Rigorous Approach to Simulate Electromagnetic Interactions in Biological Systems

Kenneth W. Allen*, William D. Hunt†, Jonathan D. Andreasen, John D. Farnum, Alex Saad-Falcon, Ryan S. Westafer, and Douglas R. Denison

***Abstract*—In this work, we explore a novel approach to quantify and predict events related to biomolecular recognition. The theoretical framework leverages mathematics from radar signal processing and communication networks in the form of a matched filter to describe the recognition of a ligand by a receptor. Based on first principles this cross correlation between the electrostatic fields of the ligand and receptor is a new approach to predict the details of interaction between the two biomolecules. The output of this matched filter is analogous to Gibbs free energy of classical thermodynamics with the added advantage that it has the potential to identify the path the ligand takes as it moves towards the receptor. The biomolecules used for demonstration purposes in this paper are C12 (ligand) and QscR (receptor). They are important to the understanding of an opportunistic pathogen, *Pseudomonas aeruginosa*, which strikes cystic fibrosis patients and burn victims, among others. This work could provide an alternative approach to more rigorous common methods, e.g. molecular dynamics simulations, for analyzing and understanding molecular recognition events.**

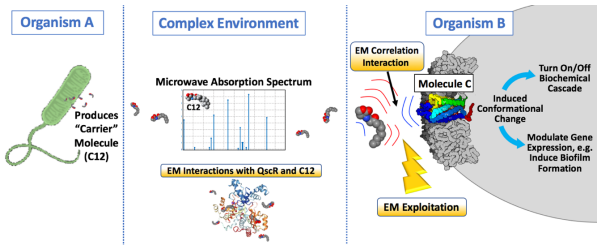

Figure 1. Illustration of potential roles of biological antenna structures in a well-known molecular signaling pathway. Organism A serves as a transmitter and the C12 molecule as a chemical carrier. The associated molecular electrostatic field is absorbed and its microwave spectrum broadened by interactions with the medium. The receiver Organism B presents a receptor molecule that reacts to the time-modulated and spatially-dependent field from its ligand C12, resulting in a purposeful response to the EM field from the C12 molecule.

I. Introduction

BIOPHYSICAL chemistry tends to utilize concepts based on 19th century thermodynamics to describe molecular interactions. These techniques have their basis in the machines of the industrial age such as the steam engine in which "billions and billions" of molecules are involved. in terms of one-on-one interactions. For instance, ligand-receptor binding is a fundamental process in a host of cellular interactions, where one molecule (a ligand) binds to another molecule (a receptor) with a potential third molecule binding elsewhere on the receptor and modifying binding by virtue of allosteric inhibition. Once the ligand-receptor binding event takes place it can trigger a cascade of events within a cell. When we focus on the ligand-receptor biomolecular recognition event more closely, we see that it has a great deal in common with matched filters in radar signal processing [1]. The receptor may be considered "matched" to a specific ligand in some way (in our case, we will consider the energy contained in the electrostatic fields as a measure of "match"). This helps the receptor discriminate between the ligand against a myriad of other molecules which are in the cell at the same time.

Currently, the simulation of biomolecular interactions is done using computational tools known generally as Molecular Dynamics (MD) [2–4]. Although MD simulations accurately model binding events between a ligand and receptor pair [5,6], there are limitations inherent in the approach. Primary among these limitations is the general complexity of MD models of the molecular interaction. In this work, we will describe an electromagnetic approach to biomolecule recognition events. The concept of the matched filter – termed here the "Molecular Ambiguity Function" – will be described, followed by an example case of the interaction between QscR and C12, as shown in Fig. 1, as a critical component in the *Pseudomonas aeruginosa* pathogen [7], with a discussion of results and conclusions.

K.W. Allen et al. are with the Georgia Institute of Technology, Georgia Tech Research Institute, Advanced Concepts Laboratory, Atlanta, GA, 30318 USA. *Corresponding author email: Kenneth.Allen@gtri.gatech.edu, †W.D. Hunt is with the Georgia Institute of Technology, School of Electrical and Computer Engineering.

K. W. Allen et al., "Rigorous Approach to Simulate Electromagnetic Interactions in Biological Systems," NAECON 2018 - IEEE National Aerospace and Electronics Conference, Dayton, OH, 2018, pp. 491-495, doi: 10.1109/NAECON.2018.8556724.

II. Molecular Ambiguity Function

A. Concept

Ligand-receptor interactions are traditionally described, in physical chemistry, from the framework of thermodynamics. In that framework, the concentrations of the bound ligand-receptor pairs ($[RL]$), and unbound ligands ($[L]$) and receptor ($[R]$) molecules are expressed in relation to binding affinity (K_a) and dissociation constant(K_d), explicitly shown in (1).

$$K_a = \frac{[RL]}{[R][L]} = \frac{1}{K_d} \tag{1}$$

The dissociation constant can then be related to the gas constant (R) and absolute temperature (T) to describe the minimization of thermodynamic potential of the chemical system, i.e. Gibbs free energy (ΔG) shown in (2).

$$\Delta G = RTln(K_d) \tag{2}$$

In this work, we seek to recast biomolecular interactions in terms of electromagnetic fields to capture the essential dynamics leading to a binding event. The current thermodynamic approach describes the interactions from a chemical system paradigm that averages the responses of billions of interactions. However, these biomolecular interactions happen on an individual basis which is not adequately explained within the existing framework, (2). In order to model biosystems for localized events, concepts from signal processing are leveraged, e.g. cross-correlations, matched filters, etc.

Molecular recognition between the ligand and receptor requires the spatial distribution of the electric fields for both the ligand ($\vec{E}_L(\vec{r}\text{-}\vec{r}_L)$) and receptor ($\vec{E}_R(\vec{r}\text{-}\vec{r}_R)$), where $\vec{r}$, $\vec{r}_L$, $\vec{r}_R$ is the observation point and location of fields for the ligand and receptor, respectively. These electric fields are the result of various sources from within the ligand and receptor such as ions ($\vec{E}_{Li}$, $\vec{E}_{Ri}$), dipoles ($\vec{E}_{Ld}$, $\vec{E}_{Rd}$), quadrupoles ($\vec{E}_{Lq}$, $\vec{E}_{Rq}$), and higher order sources ($h.o.s$), respectively.

$$\vec{E}_L(\vec{r}-\vec{r}_L) = \sum \vec{E}_{Li}(\vec{r}-\vec{r}_L) + \sum \vec{E}_{Ld}(\vec{r}-\vec{r}_L) + \sum \vec{E}_{Lq}(\vec{r}-\vec{r}_L) + h.o.s. \tag{3}$$

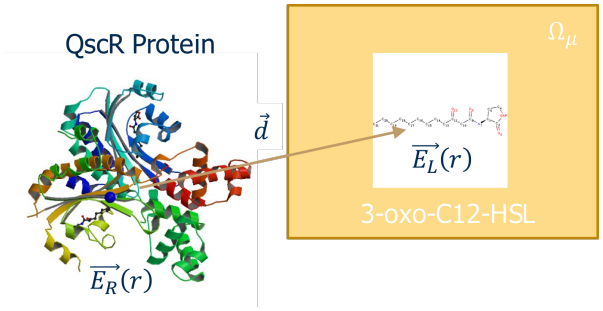

Figure 2. Schematic representation of the ligand and receptor, C12 and QscR respectively. The separation between the receptor and ligand is specified by the vector $\vec{d}$, with field $\vec{E_R}$ and field $\vec{E_L}$ representing the three-dimensional vector field distribution due to the receptor and ligand, respectively.

$$\vec{E}_R(\vec{r}-\vec{r}_R) = \sum \vec{E}_{Ri}(\vec{r}-\vec{r}_R) + \sum \vec{E}_{Rd}(\vec{r}-\vec{r}_R) + \sum \vec{E}_{Rq}(\vec{r}-\vec{r}_R) + h.o.s. \tag{4}$$

The resultant forces due to these electric fields will influence the dynamics of ligand-receptor binding events and may play a central role in molecular binding by *e.g.* spatial state preparation. The total energy associated with the electrostatic interaction between the fields generated by the ligand-receptor pair is given by

$$U_{LR} = \frac{1}{2}\epsilon \int_{\Omega_\mu} \vec{E}_L(\vec{r'}+\vec{r}) \cdot \vec{E}_R(\vec{r'}+\vec{r})\, d^3r', \tag{5}$$

where U_{LR} is the interaction energy, ϵ is the permittivity of the surrounding medium, Ω_μ is the volume surrounding the ligand and receptor, $\vec{r'}$ is the vector to the observation point, and $\vec{r}$ is the path for the ligand molecule. The integral is bound to Ω_μ instead of $\pm\infty$ to compute the correlation over a localized volume in the vicinity of the ligand-receptor to capture that interaction where the primary focus is on the effect of the protein on the ligand.

Equation (5) provides a cross-correlation between the electric fields generated by the ligand and receptor as the ligand molecule traces a spatial path relative to the receptor molecule. This cross-correlation of a single molecular recognition event can be recast in terms of Gibbs free energy, which we refer to as the Molecular Ambiguity Function (MAF).

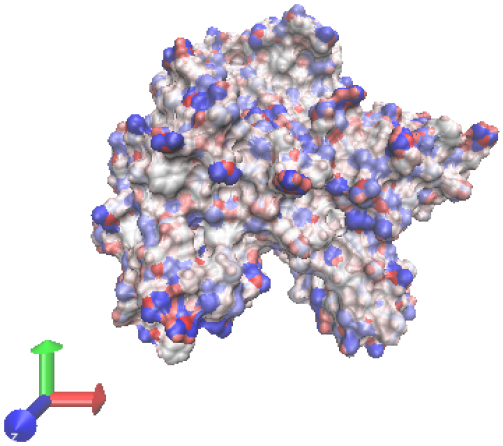

Figure 3. 3D partial charge distribution of QscR with a BWR color scale range of [-0.83, 0.64] elementary charge. The protein is solvated but water molecules are not shown.

$$\Delta G_{LR}(\vec{r}) \propto U_{LR} = \frac{1}{2}\epsilon \int_{\Omega_\mu} \vec{E}_L(\vec{r'}+\vec{r}) \cdot \vec{E}_R(\vec{r'}+\vec{r})\, d^3r' \tag{6}$$

We apply the Molecular Ambiguity Function [(6)] for the C12 ligand and QscR receptor protein, as shown in Figure 2. We calculate the spatial field distributions based on charge distributions from Nanoscale Molecular Dynamics (NAMD) simulations [2, 3], discussed in the next section. NAMD captures the aggregate of electric field components, rather than the discrete moments given in (3) and (4).

B. Spatial Electric Field Distributions

The structural form of biomolecules of interest – including their chemical composition and conformation – can be retrieved from the Protein Data Bank [8], and these parameters are used as the starting point for NAMD simulations that ultimately result in distributions of atoms within the molecule, Fig. 3. This distribution in turn provides the location of charge centers (assumed instantaneously static) due to ions within the molecule. A typical molecule can contain thousands of local charge centers that produce spatial field structure close to the molecule, but that average to a small net dipole moment (typically on the order of tens $e^-\mathring{A}$ or roughly 5 times that number in debye units) within tens of nanometers from the molecule. The local field that reveals spatial structure in the electric fields is within the region of interest when evaluating the integral in (6).

As a demonstration of the MAF approach the charge distributions for the receptor protein (QscR), Figure 3, and ligand (C12) for a single set of condition parameters (temperature, solution, etc.) are considered. The charge distributions, such as the QscR in Figure 3, are impressed onto a grid where the electric fields are calculated. The protein field values are at grid points located around the ligand which is outside but near the surface of the protein. The electric fields are determined by the partial charges calculated for each atom of the biomolecules, shown in Figures 4 and 5 for QscR and C12 respectively. The QscR protein atoms are collectively responsible for the overall

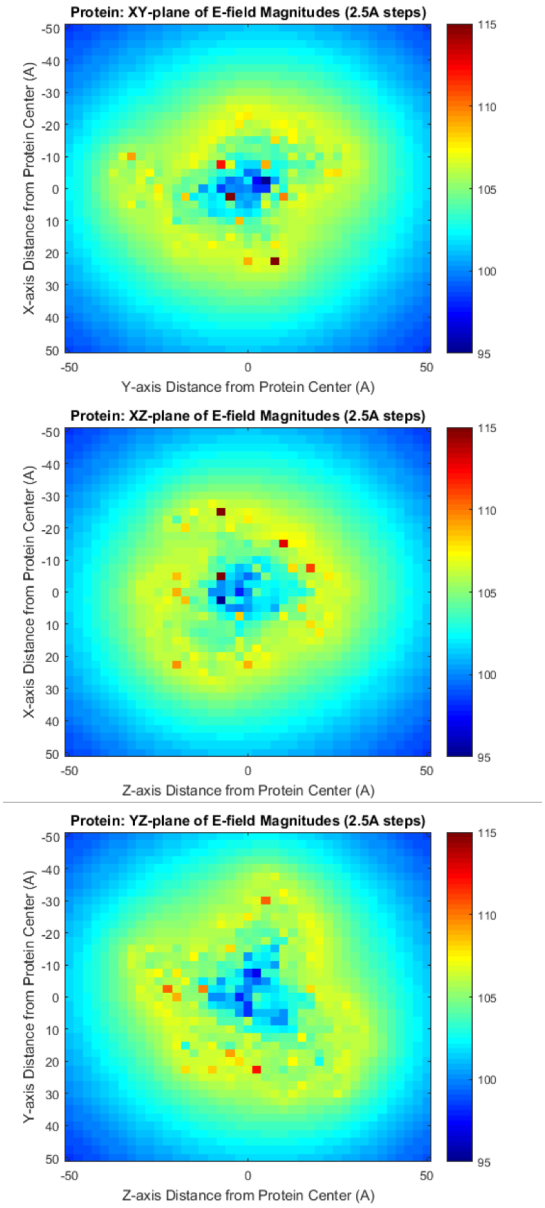

Figure 4. Receptor (QscR protein) electric field magnitude (dB V/$\mathring{A}$) distributions with 2.5 $\mathring{A}$ spatial resolution in the (top) $\hat{x}$-$\hat{y}$, (middle) $\hat{y}$-$\hat{z}$, and (bottom) $\hat{x}$-$\hat{z}$ planes.

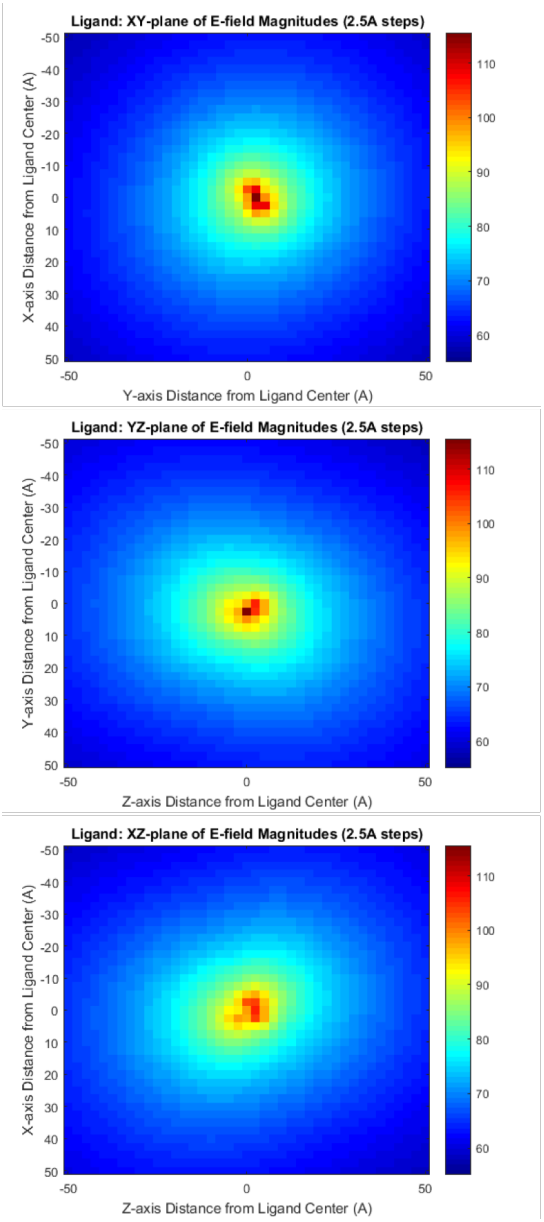

Figure 5. Ligand (C12) electric field magnitude distributions (dB V/$\mathring{A}$) with 2.5 $\mathring{A}$ spatial resolution in the (top) $\hat{x}$-$\hat{y}$, (middle) $\hat{y}$-$\hat{z}$, and (bottom) $\hat{x}$-$\hat{z}$ planes.

spatial distribution of the receptor electric field ($\vec{E_R}$), as shown in Figure 4. Similarly, the atoms comprising the C12 molecule result in the electric field distribution for the ligand ($\vec{E_L}$), as shown in Figure 5. In both cases 2.5 $\mathring{A}$ spatial resolution is used to generate the electric field maps for a 100x100 nm^2 area. To the extent that the partial charges are accurate and the partial charge model produces an accurate picture of the charge distribution on a molecule, the electric fields will be an accurate representation of the true fields produced by the molecules.

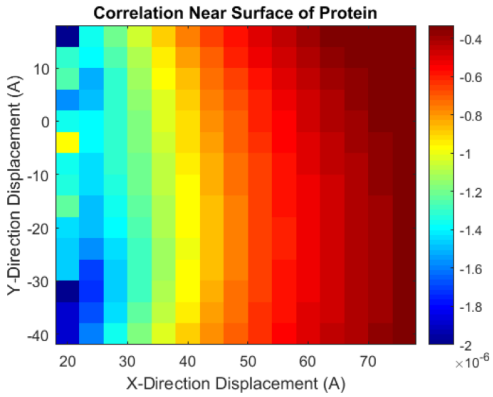

Figure 6. Correlation of the ligand and receptor electric fields for translation of the ligand with respect to the receptor in the $\hat{x}$-$\hat{y}$ plane.

C. Correlation Integral

In order to evaluate the MAF, the ligand coordinates are transformed using a translation vector $\vec{d}$ and rotation vector $\vec{\theta}$ (e.g., roll, pitch, and yaw). Integrating over all space $d\vec{r}$ is not necessary, as discussed in Section *II. A.*, since the primary interaction of interest is the effect of the protein on the ligand; therefore, a volume encompassing the ligand ($\Omega_\mu = 100x100x100$ nm^3) is considered for the correlation integral (7) which is proportional to the Molecular Ambiguity Function (6).

$$\Delta G_{LR}(\vec{r}) \propto C_{LR}(\vec{d},\vec{\theta}) = \int_{\Omega_\mu} \vec{E}_R(\vec{r}) \cdot \vec{E}_L(\vec{r}+\vec{d},\vec{\theta})\, d^3r' \tag{7}$$

The MAF is a 6-dimensional function with three spatial dimensions and three rotational dimensions. Here the MAF is calculated over $\hat{x}$ and $\hat{y}$ spatial displacements while keeping all other degrees of freedom fixed. While the grid used for numerical approximation of the integral has 2.5 Å spacing, the points at which the correlation function are evaluated can have their own spacing (4 Å spacing is used in Fig. 6).

The correlation (C_{LR}) between the electric field produced by the ligand ($\vec{E_L}$) and by the receptor ($\vec{E_R}$) was calculated by (7). The correlation function is usually defined for two scalar functions. In order to calculate the correlation between two vector fields, the dot product between the ligand and receptor electric fields are integrated. The description of how the fields are obtained can be found in Section *II. B.* and the spatial distributions of the electric fields for the QscR and C12 are shown in Figures 4 and 5 respec-

tively.

The dimensionality of the problem was restricted to the translation of the ligand with respect to the protein in the $\hat{x}$-$\hat{y}$ plane. The likely final positions of the ligand are found where $x < 30$ Å in Fig. 6. In this region, the inhomogeneous charge distribution of the protein has more effect, and there are distinct regions along the $\hat{y}$ axis that are more energetically favorable.

III. Discussion

In order to interpret the results, namely Fig. 6, we consider that configurations ($\vec{r}$,θ) having negative Gibbs free energy occur spontaneously; they are energetically favorable. A testable analogy also is made between a ligand binding event and ion trapping. In this work the integration volume surrounds the ligand, we determined through the use of a simplified conceptual model (i.e. a single charge trapped in a 1D potential well) that the correlation function should be negative and at a minimum when trapped. A correlation of the electric fields is performed, assuming no rotation (θ), and translation only in the $\hat{x}$-$\hat{y}$ plane. The result is shown in Fig. 6. All values are negative in this region of space, and the gradient shows that it is energetically favorable for the ligand to approach the protein toward the left edge of Fig. 6. As the ligand encounters the complex electric field structure near the protein, it is observed that some regions are more favorable than others. Further work in molecular dynamics and/or "docking" algorithms may corroborate the likely trajectories apparent in the MAF as shown here. This MAF conjecture can be validated by in-depth MD simulations, which are underway. In future studies, aspects of the complex environment, which could lead to screening or other significant effects, should be considered when calculating the Molecular Ambiguity Function.

IV. Conclusion

In this work, we demonstrated the mathematical and theoretical framework for a novel approach to quantify and predict events of molecular recognition. This approach leveraged concepts from radar signal processing and communication networks. Specifically, a spatial electrostatic cross-correlation between a ligand and receptor is performed to represent a spatial Gibbs free energy function, i.e. the Molecular Ambiguity Function, which describes the interaction between the biomolecules. This Molecular Ambiguity Function was evaluated for biomolecules showing specific relative orientations with minimum magnitude correlation that may indicate spatial state preparation of the ligand prior to binding with the receptor. The biomolecules C12 and QscR were studied to represent a critical link in the chain of the quorum sensing process from the *Pseudomonas aeruginosa* pathogen that gives rise to biofilm formation which impacts cystic fibrosis patients and burn victims. This work could augment computationally expensive calculations for analyzing molecular recognition events, particularly at distances beyond mechanical interaction.

Acknowledgement

This work was partially funded by DARPA/DSO Contract HR001117C0124.

References

[1] G. Turin, "An introduction to matched filters," *IRE transactions on Information theory*, vol. 6, no. 3, pp. 311–329, 1960.

[2] W. Humphrey, A. Dalke, and K. Schulten, "Vmd: visual molecular dynamics," *Journal of molecular graphics*, vol. 14, no. 1, pp. 33–38, 1996.

[3] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten, "Scalable molecular dynamics with namd," *Journal of computational chemistry*, vol. 26, no. 16, pp. 1781–1802, 2005.

[4] L. Kalé, R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan, and K. Schulten, "Namd2: greater scalability for parallel molecular dynamics," *Journal of Computational Physics*, vol. 151, no. 1, pp. 283–312, 1999.

[5] M. Balsera, S. Stepaniants, S. Izrailev, Y. Oono, and K. Schulten, "Reconstructing potential energy functions from simulated force-induced unbinding processes." *Biophysical journal*, vol. 73, no. 3, p. 1281, 1997.

[6] B. Isralewitz, M. Gao, and K. Schulten, "Steered molecular dynamics and mechanical functions of proteins," *Current opinion in structural biology*, vol. 11, no. 2, pp. 224–230, 2001.

[7] M. Schuster and E. P. Greenberg, "A network of networks: quorum-sensing gene regulation in pseudomonas aeruginosa," *International Journal of Medical Microbiology*, vol. 296, no. 2-3, pp. 73–81, 2006.

[8] F. C. Bernstein, T. F. Koetzle, G. J. Williams, E. F. Meyer Jr, M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi, "The protein data bank: A computer-based archival file for macromolecular structures," *European Journal of Biochemistry*, vol. 80, no. 2, pp. 319–324, 1977.