and pH are presented in the Appendix: pH 5 and pH 7 for BOx in Figures B.5 and B.6 respectively, and for GOx in Figures B.7 and B.8. Although these figures are built from the same data underlying Sections 4.1 and 4.2, they allow the individual microstates that make up the macrostates discussed there to be examined in detail, beyond the average behavior already presented.

We begin with the first two columns of these figures, directly related to the electrostatic behavior of the system. In the energy map of Figure 4.5 ($\Phi_e = +25.7$ mV), the interaction energy ΔE_{bind} is color coded in blue when favorable (negative), red when unfavorable (positive), and white near zero; the vast majority of the map appears in shades of blue, with only a small number of pale red points. A single yellow point marks the lowest energy orientation, at $(\theta, \varphi) = (70, 272.73)$, which corresponds exactly to the Boltzmann-optimal orientation reported for this condition in Table 4.2, surrounded by a cluster of neighboring orientations also of low energy. The highest energy orientation is located at $(100, 329.14)$, at the upper right edge of the map, likewise surrounded by a region of high energy; a second band of comparatively low energy appears along the lower edge of the plot. This apparent discontinuity, one favorable region at the top and another at the bottom, is an artifact of the angular periodicity, rotation angles close to 0° and to 360° are geometrically contiguous, but the plot places them at opposite ends of the vertical axis.

The Boltzmann probability column P_i is, as expected, an almost identical reflection of the energy map, since it is built directly from Equation 2.85; however, its exponential value accentuates the relative differences between orientations. For this condition, the most probable orientation reaches $P = 9.04\%$, which is 36 times more probable than if the 406 orientations were equally accessible ($1/406 = 0.246\%$), and more than 1000 times more probable than the least probable orientation ($P = 0.0089\%$). Considering together the three most probable orientations, which in this case are concentrated within the same neighboring angular cluster, their combined probability reaches 14.79% , approximately 539 times the combined probability of the three least favorable orientations (0.027%), consistent with reading the interaction energy in terms of contiguous zones or clusters, rather than isolated points. This extreme concentration of probability in a handful of neighboring orientations, a direct consequence of the exponential nature of the Boltzmann distribution, confirms that the Boltzmann-optimal orientation reported in Table 4.2 is not an arbitrary point within a flat landscape, but the center of a dominant region of the orientational space.

In the energy map for the negative potential (-25.7 mV) of this same Figure 4.5, the patch pattern is even more pronounced, but with a behavior opposite to that of the positive potential, the

central band of the map appears tinted red (unfavorable), while it is the upper and lower edges that concentrate the negative (favorable) energies. The minimum energy is located at (100, 339.43), with a Boltzmann probability of 2.29%, which is considerably lower than in the positive case, and the least favorable orientation at (80, 144.00), with 0.024%. This is directly reflected in the associated probability map, where the points of highest probability appear more scattered and numerous than in the positive case, where they were concentrated in a small handful of standout points, although they still cluster into recognizable angular groups.

For zero potential, the picture changes substantially, both the energy map and the probability map flatten out in terms of range, and the clearly defined cluster zones seen in the two previous cases practically disappear, giving way to a much more uniform distribution. Even so, the vast majority of the map remains in favorable blue tones, although with small differences between orientations. This is consistent with the fact that, in the absence of a significant electric field, the protein lacks a well defined, energetically preferred direction. Even so, it stands out that the most probable orientation at zero potential, (70, 272.73), with $P = 1.36\%$, coincides exactly with that of the positive case, although with a considerably lower probability. This could be due to three possible factors. First, that the nominal “zero” potential in the simulations is not exactly null, but a residual value of 1.81×10^{-7} mV (Section 3.6.2, Phase 0.5), whose influence, although extremely small, cannot be entirely ruled out; second, that this orientation could be intrinsically favorable for geometric or electrostatic reasons specific to the protein as a whole, for example, its total net negative charge ($-17e$ for BOx at pH 9), independent of both the applied potential and the specific orientation; and third, that the local charge of the particular residues that end up closest to the electrode in this orientation, as distinct from the protein’s total net charge, may specifically favor this configuration, which is discussed in detail in Section 4.3.4. Distinguishing between these explanations, and establishing whether any of them is causal or whether they merely coincide in the same orientation, exceeds the scope of this work.

Figure B.6 shows, for BOx at pH 7, a pattern practically identical to that of pH 9, with one notable difference, in the probability map at +25.7 mV, a second, non-negligible probability zone begins to appear on the right hand side of the map, corresponding to the angles (140, 93.91), separated from the main cluster that dominates at pH 9. The Boltzmann-optimal orientations for each potential are identical to those at pH 9 (Table 4.2), although the energy wells are flatter and the associated probabilities, consequently, lower. The most relevant difference with respect to pH 9 is that, within the set of three potentials itself, the globally most probable orientation at pH 7 now corresponds to the negative potential rather than the positive one, a crossover that can be verified directly in Table 4.2.

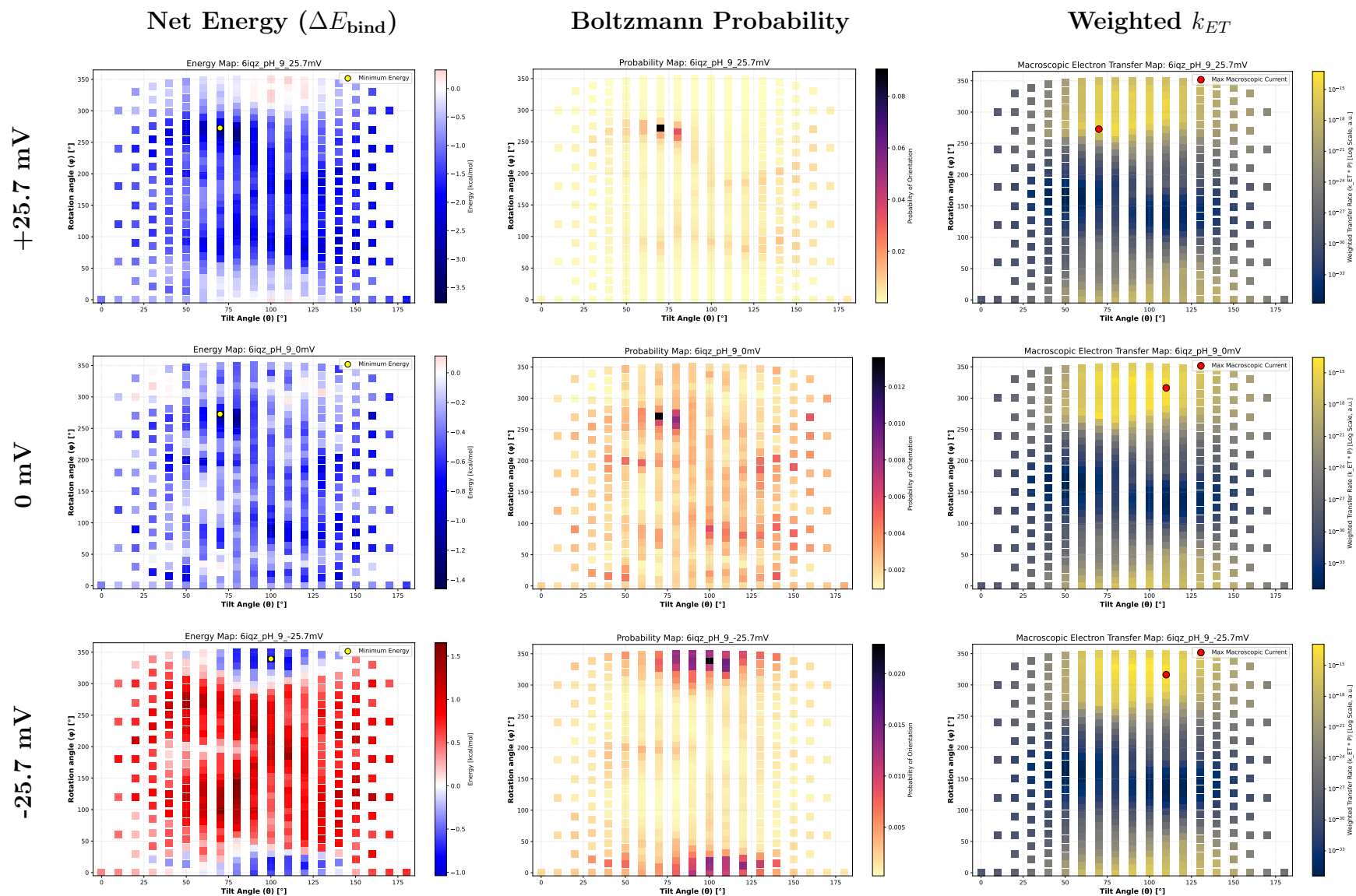


Figure 4.5: Thermodynamic and kinetic maps for Bilirubin Oxidase (BOx) at pH 9, for the three electrode potentials studied (+25.7 mV, 0 mV, -25.7 mV). Columns: net interaction energy (ΔE_{bind}), Boltzmann probability, and weighted electron transfer rate. The axes represent the relative orientation of the protein on the electrode: tilt angle (θ) and rotation angle (ϕ), with the protein electrode distance held fixed.

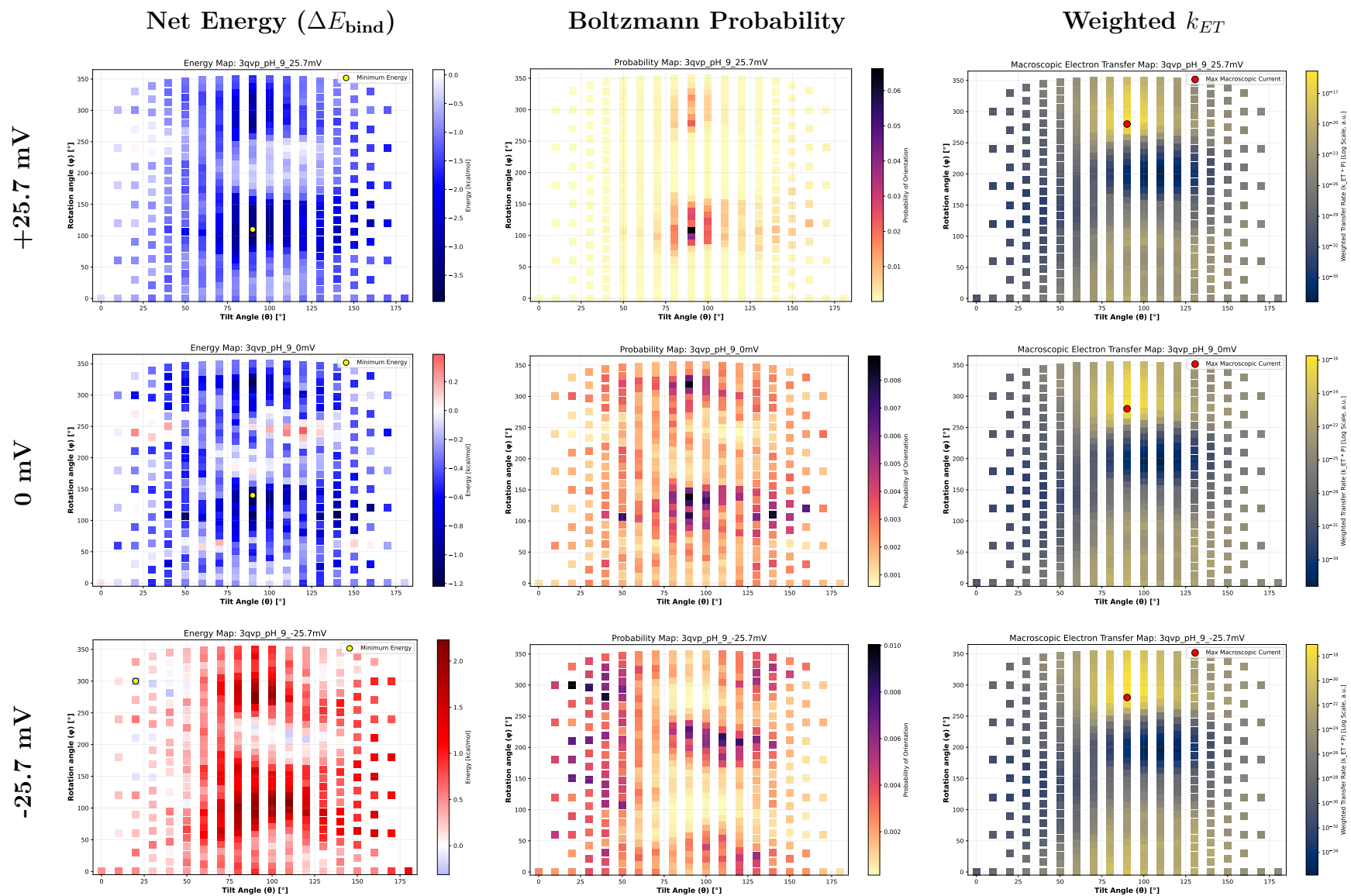


Figure 4.6: Thermodynamic and kinetic maps for Glucose Oxidase (GOx) at pH 9, for the three electrode potentials studied (+25.7 mV, 0 mV, -25.7 mV). Columns: net interaction energy (ΔE_{bind}), Boltzmann probability, and weighted electron transfer rate. The axes represent the relative orientation of the protein on the electrode: tilt angle (θ) and rotation angle (ϕ), with the protein-electrode distance held fixed.

For pH 5 (Figure B.5), the picture changes more markedly. The distribution at zero potential becomes even more dispersed and lacking in structure than at pH 7 and 9, while at negative potential a distribution similar to that of the other two pH levels is preserved, but with notably more unfavorable energies (more intense reds). The most important change is that the Boltzmann-optimal orientation at +25.7 mV is no longer (70, 272.73) but instead becomes (140, 93.91), the same secondary zone that had already emerged at pH 7, which now concentrates the greatest probability. The map at negative potential, on the other hand, is the most invariant of the three pH levels, keeping the optimal orientation at (100, 339.43) in all three cases, with an energy well that deepens slightly relative to pH 7 and 9 (Table 4.2).

For GOx at pH 9 (Figure 4.6), the positive potential (+25.7 mV) produces an almost entirely favorable energy map, with a pale central zone of near-zero energy. The corresponding probability map reveals two clearly distinguished clusters, one dominant lower cluster, where the Boltzmann-optimal orientation (90, 110.00) is located, and a secondary one with the same tilt but a larger rotation angle, around (90, 280.00).

The zero potential reproduces the same dispersed behavior observed in BOx; it is the most random map of the three, and it imitates, in a diluted way, the positive potential map, although the Boltzmann-optimal orientation does not coincide exactly, appearing instead shifted by 30 in rotation, at (90, 140.00). The negative potential map (−25.7 mV) turns out to be practically opposite to the positive potential one, most of the map shows unfavorable energies, with a central band of slightly favorable energies, and its optimal orientation is located at (20, 300.00), with a notably shallower well, more evenly spread probabilities, and high probability zones located opposite to those of the positive potential (the probability values for each orientation are reported in Table 4.2).

For GOx at pH 7 (Figure B.8), the maps reproduce the same structure as at pH 9, although with more evenly spread probabilities, which allows the secondary zones, previously minor compared to the dominant cluster at pH 9, to compete more closely with it. The relative order of dominance among potentials is preserved, the positive potential remains the most probable within this pH, unlike what was observed in BOx, where this order was reversed between pH 9 and pH 7. Furthermore, the Boltzmann-optimal orientations for each potential are identical to those at pH 9 in all three cases (Table 4.2).

The picture for GOx at pH 5 (Figure B.7) turns out to be qualitatively different from the other two pH levels, with none of its three maps reproducing the cluster structure observed previously. At positive potential, an extensive band of favorable energy occupies the lower half and the entire

right-hand side of the map, with the optimal orientation at (160, 120.00). At negative potential, the color-inversion pattern relative to the positive potential already observed in the rest of the cases is preserved, with the optimal orientation at (60, 255.48). Zero potential, on the other hand, breaks with the pattern followed so far, in all other conditions (BOx and GOx at pH 7 and 9), the zero-potential map imitated, in a diluted way, the positive potential map; here, by contrast, is the only case in which the zero potential map imitates the negative potential map instead, including the same Boltzmann-optimal orientation (60, 255.48). This could be related to the charge of the residues closest to the electrode in that orientation (Section 4.3.4), as well as to the fact that this is the only case in which the protein exhibits a positive net charge (+1 e , Table 3.1), unlike all other cases, where it is negative, consistent with the protonation discontinuity described in Section 4.2.

We now analyze the third column, corresponding to the weighted electronic current $k_{\text{ET}}(r_i) \cdot P_i$. Across the six figures, the set of clusters in this column is practically identical to that of the reference k_{ET} map (Raw k_{ET} , Section 4.3.2) of the corresponding protein. This is in agreement with the range of variation of each factor: while P_i typically varies between 10^{-2} and 10^{-5} , k_{ET} varies between 10^{-11} and 10^{-34} , several orders of magnitude more, so that in the product $k_{\text{ET}} \cdot P_i$ the spatial structure of k_{ET} dominates almost entirely over that of P_i . Even so, the zones of higher probability do slightly shift the colors of the weighted map relative to the Raw k_{ET} map, a subtle effect, barely noticeable at first glance, but consistent across the six figures.

A red dot marks, on each map, the orientation of highest weighted current. For BOx at pH 7 and pH 9 with positive potential (+25.7 mV), this point coincides exactly with the Boltzmann-optimal orientation (70, 272.73); in the remaining BOx conditions, it coincides instead with the optimal orientation of Raw k_{ET} , (110, 316.36). In GOx, by contrast, all nine conditions always coincide with the optimal orientation of Raw k_{ET} , (90, 280.00). This orientation of maximum weighted current is the one reported as θ/φ_W in Table 4.2, and its coincidence, or lack thereof, with Boltzmann or Raw k_{ET} is summarized in the Match column. The Contrib. column indicates what percentage of the total current this single orientation contributes. Across the set of 18 conditions, this value ranges from 16.59% to 69.24%, with an average of 35.93%. The total macroscopic current I_{total} , also shown in Table 4.2 and in Figure 4.7, corresponds to the sum of this product over the 406 orientations of each map. It is worth noting that, in BOx, it is precisely the conditions with Match Boltzmann (pH 7 and 9 at positive potential) that show the highest contribution percentage within their respective pH.

Looking now exclusively at the macroscopic current in Table 4.2 and Figure 4.7, GOx shows its highest currents at pH 7 and pH 9 under positive potential. This is in line with the protein's net negative charge, but it also has a more specific explanation, at both pH levels, the second

probability zone identified in the previous block coincides with the zone of highest Raw k_{ET} , which explains both the Match and the fact that the contribution exceeds 65% in both cases. In fact, at these two pH levels the GOx energy map exhibits a silhouette comparatively similar to that of the Raw k_{ET} map, an electrostatically favorable region overlaps, to a good extent, with the region of highest k_{ET} . Between these two cases, pH 7 turns out to be the best of all; this relates to what was noted in the previous block, where the energy map at pH 7 was comparatively less peaked than at pH 9, allowing the second zone of favorable probabilities to compete more closely with the most probable one, although the percentage difference between the two ($I_{\text{total}} = 2.498 \times 10^{-15}$ at pH 7 versus 2.188×10^{-15} at pH 9, a difference of $\sim 14\%$) is not particularly large.

For GOx at pH 5, the three currents turn out to be comparatively close to one another, reflecting the fact that in the energy maps the three most probable orientations are located far from the Raw k_{ET} zone. The zero potential produces a slightly higher current than the negative potential (4.016×10^{-16} versus 3.787×10^{-16}), consistent with both maps sharing the same Boltzmann-optimal orientation (60, 255.48) observed in the previous block, and with this being the only condition where the protein exhibits a positive net charge, which could explain why, in this particular case, the protein's electrostatic response does not clearly follow any of the trends observed in the rest of the conditions.

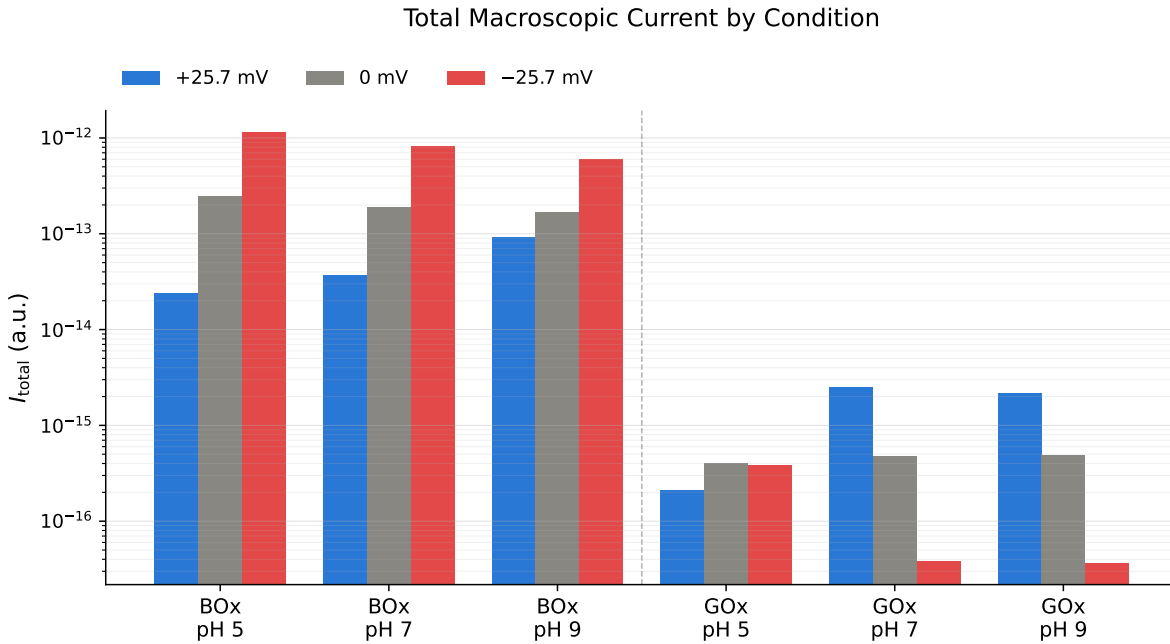


Figure 4.7: Total macroscopic current I_{total} (logarithmic scale) for the 18 conditions studied, grouped by protein and pH, with the three applied potentials shown in different colors.

Speaking specifically about BOx, the maximum currents are reached under negative potential (-25.7 mV), which is unexpected, since in the previous sections on the influence of potential and pH (Sections 4.1 and 4.2), both BOx and GOx generally favored the positive potential electrostatically, and this can easily be confirmed by observing that the $\langle \Delta E_{\text{bind}} \rangle$ column in Table 4.2 shows the most negative energies at positive electrode potential; however, when comparing this information with the weighted current, it is precisely the negative potential conditions that produce the highest current (Table 4.2, Figure 4.7). The reason can be explained by looking at the BOx energy maps at -25.7 mV, in all of them, the vast majority of the map is markedly unfavorable (red), and it is precisely the single favorable cluster, where the Boltzmann-optimal orientation is located, that falls very close to the zone of highest Raw k_{ET} . It is therefore not surprising that the weighted optimum coincides with the geometric one (Match column). Specifically, the Boltzmann-optimal angle is $(100, 339.43)$, versus the geometric optimum of Raw k_{ET} at $(110, 316.36)$; the distance Dist_Z of the Boltzmann-optimal point is $19.54 \text{ \AA}$, the closest to the absolute optimum ($17.83 \text{ \AA}$) in the whole of Table 4.2. The least favorable orientations in terms of current, by contrast, are concentrated under positive potential.

It is also interesting that the currents increase at lower pH, reaching their maximum at pH 5, 41.9% higher than the second-best case, pH 7 (Figure 4.7). This is consistent with what was observed in the previous block, at pH 5 and pH 7, the orientation of highest probability comes to correspond to the negative potential (compared to the same pH under positive or zero potential), and at pH 5 specifically the energy well deepens to -1.3520 kcal/mol, versus -1.0937 kcal/mol at pH 7, although both orientations show a practically identical Boltzmann probability, 2.53% and 2.54% respectively). This is a reminder that the Boltzmann probability of an orientation does not depend solely on the absolute depth of its own energy well, but on its energy relative to the complete set of 406 orientations sampled under that condition, a deeper energy well can correspond to a similar or even lower probability than a shallower one, if in the first case there are other nearby orientations that are almost as favorable, thus spreading the probability among several, or if in the second case the rest of the set is, on average, comparatively more unfavorable. Figure 4.1 shows the energy range without resolving the probability, and the maps in Figure 4.5 show probability and energy as a function of angle, although the color variation between nearby orientations is not always easy to appreciate visually; as a complement (not strictly necessary, but useful for appreciating this effect more clearly), Figure B.9 (Appendix) shows the complete energy and probability distribution jointly for each condition, although without angular resolution.

This result differs from what might be expected if one assumes that bilirubin oxidase shows better catalytic performance under conditions close to neutral pH. Filip et al. [36] report, via cyclic

voltammetry of BOx (*Myrothecium verrucaria*, the same species modeled in this work), that the cathodic potential of the T1 site shifts toward more positive values as pH decreases between 7.5 and 5.0 (slope of $-66 (\pm 3)$ mV/pH). Although this quantity (the experimental redox potential of T1) is not directly comparable to the electrostatic interaction energy ΔE_{bind} calculated in this work, both point in the same qualitative direction, a greater favorability of the T1 site toward the electrode under acidic conditions. This coincidence should be taken as an indication, not as a confirmed mechanistic explanation, given that the model used here captures only the transfer toward T1, governed exclusively by the donor–acceptor distance and the Boltzmann probability, without incorporating the subsequent internal electron transport toward the trinuclear cluster nor the catalytic limitations the enzyme might experience at later stages. The result obtained should therefore be interpreted as an electrostatic and geometric optimization of the initial protein–electrode coupling, and not as a complete prediction of the actual bioelectrocatalytic performance of the system.

This contrast, BOx showing higher current under negative electrode potential, and GOx under positive potential, may at first glance seem contrary to what would be expected from a spontaneously operating EFC (Section 2.2.2), where BOx should occupy the cathode (by definition, positive) and GOx the anode (by definition, negative). However, this assignment of signs arises, in a spontaneous EFC, precisely from the electrochemical need of each enzyme: bilirubin oxidase seeks to reduce oxygen to water (Equation 2.7), and to do so it needs electrons, which it draws from the nearby electrode, leaving it positively charged (cathodic); glucose oxidase, on the other hand, seeks to oxidize glucose, releasing electrons (Equation 2.6), which it delivers to the nearby electrode, charging it negatively (anodic). That is, in a real EFC the sign of each electrode is not an externally imposed condition, but rather a consequence that emerges from the enzyme’s own electrochemical reaction while adsorbed onto it.

In this work, by contrast, the potential of each electrode is imposed externally as a Dirichlet boundary condition, in a preexisting way and independently for each isolated protein. Under this imposed condition, the negative potential electrode acts, in a certain sense, as a source of electrons, so it is consistent that BOx, which seeks to reduce oxygen and therefore requires electrons, would prefer that condition. Analogously, the positive potential electrode acts as an electron sink, favoring GOx, which seeks to release electrons when oxidizing glucose. The results obtained in this work are, therefore, fully consistent with the expected direction of electron flow for each enzyme in an EFC.

4.3.4 Electrostatic Fingerprint Induced on the Electrode

The charge perturbation $\Delta\sigma$, Equation (2.47), allows the same quantity that ΔQ , Equation (2.83), reports in integrated form in Table 4.2 to be examined point by point over the electrode surface. Figures 4.8, 4.9, and 4.10 show this fingerprint for BOx and GOx, in their Boltzmann-optimal orientation, grouped by pH and for the three potentials studied. Red indicates a local increase in surface charge density attributable to the protein, blue a decrease, and white a null perturbation. The color scale is not linearly spaced with respect to $\Delta\sigma$, intervals near zero are visually stretched relative to intervals near the extremes, which improves the visibility of small perturbations that would otherwise appear too faint, two dotted lines on the color bar mark the actual maximum and minimum of $\Delta\sigma$ for that condition, facilitating its reading given this non uniform scale. On each map, using the position of the alpha carbon as the representative of each residue, the 10 residues closest to the electrode are identified, the same ones that make up Q_{10} in Table 4.2, with AA_{min} being the closest of that group, colored according to their net charge, orange for positive charge, purple for negative, and yellow for neutral. The complete list of these 10 residues, ranked by distance to the electrode, along with their net charge, for all eighteen conditions, is reported in the Appendix (Tables A.6 and A.7). Table 4.4 reports the baseline used to calculate $\Delta\sigma$ and ΔQ , symmetric in magnitude according to the sign of the potential, and null in its absence.

Table 4.4: Charge induced on the isolated electrode (without protein), $Q_{\text{electrodo}}$, for each applied potential.

Φ_e (mV)	$Q_{\text{electrodo}}$ (e)
+25.7	+15.6387
0	0.0000
-25.7	-15.6387

In all the figures, a clear correspondence is observed between the charge of the closest residue and the sign of the perturbation in its fingerprint on the electrode, negative residues generate red zones (positive perturbation) and positive residues generate blue zones (negative perturbation). Since the electrode is a conductor with free electrons, a charged residue in its vicinity locally repels or attracts those electrons, leaving behind a charge perturbation of opposite sign to that of the residue, regardless of the background potential of the electrode. The notable similarity between maps of different conditions is largely due to the fact that they share the same Boltzmann-optimal orientation, already documented in Section 4.3.3 and summarized in Table 4.2.

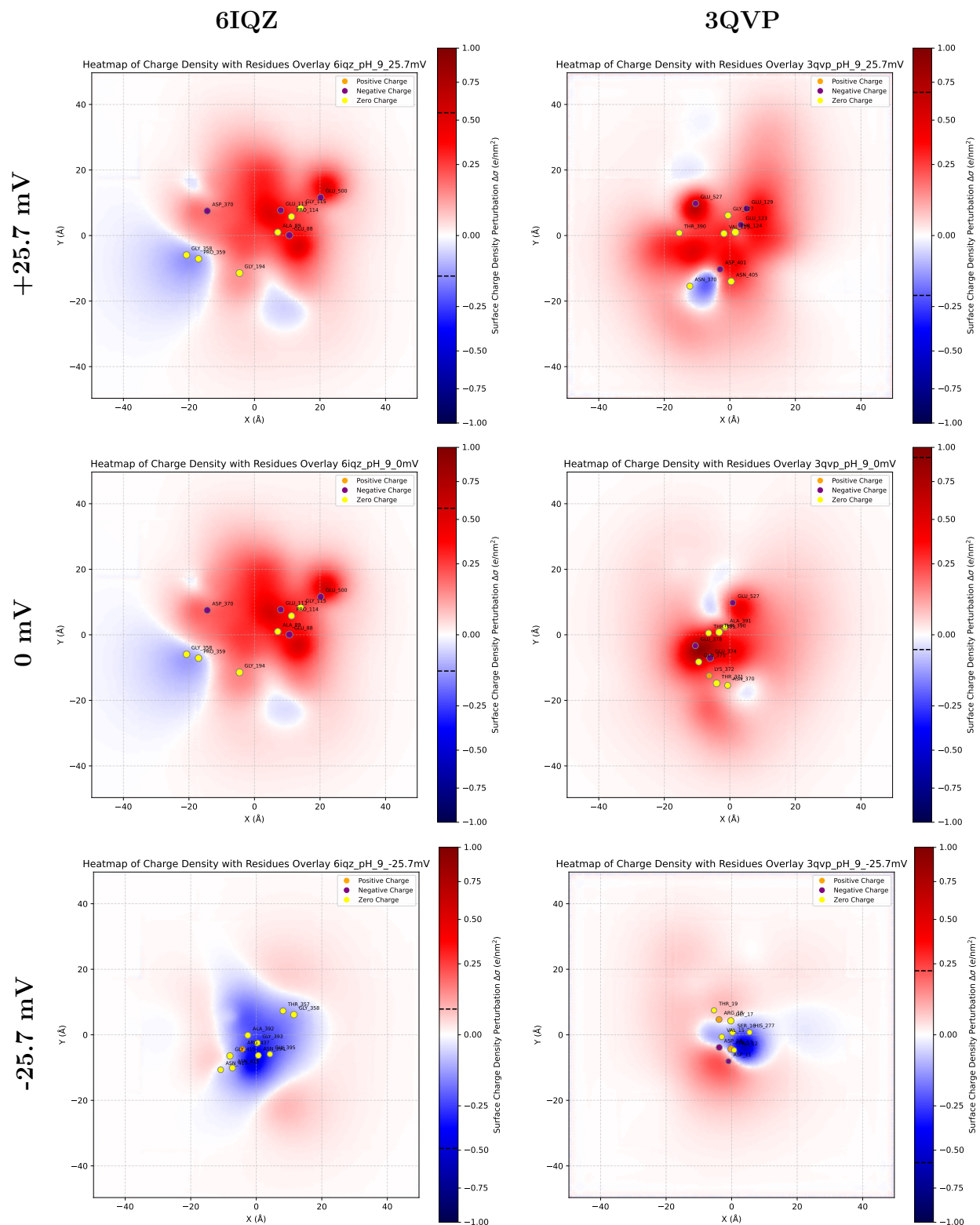


Figure 4.8: Induced surface charge density perturbation ($\Delta\sigma$) for 6IQZ (left) and 3QVP (right) at pH 9 and the three electrode potentials studied (+25.7, 0, -25.7 mV), for the highest-probability orientation of each condition. The 10 closest residues to the sensor are labeled by name and residue number.

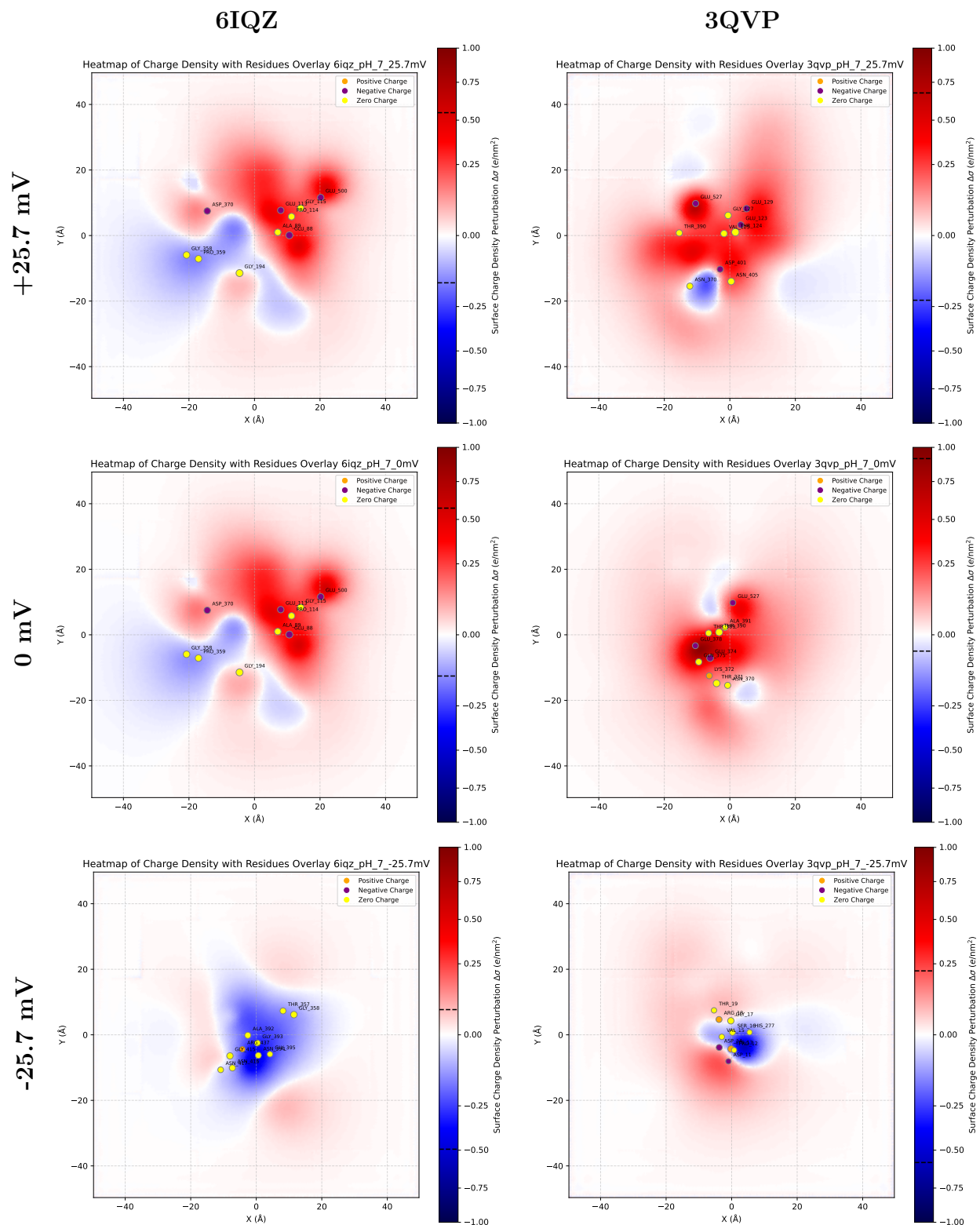


Figure 4.9: Induced surface charge density perturbation ($\Delta\sigma$) for 6IQZ (left) and 3QVP (right) at pH 7 and the three electrode potentials studied (+25.7, 0, -25.7 mV), for the highest-probability orientation of each condition. The 10 closest residues to the sensor are labeled by name and residue number.

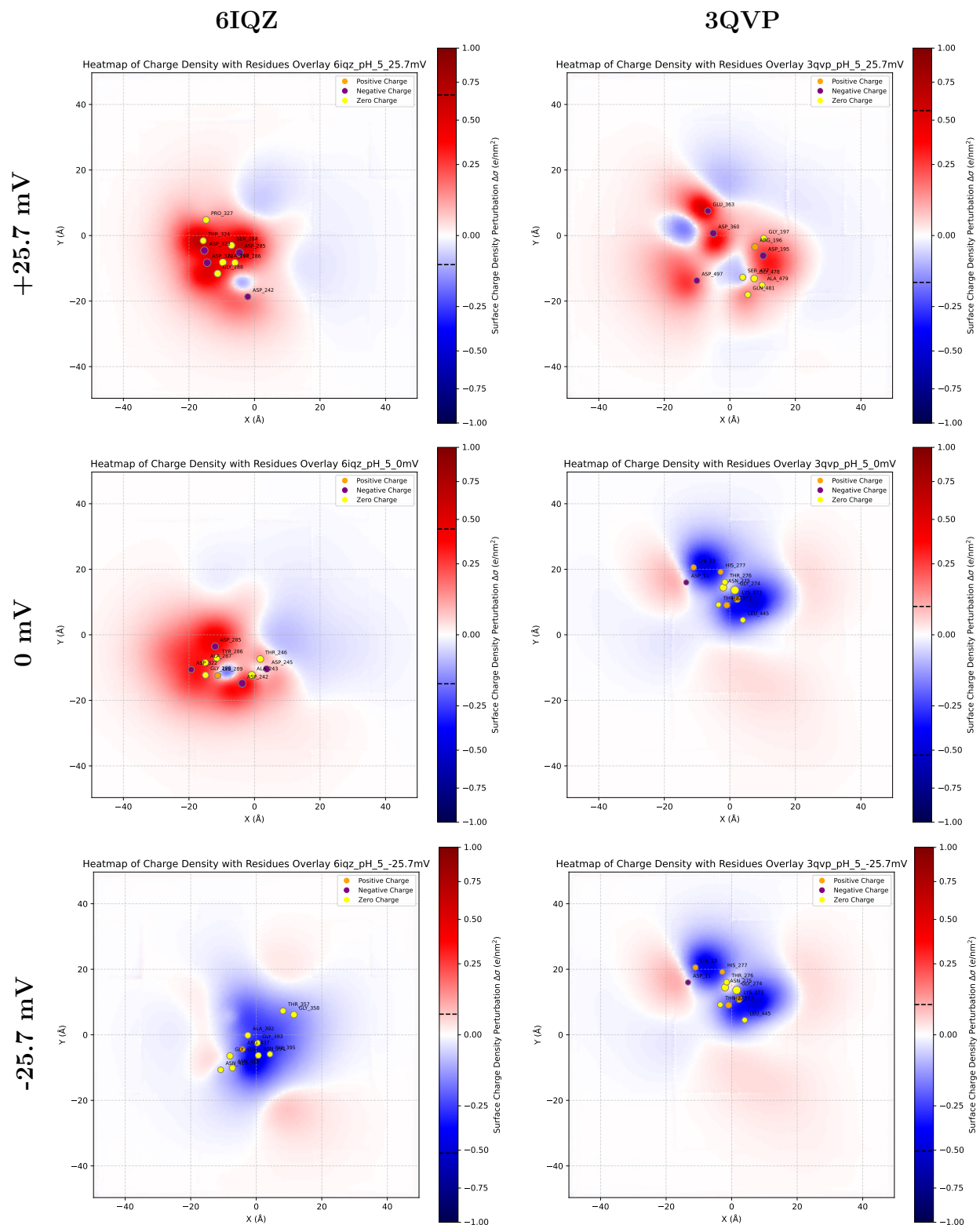


Figure 4.10: Induced surface charge density perturbation ($\Delta\sigma$) for 6IQZ (left) and 3QVP (right) at pH 5 and the three electrode potentials studied (+25.7, 0, -25.7 mV), for the highest-probability orientation of each condition. The 10 closest residues to the sensor are labeled by name and residue number.

Examining Q_{10} across all the conditions studied, two patterns emerge without exception. First, the sign of Q_{10} is always opposite to the sign of the applied potential: at $+25.7$ mV, Q_{10} is negative in all six conditions (-3 or $-4e$), while at -25.7 mV it is positive or null, in line with the macroscopic preference of the protein (of negative net charge in most conditions) for the positive electrode (Sections 4.1 and 4.2). Second, the sign of ΔQ is always opposite to that of Q_{10} , consistent with the same local effect already described. It can also be noted, comparing the position of the dotted lines on the color bar of each diagram, that the magnitude of the perturbation is likewise not symmetric between its positive and negative ends, and that the dominant side generally coincides with the sign of ΔQ for that same condition.

This could further explain the behavior at zero potential noted in Section 4.3.3, in BOx at pH 7 and pH 9, where the map at 0 mV weakly imitates that of positive potential: $Q_{10} = -4e$, the most negative value in the whole of Table 4.2, which makes that orientation especially sensitive even to the remaining residual positive field in the nominal 0 mV condition. In GOx at pH 5, by contrast, the map at 0 mV imitates that of negative potential for the opposite reason: $Q_{10} = +3e$, the most positive value in the whole table, makes that orientation reject even the remaining residual positive field, thus behaving as the diluted version of the negative potential angular map. Alternatively, or in addition, both orientations could simply be intrinsically favorable for geometric or electrostatic reasons specific to each protein, independent of the applied potential, as already raised in Section 4.3.3.

Crossing the sign of ΔQ against the sign of each protein's net charge (Table 3.1), 14/18 of the conditions show opposite signs, consistent with the protein's electrostatic preference for an electrode of opposite sign to its own charge. Of these 14, in 12 the Q_{10} has the same sign as the protein's net charge, reinforcing that result; in the remaining two (GOx at pH 7 and 9 under $\Phi_e = -25.7$ mV), $Q_{10} = 0$, neither favoring nor contradicting the net-charge trend. In the four remaining conditions, BOx at $\Phi_e = -25.7$ mV across the three pH levels and GOx at pH 5 with $\Phi_e = +25.7$ mV, ΔQ has the same sign as the protein's net charge, which appears to contradict that expected preference; in all four cases, however, this is directly explained by Q_{10} . In BOx, the protein is net negative and generally does not favor an electrode that is also negative; however, $Q_{10} = +1e$ in the Boltzmann-optimal orientation of that condition, that is, the nearby residue itself has a positive charge, opposite to that of the protein as a whole, but at the same time also opposite to that of the electrode, which is negatively charged under that condition, and it is this local complementarity that makes this orientation one of the few favorable zones available despite the protein's general rejection of that potential, thereby determining the final sign of ΔQ . The same logic, with the signs reversed, explains the case of GOx at pH 5 and $\Phi_e = +25.7$ mV, since this

is the only condition with a positive net charge, the protein generally does not favor an electrode that is also positive, but $Q_{10} = -3e$ in the Boltzmann-optimal orientation, a nearby residue of negative charge, opposite to the charge of the electrode, allowing that particular orientation to be favorable despite the contrary macroscopic trend. In no case, therefore, is the sign of ΔQ left unexplained, when it coincides with the simple capacitive model, the local distribution generally reinforces it; when it departs from that model, it is precisely the local distribution that explains why.

Restricting the comparison solely to the zero potential conditions, the only ones evaluated by Terrani in his Table 4.4, the six conditions of this work (three pH levels, two proteins) follow, without exception, the electrostatic preference expected given the protein’s net charge, whereas Terrani reports precisely the opposite for BOx at pH 7 and 9, the most stable orientation having a Q of the same sign as the protein. This comparison, restricted to the same potential condition evaluated by Terrani, puts into perspective his conclusion [1] that the local charge distribution is more decisive than the protein’s net charge and the applied potential, and that adsorption can be electrostatically unfavorable even with opposite sign charges; in this work, both mechanisms, local and global, are consistent with each other in the six conditions directly comparable with his.

This difference is markedly apparent in BOx at pH 9 and $\Phi_e = 0$ mV, the fingerprint of the orientation identified here as Boltzmann-optimal for that condition, with tilt and rotation angles (70, 272.73) (Figure 4.8), coincides in geometry, fingerprint shape, and distribution of nearby residues (not exactly, but almost completely, possibly because the angular sampling is slightly offset due to not sharing the same initial reference point) with the orientation that Terrani [1] classifies as the least stable for that same condition (his Fig. 4.17). That is, for the same condition of BOx at pH 9 and zero potential, the very same region of the protein that turns out to be the most favorable in this work is tabulated by Terrani as the least favorable. An analogous case, although not directly comparable, occurs in GOx, the orientation that Terrani identifies in his Figure 4.17 as the least stable for 3QVP at pH 9 and zero potential is almost completely similar to the one observed in this work as the most probable for 3QVP at pH 9, but under $\Phi_e = +25.7$ mV (Figure 4.8).

Looking at other aspects of the figures, when comparing BOx at pH 7 and pH 9 under the same potential (in the +25.7 mV and 0 mV cases), since they share a Boltzmann-optimal orientation and therefore a very similar electrostatic fingerprint in its general structure, the pH 7 map (Figure 4.9) shows an additional central blue thread, absent from the equivalent pH 9 map (Figure 4.8). This difference cannot be explained by looking only at the ten residues that make up Q_{10} , since these are the same under both conditions. A broader visual inspection reveals that the emerging blue zone coincides with a change in the protonation state of a somewhat more distant residue, specifically

histidine 139, corresponding to the 23rd closest residue to the electrode, only 3.3 Å farther than the closest one of all (GLY194, at 7.27 Å), and which sits right above the negative charge density zone generated, it is neutral at pH 9 but acquires a positive charge at pH 7. This illustrates that the electrostatic fingerprint is sensitive to protonation changes in residues outside the formally reported group of the 10 closest ones, this change, and its influence on the fingerprint at pH 7, could partly explain both the reduction in ΔQ between pH 9 and pH 7 (3.638 e to 2.486 e) and the shallower energy well in the same transition (−3.7709 to −2.4931 kcal/mol).

Chapter 5

Conclusions and Recommendations

5.1 Fulfillment of the Objectives

The electrostatic interaction system between BOx and GOx adsorbed on a flat electrode was modeled, evaluating each protein under nine combinations of pH and potential (18 conditions in total, 406 orientations each), using the boundary element method to solve the LPBE with PyGBe, with the results condensed in Table 4.2 (Chapter 4).

The complete mathematical formulation of electrostatics in a continuous dielectric medium was established, from Gauss's law to the LPBE, as well as the derivation of the boundary element method and its extension to multi-surface systems, including the protein-solvent interface, the flat electrode, the internal cavities, and the Stern layer (Chapter 2).

A Python code was implemented that automates the sampling of the 406 orientations and the sequential execution of PyGBe, extending the work of Urzúa et al. [3] with improvements in respecting the solid angle during sampling, the use of the principal axis of inertia as an angular reference independent of the protonation state, chirality verification at each rotation, and the coordinated rotation of the internal cavity meshes together with the protein; beyond the original objective, the code additionally incorporates the calculation of electron transfer kinetics and electrostatic fingerprints (Section 3.6).

The results were post-processed, generating energy and Boltzmann probability maps for the 18 conditions, additionally incorporating the weighted current and the electrostatic fingerprints (Chapter 4).

The adsorption stability was analyzed and the orientational sensitivity to pH and potential

was quantified (17/18 and 12/18 conditions, respectively), identifying for each condition the Boltzmann-optimal orientation and its associated probability (Table 4.2), by means of the ranges of variation reported in Table 4.1. The complete orientational energy and probability maps for the 406 microstates of each condition were also examined, identifying their organization into clusters or contiguous zones, generally opposite to one another upon reversing the sign of the applied potential, and considerably more scattered and lacking in structure under zero potential (Chapter 4).

5.2 Main Findings

The six most relevant findings of this work are the following.

First, orientation consistently dominates over pH (17/18 conditions) and over the applied potential in most cases (12/18), but when directly comparing pH and potential against each other, it is the potential that dominates over pH for both enzymes, a result opposite to that reported both by Terrani [1] and by Ruiz et al. [2]. As discussed in Section 4.2, this discrepancy is unlikely to stem from Terrani’s use of the full nonlinear Poisson-Boltzmann equation rather than its linearized form, since his own validation shows strong agreement between both formulations, nor from any general omission of the surface energy term, which he does compute, though only for a single reference orientation rather than as a Boltzmann average; even when restricting the analysis in this work to the same solvation-only, macro-averaged quantity used by Terrani, the applied potential still dominates over pH (Section 4.3.1), indicating that this asymmetry alone does not account for the discrepancy. Several methodological differences could instead explain it, both prior works include an explicit Stern layer, absent in this one, and Terrani additionally evaluates a considerably wider range of applied potentials. However, the author considers the most plausible explanation to be a numerical limitation specific to the finite-difference approach used by Terrani, in which the interpolation of point charges onto a fixed background grid introduces a position-dependent error that could not be reliably corrected when the protein is reoriented, a limitation the boundary element method used here is not subject to.

Second, the macroscopic current does not necessarily follow the electrostatic preference of each protein, although both enzymes electrostatically favor the positive potential, BOx reaches its maximum current under negative potential. This is coherently explained by the redox role of each enzyme in an enzymatic fuel cell (EFC), BOx reduces oxygen and requires electrons, while GOx oxidizes glucose and releases them, and the observed direction of electron flow, BOx favored by negative potential, GOx by positive, is consistent with the signs each enzyme would require in a real EFC, even though here the potential is imposed as an external boundary condition rather than

emerging spontaneously from the reaction. This is further reinforced by the geometric coincidence, in the BOx condition at negative potential, between the Boltzmann-optimal orientation and the optimal orientation for pure electron transfer (Section 4.3.3).

Third, the local charge distribution (quantified via Q_{10}) reinforces, in most conditions, the macroscopic net charge preference of the protein, rather than contradicting it as Terrani concludes for his own system; in 14/18 conditions the sign of ΔQ is opposite to that of the net charge, consistent with the protein’s electrostatic preference for an electrode of opposite sign to its own charge, actively reinforced by Q_{10} in 12 of them; the four exceptions are completely explained by local residues of unexpected sign. This suggests that the local charge distribution exerts a secondary but real influence on the interaction energy and the preferential orientation, in the vast majority of conditions, the macroscopic behavior (net charge) remains the determining factor, while the local distribution acts as a fine modulator and explains the specific exceptions to the general pattern, including, notably, the finding in the second point regarding the maximum current of BOx. Restricting the comparison to the zero potential conditions, the only ones evaluated by Terrani, this relationship holds without exception in this work, whereas he reports precisely the opposite for BOx at pH 7 and 9 (Section 4.3.4).

Fourth, BOx transfers electrons several orders of magnitude faster than GOx in its optimal orientation (17.83 versus 21.67 Å), with qualitatively distinct angular patterns of k_{ET} , a smooth transition between favorable and unfavorable zones in BOx, versus a somewhat more abrupt transition in GOx, for which two unconfirmed hypotheses are proposed, an oligosaccharide acting as a geometric arm, and a greater relative asphericity of the protein. These optimal distances also fall within the viable direct electron transfer (DET) regime reported in the literature, although at the lower limit of efficiency for both enzymes, BOx within the range but above the efficient transfer threshold, and GOx already at the limit beyond which mediated transfer (MET) is typically required (Section 4.3.2).

Fifth, the protonation discontinuity between pH 5 and pH 7, consistent with the pKa values of ionizable residues predicted by PROPKA in the original preparation of the structures [1], dominates over the pH 7–9 transition in most of the analyses carried out in terms of macroscopic energy, with the notable exception of BOx at $\Phi_e = +25.7$ mV, where the pH 7–9 jump turns out to be larger than the pH 5–7 jump (Section 4.2). In particular, pH 5 turns out to be the qualitatively most atypical condition of the set for both enzymes, although more markedly so in GOx. This atypicality is particularly noticeable in the orientational maps of GOx at pH 5, where none of the three potentials reproduces the same cluster structure observed under the other conditions, unlike

BOx at pH 5, whose map at negative potential remains practically invariant relative to the other two pH levels (Section 4.3.3).

Sixth, using only ΔE_{solv} as an approximation both for the orientational ranking and for the macroscopic trend of preferred charge turns out to be reasonable, although not exact (15/18 agreement in optimal orientation, Spearman correlation $\rho = 0.83\text{--}1.00$), quantitatively confirming a qualitative intuition long held by the research group, but never rigorously verified before (Section 4.3.1).

5.3 Recommendations

This work presents limitations that open up concrete directions for future work. The rigid body approximation does not capture conformational fluctuations or protonation changes induced by proximity to the electrode, this could be addressed with molecular dynamics. The absence of a Stern layer is a real physical difference with respect to Terrani and Ruiz et al. whose effect could not be isolated; incorporating it would make it possible to determine whether it explains the discrepancy observed in the dominance of pH and potential. The partial treatment of internal cavities and the potential range restricted to the validity of the LPBE are also approximations whose quantitative effect remains to be evaluated in greater detail.

The simplified Marcus model does not incorporate the complete redox cascade or the catalytic limitations of each enzyme, extending it would allow more direct comparisons with experimental activity. Given that the donor–acceptor distances obtained lie close to the limits of the viable DET regime, it is recommended to determine the decay parameter β specific to each protein instead of the generic consensus value used here ($1.4 \text{ \AA}^{-1}$, particularly severe); Ruiz et al. [2] uses $\beta = 0.45 \text{ \AA}^{-1}$ for his hydrogenase, which suggests that a protein-specific value could significantly change the viability of the DET regime compared to the one adopted here.

The hypotheses regarding the greater angular sensitivity of GOx (oligosaccharide, asphericity) are based on only two proteins and require a broader set for their confirmation. Given the geometric nature of this finding itself, and since only a flat electrode was used here, it would be of interest to explore electrodes of different geometry. Finally, it is recommended to explore larger potentials using the complete nonlinear PBE, as well as to experimentally validate the preferential orientations predicted here.

Regarding the pH treatment, all PQR files used in this work were reused directly from Terrani [1], without independently regenerating or reparametrizing them. While the resulting files were carefully inspected (verifying residue counts, protonation patterns, and net charge consistency) this reliance means that any undetected error in the original parametrization would propagate unnoticed into this work. Regenerating the PQR files independently, using PDB2PQR and PROPKA under the same conditions, and comparing PROPKA against alternative pKa estimators, would provide an additional layer of validation and help assess how sensitive the orientational and energetic conclusions of this work are to this particular choice.

From a coding perspective, the automation code could be made more robust and user friendly. Part of the PQR parsing reads fields from fixed positions in each line, instead of separating them by blank spaces, which makes that part of the pipeline sensitive to how the file was generated; a more consistent parsing strategy across all scripts would reduce this dependency. More broadly, simulation parameters, such as solid-angle sampling, directory naming, computational environment, or whether to save the `.phi` output, are scattered across multiple scripts and shell files rather than centralized; consolidating them into a single configuration file would reduce redundancy and ease future use.

Bibliography

- [1] Leonardo Luca Terrani. Poisson-boltzmann approach to the bioelectrochemistry of adsorbed enzymes. Master's thesis, Politecnico di Milano, Milan, Italy, 2024.
- [2] Manuel Antonio Ruiz-Rodríguez, Christopher D. Cooper, Walter Rocchia, Mosè Casalegno, Yossef López de los Santos, and Guido Raos. Modeling of the electrostatic interaction and catalytic activity of [nife] hydrogenases on a planar electrode. *The Journal of Physical Chemistry B*, 126(43):8777–8790, 2022. PMID: 36269122.
- [3] Sergio A. Urzúa, Perla Y. Saucedo-Oloño, Carlos D. García, and Christopher D. Cooper. Input files for "predicting the orientation of adsorbed proteins steered with electric fields using a simple electrostatic model", July 2022.
- [4] Andreas S Bommarius and Bettina R Riebel-Bommarius. *Biocatalysis: fundamentals and applications*. John Wiley & Sons, 2004.
- [5] Zhirong Sun. *Foundations for the Study of Structure and Function of Proteins*, pages 303–336. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.
- [6] Tamar Schlick. *Molecular modeling and simulation: an interdisciplinary guide*, volume 2. Springer, 2010.
- [7] Jianting Yang, Qian Li, and Liujiao Bian. Spectroscopic analysis and docking simulation on the recognition and binding of tem-1 β -lactamase with β -lactam antibiotics. *Experimental and Therapeutic Medicine*, 14(4):3288–3298, 2017.
- [8] A.R. Leach. *Molecular Modelling: Principles and Applications*. Prentice Hall, 2001.
- [9] Sergio A. Urzúa, Perla Y. Saucedo-Oloño, Carlos D. García, and Christopher D. Cooper. Predicting the orientation of adsorbed proteins steered with electric fields using a simple electrostatic model. *The Journal of Physical Chemistry B*, 126(28):5231–5240, 2022. PMID: 35819287.

- [10] Hui Chen, Olja Simoska, Koun Lim, Matteo Grattieri, Mengwei Yuan, Fangyuan Dong, Yoo Seok Lee, Kevin Beaver, Samali Weliwatte, Erin M Gaffney, et al. Fundamentals, applications, and future directions of bioelectrocatalysis. *Chemical Reviews*, 120(23):12903–12993, 2020.
- [11] Assil Guessab, I Mazurenko, E Lojou, and Anne de Poulpiquet. Electrochemistry of redox enzymes: from functional enzyme immobilization to enzymatic bio-electrochemical devices (tutorial). *Bioelectrochemistry*, page 109265, 2026.
- [12] Eileen Hao Yu and Keith Scott. Enzymatic biofuel cells—fabrication of enzyme electrodes. *Energies*, 3(1):23–42, 2010.
- [13] Yoo Seok Lee, Koun Lim, and Shelley D Minter. Cascaded biocatalysis and bioelectrocatalysis: overview and recent advances. *Annual Review of Physical Chemistry*, 72(1):467–488, 2021.
- [14] Benoit Roux and Thomas Simonson. Implicit solvent models. *Biophysical Chemistry*, 78(1):1–20, 1999.
- [15] David Andelman. Introduction to electrostatics in soft and biological matter. *Soft condensed matter physics in molecular and cell biology*, 6:97–122, 2006.
- [16] Christopher Cooper Villagran. *Biomolecular electrostatics with continuum models: A boundary integral implementation and applications to biosensors*. Boston University, 2015.
- [17] Carlos Alberto Brebbia and Jose Dominguez. *Boundary elements: an introductory course*. WIT press, 1994.
- [18] Rudolph A Marcus and Norman Sutin. Electron transfers in chemistry and biology. *Biochimica et Biophysica Acta (BBA)-Reviews on Bioenergetics*, 811(3):265–322, 1985.
- [19] Jennifer R DeRosa, Tian Qiu, D Vale Cofer-Shabica, and Joseph E Subotnik. Marcus theory and the condon approximation revisited ii: The horror of triplet energy transfer. *Journal of Chemical Theory and Computation*, 22(4):1584–1593, 2026.
- [20] Jorge Ortiz Pinilla and Andrés Felipe Ortiz Rico. ¿pearson y spearman, coeficientes intercambiables? *Comunicaciones en Estadística*, 14(1):53–63, 2021.
- [21] Christopher D. Cooper, Natalia C. Clementi, Gilbert Forsyth, and Lorena A. Barba. PyGBe: Python, GPUs and boundary elements for biomolecular electrostatics. *JOSS*, 1(4):43, aug 2016.

- [22] Phaneeswara-Rao Kommoju, Zhi-wei Chen, Robert C. Bruckner, F. Scott Mathews, and Marilyn Schuman Jorns. Probing oxygen activation sites in two flavoprotein oxidases using chloride as an oxygen surrogate. *Biochemistry*, 50(24):5521–5534, 2011. PMID: 21568312.
- [23] Elizabeth Jurrus, Dave Engel, Keith Star, Kyle Monson, Juan Brandi, Lisa E Felberg, David H Brookes, Leighton Wilson, Jiahui Chen, Karina Liles, et al. Improvements to the apbs biomolecular solvation software suite. *Protein science*, 27(1):112–128, 2018.
- [24] Chresten R Søndergaard, Mats HM Olsson, Michał Rostkowski, and Jan H Jensen. Improved treatment of ligands and coupling effects in empirical calculation and rationalization of p k a values. *Journal of chemical theory and computation*, 7(7):2284–2295, 2011.
- [25] David A Case, Hasan Metin Aktulga, Kellon Belfon, David S Cerutti, G Andrés Cisneros, Vinícius Wilian D Cruzeiro, Negin Forouzesh, Timothy J Giese, Andreas W Gotz, Holger Gohlke, et al. Ambertools. *Journal of chemical information and modeling*, 63(20):6183–6191, 2023.
- [26] Araz Jakalian, David B Jack, and Christopher I Bayly. Fast, efficient generation of high-quality atomic charges. am1-bcc model: Ii. parameterization and validation. *Journal of computational chemistry*, 23(16):1623–1641, 2002.
- [27] Pietro GA Aronica, Stephen J Fox, and Chandra S Verma. Comparison of charge derivation methods applied to amino acid parameterization. *ACS omega*, 3(4):4664–4673, 2018.
- [28] Junmei Wang, Wei Wang, Peter A Kollman, and David A Case. Automatic atom type and bond type perception in molecular mechanical calculations. *Journal of molecular graphics and modelling*, 25(2):247–260, 2006.
- [29] Stephen M Jones and Edward I Solomon. Electron transfer and reaction mechanism of laccases. *Cellular and molecular life sciences*, 72(5):869–883, 2015.
- [30] Takeshi Sakurai and Kunishige Kataoka. Basic and applied features of multicopper oxidases, cueo, bilirubin oxidase, and laccase. *The Chemical Record*, 7(4):220–229, 2007.
- [31] Sergio Decherchi and Walter Rocchia. A general and robust ray-casting-based algorithm for triangulating surfaces at the nanoscale. *PloS one*, 8(4):e59744, 2013.
- [32] Mahfuza Akter, Takaki Tokiwa, Mitsuo Shoji, Koji Nishikawa, Yasuteru Shigeta, Takeshi Sakurai, Yoshiki Higuchi, Kunishige Kataoka, and Naoki Shibata. Redox potential-dependent formation of an unusual his–trp bond in bilirubin oxidase. *Chemistry – A European Journal*, 24(68):18052–18058, 2018.

- [33] Pablo Ramírez, Nicolas Mano, Rafael Andreu, Tautgirdas Ruzgas, Adam Heller, Lo Gorton, and Sergey Shleev. Direct electron transfer from graphite and functionalized gold electrodes to t1 and t2/t3 copper centers of bilirubin oxidase. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, 1777(10):1364–1369, 2008.
- [34] MDRG González, Virgilio Bocanegra García, Gildardo Rivera, and Isaías Balderas Rentería. Biotecnología y enzimas de interés en alimentos: Glucosa oxidasa, mecanismo de acción, importancia industrial y aplicaciones. *Avances deficiencia y tecnología alimentaria en México*, pages 617–633, 2012.
- [35] Muhamed Amin and Jochen Küpper. Variations in proteins dielectric constants. *ChemistryOpen*, 9(6):691–694, 2020.
- [36] Jaroslav Filip and Jan Tkac. The ph dependence of the cathodic peak potential of the active sites in bilirubin oxidase. *Bioelectrochemistry*, 96:14–20, 2014.

Appendix A

Supplementary Tables

Table A.1: Absolute free energy components of the isolated protein, for each protein and pH.

Protein	pH	E_{solv} (kcal/mol)	E_{surf} (kcal/mol)	E_{Coulomb} (kcal/mol)
BOx (6IQZ)	5	-1307.0134	0	-50071.0814
	7	-1464.6497	0	-50190.6112
	9	-1660.1130	0	-50056.0824
GOx (3QVP)	5	-1403.9674	0	-54005.2085
	7	-2416.0516	0	-53371.3118
	9	-2791.6220	0	-52971.9327

Table A.2: Absolute free energy components of the isolated electrode, for each applied potential.

Φ_e (mV)	E_{solv} (kcal/mol)	E_{surf} (kcal/mol)	E_{Coulomb} (kcal/mol)
+25.7	0	-11.8243	0
0	0	≈ 0	0
-25.7	0	-11.8243	0

Table A.3: Range of variation in interaction energy ΔE_{bind} induced by orientation, potential, and pH, for each of the 18 conditions studied (all values in kcal/mol). The last column indicates which of the two variables (potential, pH, or both) the orientation range exceeds.

Protein	pH	Φ_e (mV)	Orient. range	Potential range	pH range	Dominates over
BOx	5	+25.7	2.86	0.92	0.95	Both
BOx	5	0	0.97	0.92	0.31	Both
BOx	5	-25.7	2.81	0.92	0.32	Both
BOx	7	+25.7	3.07	1.51	0.95	Both
BOx	7	0	1.07	1.51	0.31	pH only
BOx	7	-25.7	2.72	1.51	0.32	Both
BOx	9	+25.7	4.10	2.04	0.95	Both
BOx	9	0	1.57	2.04	0.31	pH only
BOx	9	-25.7	2.69	2.04	0.32	Both
GOx	5	+25.7	2.32	0.33	2.39	Potential only
GOx	5	0	1.23	0.33	0.56	Both
GOx	5	-25.7	2.54	0.33	0.51	Both
GOx	7	+25.7	3.65	2.65	2.39	Both
GOx	7	0	1.57	2.65	0.56	pH only
GOx	7	-25.7	2.52	2.65	0.51	pH only
GOx	9	+25.7	4.05	3.05	2.39	Both
GOx	9	0	1.61	3.05	0.56	pH only
GOx	9	-25.7	2.54	3.05	0.51	pH only

Table A.4: Three orientations with the highest and lowest electron transfer rate constant k_{ET} for both enzymes.

Protein	Range	θ (°)	φ (°)	r (Å)	k_{ET} (a.u.)
BOx (6IQZ)	Top 3	110.00	316.36	17.8317	1.4391×10^{-11}
		100.00	329.14	18.2315	8.2226×10^{-12}
		110.00	305.45	18.3157	7.3084×10^{-12}
	Bottom 3	50.00	160.00	53.6356	2.4485×10^{-33}
		50.00	173.33	53.3113	3.8554×10^{-33}
		60.00	162.58	53.1754	4.6636×10^{-33}
GOx (3QVP)	Top 3	90.00	280.00	21.6732	6.6441×10^{-14}
		90.00	290.00	22.3823	2.4622×10^{-14}
		100.00	288.00	22.6377	1.7218×10^{-14}
	Bottom 3	110.00	196.36	55.0482	3.3887×10^{-34}
		110.00	207.27	54.7316	5.2787×10^{-34}
		100.00	195.43	54.5594	6.7178×10^{-34}

Table A.5: Optimal orientation, Boltzmann probability, and expected value $\langle \Delta E_{\text{solv}} \rangle^*$, calculated using a Boltzmann distribution defined exclusively by the solvation energy ΔE_{solv} (without the surface term), for the 18 conditions studied. The asterisk distinguishes this quantity from $\langle \Delta E_{\text{solv}} \rangle$ (Section 4.1), where the Boltzmann weights come from the complete binding energy; this table evaluates the robustness of the orientational ranking against the inclusion of the surface term. The last column reports the Spearman correlation by Equation (2.91), between the ranking of the 406 orientations induced by ΔE_{solv} and by ΔE_{bind} (Table 4.2).

Protein	pH	Φ_e (mV)	$\theta/\varphi_{B,\text{solv}}$ (deg)	$\Delta E_{\text{solv}}^{\text{min}}$ (kcal/mol)	$P_{B,\text{solv}}$ %	$\langle \Delta E_{\text{solv}} \rangle^*$ (kcal/mol)	ρ
6IQZ	5	+25.7	160/240.00	-1.2555	0.93	-0.5959	0.9825
6IQZ	5	0	150/63.53	-0.6546	0.57	-0.1911	1.0000
6IQZ	5	-25.7	100/339.43	-0.9311	1.17	-0.1064	0.9195
6IQZ	7	+25.7	70/272.73	-1.7558	1.28	-0.8991	0.9868
6IQZ	7	0	70/272.73	-0.8777	0.66	-0.3264	1.0000
6IQZ	7	-25.7	100/339.43	-0.8302	1.10	-0.0174	0.8828
6IQZ	9	+25.7	70/272.73	-2.6931	4.13	-1.2147	0.9828
6IQZ	9	0	70/272.73	-1.4620	1.36	-0.5022	1.0000
6IQZ	9	-25.7	100/339.43	-0.8386	0.97	-0.1033	0.8509
3QVP	5	+25.7	160/120.00	-0.7493	0.86	-0.0859	0.9342
3QVP	5	0	60/255.48	-0.4568	0.60	+0.0357	1.0000
3QVP	5	-25.7	60/255.48	-0.8029	1.07	-0.0098	0.9339
3QVP	7	+25.7	90/110.00	-2.3661	2.37	-1.2705	0.9769
3QVP	7	0	90/140.00	-1.1514	0.87	-0.4630	1.0000
3QVP	7	-25.7	50/106.67	-0.5251	0.66	+0.0113	0.8498
3QVP	9	+25.7	90/110.00	-2.6262	2.84	-1.4734	0.9769
3QVP	9	0	90/140.00	-1.2186	0.89	-0.5281	1.0000
3QVP	9	-25.7	50/106.67	-0.5889	0.75	+0.0201	0.8368

Table A.6: Ten residues closest to the electrode, ranked by distance, for each of the nine conditions of BOx (6IQZ). Each cell shows residue name, number, and net charge in parentheses.

Rank	+25.7	pH 5 0	-25.7	+25.7	pH 7 0	-25.7	+25.7	pH 9 0	-25.7
1	ASP285 (-1)	ALA243 (0)	GLY358 (0)	GLY194 (0)	GLY194 (0)	GLY358 (0)	GLY194 (0)	GLY194 (0)	GLY358 (0)
2	ASP322 (-1)	ASP242 (-1)	GLY415 (0)	GLU88 (-1)	GLU88 (-1)	GLY415 (0)	GLU88 (-1)	GLU88 (-1)	GLY415 (0)
3	ASP323 (-1)	ASP285 (-1)	ASN416 (0)	PRO359 (0)	PRO359 (0)	ASN416 (0)	PRO359 (0)	PRO359 (0)	ASN416 (0)
4	ALA287 (0)	THR246 (0)	ALA392 (0)	GLY358 (0)	GLY358 (0)	ALA392 (0)	GLY358 (0)	GLY358 (0)	ALA392 (0)
5	GLY288 (0)	TYR286 (0)	GLY393 (0)	GLY115 (0)	GLY115 (0)	GLY393 (0)	GLY115 (0)	GLY115 (0)	GLY393 (0)
6	TYR286 (0)	GLY288 (0)	THR357 (0)	PRO114 (0)	PRO114 (0)	THR357 (0)	PRO114 (0)	PRO114 (0)	THR357 (0)
7	PRO327 (0)	ALA287 (0)	ASN394 (0)	ASP370 (-1)	ASP370 (-1)	ASN394 (0)	ASP370 (-1)	ASP370 (-1)	ASN394 (0)
8	SER284 (0)	ASP245 (-1)	ASN417 (0)	GLU500 (-1)	GLU500 (-1)	ASN417 (0)	GLU500 (-1)	GLU500 (-1)	ASN417 (0)
9	THR324 (0)	ASP322 (-1)	ARG437 (1)	ALA89 (0)	ALA89 (0)	ARG437 (1)	ALA89 (0)	ALA89 (0)	ARG437 (1)
10	ASP242 (-1)	LYS289 (1)	GLY395 (0)	GLU113 (-1)	GLU113 (-1)	GLY395 (0)	GLU113 (-1)	GLU113 (-1)	GLY395 (0)
Q_{10} (e)	-4	-3	+1	-4	-4	+1	-4	-4	+1

Table A.7: Ten residues closest to the electrode, ranked by distance, for each of the nine conditions of GOx (3QVP). Each cell shows residue name, number, and net charge in parentheses.

Rank	+25.7	pH 5 0	-25.7	+25.7	pH 7 0	-25.7	+25.7	pH 9 0	-25.7
1	ASP195 (-1)	GLY274 (0)	GLY274 (0)	THR124 (0)	THR390 (0)	LYS13 (1)	THR124 (0)	THR390 (0)	LYS13 (1)
2	ASP360 (-1)	ASN275 (0)	ASN275 (0)	ASN405 (0)	THR371 (0)	GLY17 (0)	ASN405 (0)	THR371 (0)	GLY17 (0)
3	GLY478 (0)	LYS273 (1)	LYS273 (1)	VAL125 (0)	GLU374 (-1)	ASP14 (-1)	VAL125 (0)	GLU374 (-1)	ASP14 (-1)
4	ARG196 (1)	HIS272 (1)	HIS272 (1)	GLY127 (0)	GLN375 (0)	ARG18 (1)	GLY127 (0)	GLN375 (0)	ARG18 (1)
5	GLY197 (0)	THR276 (0)	THR276 (0)	GLU129 (-1)	ASN370 (0)	SER16 (0)	GLU129 (-1)	ASN370 (0)	SER16 (0)
6	GLU363 (-1)	LYS13 (1)	LYS13 (1)	THR390 (0)	THR389 (0)	VAL15 (0)	THR390 (0)	THR389 (0)	VAL15 (0)
7	ASP497 (-1)	HIS277 (1)	HIS277 (1)	ASP401 (-1)	LYS372 (1)	THR19 (0)	ASP401 (-1)	LYS372 (1)	THR19 (0)
8	SER477 (0)	THR271 (0)	THR271 (0)	GLU527 (-1)	GLU378 (-1)	ASP11 (-1)	GLU527 (-1)	GLU378 (-1)	ASP11 (-1)
9	GLN481 (0)	LEU445 (0)	LEU445 (0)	GLU123 (-1)	GLU527 (-1)	PRO12 (0)	GLU123 (-1)	GLU527 (-1)	PRO12 (0)
10	ALA479 (0)	ASP11 (-1)	ASP11 (-1)	ASN370 (0)	ALA391 (0)	HIS277 (0)	ASN370 (0)	ALA391 (0)	HIS277 (0)
Q_{10} (e)	-3	+3	+3	-4	-2	0	-4	-2	0

Appendix B

Supplementary Figures

Macrostate $\langle \Delta E_{solv} \rangle$ vs applied potential

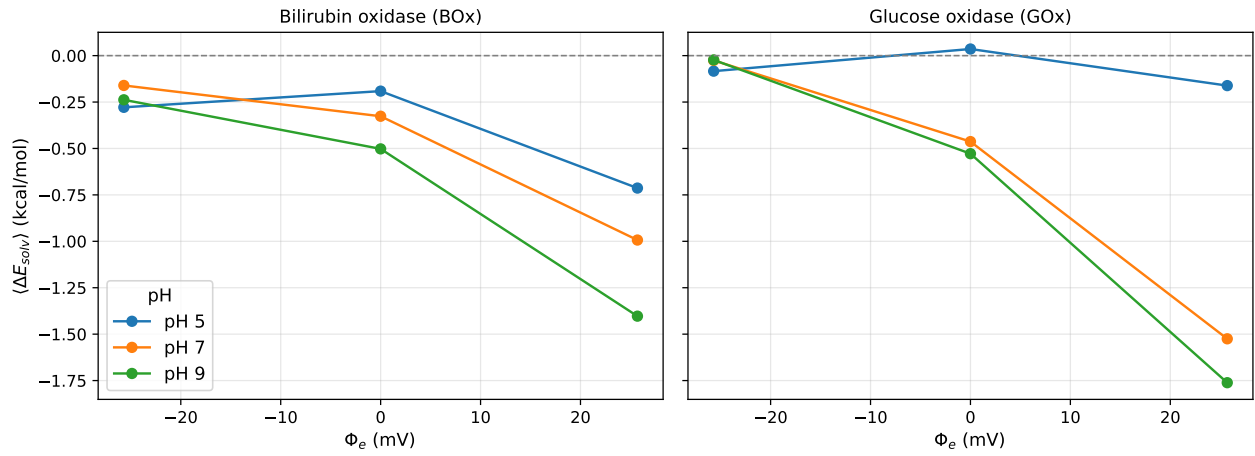


Figure B.1: Solvation contribution to the expected value of the interaction energy, $\langle \Delta E_{solv} \rangle$, as a function of the applied potential Φ_e , for each pH and both enzymes.

Macrostate $\langle \Delta E_{surf} \rangle$ vs applied potential

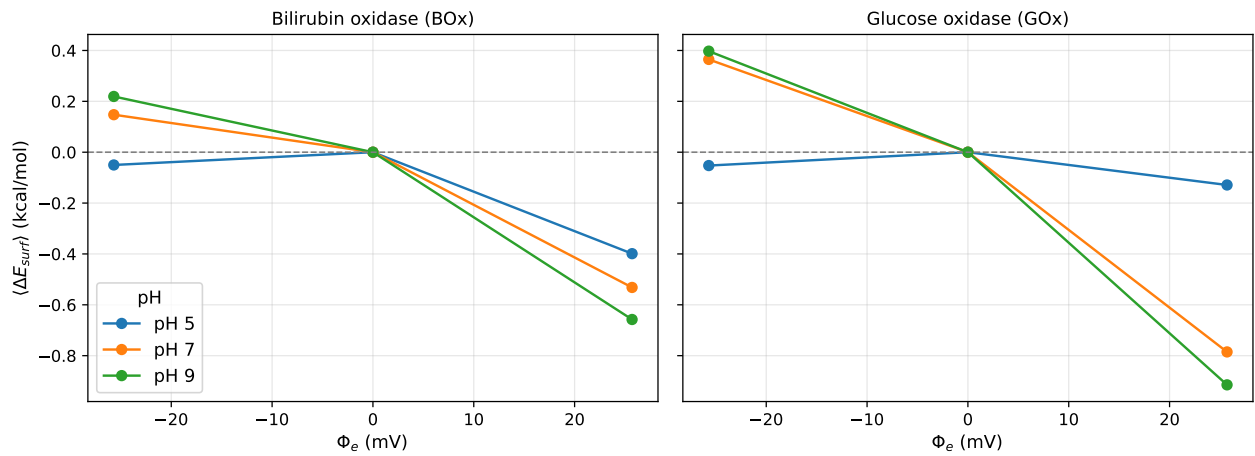


Figure B.2: Surface contribution to the expected value of the interaction energy, $\langle \Delta E_{surf} \rangle$, as a function of the applied potential Φ_e , for each pH and both enzymes.

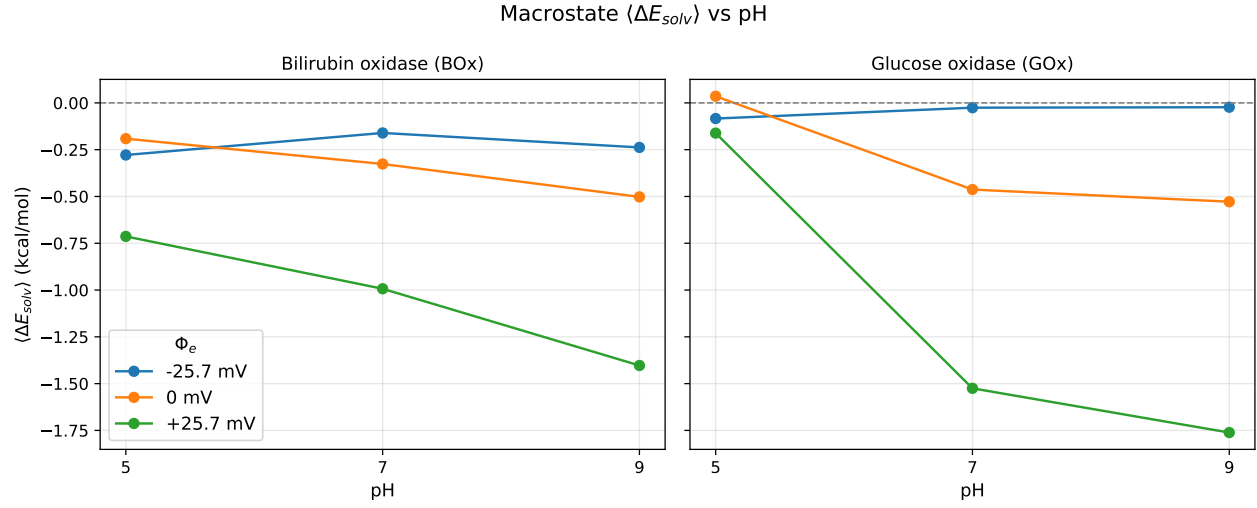


Figure B.3: Solvation contribution to the expected value of the interaction energy, $\langle \Delta E_{solv} \rangle$, as a function of pH, for each applied potential Φ_e and both enzymes.

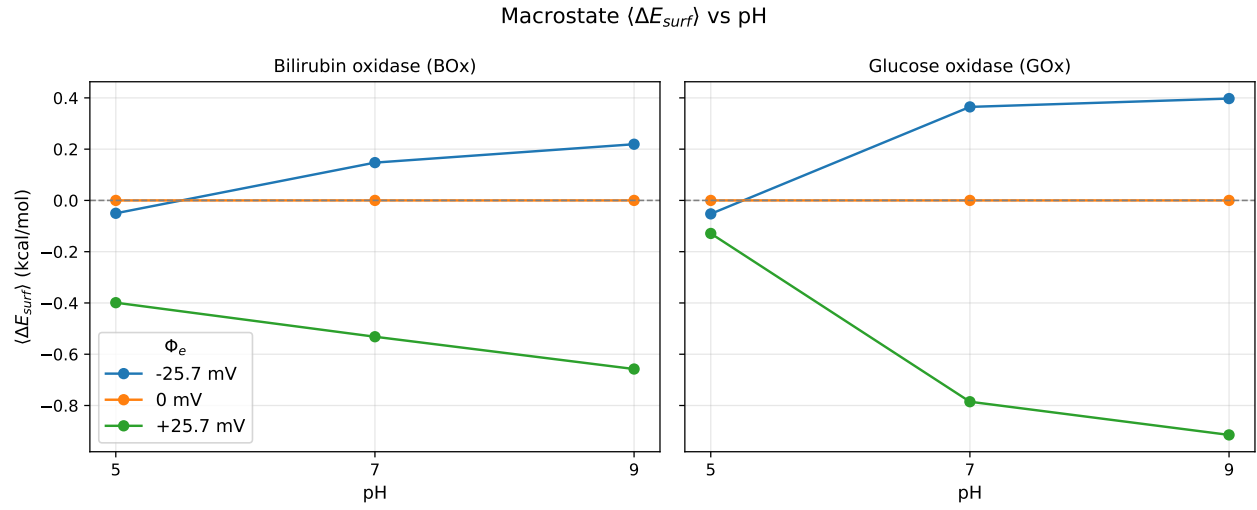


Figure B.4: Surface contribution to the expected value of the interaction energy, $\langle \Delta E_{surf} \rangle$, as a function of pH, for each applied potential Φ_e and both enzymes.

+25.7 mV

0 mV

-25.7 mV

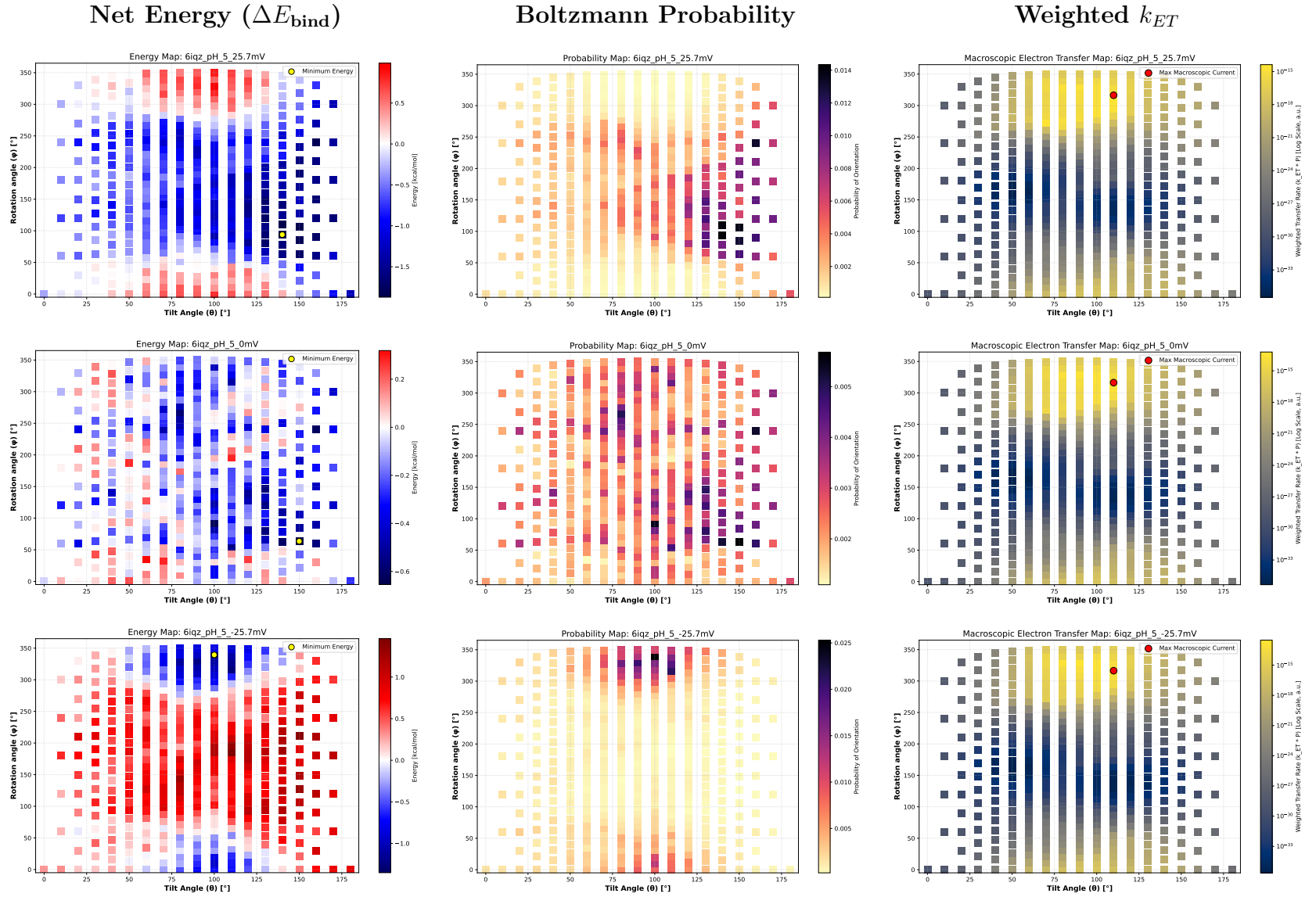


Figure B.5: Thermodynamic and kinetic maps for Bilirubin Oxidase (BOx) at pH 5, for the three electrode potentials studied (+25.7 mV, 0 mV, -25.7 mV). Columns: net interaction energy (ΔE_{bind}), Boltzmann probability, and weighted electron transfer rate. The axes represent the relative orientation of the protein on the electrode: tilt angle (θ) and rotation angle (ϕ), with the protein electrode distance held fixed.

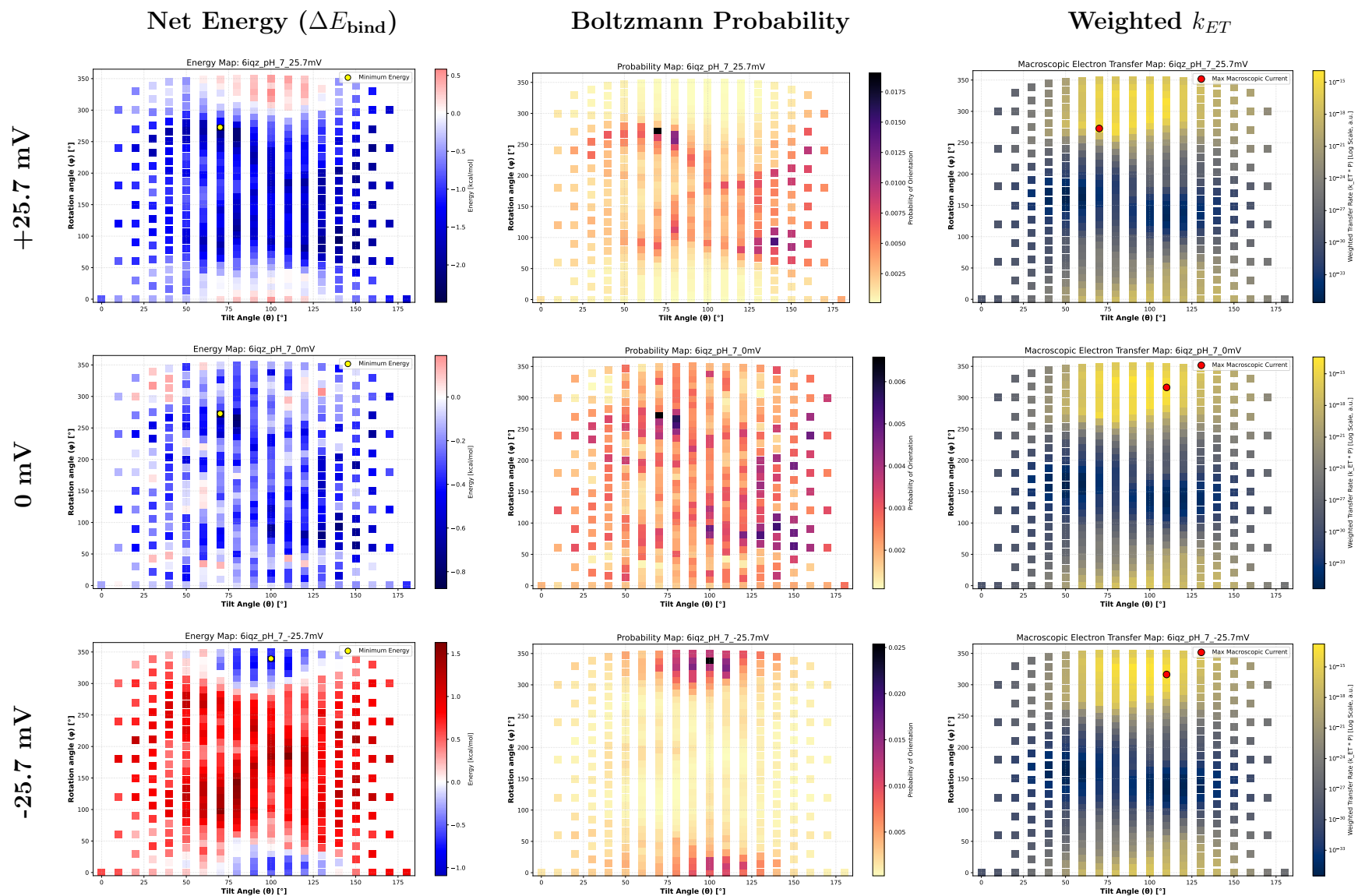


Figure B.6: Thermodynamic and kinetic maps for Bilirubin Oxidase (BOx) at pH 7, for the three electrode potentials studied (+25.7 mV, 0 mV, -25.7 mV). Columns: net interaction energy (ΔE_{bind}), Boltzmann probability, and weighted electron transfer rate. The axes represent the relative orientation of the protein on the electrode: tilt angle (θ) and rotation angle (ϕ), with the protein electrode distance held fixed.

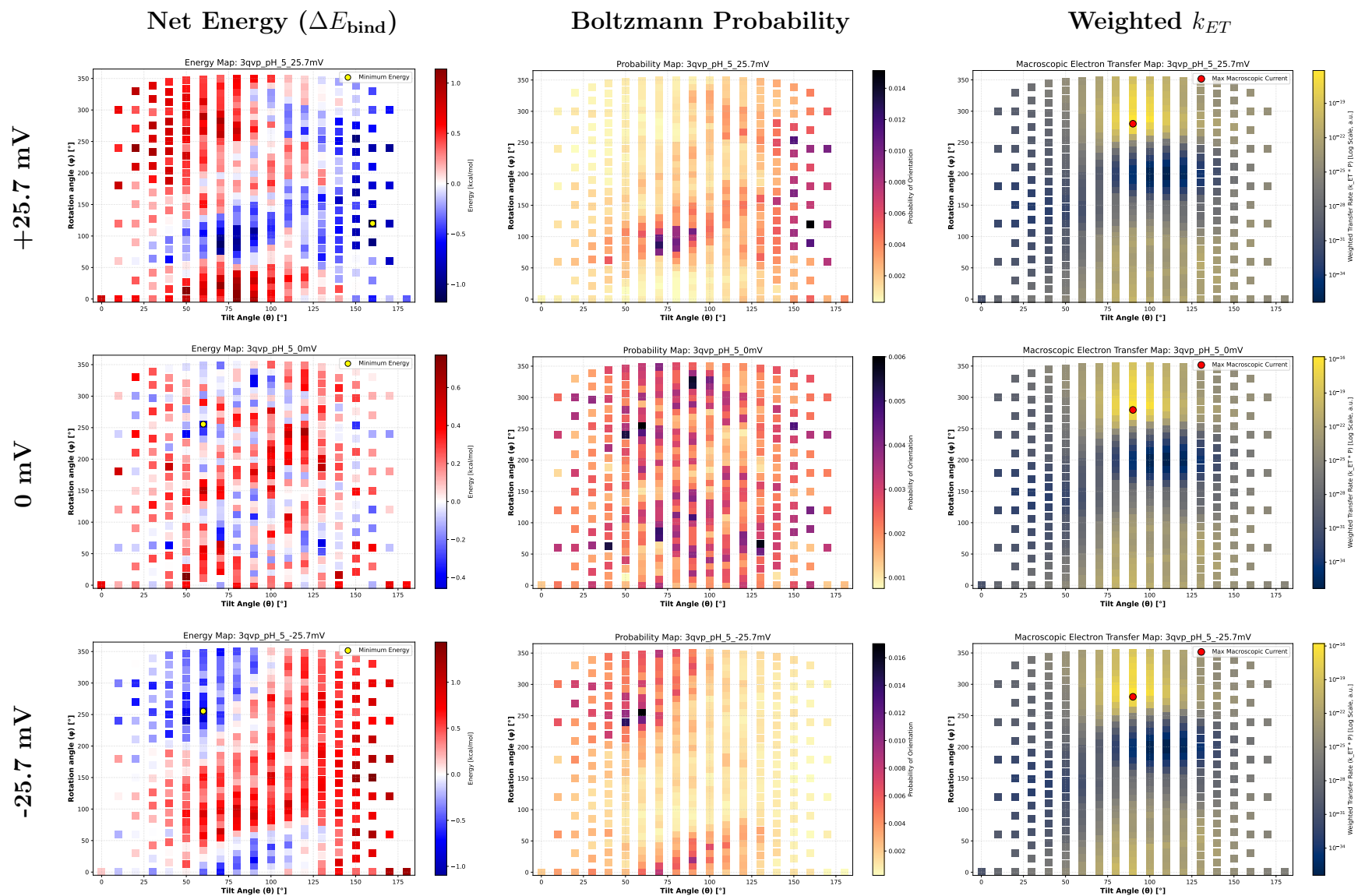


Figure B.7: Thermodynamic and kinetic maps for Glucose Oxidase (GOx) at pH 5, for the three electrode potentials studied (+25.7 mV, 0 mV, -25.7 mV). Columns: net interaction energy (ΔE_{bind}), Boltzmann probability, and weighted electron transfer rate. The axes represent the relative orientation of the protein on the electrode: tilt angle (θ) and rotation angle (ϕ), with the protein electrode distance held fixed.

X

+25.7 mV

0 mV

-25.7 mV

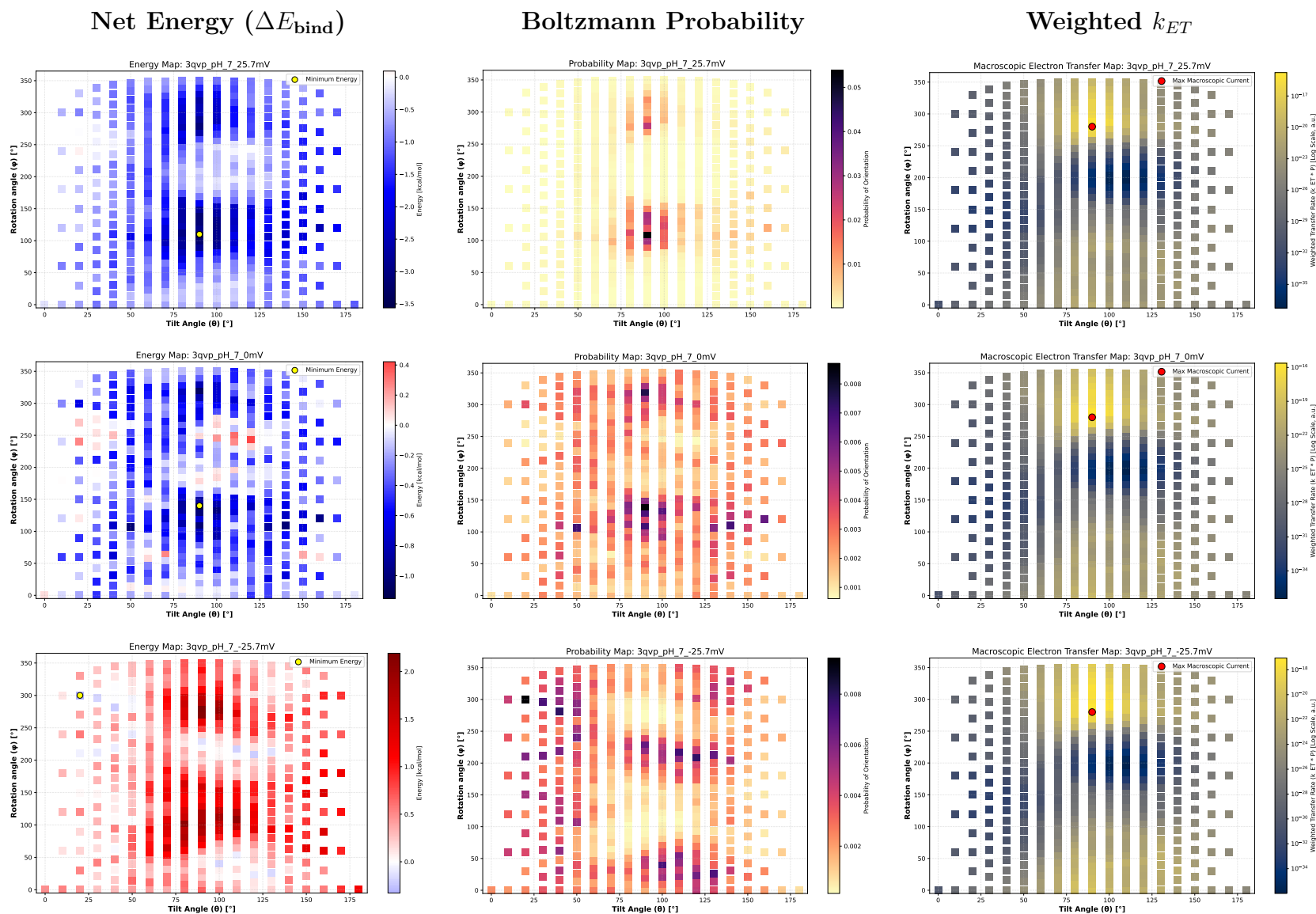


Figure B.8: Thermodynamic and kinetic maps for Glucose Oxidase (GOx) at pH 7, for the three electrode potentials studied (+25.7 mV, 0 mV, -25.7 mV). Columns: net interaction energy (ΔE_{bind}), Boltzmann probability, and weighted electron transfer rate. The axes represent the relative orientation of the protein on the electrode: tilt angle (θ) and rotation angle (ϕ), with the protein electrode distance held fixed.

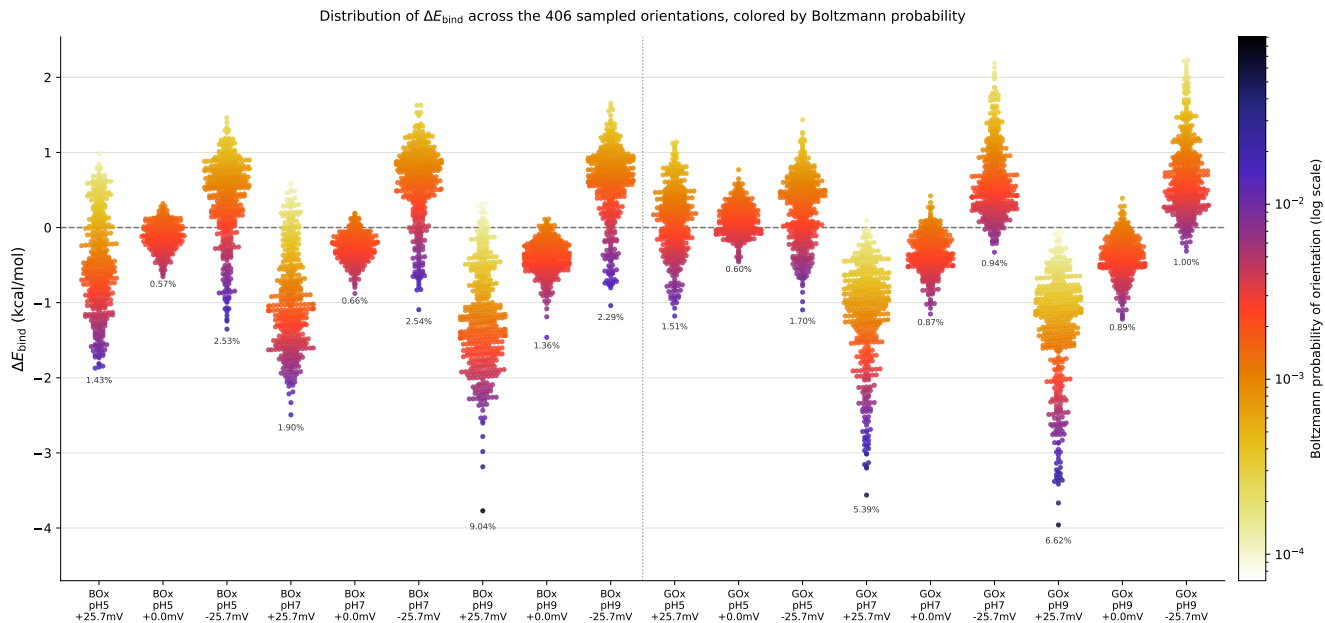


Figure B.9: Distribution of ΔE_{bind} for the 406 sampled orientations across each of the eighteen studied conditions, colored according to the Boltzmann probability of each orientation (logarithmic scale). The label below each distribution indicates the maximum probability achieved under that condition, corresponding to the Boltzmann-optimal orientation reported in Table 4.2. The horizontal dashed line marks $\Delta E_{\text{bind}} = 0$; the vertical line separates BOx (left) from GOx (right).

