BIOTECHNO 2026

The Eighteenth International Conference on Bioinformatics, Biocomputational Systems and Biotechnologies

ISBN: 978-1-68558-361-3

March 8th –12th, 2026

Valencia, Spain

BIOTECHNO 2026 Editors

Constantine Kotropoulos, Aristotle University of Thessaloniki, Greece

BIOTECHNO 2026

Foreword

The Eighteenth International Conference on Bioinformatics, Biocomputational Systems and Biotechnologies (BIOTECHNO 2026), held between March 8 - 12, 2026, covered these three main areas: bioinformatics, biomedical technologies, and biocomputing.

Bioinformatics deals with the system-level study of complex interactions in biosystems providing a quantitative systemic approach to understand them and appropriate tool support and concepts to model them. Understanding and modeling biosystems requires simulation of biological behaviors and functions. Bioinformatics itself constitutes a vast area of research and specialization, as many classical domains such as databases, modeling, and regular expressions are used to represent, store, retrieve and process a huge volume of knowledge. There are challenging aspects concerning biocomputation technologies, bioinformatics mechanisms dealing with chemoinformatics, bioimaging, and neuroinformatics.

Biotechnology is defined as the industrial use of living organisms or biological techniques developed through basic research. Bio-oriented technologies became very popular in various research topics and industrial market segments. Current human mechanisms seem to offer significant ways for improving theories, algorithms, technologies, products and systems. The focus is driven by fundamentals in approaching and applying biotechnologies in terms of engineering methods, special electronics, and special materials and systems. Borrowing simplicity and performance from the real life, biodevices cover a large spectrum of areas, from sensors, chips, and biometry to computing. One of the chief domains is represented by the biomedical biotechnologies, from instrumentation to monitoring, from simple sensors to integrated systems, including image processing and visualization systems. As the state-of-the-art in all the domains enumerated in the conference topics evolve with high velocity, new biotechnologes and biosystems become available. Their rapid integration in the real life becomes a challenge.

Brain-computing, biocomputing, and computation biology and microbiology represent advanced methodologies and mechanisms in approaching and understanding the challenging behavior of life mechanisms. Using bio-ontologies, biosemantics and special processing concepts, progress was achieved in dealing with genomics, biopharmaceutical and molecular intelligence, in the biology and microbiology domains. The area brings a rich spectrum of informatics paradigms, such as epidemic models, pattern classification, graph theory, or stochastic models, to support special biocomputing applications in biomedical, genetics, molecular and cellular biology and microbiology. While progress is achieved with a high speed, challenges must be overcome for large-scale bio-subsystems, special genomics cases, bio-nanotechnologies, drugs, or microbial propagation and immunity.

We take here the opportunity to warmly thank all the members of the BIOTECHNO 2026 Technical Program Committee, as well as the numerous reviewers. The creation of such a high quality conference program would not have been possible without their involvement. We also kindly thank all the authors who dedicated much of their time and efforts to contribute to BIOTECHNO 2026.

Also, this event could not have been a reality without the support of many individuals, organizations, and sponsors. We are grateful to the members of the BIOTECHNO 2026 organizing committee for their help in handling the logistics and for their work to make this professional meeting a success.

We hope that BIOTECHNO 2026 was a successful international forum for the exchange of ideas and results between academia and industry and for the promotion of progress in the fields of bioinformatics, biocomputational systems and biotechnologies.

We are convinced that the participants found the event useful and communications very open. We also hope that Valencia provided a pleasant environment during the conference and everyone saved some time for exploring this beautiful city

BIOTECHNO 2026 Chairs:

BIOTECHNO 2026 Steering Committee
Petre Dini, IARIA, USA/EU
Hesham H. Ali, University of Nebraska at Omaha, USA
Xiaobo Zhou, University of Texas Health Science Center – Houston, USA

BIOTECHNO 2026 Publicity Chairs
Francisco Javier Díaz Blasco, Universitat Politècnica de València, Spain
Ali Ahmad, Universitat Politècnica de València, Spain
Sandra Viciano Tudela, Universitat Politecnica de Valencia, Spain
José Miguel Jiménez, Universitat Politecnica de Valencia, Spain

BIOTECHNO 2026

Committee

BIOTECHNO 2025 Steering Committee
Petre Dini, IARIA, USA/EU
Hesham H. Ali, University of Nebraska at Omaha, USA
Xiaobo Zhou, University of Texas Health Science Center – Houston, USA

BIOTECHNO 2026 Publicity Chairs
Francisco Javier Díaz Blasco, Universitat Politècnica de València, Spain
Ali Ahmad, Universitat Politècnica de València, Spain
Sandra Viciano Tudela, Universitat Politecnica de Valencia, Spain
José Miguel Jiménez, Universitat Politecnica de Valencia, Spain

BIOTECHNO 2026 Technical Program Committee

Behrooz Abbaszadeh, University of Ottawa, Canada
Antonino Abbruzzo, University of Palermo, Italy
A M Abirami, Thiagarajar College of Engineering, Madurai, India
Don Adjeroh, West Virginia University, USA
Aftab Ahmad, City University of New York, USA
Jens Allmer, Hochschule Ruhr West - University of Applied Sciences, Germany
Joel P. Arrais, University of Coimbra, Portugal
Simone Avesani, Università di Verona, Italy
Yoseph Bar-Cohen, Electroactive Technologies / NDEAA Lab - Jet Propulsion Laboratory (JPL), USA
Kais Belwafi, King Saud University, Saudi Arabia
Boubaker Ben Ali, University of Bordeaux, France / University of Manouba, Tunisia
Razvan Bocu, Transilvania University of Brasov, Romania
Vincenzo Bonnici, Università di Parma, Italy
Paolo Cazzaniga, University of Bergamo, Italy
Peter Clote, Boston College, USA
Matteo Comin, University of Padova, Italy
Giovanni Cugliari, University of Turin, Italy
Adithi Deborah Chakravarthy, University of Nebraska at Omaha, USA
Maria Evelina Fantacci, University of Pisa, Italy
Abdolhossein Fathi, Razi University, Iran
Silvana Giuliatti, University of São Paulo, Brazil
Henry Griffith, University of Texas at San Antonio, USA
Md Shafaeat Hossain, Southern Connecticut State University, USA
Chih-Cheng Hung, Kennesaw State University, USA
Dimitrios G. Katehakis, Center for eHealth Applications and Services FORTH | Institute of Computer Science, Greece
Jan Kubicek, VSB - Technical University of Ostrava, Czech Republic
Antonio LaTorre, Universidad Politécnica de Madrid, Spain
Chen Li, Monash University, Australia

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission or reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article is does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

Evaluation of a Galvanic Vestibular Stimulation System to Reduce Cyberscikness in Virtual Reality 1
Clarice da Costa Correa, Emmanuelle Chouin, Matthieu Denoual, Romulo Gabriel Pinheiro Moura, Ruth Pastora Saraiva Leao, Raimundo Furtado Sampaio, Giovanni Cordeiro Barroso, Daniel Reboucas Jaguaribe, Yann Jezequel, and Herve Pochat

Attractive Casimir–Lifshitz Forces as a Universal Driver of Prebiotic Protocell Aggregation and Cluster Formation 8
Michael Massoth

Emergent Information Formation in Prebiotic Protocell Clusters: A Computational Mechanics Framework of ?-Machines and Attractor Memory 17
Michael Massoth

From Physical Difference to Meaning: A Constructor-Theoretic Framework for Prebiotic Information in Casimir–Lifshitz-Coupled Protocell Clusters 24
Michael Massoth

Physical Origin of Proto-Information: Syntax, Semantics and Pragmatic Emergence in Prebiotic Protocell Clusters 32
Michael Massoth

Quantum Field-induced Proton H? Gradients in Prebiotic Protocell Clusters 40
Michael Massoth

Prebiotic Sustainability in the Nanoworld: A Physical Framework for Early Protocell Cluster Stability 48
Michael Massoth

Network-Based Analysis of Gut Microbiome Profiles in Autism Spectrum Disorder 57
Shado Badr Basem, Hesham Ali, and Andreas Pester

From Prebiotic Sustainability to Proto-Liveliness: A Biocomputational Framework with Structural, Energetic and Informational Metrics for Protocell Cluster Evolution 64
Michael Massoth

Computational Discovery of Inhibitors Targeting Alphavirus nsP2 Proteases 74
Jose Augusto Leoncio Gomide, Gabriel de Souza Segadaes, Victor Rodrigues Terra, Yasmim Nunes Mesquita, and Nilson Nicolau Junior

Evaluation of a Galvanic Vestibular Stimulation System to Reduce Cyberscikness in Virtual Reality

Clarice da Costa C.
GREYC UMR 6072 and Neural Balance Innovation
University of Caen Normandie, ENSICAEN, CNRS
14000 Caen, France
E-mail: clarice.dacosta-correa@unicaen.fr

Matthieu Denoual
GREYC UMR 6072
University of Caen Normandie, ENSICAEN, CNRS
14000 Caen, France
E-mail: matthieu.denoual@ensicaen.fr

Emmanuelle Chouin, Yann Jezequel, Hervé Pochat
Neural Balance Innovation
14000 Caen, France
E-mail: ec@nbi.fr, it@nbi.fr, hp@nbi.fr

Rômulo Gabriel P. Moura, Ruth P. S. Leão, Raimundo F. Sampaio, Giovanni C. Barroso, Daniel R. Jaguaribe
Department of Electrical Engineering
Federal University of Ceará
60455-760 Fortaleza, Brazil
E-mail : romulogabriel@alu.ufc.br, rleao@dee.ufc.br, rfurtado@dee.ufc.br, gcb@fisica.ufc.br, djaguaribe@alu.ufc.br

Abstract — **This work presents the evaluation of a Galvanic Vestibular Stimulation (GVS) system designed to reduce cybersickness and enhance the sense of immersion during Virtual Reality (VR) exposure. The system integrates a dual-channel Howland current source, an ESP32-based control platform, and embedded safety mechanisms to ensure compliance with established safety guidelines. A complete experimental protocol was developed and applied within the Virtual Reality Substation Laboratory (VRSL), combining technical validation and a user-based evaluation with 20 participants. The results showed that prior exposure, without an adequate washout period, can affect the effectiveness of the stimulation and, consequently, the severity of cybersickness symptoms. Overall, the findings support the preventive potential of controlled vestibular stimulation for mitigating cybersickness and improving user comfort during VR immersion.**

Keywords- Cybersickness; Galvanic Vestibular Stimulation; Virtual Reality; Motion sickness; Washout;

I. INTRODUCTION

This section presents the background and motivation for investigating Galvanic Vestibular Stimulation (GVS) as a potential approach to reducing cybersickness in Virtual Reality (VR).

A. State of the Art in Galvanic Vestibular Stimulation and Virtual Reality Immersion

In recent years, the use of VR has expanded rapidly, beyond video games, VR is now used for professional training, simulation, education [1], rehabilitation and remote collaboration. In the entertainment field, VR has transformed user interaction by providing highly immersive and realistic experiences. However, despite its potential benefits the exposure to VR environments has been associated with certain side effects, supporting the hypothesis that VR may induce a conflict between sensory and spatial integration [2]. Simón-Vicente *et al.* [2], through a systematic review, reported that the most frequent symptoms following VR immersion are disorientation, nausea, and oculomotor disturbances.

The most widely accepted hypothesis regarding the origin of these symptoms is that they result from a mismatch between sensory and spatial integration, in which the user perceives motion visually while the vestibular system signals no corresponding movement [3]. The term commonly associated with these symptoms in VR is cybersickness, traditionally considered a form of visually induced motion sickness that occurs when sensory information and internal predictions about orientation and self-movement are inconsistent [4].

To explore whether it is possible to physically mitigate this issue, it is necessary to consider the vestibular system, the sensory organ responsible for spatial orientation and balance. The vestibular nerve transmits information about self-motion and orientation relative to Earth's gravity from the vestibular organs (the semicircular canals, saccule, and utricle) to the brain [5].

As proposed by Weech *et al.* [6], the addition of noisy Galvanic Vestibular Stimulation (nGVS) to the vestibular system can reduce vestibular reliability, thereby decreasing the conflict between sensory and spatial cues. When processing sensory information, the brain re-weights these inputs, assigning greater trust to visual self-motion cues rather than to the vestibular information [7]. This mechanism suggests that controlled vestibular stimulation could enhance the sense of immersion by reducing cybersickness during VR exposure.

B. Objectives and Contributions of the Work

The main contribution of this work is to validate a GVS system. This system is designed to enhance the sense of immersion and to avoid cybersickness in a VR environment, specifically in a medium-voltage substation simulation, which will be further explained in the next section. The proposed system aims to stimulate the vestibular system in a controlled

and safe manner, with the goal of reducing the symptoms of cybersickness.

To provide a clear overview of this work, the remainder of the paper is organized as follows. Section II describes the development of the GVS system, while Section III presents the VR environment used in this study. Section IV details the experimental methodology and evaluation procedures, and Section V presents the results and discussion. Finally, Section VI provides the conclusions and future perspectives.

II. Design and Development os the Galvanic Vestibular Stimulation System

This section details the technical development of the proposed GVS system, including its hardware architecture, safety mechanisms, and software implementation.

A. System Specifications and Hardware Architecture

The general concept of the circuit is a voltage-to-current converter, commonly known as a *Howland Current Source* [9], and seen in Figure 1. In this work, the converter is integrated with an ESP32 microcontroller and a broader electronic circuit to ensure precise signal control and amplification.

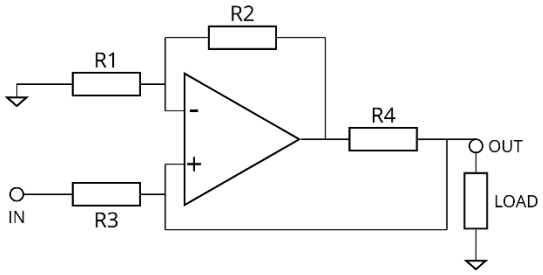


Figure 1. Howland Current Source

The system responsible for generating GVS stimulation is composed of two primary components:

- an ESP32 microcontroller responsible for digital control and generating voltage signals.
- two independent Howland current sources capable of generating distinct signal types across the electrodes.

The setup allows for the connection of the electrodes to the participant. Different GPIO (General-Purpose Input/Output) ports of the ESP32 are used to independently control each Howland source and, consequently, each pair of stimulation electrodes (active electrode – reference electrode).

As highlighted by Liu *et al* [10], previous work often provides insufficient technical information for reproducibility and rarely reports detailed rigorous performance assessments. Although a known circuit topology was employed, our implementation differs from others, such as the one proposed by Liu *et al.* [10], by including two protection mechanisms: one electrical and one software based. These safeguards ensure that the user does not experience any leakage current either during power-up or operation.

Regarding user safety, circuit design follows the guidelines proposed by Antal *et al.* [11], which recommend limiting the current output to ±4 mA. In this work, the system restricts the output to ±2.6 mA. Several studies have reported variations in human impedance between the mastoids [10], which depend on factors, such as electrode placement, physiology, and tissue moisture. To ensure the feasibility and reliability of the system under this variability, the electronics were designed to adapt to impedances of up to 9 kΩ.

B. Software Implementation and Code Development

The software architecture for GVS stimulation was developed to control the key features of the signals, including amplitude, frequency, and waveform type. The ESP32 microcontroller was programmed using the PlatformIO extension in VS Code, which offers greater flexibility for integrating both analog and digital control functions. The firmware manages the active electrodes independently, enabling the generation of simultaneous stimulation patterns across the user's vestibular system.

To make the system more practical and functional, both during testing and in future industrial applications, such as the commercialization phase, a dedicated application was developed to provide a visual and user-friendly interface for programming the stimulation parameters.

As part of the data collection process, an Inertial Measurement Unit (IMU) was incorporated to capture movement data. Specifically, the MPU6886 sensor was used, which integrates a 3-axis gyroscope and a 3-axis accelerometer, both equipped with a 16-bit Analog-to-Digital (ADC). The IMU was programmed using the Arduino IDE to control data acquisition and analyze how body sways varied across the axes. The sensitivities were configured as:

- Gyroscope Sensitivity: 32.8 LSB per °/s (with a ±1000 °/s range)
- Accelerometer Sensitivity: 4096 LSB per g (with a ±8 g range).
- The sampling frequency was set at 200 Hz, following the approach described by Goel *et al*. [8].

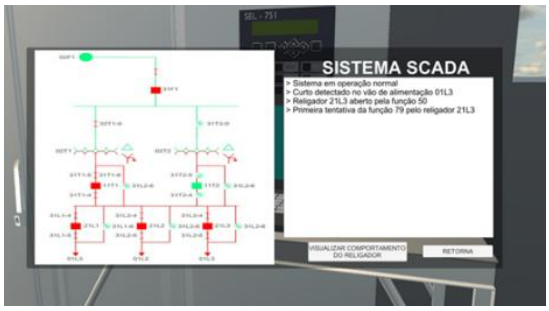
III. Description of the Virtual Reality Application

For these tests, the software "Virtual Reality Substation Laboratory (VRSL)" [1] was used. It was developed by the Smart Electrical Grids Laboratory (GREI) at the Federal University of Ceará (UFC). The software was designed to accurately reproduce the power distribution substation located on the UFC campus. The virtual environment is composed of two main areas:

- Outdoor Yard: Users can visualize and interact with various components, such as current transformers, potential transformers, and other equipment, as shown in Figure 2c.
- Control House: Users can access the protection and automation panels, which include interactive components and graphical interfaces displaying real data acquired from the physical substation.

All interactions are performed in first-person mode using a VR headset. Users can explore the environment through three different modes of navigation: Walking using console

controls; Flying through the area using an integrated drone; Watching simulations of real operational events and accidents within the substation (as shown in Figure 2a).



(a)


(b)


(c)

Figure 2. Virtual Reality Substation Laboratory (VRSL) interface. (a) View of simulation, (b) Done-mode navigation, (c) First-person view

The VRSL software was selected because its three distinct modes of navigation (walking, video playback, and drone flight) offer a wide range of motion experiences for the user. To ensure proper operation with the system, the software requires a computer and a VR headset meeting the recommended specifications in Table I.

TABLE I. RECOMMENDED HARDWARE SPECIFICATIONS

Component	Specification
RAM Memory	16 GB DDR4
Graphics Card	4 GB
Processor	Intel i5-13450HX
Storage	256 GB SSD
VR Headset	Meta Quest 3

IV. EXPERIMENTAL METHODOLGY AND EVALUATION PROTOCOL

This section presents the experimental results and discusses their implications for immersion tolerance and cybersickness reduction.

A. Creation of the Experimental Protocol

During the development of the experimental protocol, two main research questions were defined:

- Does the use of GVS stimulation improve the sense of immersion?
- Can the combination of GVS stimulation with a VR headset reduce symptoms of cybersickness?

To address these questions, five questionnaires were selected to ensure comprehensive data collection. First, a health questionnaire was administered to identify any pre-existing medical conditions or regular medication use. Participants reporting relevant health issues were excluded from the experiment to ensure that only healthy individuals were tested.

The System Usability Scale (SUS), developed by Brooke in 1986 [12], consisting of ten items, was used to evaluate participants' perceptions of the overall system, in this case, the VRSL and the Stimbox. Additionally, the Simulator Sickness Questionnaire (SSQ) [13] was applied to measure the level of sickness symptoms and compute individual subscale scores.

To finalize the experimental protocol, participants were given 15 minutes of free exploration within the VR environment. They received minimal instructions, being allowed to move freely and interact naturally. The only requirement was that, during each immersion, participants should test the three available modes of interaction, video simulation, first-person navigation, and drone control.

B. Validation and Testing Procedures

The validation of the proposed experimental protocol was conducted in two complementary phases: a technical validation to assess the performance of the stimulation, and a user-based evaluation to examine the perceptual and usability aspects of the system.

During the technical validation, the reliability of the electrical signals generated by the device was verified prior to the experiments. Both sinusoidal and white noise waveforms were tested to confirm amplitude stability and to ensure that signal values did not change abruptly, thereby guaranteeing participant safety.

Before the test each participant completed a set of pre-screening questionnaires to confirm eligibility for the experiment and signed a consent form, accordingly to the Declaration of Helsinki, and then proceeded through two stages: the initial assessment phase and the testing phase, both carried out using the GVS system.

For each participant, the application synchronized the stimulation amplitude data with the inertial measurements, generating normalized graphs for each amplitude indicator.

Similarly to Goel *et al.* [14], the electrodes were positioned symmetrically over the mastoids. The skin surface where the electrodes were placed was cleaned and dried, followed by the application of a thin layer of conductive gel to ensure uniform current density, improve electrode fixation, and minimize skin irritation.

Each participant underwent two VR immersion sessions (using the VRSL + GVS system), each lasting 15 minutes: one without stimulation (sham) and one with white noise

stimulation (nGVS), with the order of exposure being randomized (A: sham → nGVS; B: nGVS → sham), using the system presented at Figure 3.

Figure 3. Complete system for the test; VRSL + GVS system

To analyze the efficacy of the nGVS stimulation in reducing cybersickness, two complementary indicators were evaluated: the symptoms measured by the SSQ and its sub-scores, as well as the tolerance to immersion, measured by the duration of time each participant was able to remain in the VR environment (maximum 15 minutes). The total SSQ score was computed using (1).

$$SSQ_{total} = \frac{(9.54*nausea)+(7.58*oculomotor)+(13.92*disorientation)}{3.74} \quad (1)$$

Asymmetric continuous variables were analyzed using the Wilcoxon signed-rank test [15]. Ordinal variables were examined using either the sign test or Bowker's test of symmetry [16], while binary variables were analyzed with McNemar's exact test [17]. Paired risk differences (RD) and their exact 95% confidence intervals (CI) were calculated for clinical interpretation and effect size estimation.

V. Results and Discussion

This section outlines the main limitations of the present study and discusses future perspectives for improving the experimental design and expanding the application of nGVS in VR environments.

A. Quantitative and Qualitative Results

Firstly, the age distribution of the participants was analyzed. As shown in Figure 4, the 20 participants ranged from 19 to 65 years, with an average of 33 years. The distribution reflects the availability of volunteers and the location where the tests were conducted.

Table II presents the first parameter evaluated: whether participants were able to complete the full 15-minute VR immersion. During the sham condition, 15% of the participants did not complete the entire session. In contrast, under nGVS stimulation, all participants reached the 15-minute target. The paired contingency table showed 3 discordant pairs in favor of nGVS and none in the opposite direction, corresponding to a +15% paired improvement (RD = +0.15). Although this difference did not reach statistical significance in the McNemar test ($\chi^2 = 1.33$, $p = 0.25$), the discordance pattern and the absence of any dropouts under nGVS indicate a clear positive trend in immersion tolerance. This preliminary analysis indicates that nGVS enabled all participants to tolerate the full 15-minute immersion, whereas 15% failed to complete the session under sham stimulation. This represents a +15% paired improvement in immersion tolerance, suggesting that the stimulation may meaningfully mitigate cybersickness symptoms.

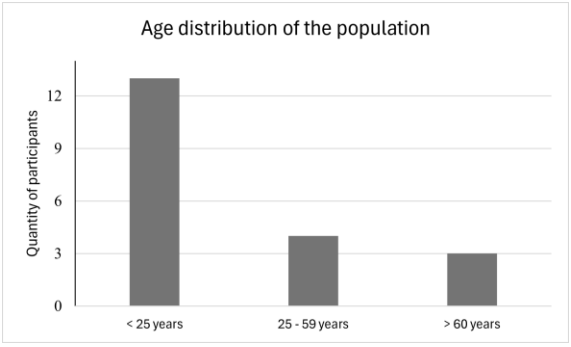


Figure 4. Age distribution of the studied population

TABLE II. Immersion Period of the Participants

Condition	Did not complete	Completed
Sham	3 (15 %)	17 (85 %)
nGVS	0	20 (100%)

After the trials, when examining all the information collected, it was observed that, due to participants' availability, it was possible to divide them into two groups based on the time interval between immersions, as described in Table III. According to Fujimoto *et al.* [18], avoiding crossover effects between sham and nGVS requires a minimum washout period, where the subject didn't received any stimulation or was using the VR system, of approximately 4 hours. Therefore, the results were reanalyzed using this new grouping to better quantify the influence of this interval within the ISF group (≤ 1 hour) and the SF group (> 1 hour).

TABLE III. Divison of the Participants Based on the Washout

Group	Quantity of participants
ISF (<1h)	13
SF (>1h)	7

The analysis of the ten items of the SUS questionnaire was performed using box-plot representations. For the ISF group, shown in Figure 5a, the SUS scores displayed very similar distributions between the sham and nGVS conditions. Both median and mean values remained close to 80, indicating consistently high perceived usability in both cases. Since the SUS reflects only the overall usability of the system, a paired Wilcoxon test was also conducted, yielding $W = 47.0$ and $p = 0.7609$, indicating no significant difference between the two immersions.

For the SF group, shown in Figure 5b, the SUS scores demonstrated greater dispersion under the nGVS condition compared to sham, indicating increased variability in perceived usability. Under nGVS, both the mean and median were slightly lower, with a broader interquartile range shifting toward lower ratings, whereas in the sham condition the median and mean remained around 80. Despite this increased

variability, the paried Wilcoxon test revealead significant difference between conditions, with W = 3.5 and p = 0.2763. These patterns suggest that, for the SF group, nGVS may influence usability perceptions, though the variability indicates that the effect is not uniform across participants.

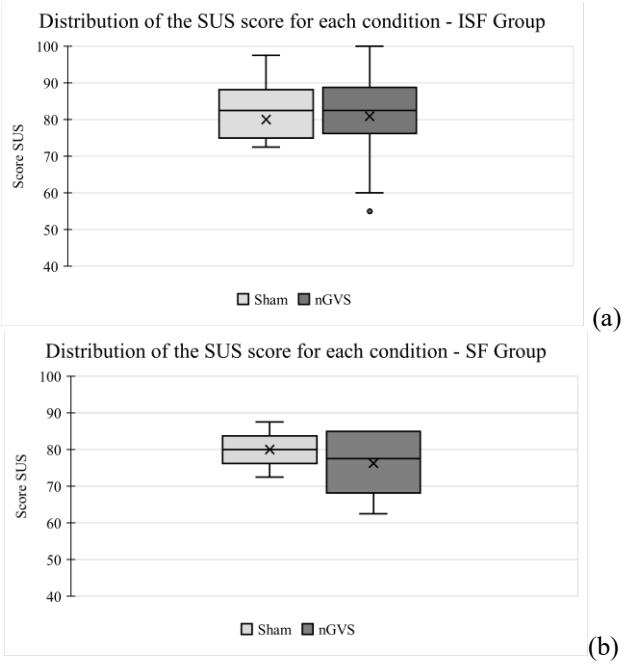


Figure 5. Box-plot of the SUS questionnaire scores (a) ISF (b) SF Group

For the SSQ, the total score and its sub-scores were analyzed, to determine whether the time between immersions affected the results. The difference between washout categories was not statistically significant (W = 31.5, p = 0.41). However, the sub-scores defined in (1) revealed a more specific pattern centered on the nausea sub-score. The distribution of values was more positive in the SF group (W = 20, p = 0.067), whereas the oculomotor (W = 26, p = 0.20) and disorientation (W = 49, p = 0.59) sub-scores did not show significant differences. Overall, within this study, the potential benefits of nGVS appear primarily in the nausea dimension.

When considering the intra-subject analysis and the order of exposure, the pattern of responses becomes clearer, as illustrated in Figure 6. In Group B (nGVS→sham), nausea tended to be reduced (HL > 0), with confidence intervals generally favoring the nGVS immersion, even with descriptive p-values in the 0.10–0.20 range. In contrast, in Group A (sham→nGVS), nausea did not improve and, in some cases, worsened, consistent with a carryover effect between immersions.

In conclusion, the different analyses of the SSQ data and immersion tolerance converge toward the interpretation that nGVS may primarily exert a *preventive* effect on nausea during VR immersion. This effect is most evident in conditions where stimulation is applied before or early in exposure and when sufficient washout time is allowed between sessions, as suggested by the higher completion rate and the more favorable nausea sub-scores in the SF group. In contrast, when nausea is already present from the first immersion or when the washout interval is short, nGVS does not reliably reduce symptoms and may, in some cases, fail to counteract carryover effects from the initial exposure. The oculomotor and disorientation dimensions show weaker, non-significant changes, but their overall trends remain directionally consistent with the nausea findings, supporting the interest of nGVS for future, larger-scale testing phases.

B. Limitations and Perspectives

Even with positive results, this work was subject to several limitations. The number of participants was restricted due to the limited availability of volunteers and the short experimental timeframe. Therefore, the results had to be analyzed according to the washout interval between immersions, which was essential to determine whether vestibular stimulation could truly improve immersion quality and reduce cybersickness symptoms. Short washout periods introduced potential carryover effects, particularly in the nausea dimension, which may have attenuated or reversed the expected benefits of nGVS in some participants.

Moreover, because the study used the VRSL environment, participation was limited to individuals with prior knowledge of electrical distribution substations. This ensured that participants could properly evaluate the realism and immersion of the virtual substation but, at the same time, reduced the diversity of the sample and limited the generalizability of the findings. The use of a professional-training VR environment may also have produced task-specific sensory demands, which could differ from those observed in entertainment-based or non-technical VR applications.

Looking ahead, future work will include a new round of experiments using a different VR environment, allowing for a larger and more heterogeneous sample of participants. Expanding the diversity of VR contexts will help determine whether the preventive effect of nGVS on nausea generalizes across various types of immersive content. Additional test sessions distributed over multiple days and with controlled washout intervals are also planned to better evaluate the reproducibility, robustness, and long-term effects of nGVS stimulation on comfort, posture stability, and immersion during VR exposure.

VI. CONCLUSION AND FUTURE WORK

This work presented the development and validation of a GVS system designed to improve immersion and comfort in VR environments. The system integrates hardware and software capable of delivering controlled vestibular stimulation and was successfully implemented within the VRSL platform for experimental evaluation. Technical tests confirmed that the device operates safely and reliably, providing stable current output and respecting user safety constraints.

The experimental results showed encouraging trends in this study. Participants exposed to nGVS showed lower SSQ scores, particularly in the nausea dimension, and demonstrated a higher tolerance to VR exposure, remaining longer in the virtual environment under stimulation than under sham. Although the SUS did not reveal statistically significant

differences in perceived usability, the distribution of scores suggested a slight shift toward more positive evaluations during nGVS immersion. Importantly, the analysis considering washout duration and exposure order indicated that the benefits of nGVS are primarily preventive, emerging when stimulation is applied before VR exposure, whereas a short washout interval can attenuate or even reverse these effects.

Overall, the findings support the hypothesis that controlled vestibular stimulation can help mitigate cybersickness and enhance immersive experience in VR environments. Future work will involve testing the system with a more diverse participant pool, exploring different VR contexts, and conducting extended multi-day experiments to improve reproducibility and refine stimulation parameters. Such developments will help determine the broader applicability of nGVS as a practical, low-cost tool for improving comfort and stability in immersive VR applications.

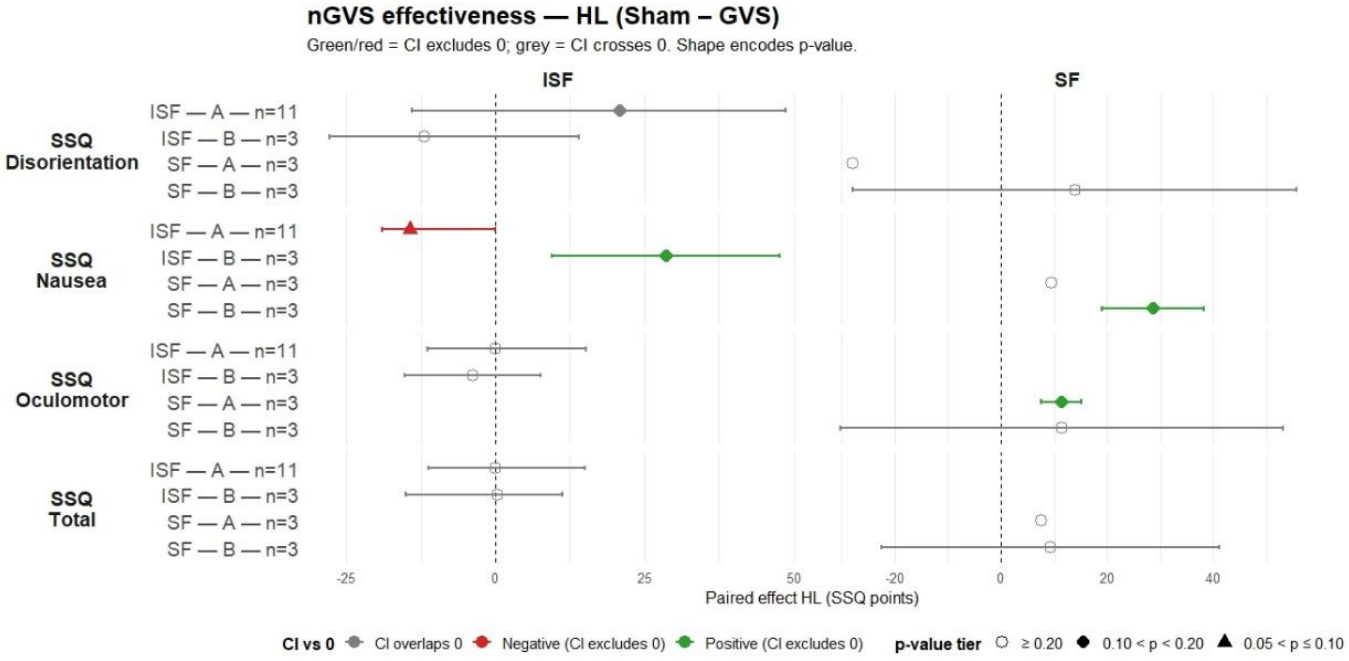


Figure 6. nGVS effectiveness when considering the order of exposure A or B

ACKNOWLEDGMENT

This study is part of a PhD thesis supported by Neural Balance Innovation (NBI).

All experiments were conducted in accordance with the Declaration of Helsinki, and all participants provided written informed consent prior to the beginning of the study.

REFERENCES

[1] D. R. Jaguaribe *et al*., "Virtual reality-based substation laboratory applied to electrical engineering education", Proc. XIX Encontro Regional Ibero-Americano do CIGRE (ERIAC 2023), Foz do Iguaçu, Brazil, CIGRE-Brasil, 2023.

[2] L. Simón-Vicente, S. Rodríguez-Cano, V. Delgado-Benito, V. Ausín-Villaverde, and E. Cubo Delgado, "Cybersickness: A systematic literature review of adverse effects related to virtual reality," Neurología, vol. 39, no. 8, pp. 701–709, 2024, doi:10.1016/j.nrl.2022.04.009.

[3] B. Guerrero, L. Cuevas, and L. Valero, "Side effects after using immersive virtual reality in a video game", Int. J. Psychol. Psychol. Ther., vol. 13, pp. 163–178, 2013.

[4] H. Oh and W. Son, "Cybersickness and its severity arising from virtual reality content: A comprehensive study," Sensors, vol. 22, no. 4, p. 1314, 2022, doi:10.3390/s22041314.

[5] Y. Valter *et al*., "A review of parameter settings for galvanic vestibular stimulation in clinical applications," Front. Hum. Neurosci., vol. 19, p. 1518727, 2025, doi:10.3389/fnhum.2025.1518727.

[6] S. Weech, J. Moon, and N. F. Troje, "Influence of bone-conducted vibration on simulator sickness in virtual reality," PLoS One, vol. 13, no. 3, p. e0194137, 2018, doi:10.1371/journal.pone.0194137.

[7] S. Weech, T. Wall, and M. Barnett-Cowan, "Reduction of cybersickness during and immediately following noisy galvanic vestibular stimulation," Exp. Brain Res., vol. 238, no. 2, pp. 427–437, 2020, doi:10.1007/s00221-019-05718-5.

[8] R. Goel, M. J. Rosenberg, H. S. Cohen, J. J. Bloomberg, and A. P. Mulavara, "Calibrating balance perturbation using electrical stimulation of the vestibular system," J. Neurosci. Methods, vol. 311, pp. 193–199, 2019, doi:10.1016/j.jneumeth.2018.10.012.

[9] V. G. Sirtoli, K. F. Morcelles, and V. C. Vincense, "Design of current sources for load common mode optimization," J. Electr. Bioimpedance, vol. 9, no. 1, pp. 59–71, 2018, doi:10.2478/joeb-2018-0011.

[10] Z. Liu, S. Suzuki, T. Fushimi, and Y. Ochiai, "Design and evaluation of a voltage-controlled current source for galvanic vestibular stimulation research," HardwareX, vol. 22, p. e00647, 2025, doi:10.1016/j.ohx.2025.e00647.

[11] A. Antal *et al*., "Low intensity transcranial electric stimulation: Safety, ethical, legal regulatory and application guidelines," Clin. Neurophysiol., vol. 128, no. 9, pp. 1774–1809, 2017, doi:10.1016/j.clinph.2017.06.001.

[12] Brooke, John. (1996). SUS -- a quick and dirty usability scale.

[13] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, “Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness,” Int. J. Aviat. Psychol., vol. 3, no. 3, pp. 203–220, 1993, doi:10.1207/s15327108ijap0303_3.

[14] R. Goel, I. Kofman, J. Jeevarajan, Y. De Dios, H. S. Cohen, J. J. Bloomberg, and A. P. Mulavara, “Using low levels of stochastic vestibular stimulation to improve balance function,” PLoS One, vol. 10, no. 8, p. e0136335, 2015, doi:10.1371/journal.pone.0136335.

[15] B. Rosner, R. J. Glynn, and M. L. Lee, “The Wilcoxon signed rank test for paired comparisons of clustered data,” Biometrics, vol. 62, no. 1, pp. 185–192, 2006, doi:10.1111/j.1541-0420.2005.00389.x.

[16] A. Krampe and S. Kuhnt, “Bowker’s test for symmetry and modifications within the algebraic framework,” Comput. Stat. Data Anal., vol. 51, pp. 4124–4142, 2007, doi:10.1016/j.csda.2007.01.021.

[17] M. Smith and G. Ruxton, “Effective use of the McNemar test,” Behav. Ecol. Sociobiol., vol. 74, p. 133, 2020, doi:10.1007/s00265-020-02916-y.

[18] C. Fujimoto *et al*., “Multicenter randomized double-blind placebo-controlled crossover study of the effect of prolonged noisy galvanic vestibular stimulation on posture or gait in vestibulopathy,” PLoS One, vol. 20, no. 1, p. e0317822, 2025, doi:10.1371/journal.pone.0317822.

Attractive Casimir–Lifshitz Forces as a Universal Driver of Prebiotic Protocell Aggregation and Cluster Formation

Michael Massoth
Department of Computer Science, Hochschule Darmstadt (h_da)
University of Applied Sciences Darmstadt, member of European University of Technology (EUt+)
Darmstadt, Germany
e-mail: michael.massoth@h-da.de

***Abstract*—How fragile, RNA-free protocells could have formed stable clusters in a hot, saline and dynamically fluctuating prebiotic environment remains unresolved. Classical explanations—such as hydrophobic effects, Derjaguin–Landau–Verwey–Overbeek (DLVO) interactions and thermal noise—fail in the relevant 5–200 nm separation range. Here we propose that attractive Casimir–Lifshitz forces act as a universal and physically unavoidable coupling mechanism between protocells. Starting from the exact sphere–sphere solution in the classical (high-T) limit and a Derjaguin (PFA) approximation, we show that Casimir–Lifshitz attraction decays algebraically as $1/L^2$ under prebiotically plausible conditions (protocell radii 100–1,000 nm; separations 5–200 nm; 50–200 mMol salt; 20–90 °C) and can exceed exponentially suppressed DLVO contributions beyond a few nanometres. From this framework we derive testable predictions—including stronger clustering of PMBCs relative to fatty-acid vesicles, enhanced effective adhesivity of larger protocells and contact lifetimes on the order of minutes—and outline a three-stage experimental roadmap to evaluate Casimir–Lifshitz forces as a realistic driver of prebiotic protocell cluster formation.**

***Keywords*—Casimir–Lifshitz Forces; Prebiotic Protocell Aggregation; Mesoscale Fluctuation-Induced Interactions; Non-Chemical Cooperation Mechanisms; Proto-Cluster Formation and Stability; Experimentally Testable Origin-of-Life Framework.**

I. INTRODUCTION

The emergence of life is commonly framed in terms of membranous protocells that provide confined reaction spaces, support concentration gradients, and enable molecular retention—features widely regarded as physical prerequisites for metabolism, information stabilization, and evolvable selection processes [4]. A broad range of experimental studies has demonstrated that simple amphiphilic systems can spontaneously assemble into membrane-bound vesicles under prebiotic conditions [5] and undergo primitive growth, fusion, and division dynamics.

What remains unresolved, however, is how mechanically fragile, RNA-free protocells could have formed stable dimers, trimers, and higher-order clusters in thermally active, saline, and dynamically fluctuating early environments. Classical interaction mechanisms offer no robust explanation in the mesoscale regime: hydrophobic forces act only at molecular contact, while electrostatic and van der Waals contributions described by Derjaguin–Landau–Verwey–Overbeek (DLVO) theory [3] are strongly screened under realistic ionic conditions and rapidly overwhelmed by thermal noise at the k_BT scale.

This gap motivates a central question: which physical mechanism could have supported reproducible, mesoscale, and non-chemical encoded protocell cooperation prior to genetic or metabolic specialization? Here we propose that attractive Casimir–Lifshitz (CL) interactions [1][2] —arising from quantum and thermal electromagnetic field fluctuations—constitute a physically unavoidable, material-dependent aggregation mechanism operating precisely in the regime where classical colloidal forces fail. Unlike DLVO-type interactions, CL forces do not rely on surface charge or molecular specificity and persist across nanometre-to-submicrometre separations in electrolyte environments.

In this work, we develop a quantitative theoretical framework for CL-mediated protocell interactions, establish their relevance relative to classical screened forces, and translate the resulting scaling laws into experimentally testable predictions for protocell clustering. By grounding early cooperative organization in fluctuation-induced physics rather than biochemical functionality, this study reframes protocell aggregation as a physically emergent precursor to later chemical and informational complexity.

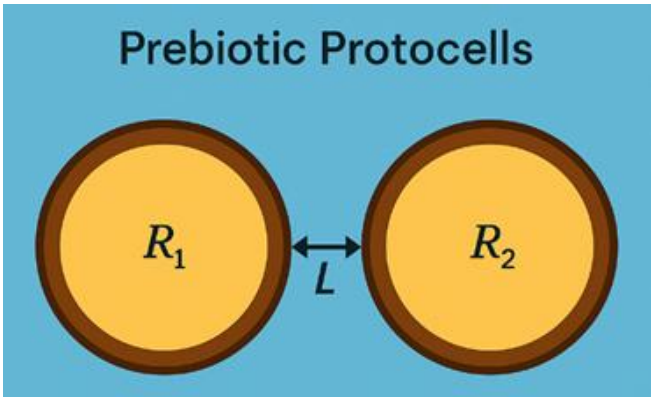


Figure 1. *Schematic of two prebiotic protocells with radii R_1 and R_2 and minimal surface-to-surface separation L in saline water*

In Figure 1 two prebiotic protocells with radii R_1 and R_2 are depicted at a minimal surface-to-surface separation L in saline

water (primordial soup), defining the sphere–sphere geometry underlying the Casimir–Lifshitz interaction model.

Contribution Summary: In this study, we propose that attractive Casimir–Lifshitz (CL) forces provide a universal, physically unavoidable mechanism for mesoscale protocell aggregation under prebiotically plausible conditions where classical Derjaguin–Landau–Verwey–Overbeek (DLVO) interactions fail. Building on the exact sphere–sphere formulation in the classical (high-temperature) limit and a Derjaguin proximity-force approximation, we derive an algebraic interaction scaling that remains operative in regimes inaccessible to exponentially screened electrostatic forces. We identify the resulting dominance window relative to DLVO contributions and show that CL attraction can reach thermal relevance at the level of individual protocell contacts. Finally, we translate these scaling results into experimentally testable predictions—including material-dependent clustering, radius-dependent effective adhesivity, and finite contact lifetimes—and outline a staged experimental roadmap to evaluate fluctuation-induced forces as a candidate driver of early protocell cluster formation.

The structure of this paper is as follows:

In Section II, we contrast classical DLVO theory with Casimir–Lifshitz theory, highlighting their distinct physical origins, interaction ranges and relevance under ion-rich prebiotic conditions.

In Section III, we develop a quantitative Casimir–Lifshitz framework for spherical protocells, progressing from the fundamental Lifshitz formulation to exact sphere–sphere results and biologically tractable Derjaguin-type approximations, and directly compare these predictions with DLVO and van der Waals interactions.

In Section IV, we delineate a prebiotically realistic parameter window—geometry, dielectric properties, temperature and salinity—within which fluctuation-induced forces become physically and geochemically plausible.

In Section V, we derive experimentally testable predictions regarding material dependence, size scaling, temperature sensitivity and measurable deviations from DLVO behaviour.

Sections VI and VII integrate these physical results into a broader prebiotic evolutionary perspective, discussing implications for early protocell clustering, proto-cooperation and emergent organization.

Section VIII concludes with a summary and outlines directions for future experimental and theoretical work.

II. THEORETICAL BACKGROUND: CASIMIR–LIFSHITZ VERSUS DLVO

The physical interactions between protocells can be described largely by two frameworks: classical DLVO theory and the Casimir–Lifshitz theory [1][2] of quantum- and thermally induced fluctuation forces. Both include attractive and repulsive contributions at nanometre scales, but they differ fundamentally in physical origin, range and environmental robustness. For prebiotic scenarios this distinction is crucial, as early environments were ion-rich, thermally dynamic and chemically heterogeneous—conditions under which classical colloidal idealizations apply only in a limited sense.

A. DLVO Theory: Limitations in a Prebiotic Primordial Context

Classical DLVO theory models the force balance between colloidal objects as the sum of an electrostatic double-layer repulsion and a short-range van der Waals attraction [3]. The electrostatic component is governed by the Debye length, which shrinks exponentially with increasing ionic strength. In saline solutions—realistic for prebiotic environments—the Debye length is typically only ~1–2 nm, rendering electrostatic repulsion extremely short-ranged.

The van der Waals contribution is likewise confined to very small separations. Direct AFM force measurements in electrolyte solutions show that additional attractive components—depending on ion valency and surface chemistry—are typically significant only within ~0.3–1.0 nm, at most 1–3 nm, and fall below k_BT beyond these distances.

Under prebiotic ionic conditions, and for separations L of 2–200 nm relevant to protocell clustering, classical van der Waals and DLVO interactions therefore provide no robust mechanism for stable mesoscale association.

B. Casimir–Lifshitz Theory

The Casimir–Lifshitz theory [1][2] is a field-theoretic generalization of van der Waals interactions. It arises from quantum and thermally induced electromagnetic field fluctuations. Crucially, these forces do not depend on real charges, receptors or chemical bonds, but on how the fluctuating electromagnetic field couples to the material-dependent reflection properties of the interacting interfaces. The central quantity is therefore the dielectric response function $\varepsilon(i\xi_n)$, rather than surface charge.

Casimir–Lifshitz forces act irrespective of salt concentration or molecular specificity and operate over distances of 2–200 nm—the very regime in which prebiotic protocells would have interacted. This mechanism is present both at T=0 through quantum vacuum fluctuations and at finite temperature through additional thermal contributions.

C. Relevance for Prebiotic Protocell Clusters

Protocell membranes exhibit a strong dielectric contrast [12] relative to saline water: $\varepsilon_{Membran} \approx 2$ to $8 \ll \varepsilon_{Water}$ (at 50–200 mMol) $\approx$ 75 to 78. Such dielectric-contrast–driven fluctuation forces were first described by Casimir [1] and later generalized to real materials and media by Lifshitz [2]. This contrast satisfies the sign condition for attractive Casimir–Lifshitz interactions. Thermal activity further enhances the relevant fluctuation modes, yielding a robust, non-chemical aggregation mechanism capable of stabilizing mesoscale

protocell assemblies without any genetic, metabolic or enzymatic specialization.

Empirical support for long-range, fluctuation-related forces in electrolyte environments includes AFM studies [13] on multivalent-ion-induced attraction (Moazzami-Gudarzi et al., 2016) [18], force measurements between silica particles in electrolytes (Valmacco et al., 2016) [19] and detailed AFM analyses of van der Waals [14] and DLVO contributions in saline media (Butt et al.; 1991) [20]. Together, these findings reinforce the plausibility of fluctuation-induced forces as contributors to mesoscale stability under prebiotic conditions.

III. QUANTITATIVE CASIMIR–LIFSHITZ THEORY AND APPROXIMATION METHODS

To assess the strength, range and biological relevance of Casimir–Lifshitz interactions in prebiotic protocell systems, we require a theoretical framework that links exact field-theoretic formulations with biologically tractable approximations. Protocells are mesoscale objects with radii of ~200 nm to 1000 nm; their interactions therefore cannot be captured by simple plate geometries but must be described using a sphere–sphere configuration that explicitly incorporates temperature, ionic strength and material properties.

In this section, we proceed from the universal Lifshitz formulation (*A.*) to an exact sphere–sphere description in the classical (high-T) limit (*B.*), derive an asymptotic Derjaguin (PFA) approximation (*C.*), and finally compare the resulting algebraic interaction range with the exponentially screened DLVO predictions (*D.*).

To ensure a clear distinction between physical results, model-dependent conclusions, and interpretative extensions, we explicitly differentiate three epistemic levels throughout this work.

Level I (Established Physics) refers to results derived directly from Casimir–Lifshitz theory and primary experimental literature.

Level II (Model-Based Deductions) comprises conclusions that follow from the specific geometries, parameters, and approximations introduced here.

Level III (Prebiotic Hypotheses) denotes interpretative extensions that apply these physical results to early protocell systems and evolutionary contexts. This epistemic distinction is independent of the formal subsection labeling used throughout the manuscript.

A. Casimir–Lifshitz Fundamental Formulation

The Casimir–Lifshitz interaction [1][2] between two bodies in a medium arises from modifications of the electromagnetic fluctuation spectrum between their interfaces. In the field-theoretic formulation, the free energy F(L) at a surface separation L is expressed as a Matsubara sum. The standard Lifshitz–Matsubara expression for F(L) follows the classical works of Lifshitz (1956) [2] and the formulation in Boström and Sernelius (2000) [17]. The frequency-dependent dielectric response enters through $\varepsilon(i\xi_n)$, with full specifications provided in these references. The summation runs over discrete Matsubara frequencies $\xi_n=2\pi n k_B T/\hbar$. The terms $r_1^{(n)}$ and $r_2^{(n)}$ represent reflection-coefficient-like material parameters determined by the dielectric response $\varepsilon(i\xi_n)$. Casimir–Lifshitz forces have been experimentally demonstrated [6].

(Level-I:) This formulation is universal and applies to arbitrary material combinations, geometries and intervening media. For biological systems, it implies that the resulting force does not rely on specific chemical bonds but emerges solely from the fundamental interfacial electromagnetic properties of membranes and their surrounding medium.

B. Exact Sphere–Sphere Geometry in the Classical Limit

Because prebiotic protocells are approximately spherical compartments suspended in electrolyte solutions, the sphere–sphere configuration [7][8] represents the physically most realistic interaction geometry. The exact description of the Casimir–Lifshitz interaction energy in the classical (high-T) limit was developed within the scattering-field formalism of Rahi et al. [9] and analytically evaluated by Bimonte and Emig [10] for two spheres of radii R_1 and R_2, separated by a surface-to-surface distance L, including Debye screening with $\kappa=1/\lambda_D$. Because the full expressions are not part of the methodological advance presented here, we refer to Eq. (3) in Bimonte and Emig [10] and use this formulation as the theoretical foundation for the practically relevant PFA scaling discussed in Section *III.C.*

(Level-I:) This formulation captures the exact curvature dependence of the interaction and does not rely on idealized parallel plates. In the classical limit, the n=0 Matsubara term dominates; in electrolyte media this contribution is partially screened by Debye damping, $\Lambda \rightarrow \Lambda e^{-\kappa L}$. Higher-order Matsubara modes ($n \geq 1$) remain present and are fully included in the Lifshitz formalism [11].

(Level-II:) As a consequence, the interaction is only partially attenuated in saline environments rather than eliminated, because the higher modes persist. Thus, Casimir–Lifshitz forces remain operative over 5–200 nm even under realistic prebiotic conditions.

C. Derjaguin Approximation as a Practical Scaling

In the biologically relevant regime $L \ll R$ (here L=2–100 nm and R=200–1000 nm), the exact Casimir–Lifshitz description for a sphere–medium–sphere system can be reduced, via the Derjaguin proximity-force approximation (PFA), to a simple scaling form: $F_{CL}(L) \propto A_{eff} * R_{eff}/L^2$. Here, the effective curvature radius R_{eff} acts as the dominant amplification factor for biological cluster stability. A_{eff} is an effective, medium-dependent Hamaker constant derived from the full Lifshitz spectrum. Larger protocells therefore generate systematically stronger coupling at identical material parameters.

The following applies:
$F_{CL}(L) \approx -(A_{eff}/6) * (R_{eff}/L^2)$ with $R_{eff} = (R_1 * R_2)/(R_1 + R_2)$.
(Level-I/II:) Here, A_{eff} denotes an effective Hamaker constant that integrates the spectral dielectric response of the membrane–water system. The force scales linearly with the effective curvature radius R_{eff} and decays algebraically as $1/L^2$. (Level-II:) For the parameter ranges considered below, the resulting potential wells reach several k_BT, making them relevant for mesoscale cluster stabilization.
The corresponding effective binding potential $U_{CL}(L)$ follows directly from integration: $U_{CL}(L) \propto -(A_{eff}/6) * (R_{eff}/L)$, with A_{eff} the membrane–water Hamaker constant, R_{eff} the reduced curvature of the two spheres, and L the minimal surface-to-surface distance. The approximation holds for $L \ll R_i$, L larger than the membrane thickness, and smooth, non-adhesively functionalized interfaces.
In Figure 2, the attractive Casimir–Lifshitz force $F_{CL}(L)$ between two PMBC-like protocells with $R_1 = R_2 = 500$ nm is shown as a function of separation L (logarithmic x-axis). The algebraic decay $F_{CL}(L) \propto 1/L^2$ yields a pronounced mesoscale attraction over distances of 5–200 nm. All numerical results are obtained using an effective Hamaker constant $A_{eff} = 5 \times 10^{-21}$ Joule. A_{eff} should be understood as an effective, medium-dependent Hamaker constant derived from the full Lifshitz spectrum, not a fitted free parameter.

Biophysical implication: Because $U_{CL}(L) \propto -R_{eff}/L$, large protocell radii and moderate separations favour mesoscale cluster stability.
Core statement: Fluctuation-induced attraction strengthens with protocell curvature R and decays only algebraically with distance L, enabling non-chemical mesoscale stabilization of protocell assemblies.

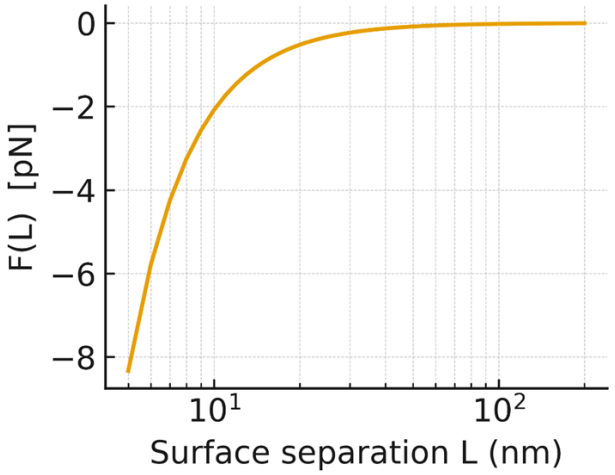


Figure 2. *Attractive Casimir–Lifshitz force $F_{CL}(L)$ between two PMBC-like protocells (R_1=R_2=500 nm) as a function of separation L.*

Quantum, crossover and classical regimes: For separations $L \leq 10$ nm, the interaction is dominated by the quantum-fluctuation (T = 0) term, and a non-retarded PFA form $F_{CL}(L) \approx -(A_0/6) * (R_{eff}/L^2)$ is appropriate. Between 10–30 nm, a crossover regime appears in which the full Matsubara summation must be retained. Practically, this behaviour can be described by an effective Hamaker constant $A_{eff}(L,T) = A_0 + \Delta A_T(L)$, while the scaling $F_{CL} \propto R_{eff}/L^2$ remains unchanged. Above ~30–50 nm, the classical (high-T) contribution dominates, with thermal modes setting the leading term.
Across the relevant prebiotic range, the algebraic dependence $F_{CL} \propto R_{eff}/L^2$ is preserved, and the effective strength $A_{eff}(L,T)$ simply reflects the combined quantum and thermal contributions. No exponential screening occurs, in stark contrast to DLVO components.

D. Comparison: Brownian Motion vs. Casimir–Lifshitz vs. DLVO vs. Van-der-Waals

In a thermally active, ion-rich prebiotic environment at ~25 °C, stable protocell clusters require interaction wells deeper than ~3–5 k_BT per contact to persist on mesoscale timescales. Forces of order $\leq k_BT$ or with ranges shorter than ~3 nm are insufficient to stabilize assemblies against Brownian disruption over separations of 5–200 nm.
The key question, therefore, is which interaction dominates at prebiotically plausible distances. DLVO theory predicts an electrostatic contribution decaying as $V_{DLVO}(L) \propto \exp(-L/\lambda_D)$ with Debye lengths $\lambda_D \approx 0.7$–1.4nm for ionic strengths of 50–200 mM. Under these conditions the DLVO range is unavoidably exponential and extremely short.
(Level-I:) Classical van der Waals interactions are likewise restricted to sub-nanometre to few-nanometre separations. By contrast, the PFA scaling of the Casimir–Lifshitz force yields $F_{CL} \propto R_{eff}/L^2$, with no exponential collapse.
(Level-II:) For parameter ranges R ≈ 500–1000 nm and L ≈ 2–200 nm, a regime emerges in which the Casimir–Lifshitz contribution either dominates or is comparable to the residual DLVO terms. For realistic prebiotic distances (L > 2–100 nm) applies $|F_{CL}(L)| \geq |F_{DLVO}(L)|$.

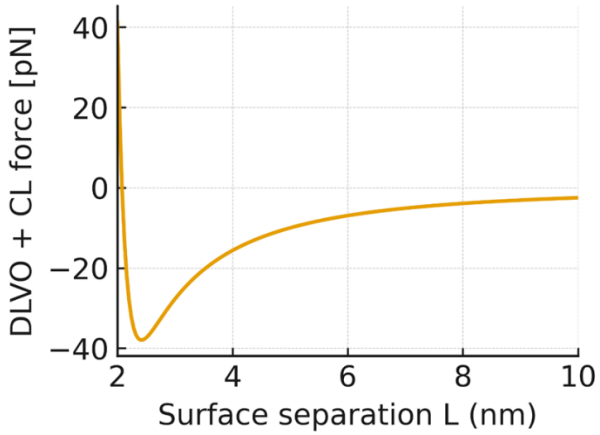


Figure 3. *Resulting total force from F_{DLVO} and F_{CL} contributions over L=2–10 nm.*

In Figure 3, the resulting total force from combined F_{DLVO} and F_{CL} contributions is shown over separations of 2–10 nm.

The algebraic Casimir–Lifshitz component generates a residual, non-DLVO-compatible attraction as potential well that remains experimentally detectable. The resulting attractive force (down to −38 pN) stabilizes the bound protocell cluster configuration.

(Level-III:) Under plausible prebiotic conditions, there therefore exist distance intervals where Casimir–Lifshitz interactions match or exceed the remaining DLVO contributions—without implying universal dominance. This supports the interpretation that Casimir–Lifshitz forces represent a general physical aggregation mechanism in saline, thermally dynamic early-Earth environments.

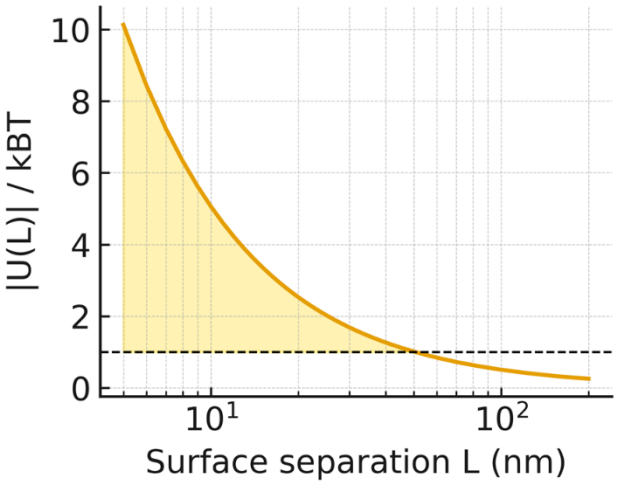


Figure 4. *Distance-dependent interaction potential /U(L)/ / k_BT between two PMBC protocells (R=500 nm) as function of separation L.*

The quantity |U(L)| / k_BT is dimensionless and measures the strength of the interaction relative to thermal fluctuations at 25 °C (298.15 K). Values > 1 k_BT imply that the interaction can suppress Brownian separation and stabilize protocell contacts.

In Figure 4, the distance-dependent interaction potential |U(L)| / k_BT between two PMBC protocells with R = 500 nm is shown as a function of separation L. Values exceeding 1 k_BT define the *thermal relevance zone*, in which Casimir–Lifshitz interactions overcome Brownian motion and stabilize nanoscale protocell cluster building over separations L of 2–50 nm.

Sensitivity analysis:

The sensitivity analyses performed within this framework show that Casimir–Lifshitz–mediated protocell coupling is most strongly controlled by the inter-compartment separation L, reflecting the algebraic scaling $F_{CL} \propto L^{-2}$ and $U_{CL} \propto L^{-1}$. Variations in protocell size enter linearly via the effective curvature radius R_{eff}, such that larger compartments systematically exhibit stronger effective adhesion at fixed separation. Material-dependent parameters, captured through the effective Hamaker constant A_{eff}, scale the absolute interaction strength but do not alter the qualitative distance dependence. Temperature primarily modulates the relative contribution of quantum and thermal fluctuation modes, without changing the underlying algebraic scaling. In contrast, variations in ionic strength predominantly affect DLVO contributions through exponential screening, while Casimir–Lifshitz interactions remain only partially attenuated. Together, these sensitivities demonstrate that the dominance window of fluctuation-induced attraction is robust across realistic variations in geometry, material properties, temperature, and salinity.

IV. PREBIOTICALLY REALISTIC PARAMETER WINDOW

For the Casimir–Lifshitz interactions derived in Section 3 to function as realistic drivers of prebiotic self-organization, they must operate within a physically and geochemically plausible parameter space. We therefore consolidate ranges supported by geochemical considerations and experimental protocell research: geometry (*A.*), dielectric properties (*B.*), temperature (*C.*) and ionic environment (*D.*).

A. Geometry: Protocell Radii and Inter-Compartment Distances L

(Level-II:) Experimental and theoretical studies of prebiotic vesicles [5] suggest that early protocells commonly exhibited radii on the order of 200–1000 nm. At these sizes, effective compartmentalization is feasible and the Derjaguin-PFA interaction window of Casimir–Lifshitz forces becomes particularly relevant. Inter-protocell gaps within clusters of ~2–200 nm are plausible: smaller than the vesicle diameter but larger than hydration shells. These distances are experimentally accessible (e.g., cryo-TEM, confocal microscopy) and lie precisely in the regime where fluctuation-induced forces are not thermally overwhelmed, while classical Coulomb and DLVO contributions are already strongly attenuated.

(Level-III:) Geometrically, this yields a favourable region in which protocells are large enough to generate substantial Casimir–Lifshitz attraction through R_{eff}, yet small enough for thermal and chemical gradients within clusters to remain functionally effective.

B. Membrane Materials: Fatty-Acid Vesicles and PMBCs

(Level-II:) Prebiotic models commonly consider fatty-acid vesicles and protein-/polymer-based compartments (PMBCs). Simple fatty-acid membranes exhibit low effective permittivities $\varepsilon_{membrane} \approx 2–4$, far below that of saline water $\varepsilon_{water} \approx 75–78$ at 50–200 mMol ionic strength. This contrast satisfies the sign condition for attractive Casimir–Lifshitz interactions: two vesicles in water generally attract. PMBCs are thicker, mechanically more robust and have effective permittivities of $\varepsilon_{membrane} \approx 3–8$. Owing to their higher polarizability and stability, PMBCs represent particularly

suitable systems in which Casimir–Lifshitz forces should be pronounced and experimentally accessible.
(Level-III:) Both classes—fatty-acid vesicles and PMBCs—thus provide realistic membrane architectures yielding the dielectric contrast required for attractive fluctuation-induced coupling.

C. Temperature: 20–80 °C as an Activation Window

Many geochemical scenarios [5] place early compartments in hydrothermal or volcanic environments with temperatures between ~20 and 90 °C. (Level-I/II:) Temperature enters Casimir–Lifshitz theory explicitly through the classical contribution to the free energy, which scales $\propto T$ in the high-T limit, thereby enhancing the thermal component of the attraction as long as membrane integrity is maintained. Temperature also modulates membrane fluidity, influencing the likelihood of fusion, hemifusion and transient pore formation.
(Level-II/III:) Fatty-acid vesicles typically remain stable between ~20 and 60 °C, whereas PMBCs can remain intact up to ~80 °C. This defines a temperature window in which Casimir–Lifshitz forces are sufficiently strong without thermal destruction dominating.

D. Salt and Ionic Strength: 50–200 mMol as the Screening Regime

(Level-II:) Prebiotic waters [5] likely exhibited ionic strengths of ~50–200 mMol [15][16], comparable to modern marine or brackish conditions. In this regime, the Debye length is only ~0.7–1.4 nm, limiting DLVO electrostatic double-layer repulsion to a few nanometres. Casimir–Lifshitz interactions behave differently: the zero-frequency term is partially screened by ions, but higher fluctuation modes largely persist. This leads to attenuation, but not elimination, of the total force over 2–100 nm.
(Level-III:) Even in saline prebiotic environments, an attractive, universal coupling remains active that can stabilize protocell clusters on mesoscale distances.

E. Consolidated Window Statement

(Level-III:) Protocell radii of 200–1000 nm, separations of 2–100 nm, temperatures of 20–80 °C and ionic strengths of 50–200 mMol [5] define a physically and geochemically plausible parameter window in which Casimir–Lifshitz interactions [1][2] arise and can contribute to non-chemical mesoscale aggregation of protocells.

V. PREDICTABLE EFFECTS AND EXPERIMENTALLY TESTABLE HYPOTHESES

The theoretical framework developed above leads to specific, quantitative and experimentally testable predictions. At its core, we propose that Casimir–Lifshitz (CL) interactions generate reproducible mesoscale stability and aggregation effects within prebiotically plausible parameter ranges. The following hypothesis groups constitute an empirically accessible test program for biophysical, prebiotic and synthetic-protolife experiments.

A. Material-dependent Clustering: PMBCs > Fatty-Acid Vesicles

Because of their higher polarizability and greater membrane thickness, PMBC compartments are expected to form deeper CL potential wells and more stable clusters than pure fatty-acid vesicles. PMBCs should therefore assemble into clusters more frequently, with larger diameters, longer-lived geometries and extended contact times. The underlying mechanism is the proportional enhancement of fluctuation coupling through the membrane's dielectric response: higher polarizability yields deeper potential wells and stronger *physical adhesion* without chemical bonding.

Experimental signature: increased co-residence probability ($\leq$100 nm) of two labelled compartments in FRET, dual-fluorescence or confocal time-trace analyses.

B. Radius-dependent Adhesion: Larger Compartments Couple More Strongly

From the Derjaguin-PFA approximation (Section 3), the force contribution scales linearly with the effective radius R_{eff}: $F_{CL} \propto R_{eff}$. Larger protocells should therefore appear effectively more "adhesive'' at identical separations, exhibiting reduced relative drift, limited rotational freedom and enhanced contact persistence.
Biologically, this implies that volumetric growth or swelling processes may have conferred early selective advantages—not because of internal chemistry alone, but due to more stable physical coupling.

Experimental signature: systematic radius dependence in automated trajectory analysis (particle tracking, µPIV, confocal or epifluorescence time series).

C. Temperature Window for Maximal Coupling

Casimir–Lifshitz interactions contain a temperature-dependent component that scales $\propto T$ in the classical regime. This yields the prediction of an optimal stability window around ~40–80 °C, where thermal amplification and membrane integrity are balanced. Below ~20 °C, viscoelastic membrane processes slow and may become too rigid; above ~80 °C, structural failures emerge more rapidly. This predicted window aligns remarkably well with hydrothermal and vulcanolimnic habitats [5] invoked in prebiotic models.

Experimental signature: maximal cluster persistence, contact time or dissociation half-life in thermostated microfluidic assays.

D. Minute-scale Contact Times from Shallow Potential Wells

A characteristic feature of weak but sustained attraction is the formation of shallow potential wells with depths of several k_BT. Such wells permit semi-stable binding without irreversible fusion—an evolutionarily favourable regime that facilitates exchange, fusion events and lateral material transfer. We therefore predict that, under prebiotic conditions, two protocells should display contact times on the order of minutes before separation or remobilization occurs.

Experimental signature: dwell-time distributions (sub-100-nm regime) measured via Total Internal Reflection Fluorescence Microscopy (TIRF), Fluorescence Recovery After Photobleaching (FRAP) or single-particle dwell-time analysis.

E. Evidence for Algebraic Residual Attraction beyond DLVO Predictions

Because DLVO potentials decay exponentially with the Debye length, whereas CL forces decay algebraically, measurable discrepancies are expected between observed and DLVO-predicted force profiles—particularly within the 2–200 nm window. The critical finding would be the detection of residual attraction even when DLVO models predict neutrality or repulsion.

Experimental signature: non-zero adhesion or pull-off forces in optical or magnetic tweezers, AFM force spectroscopy or micro-traction assays.

VI. PREBIOTIC EVOLUTIONARY IMPLICATIONS

The results presented here demonstrate that Casimir–Lifshitz attraction, under prebiotically plausible conditions, provides a physically unavoidable and energetically relevant contribution to protocell stabilization and aggregation. Unlike DLVO-derived interactions, CL forces remain sensitive to material and geometric properties, act independently of metabolic or genetic mechanisms, and therefore precede classical biochemical modes of cooperation. The resulting cluster formation constitutes an early form of spatial coupling from which selectable chemical and informational organization could emerge.

A. Physical Emergence Preceding Chemical Specialization

Traditional origin-of-life models often assume that chemical functionality is a prerequisite for cooperation. Our findings indicate that attractive Casimir–Lifshitz coupling can already generate stability, cohesion and spatial coordination in the absence of biochemical instruction. Cooperation thus appears first as a physical order state, not as a product of pre-existing biofunction. This perspective aligns with dissipative self-organization frameworks in which mesoscale structuring arises deterministically from fluctuation-driven dynamics and forms an evolutionary precursor to chemical specialization.

B. Protocell Clusters as Precursors of Functional Micro-ecosystems

Metastable protocell clusters couple chemically relevant microenvironments, enhancing molecular retention, local concentration increases and stable microgradients—without requiring specialized transport machinery. The resulting exchange and recombination processes endow protocell ensembles with features of pre-ecological functional architectures, whose mesoscale connectivity may provide adaptive benefits independent of metabolic or genetic elaboration.

C. Preconditions for Proto-Informational Emergence

Physical coupling establishes a stable interaction space in which repeatable contact patterns, persistence and rudimentary memory effects arise. Such statistically non-random state differentiations constitute a minimal form of proto-informational structure, emerging long before sequence-based biopolymers existed. Information formation therefore appears as an emergent by-product of structurally stabilized interaction states rather than exclusively as an output of genetic systems.

D. Persistence and Resilience as Prebiotic Selection Factors

In prebiotic contexts, persistence rather than replication represents the primary mode of selection. CL-stabilized protocell clusters meet this requirement: they withstand environmental fluctuations, remain reversibly reconfigurable, passively harness external energy flows and form mesoscale networked structures. Casimir–Lifshitz attraction thus constitutes a coherent candidate for an early evolutionary selection filter, operating prior to genetic information systems and enabling structural durability in a noise-dominated environment.

E. Limitations and Scope Clarification

This work does not claim that attractive Casimir–Lifshitz forces alone account for protocell cooperation or aggregation. Rather, we propose that they represent a physically unavoidable under realistic material and environmental conditions, non-chemical baseline mechanism whose magnitude has been underestimated in prebiotic models. Additional chemical or environmental contributions may coexist, but CL forces provide a universal physical floor on which further stabilization processes may build.

VII. DISCUSSION

In this work, we develop a theoretically coherent and experimentally actionable framework in which Casimir–Lifshitz (CL) forces act as a prebiotically relevant, physically determined aggregation mechanism between protocells. To clarify why physical coupling dominates under realistic

early-Earth conditions (primordial soup), the interaction landscape is summarized schematically in Figure 5.

Physical vs. Chemical Interaction Landscape

Interaction	Range	Salt sensitivity	Chemistry required
Hydrophobic	< 1 nm	–	yes
DLVO	< 2 nm	high	yes
Casimir–Lifshitz	2–100 nm	low	no

Figure 5. Physical interaction regimes relevant to protocell clustering

Figure 5 compares interaction range, salt sensitivity, and chemical specificity of hydrophobic forces, DLVO interactions, and Casimir–Lifshitz forces. The schematic highlights that only Casimir–Lifshitz interactions combine mesoscale range (2–100 nm), low ionic sensitivity, and chemistry-independent operation, identifying them as a uniquely robust physical coupling mechanism under saline prebiotic conditions.

Based on field-theoretically consistent sphere–sphere models and biologically scalable parameter ranges, we show that these interactions operate with significant range and energetic relevance under plausible early-Earth conditions. They therefore constitute not merely a theoretical possibility but a functional mesoscale contribution to pre-cooperative organization.

The algebraic distance dependence of Casimir–Lifshitz forces provides a key plausibility advantage over exponentially screened DLVO components: above the DLVO damping regime, fluctuation-induced attraction remains dominant or energetically comparable, even in saline, thermally active, and dynamically perturbed early habitats. This addresses a long-standing gap in origin-of-life models that have traditionally explained mesoscale structural stability primarily through chemical, hydrophobic, or stochastic processes.

The resulting research logic reframes the emergence of cooperative protocell ensembles from a chemically metaphorical narrative into a metrically quantifiable mechanistic framework, as all relevant system parameters (permittivities, radii, ionic strengths, distances, temperatures) are experimentally tunable, measurable, and modelable. Cooperation thus appears not as a late by-product of biochemical innovation but as a physically grounded starting architecture upon which metabolic, genetic, and semantic complexity could later evolve.

VIII. CONCLUSION AND FUTURE WORK

This work establishes Casimir–Lifshitz forces as a physically unavoidable and prebiotically relevant baseline attraction between protocells operating at the mesoscale. Their algebraic distance dependence confers greater stability and range robustness than electrostatics-based models under realistic early-habitat conditions. As a result, the emergence of cooperative protocell assemblies becomes metrically modelable, experimentally accessible and decoupled from purely chemical or stochastic explanations. The model positions physical cooperation architecture as an evolutionary starting point rather than a late biochemical outcome, forming the foundation for a systematic research programme integrating physics, proto-ecology and proto-information. This framework directly motivates quantitative experimental tests of protocell clustering under controlled salinity, temperature, and material conditions.

Figure 6 translates the physical framework developed here into a concrete, staged experimental roadmap for future validation.

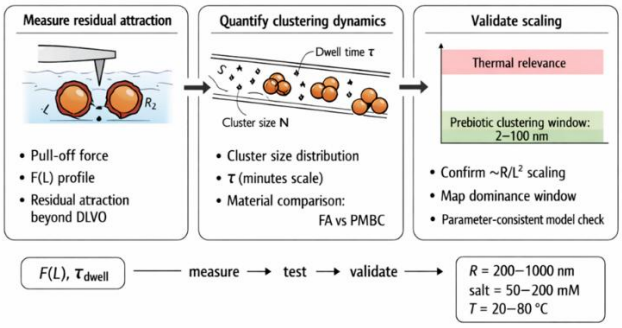


Figure 6. Experimental roadmap for physical protocell clustering validation

In Figure 6, a three-stage experimental roadmap is presented that operationalizes the physical framework introduced in this work. The schematic describes how residual fluctuation-induced forces, clustering dynamics, and scaling behavior can be systematically measured and validated, defining a practical experimental agenda for testing Casimir–Lifshitz–driven protocell cooperation under prebiotic conditions.

If prebiotic protocells did not drift alone in the chaos of the early oceans but were joined into semi-stable clusters by universal fluctuation forces, then the first evolutionary unit was not the isolated protocell, but the cluster (as dimers and tetrahedrons)—a naturally arising, collectively stabilized nanoscale assembly, shaped by physics into the earliest proto-ecosystem.

REFERENCES

[1] H. B. G. Casimir, "On the attraction between two perfectly conducting plates," *Proc. K. Ned. Akad. Wet.*, vol. 51, pp. 793–795, 1948.

[2] E. M. Lifshitz, "The theory of molecular attractive forces between solids," *Sov. Phys. JETP*, vol. 2, pp. 73–83, 1956.

[3] J. N. Israelachvili, *Intermolecular and Surface Forces*, 3rd ed., Academic Press, London, 2011.

[4] **[4]** P. C. W. Davies, "Does quantum mechanics play a non-trivial role in life?," *BioSystems*, vol. 78, pp. 69–79, 2004.

[5] I. Gözen *et al*., "Protocells: Milestones and recent advances," *Small*, vol. 18, pp. 2106624-1–2106624-12, 2022.

[6] S. K. Lamoreaux, "Demonstration of the Casimir force in the 0.6 to 6 μm range," *Phys. Rev. Lett.*, vol. 78, pp. 5–8, 1997.

[7] J. L. Garrett, D. A. T. Somers, and J. N. Munday, "Measurement of the Casimir force between two spheres," *Phys. Rev. Lett.*, vol. 120, pp. 040401-1–040401-5, 2018.

[8] P. Rodriguez-Lopez, "Casimir energy and entropy in the sphere–sphere geometry," *Phys. Rev. B*, vol. 84, pp. 075431-1–075431-9, 2011.

[9] S. J. Rahi, T. Emig, N. Graham, R. L. Jaffe, and M. Kardar, "Scattering theory approach to electrodynamic Casimir forces," *Phys. Rev. D*, vol. 80, pp. 085021-1–085021-29, 2009.

[10] G. Bimonte and T. Emig, "Exact results for classical Casimir interactions: Dirichlet and Drude model in the sphere–sphere and sphere–plane geometry," *Phys. Rev. Lett.*, vol. 109, pp. 160403-1–160403-5, 2012.

[11] A. Canaguier-Durand, G.-L. Ingold, M.-T. Jaekel, A. Lambrecht, P. A. Maia Neto, and S. Reynaud, "Classical Casimir interaction in the plane–sphere geometry," *Phys. Rev. A*, vol. 85, pp. 052501-1–052501-10, 2012.

[12] B. B. Machta, S. L. Veatch, and J. P. Sethna, "Critical Casimir forces in cellular membranes," *Phys. Rev. Lett.*, vol. 109, pp. 138101-1–138101-5, 2012.

[13] D. S. Ether, I. S. Nogueira, and H. M. Nussenzveig, "Probing the Casimir force with optical tweezers," *Europhys. Lett.*, vol. 112, pp. 44001-1–44001-6, 2015.

[14] S. Pei *et al*., "Direct measurement of the van der Waals force between a pair of microspheres based on photonic force microscopy," *Opt. Eng.*, vol. 60, pp. 084101-1–084101-10, 2021.

[15] B. Liu, B. Poolman, and A. J. Boersma, "Ionic strength sensing in living cells," *ACS Chem. Biol.*, vol. 12, pp. 2510–2514, 2017.

[16] A. J. Boersma, I. S. Zuhorn, and B. Poolman, "A sensor for quantification of macromolecular crowding in living cells," *Nat. Methods*, vol. 12, pp. 227–229, 2015.

[17] M. Boström and B. E. Sernelius, "Thermal effects on the Casimir force in the 0.1–5 μm range," *Phys. Rev. Lett.*, vol. 84, pp. 4757–4760, 2000.

[18] M. Moazzami-Gudarzi, G. Trefalt, I. Szilagyi, P. Maroni, and M. Borkovec, "Nanometer-ranged attraction induced by multivalent ions between similar and dissimilar surfaces probed using an atomic force microscope (AFM)," *Phys. Chem. Chem. Phys.*, vol. 18, pp. 8739–8751, 2016.

[19] V. Valmacco, G. Trefalt, P. Maroni, and M. Borkovec, "Forces between silica particles in the presence of multivalent cations," *J. Colloid Interface Sci.*, vol. 472, pp. 108–115, 2016.

[20] H.-J. Butt, "Measuring electrostatic, van der Waals, and hydration forces in electrolyte solutions with an atomic force microscope," *Biophys. J.*, vol. 60, pp. 1438–1444, 1991.

Emergent Information Formation in Prebiotic Protocell Clusters: A Computational Mechanics Framework of ε-Machines and Attractor Memory

Michael Massoth
Department of Computer Science, Hochschule Darmstadt (h_da)
University of Applied Sciences Darmstadt, member of European University of Technology (EUt+)
Darmstadt, Germany
e-mail: michael.massoth@h-da.de

***Abstract*- Casimir–Lifshitz forces generate an unavoidable, long-range attraction between protocells under prebiotically realistic conditions. This interaction stabilizes mesoscale clusters such as tetrahedra, octahedra, and 13-cell icosahedra. These highly symmetric assemblies act as persistent macrostates whose transitions remain reproducible despite microscopic noise. A physics-guided coarse-graining yields a well-defined mesodynamics that can be represented as an ε-machine: a small deterministic automaton whose causal states correspond to cluster attractors and whose transitions encode ordered reconfiguration pathways. The theory of Rosas et al. ("Software in the natural world") shows that such systems can become informationally, causally, and computationally closed, thereby forming an autonomous proto-software layer. In this framework, prebiotic information does not arise from polymers but from attractor-based memory and structured transition dynamics in a purely physical cluster process.**

Keywords—Casimir–Lifshitz forces; Protocell clustering; Computational mechanics; ε-machines; Attractor-based memory; Prebiotic information; Mesoscale self-organization.

I. INTRODUCTION

This is the second of seven papers in the series: "A Constructivist Proto-Bio-Information Theory: A Physically Grounded Nano-Systems Architecture for Prebiotic Emergence, Information, Proto-Semantic Function, and Sustainability of Protocell Aggregation and Cluster Formation".

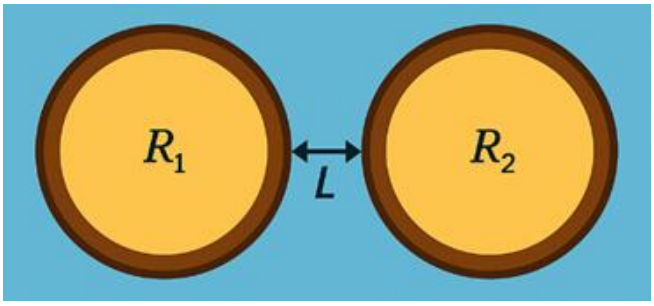


Figure 1. *Schematic of two prebiotic protocells with radii R_1 and R_2 and minimal surface-to-surface separation L in saline water (primordial soup).*

In Figure 1, a schematic representation of two spherical protocells with radii R_1 and R_2 is shown, separated by a minimal surface-to-surface distance L in saline aqueous solution (primordial soup). This sphere–sphere geometry defines the fundamental configuration used to model Casimir–Lifshitz interactions between protocell membranes under prebiotic conditions.

Casimir–Lifshitz forces arise from quantum and thermal fluctuations of the electromagnetic field between material interfaces, as first described by Hendrik Casimir [3] and generalized by Evgeny Lifshitz [4]. Unlike chemical bonds or electrostatic interactions, they do not rely on charges, receptors, or molecular specificity, but on the dielectric properties of the interacting materials and medium. In saline, thermally active environments, these forces act over nanometre-to-submicrometre distances and decay algebraically rather than exponentially. They therefore provide a physically unavoidable, environment-robust interaction mechanism precisely in the mesoscale regime where classical colloidal forces fail.

Figure 2. Thermal and quantum field fluctuations (left) generate Casimir-Lifshitz forces (right), which stabilize neighboring protocell membranes at nanoscale distances (2–100 nm) as clusters.

In Figure 2, quantum and thermal electromagnetic field fluctuations (left) give rise to Casimir–Lifshitz forces (right) between neighboring protocell membranes. These fluctuation-induced interactions act over nanometre-to-submicrometre separations (≈ 2–100 nm) in saline environments and provide a robust, non-chemical mechanism for stabilizing mesoscale protocell clusters where classical colloidal forces fail.

Massoth [1] showed that attractive Casimir–Lifshitz forces represent a universal and unavoidable aggregation mechanism under prebiotically realistic conditions. For protocell radii of R = 200–1000 nm, separations L = 5–100 nm, salt concentrations of 50–200 mMol, and temperatures of 20–90

°C, the classical DLVO electrostatic repulsion is almost fully suppressed. This is due to the short Debye length (< 2 nm) in saline environments. Van der Waals interactions act only at 0.3–3 nm and therefore cannot stabilize larger mesoscale assemblies. In contrast, Casimir – Lifshitz forces (Figure 3) retain an algebraic range of approximately ~$1/L^2$ and remain effective even in highly ionic and thermally agitated media. The strong dielectric contrast between protocell membranes ($\varepsilon \approx 2-8$) and water ($\varepsilon \approx 75$) produces a robust, attractive fluctuation-induced force that persists throughout the prebiotic parameter space.

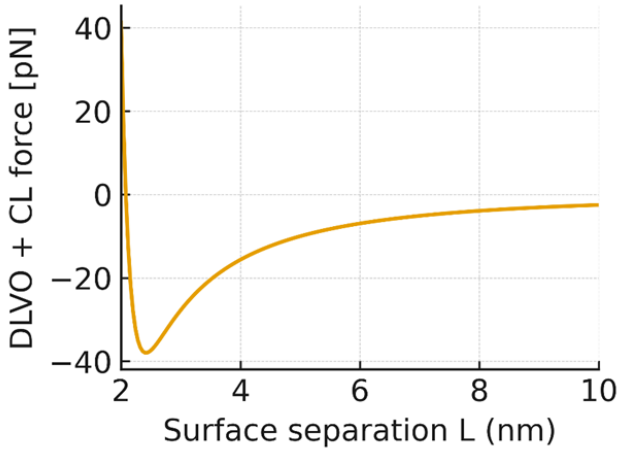


Figure 3. Illustrates the resulting net interaction force in the range of *2–10 nm*. The algebraic Casimir–Lifshitz contribution creates a residual, non-DLVO attraction that remains detectable experimentally.

The corresponding potential wells $U_{CL}(L)$ are sufficiently deep to stabilize dimers, trimers, and higher-order clusters. Larger protocells couple more strongly because of their increased effective curvature radius R_{eff}. Membrane types with higher polarizability—such as PMBCs—generate deeper potential wells than simple fatty-acid vesicles and should therefore cluster more readily. Expected mesoscale contact times fall in the minute range, and a temperature window of 40–80 °C maximizes stability.

These results indicate that attractive Casimir–Lifshitz forces provided a purely physical mode of early cooperation long before genetic or metabolic mechanisms existed. Protocell clusters may thus have acted as proto-ecological units from which functional and information-bearing structures later emerged.

Structure of the paper:

In Section 2, we derive Casimir–Lifshitz interaction scaling for spherical protocells and quantify cluster stabilization energies across “magic-number” geometries.

In Section 3, we position this mechanism within prior work on prebiotic self-organization, semantic information, and computational mechanics.

In Section 4, we introduce a coarse-grained description $Z_t=f(X_t)$ and show how Casimir-stabilized cluster attractors define recurrent macrostates with enhanced predictability.

In Section 5, we apply the Rosas framework [2] to establish informational, causal, and computational closure and identify the corresponding ε\varepsilonε-machine structure.

In Section 6, we formalize the micro-to-macro reduction from X_t and connect dimensionality reduction to conditional-entropy compression.

In Section 7, we analyze attractor-based memory and retrieval dynamics in protocell clusters and relate return paths to robust mesoscale storage.

In Section 8, we construct the explicit ε-machine representation of cluster-state dynamics and its transition structure.

In Section 9, we synthesize these results into a physically grounded account of prebiotic information prior to genes or polymer-based coding.

In Section 10, we unify the closure concepts and ε-machines into a single multilevel framework for prebiotic emergence.

In Section 11, we summarize implications and outline testable predictions and future theoretical and experimental directions.

II. Mesoscale Protocell Clusters Stabilized by Casimir–Lifshitz Forces

For protocells with radii $R \approx 200–1000\ nm$, Casimir–Lifshitz attraction at separations of a few nanometers drives robust self-assembly into densely packed mesoscale clusters. For two spherical protocells (a dimer) with radii R_1 and R_2 and minimal surface separation L, the effective potential is: $U_{CL}(L) \approx -(A_{eff}/6) * (R_{eff}/L)$, with A_{eff} the effective Hamaker constant of the membrane-water-membrane system and $R_{eff} = (R_1 * R_2)/(R_1 + R_2)$. For equal-sized protocells ($R_1 = R_2 = R$) this reduces to: $U_{CL}(L) \approx -(A_{eff}/12) * (R/L)$.

The natural scale is given by the dimensionless coupling strength: $\Lambda(R,L) = |U_{CL}(L)| / k_BT = (A_{eff} * R) / (12\ k_BT * L)$. For realistic values ($A_{eff} \approx 5\times10^{-21}\ J$, $R \approx 500\ nm$, $L = 3–30\ nm$), one obtains $\Lambda \approx 1–10$. The Casimir–Lifshitz attraction therefore exceeds thermal energy in the relevant distance range and can effectively suppress Brownian fluctuations.

The total energy of an N-particle cluster in near-contact follows from the number of particle–particle bonds $N_{bonds}(N)$: $E_N(L) \approx -N_{bonds}(N) * |U_{CL}(L)| = -N_{bonds}(N) * (A_{eff}/6) * (R_{eff}/L) = -N_{bonds}(N) * \Lambda(R,L) * k_BT$.

TABLE I. TOTAL CASIMIR–LIFSHITZ INTERACTION ENERGIES

N	Protocell Structure	N_{bonds}	$E_N/k_BT \approx$
2	Dimer	1	−5.1
3	Triangularer Trimer	3	−15.2
4	Tetrahedron	6	−30.5
6	Oktahedron	12	−61.0
7	Pentagonal Bipyramid	15	−76.2
13	Icosahedral 13-Cluster	42	−213.4

Table I, shows the total Casimir–Lifshitz interaction energies E_N/k_BT for representative N-protocell cluster geometries,

illustrating how the number of pairwise bonds Nbonds drives mesoscale stabilization.

For isotropic, short-range attraction, the system favors "magic-number" clusters that maximize contacts: Dimer (N = 2, N_{bonds} = 1), Triangular Trimer (N = 3, 3 bonds), Tetrahedron (N = 4, 6 bonds), Octahedron (N = 6, 12 bonds), Pentagonal bipyramid (N = 7, 15 bonds). On the next mesoscale level, a 13-protocell icosahedral cluster (12 around 1 center) forms with $N_{bonds} = 42$. This structure is exceptionally symmetric and tightly packed.

Using $R = 500\ nm$, $L = 10\ nm$ and $A_{eff} = 5\times10^{-21}\ Joule$, cluster binding energies range from approximately $-15.2\ k_BT$ (triangular trimer) to more than $-200\ k_BT$ (icosahedral 13-cluster). These values show that Casimir–Lifshitz attraction is strong enough not only to stabilize dimers and oligomers but to enforce the formation of robust mesoscale clusters that can serve as seeds for further hierarchical structure formation. Experimental studies of primitive vesicle compartments further support the plausibility of such prebiotic mesoscale assemblies [13].

III. Related work

Classical approaches to prebiotic organization emphasize chemical self-organization or polymer-based information, while physical interactions between early compartments are often treated as transient aggregation effects. In parallel, Computational Mechanics (J. P. Crutchfield & C. R. Shalizi, *"Computational Mechanics: Pattern and Prediction, Structure and Simplicity"*) [8][11] formalizes syntactic information via ε-machines, yet leaves the physical origin of causal states unspecified.

Recent theoretical approaches to semantic information—most notably the work of Kolchinsky and Wolpert ("*Semantic Information, Autonomous Agency and Non-Equilibrium Statistical Physics*", 2018) [16] and of Ruzzante et al. ("*Synthetic Cells Extract Semantic Information From Their Environment*", 2023) [17][18] - demonstrate that information is not intrinsically tied to genes or replication. However, they leave open which concrete physical structures carry semantically relevant states. The models operate predominantly at an abstract agent–environment level or focus on individual synthetic cells.

This paper advances these frameworks by identifying Casimir–Lifshitz–stabilized protocell clusters as a concrete physical substrate for syntactic information. Highly symmetric mesoscale clusters form robust attractors whose transitions remain reproducible despite microscopic noise. A physics-guided coarse-graining reconstructs these attractors as ε-machine states, yielding informational closure in the sense of Rosas et al., "Software in the Natural World" (2024) [2]. Unlike abstract models, information here arises directly from physical self-organization, without genes, polymers, or symbolic encoding.

IV. Emergence of Mesoscale Information in Casimir–Lifshitz–Stabilized Protocell Clusters

Early prebiotic compartments lacked genes, enzymes, and regulatory networks. Their organization depended entirely on physical processes. Yet some structures must already have been stable over time, recurrent, and minimally coherent. A natural candidate for such early organization is the set of mesoscale cluster states formed by aggregating protocells.

We now apply the computational-mechanics framework of ε-machines and attractor memory [2] to formalize how Casimir–Lifshitz–driven protocell coupling gives rise to emergent informational states.

Casimir–Lifshitz forces stabilize a small number of highly symmetric, recurrent configurations on the mesoscale. Typical examples include dimers, trimers, tetrahedra, octahedra, and 13-protocell icosahedra. These assemblies act as discrete macro-states Z_t that reappear reliably despite microscopic noise and persist over mesoscale timescales. Contact times are typically on the order of minutes, making these geometries robust attractors in configuration space.

These macro-states are not static objects. Clusters can rearrange and transition between configurations, yet remain confined to a small, stable state space. This yields a dynamic but structured mesodynamics, in which only a limited set of configurations dominates.

This situation naturally suggests a coarse-graining in which the high-dimensional microdynamics Xt—positions, forces, membrane fluctuations, and hydrodynamic interactions—collapse onto a finite set of functional states: $Z_t = f(X_t)$. The identification of such predictive macro-states directly corresponds to the construction of causal states in the sense of Crutchfield's statistical complexity framework [8]. A key consequence of this reduction is a strong increase in predictability. Even though the micro-dynamics are noisy, the meso-dynamics follow clear patterns. The system's future is far better predicted from Z_t than from the full microstate. Formally, this appears as a marked decrease in conditional entropy: $H(Z_{t+1} \mid Z_t) \ll H(Z_{t+1} \mid X_t)$, which implies informational closure. The macro-level acquires its own lawful dynamics, largely independent of microscopic details.

Within Computational Mechanics, this organized behavior is represented as an ε-machine. Its causal states E_i group all histories that yield identical future distributions. Each causal state corresponds to a meaningful cluster geometry or transition motif. Transitions $E_i \rightarrow E_j$ represent allowed transformations, such as the relaxation of a distorted octahedron or the completion of a metastable 11-cluster into a 13-icosahedral configuration.

A defining feature of this dynamics is its attractor-based memory. Cluster geometries act as robust storage states that the system reliably returns to after perturbations. These deterministic return paths match the "memory retrieval dynamics" described by Rosas et al.: the spatial configuration itself encodes past trajectories and functions as a physical information unit without genes or polymers.

This yields a physically grounded and mathematically consistent model of prebiotic information formation. Protocell clusters exhibit a software-like meso-dynamics characterized by discrete macro-states, stable attractors, and an ε-machine structure. Prebiotic information arises not from symbolic coding but from mesoscale, dissipative, attractor-based structural persistence—a necessary precursor to later molecular complexity.

V. Application of the Computational Mechanics Framework: Emergence, Closure, and ε-Machines in Protocell Clusters

The computational mechanics framework of Rosas et al. [2] provides a precise method for quantifying emergence in natural systems. It distinguishes three levels of autonomous organization—informational, causal, and computational closure—which capture how a macro-level becomes independent of microscopic details. This structure is well suited for prebiotic protocell clusters, whose Casimir–Lifshitz–stabilized attractors generate a coherent, software-like meso-dynamics.

Informational closure occurs when the future of a macro-state Z_t is better predicted from Z_t itself than from the full micro-dynamics X_t. In protocell clusters, transitions between stable geometries—such as tetrahedron → octahedron or an 11-cluster → 13-icosahedron—are largely insensitive to local fluctuations. Formally: $H(Z_{t+1} \mid Z_t) \ll H(Z_{t+1} \mid X_t)$.

Even under thermal noise, the system reliably returns to the same attractors. This marks the first level of emergent autonomy.

Causal closure requires that the macro-level causes its own future. This notion aligns with the theory of causal emergence, which shows that macro-descriptions can exert greater effective causal influence than their underlying microstates [9]. The global cluster structure—geometry, contact graph, coupling strengths—governs the dynamics, not the precise microstate. This is reflected in the equivalence of micro- and macro-interventions:

$P(Z_{t+1} \mid do(Z_t)) = P(Z_{t+1} \mid do(X_t \in pre(Z_t)))$

where $pre(Z_t)$ is the set of microstates compatible with the same cluster pattern. Thus, the macroprocess exhibits genuine causal power, not merely statistical regularity.

Computational closure denotes the ability of the macro-level to form its own internal representation. In Computational Mechanics, this is the *ε-machine*: a minimal deterministic automaton whose causal states E_i collect all pasts with identical future distributions. For protocell clusters, these E_i correspond to stable and metastable geometries—dimer, trimer, tetrahedron, octahedron, 13-icosahedron—and their transitions. The transition structure $P(E_i \rightarrow E_j)$ defines a small finite-state graph capturing the cluster's complete dynamic rules.

A core result of the Rosas framework [2] is that informational closure automatically implies computational closure. Once the future can be reliably inferred from Z_t, a consistent ε-machine emerges. For the prebiotic world, this means that the physical attractors of protocell clusters form an emergent rule-level analogous to an early "software layer." Clusters store history in their geometry, follow reproducible transition paths, and possess a robust dynamical memory without genes or polymers.

Overall, Casimir-stabilized protocell clusters generate not only stable mesoscale structures, but also an autonomous macro-level with its own causality and internal representation. Emergence, closure, and ε-machines together offer a coherent theoretical basis for a physically grounded model of prebiotic information—long before genetic coding, enzymatic control, or molecular replication existed.

VI. Coarse-graining Protocell Dynamics: From Microstates X_T to Macrostates Z_T

The microphysical dynamics of prebiotic protocells are high-dimensional. Positions, radii, membrane fluctuations, Casimir–Lifshitz forces, hydrodynamic couplings, and thermal noise create a deterministic–stochastic system with a large phase space. A microstate X_t includes spatial coordinates, force fields, local membrane structure, and solvent fluctuations - a level of detail that obscures emergent order.

A biologically meaningful coarse-graining reduces this complexity by grouping microstates into functional macro-states Z_t. These macro-states describe protocells not individually, but as reconfiguring clusters with typical sizes $N=\{2,3,4,6,7,13,\ldots\}$, characteristic geometries (tetrahedral, octahedral, icosahedral), coupling strengths $\Lambda(R,L)$, contact persistence's, and stable neighborhood topology. Such variables are insensitive to microscopic fluctuations and remain coherent over mesoscale times.

Formally, the reduction is given by $Z_t = f(X_t)$, where distinct microstates X_t and $X_{t'}$ map to the same macro-state and yield nearly identical future behavior. The macro-states thus take on a *semantic* role: they are the dynamically relevant units of cluster organization.

This coarse-graining reduces dimensionality dramatically. The emergence of low-dimensional macro-states from high-dimensional microscopic physics is a hallmark of dissipative structure formation in non-equilibrium systems [12]. Although X_t is continuous and high-dimensional, the Casimir–Lifshitz energy landscape supports only a small set of stable Z_t. Dimers, tetrahedra, and 13-icosahedra act as attractors of the state space. An attractor A satisfies: $P(Z_{t+1} \in A \mid Z_t \notin A) \rightarrow$ high, and $P(Z_{t+1} \in A \mid Z_t \in A) \rightarrow$ very high.

These attractors correspond to real physical configurations. They minimize energy, absorb fluctuations, and restrict the macro-state space to a small, recurring set with robust transitions.

From an information-theoretic perspective, macroscopic dynamics occupy a reduced, lower-dimensional state space and thus constitute a compressed representation of the underlying microscopic fluctuations, exhibiting substantially higher predictability than the far more complex micro-dynamics. This follows from the reduced conditional entropy:

$H(Z_{t+1} \mid Z_t) < H(X_{t+1} \mid X_t)$,

indicating a structured, attractor-based dynamics similar to neural systems and other dissipative self-organizing media.

The result is a set of robust, repeatable, and causally effective macro-states - the foundation of a proto-informational dynamics in which structure and function are physically coupled without genes or molecular codes. The macro-level of protocell clusters therefore represents an early form of physical organization: a prebiological "software layer" emerging from energetic and geometric self-organization, preparing the ground for later evolutionary complexity.

VII. Memory Retrieval and Attractor Dynamics in Protocell Clusters

A key result of Rosas et al. [2] is the presence of memory in systems with multiple stable attractors. Case Study F shows that noise-driven systems fall into energetic minima and form dynamic memory: past states bias transitions, positions in attractor space define functional states, and return paths after perturbations are reproducible. This "memory-retrieval" behavior appears when the macro-level is resilient to microscopic noise.

Casimir–Lifshitz-stabilized protocell clusters show the same pattern. Their fluctuation forces create a discrete set of energetic attractors—dimers, tetrahedra, octahedra, and the deep 13-icosahedral minimum (below $-200\ k_BT$). These geometries act as local minima and remain highly likely under prebiotic conditions. This behavior reflects classical principles of self-organization and attractor landscapes in far-from-equilibrium systems, as described by Kauffman [10].

The transition from a microstate X_t into an attractor can be described as motion within a basin-of-attraction structure. Let A_i denote the set of microstates leading to attractor. Under prebiotic parameters: $P(Z_{t+1} = i \mid X_t \in A_i) \approx 1$.

Small deformations from hydrodynamic or thermal noise drive the cluster along characteristic return paths back into the same attractor. Such energy-minimizing return dynamics closely parallel attractor-based memory retrieval in Hopfield networks [15]. These relaxation trajectories correspond directly to the "retrieval trajectories" of Rosas et al. [2].

Memory dynamics have two components. First, the cluster geometry acts as a physical storage medium: a distorted octahedron relaxes back to its native form, minimizing the conditional entropy along the return path,

$H(Z_{t+1} \mid Z_t \in \text{Attractor}_i) \rightarrow \text{minimal}$.

Repeated transitions also reinforce future behavior. Frequent returns to the same attractor reshape the ε-machine's transition matrix and increase recurrence. Attractors become stable patterns that store "proto-historical information" across many fluctuation cycles.

This memory needs no replication, templates, or polymers. The information is not symbolic; it is encoded in the cluster's geometry and topology. Because transitions are constrained and macrostates persist, the ε-machine forms a proto-software layer: attractor states act as nodes, retrieval paths as directed edges.

The macro-level behaves like a physical memory system. The cluster "knows" its attractor and reliably returns after perturbations. Past shapes and relaxation paths bias future transitions. This mutual dependence is characteristic of attractor-based memory and the proto-software systems described by Rosas et al. [2].

Thus, Casimir–Lifshitz-stabilized protocell clusters show a definable memory mechanism. Attractors serve as storage states, return trajectories as retrieval functions. Prebiotic matter therefore possessed a purely physical principle of information and memory — a software-like layer without genes or molecules, forming a foundation for later biological evolution.

VIII. Protocell Clusters as an ε-Machine

The emergent dynamics of prebiotic protocell clusters can be described not only qualitatively through attractors but also quantitatively using Computational Mechanics. This framework models the mesoscale cluster states and their transitions as an ε-machine — a minimal deterministic automaton that captures the full causal structure of a process. Casimir–Lifshitz-stabilized protocells provide ideal conditions: they form a small, discrete set of long-lived macro-states whose transitions remain largely independent of microscopic variations. This yields a clear, formally defined notion of "proto-software" arising purely from physical self-organization.

The system is first represented as a sequence of macro-states Z_t. These include stable cluster geometries — dimer, trimer, tetrahedron, octahedron, 13-icosahedron — and metastable transitions such as partial recombinations or fragmentations:

$Z_0, Z_1, Z_2, \ldots, Z_t \in$ {dimer, trimer, tetrahedron,
octahedron, …, decay, recombination, …}

Each macro-state summarizes many microconfigurations X_t that are functionally equivalent. Although the micro-dynamics is high-dimensional, the future is far more predictable from Z_t than from X_t — a key signature of macro-level stability.

Next, causal states E_i of the ε-machine are constructed. This definition follows the formalism of Shalizi & Crutchfield, who defined ε-machines as the minimal sufficient representation of a process [11]. Two past trajectories $z(t|-\infty{:}0)$ and $z'(t|-\infty{:}0)$ are causally equivalent when they generate identical future distributions: $P(Z_{t+1} \mid z(t|-\infty{:}0)) = P(Z_{t+1} \mid z'(t|-\infty{:}0))$.

Transitions between causal states are given by $T_{ij} = P(E_j \mid E_i)$.

Because cluster attractors are energetically stable, the automaton contains only few states and exhibits characteristic attractor cycles — a hallmark of emergent software-like dynamics in the sense of Rosas et al.

Testing closure properties confirms that the ε-machine defines an autonomous macro-level:

Informational closure: $H(Z_{t+1} \mid Z_t) \ll H(Z_{t+1} \mid X_t)$, showing that the macrostate predicts the future better than the microstate.

Causal closure: $P(Z_{t+1} \mid do(Z_t)) = P(Z_{t+1} \mid do(X_t \in pre(Z_t)))$, demonstrating that macro-interventions and equivalent micro-interventions yield the same transitions.

Computational closure: $\varepsilon(f(X_t)) = \pi(\varepsilon(X_t))$, where f is the macro-mapping and π the projection onto causal classes. This shows that coarse-graining produces the same causal architecture as full microanalysis.

Thus, the cluster dynamics establishes an autonomous computational layer — a proto-software system not based on symbolic coding but on stable attractors and reproducible transitions. In this strict formal sense, the ε-machine represents prebiotic information: a closed set of causal rules arising entirely from physical self-organization in mesoscale protocell ensembles, forming the foundation of later biological information processing.

IX. Physically Grounded Origin of Prebiotic Information

The origin of information is a central question in origins-of-life research. Compartment-based models of early cellular organization emphasize that structural stability itself may precede genetic information [14]. Classical models tie information to polymers, such as RNA or DNA. Yet before such molecules existed, systems still needed structures that were stable, recurrent, and capable of minimal functional persistence. The physical dynamics of protocell clusters offer a precise starting point.

Information emerges whenever a system forms stable, distinguishable, and reproducible macro-states whose transitions follow their own dynamics. A macro-state Z_t is informative when it predicts the future better than the underlying microstate X_t: $H(Z_{t+1} \mid Z_t) < H(Z_{t+1} \mid X_t)$.

Casimir–Lifshitz-stabilized protocell clusters satisfy this condition strongly. Their attractors — tetrahedra, octahedra, 13-icosahedra — are stable over many k_BT, recur frequently, and are largely insensitive to microscopic fluctuations. Transitions follow ordered paths: tetrahedra often relax into octahedra, incomplete 11-clusters complete into 13-icosahedra, and decay trajectories follow consistent routes.

These attractors also store memory. A slightly deformed octahedron relaxes deterministically back into the same attractor. Formally, this appears as a contraction in attractor space: $\|Z_{t+1} - A\| < \|Z_t - A\|$.

Thus, attractors function as both stable and memory-bearing states — an *attractor-based memory*, in the sense of Rosas et al. [2]. The ε-machine makes this structure explicit. The macro-states Z_t form a finite set of causal states E_i, and the transition probabilities T_{ij} define an internal, proto-symbolic rule structure. No polymers are required. Instead, information is stored in the geometry, topology, and allowed transitions of the cluster.

This yields a prebiotic, non-gentic form of information arising purely from physical self-organization. Protocell clusters behave as physical information carriers whose attractor-based transition dynamics generate an early "software layer" — an information-processing mechanism without genes or enzymes, yet with clear functional structure and evolutionary potential.

X. ε-Machines and Closure Concepts as a Unified Framework for Prebiotic Emergence

The combination of Casimir–Lifshitz stabilization, mesoscale cluster attractors, and the theory of Computational Mechanics provides a coherent model for the emergence of prebiotic information. The framework links micro-, meso-, and macro-levels and shows how protocell clusters can develop a software-like dynamics generated purely by physical self-organization. Central to this model are the three closure concepts of Rosas et al.—informational, causal, and computational closure—together with the ε-machine formalism. The resulting hierarchy is:

TABLE II. MULTI-LEVEL HIERARCHY OF PREBIOTIC PROTOCELL DYNAMICS

Level	Description	Formalization
Micro	Forces, positions, thermal noise	Stochastic dynamics X_t
Meso	Cluster attractors	Macroprocess: $Z_t = f(X_t)$
Macro	Information dynamics	ε-machine E_t, transitions T_{ij}
Software layer	Attractor-coded memory, pattern retrieval	Causal / computational closure

Table II shows the hierarchical organization of protocell dynamics from microphysical interactions through mesoscale cluster attractors to an emergent macro-level ε-machine and a proto-informational software layer characterized by causal and computational closure.

On the micro level, Casimir–Lifshitz forces, hydrodynamic coupling, and thermal noise shape protocell motion. The dynamics X_t is high-dimensional and difficult to predict.

A biologically motivated coarse-graining maps many micro-configurations onto a small set of stable macro-states Z_t. These include dimers, trimers, tetrahedra, octahedra, and 13-icosahedra. Such cluster forms have fewer degrees of freedom, show strong energetic stability, and occupy well-defined regions of state space.

On the macro level, the dynamics becomes an ε-machine. Causal states E_i group past trajectories that yield identical future distributions: $P(Z_{t+1} \mid z(t|-\infty:0)) = P(Z_{t+1} \mid z'(t|-\infty:0))$. The transitions $T_{ij}=P(E_j|E_i)$ describe the full process. Because prebiotic cluster landscapes possess only a few deep attractors, the resulting ε-machine is small, near-deterministic, and exhibits characteristic attractor cycles and retrieval trajectories. ε-machines thus become natural tools for describing proto-informational dynamics: stable cluster geometries act as "symbol-like" states, and transitions encode the rules of an early prebiotic automaton. All three closure types integrate naturally into this layered model:

Informational closure: The future is predicted better from Z_t than from X_t: $H(Z_{t+1} \mid Z_t) << H(Z_{t+1} \mid X_t)$.

Causal closure: Macro-level interventions produce the same transitions as micro-interventions that establish the same macro-state: $P(Z_{t+1} \mid do(Z_t)) = P(Z_{t+1} \mid do(X_t \in pre(Z_t)))$.

Computational closure: The ε-machine derived from the macroprocess has the same causal structure as the one obtained from the full micro-dynamics: $\varepsilon(f(X_t)) = \pi(\varepsilon(X_t))$.

Together, these elements form a unified four-level model of emergent prebiotic organization: physical forces act on the micro level; stable attractors arise on the meso level; these attractors form an ε-machine on the macro level; and attractor-coded memory with pattern retrieval yields a proto-informational "software layer" long before genes or enzymes existed.

XI. CONCLUSION AND FUTURE WORK

Our analysis shows that Casimir–Lifshitz forces drive protocells into a small set of stable mesoscale attractors—dimers, tetrahedra, octahedra, and 13-icosahedra. These attractors persist despite microscopic noise and define a structured macro-dynamic. When coarse-grained, this dynamic becomes an ε-machine that is informationally and computationally closed in the sense of Rosas et al. [2]. The cluster geometries act as symbol-like states, and their transitions form a proto-algorithmic rule set.

This establishes a physically grounded route to the emergence of prebiotic information. The autonomous macro-dynamics of protocell clusters constitutes an early "software layer" encoded not in polymers but in attractor structure and pattern retrieval. In this view, information arises from the geometry, stability, and transition rules of the cluster state space.

Prebiotic protocell aggregates therefore offered a purely physical form of memory, function, and predictive structure—providing a substrate for later chemical and genetic evolution. Long before RNA or enzymes existed, mesoscale cluster dynamics already supported many of the informational properties required for biological complexity.

REFERENCES

[1] M. Massoth, "Attractive Casimir-Lifshitz Forces as a Universal Driver of Prebiotic Protocell Aggregation and Cluster Formation," *BIOTECHNO*, 2026.

[2] F. E. Rosas, *et al.*, "Software in the natural world: A computational approach to hierarchical emergence," *arXiv* preprint, 2024, doi:10.48550/arXiv.2402.09090.

[3] H. B. G. Casimir, "On the attraction between two perfectly conducting plates," *Proc. K. Ned. Akad. Wet.*, vol. 51, pp. 793–795, 1948.

[4] E. M. Lifshitz, "The theory of molecular attractive forces between solids," *Sov. Phys. JETP*, vol. 2, pp. 73–83, 1956.

[5] J. N. Israelachvili, *Intermolecular and Surface Forces*, 3rd ed., Academic Press, London, 2011.

[6] P. C. W. Davies, "Does quantum mechanics play a non-trivial role in life?" *Biosystems*, vol. 78, pp. 69–79, 2004, doi:10.1016/j.biosystems.2004.07.001.

[7] I. Gözen, *et al.*, "Protocells: Milestones and recent advances," *Small*, vol. 18, Art. no. 2106624, 2022, doi:10.1002/smll.202106624.

[8] J. P. Crutchfield, "Inferring statistical complexity," *Phys. Rev. Lett.*, vol. 63, no. 2, pp. 105–108, 1989, doi:10.1103/PhysRevLett.63.105.

[9] E. P. Hoel, L. Albantakis, and G. Tononi, "Quantifying causal emergence," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 110, no. 49, pp. 19790–19795, 2013, doi:10.1073/pnas.1314922110.

[10] S. A. Kauffman, *The Origins of Order: Self-Organization and Selection in Evolution*, Oxford Univ. Press, Oxford, 1993.

[11] C. R. Shalizi and J. P. Crutchfield, "Computational mechanics: Pattern and prediction, structure and simplicity," *J. Stat. Phys.*, vol. 104, nos. 3–4, pp. 817–879, 2001, doi:10.1023/A:1010388907793.

[12] G. Nicolis and I. Prigogine, *Self-Organization in Nonequilibrium Systems*, Wiley, New York, 1977.

[13] M. M. Hanczyc, S. M. Fujikawa, and J. W. Szostak, "Experimental models of primitive cellular compartments," *Science*, vol. 302, no. 5645, pp. 618–622, 2003, doi:10.1126/science.1089904.

[14] D. Deamer, *Assembling Life: How Can Life Begin on Earth and Other Habitable Planets?*, Oxford Univ. Press, Oxford, 2017.

[15] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 79, no. 8, pp. 2554–2558, 1982, doi:10.1073/pnas.79.8.2554.

[16] A. Kolchinsky and D. H. Wolpert, "Semantic Information, Autonomous Agency and Non-Equilibrium Statistical Physics," Entropy, vol. 20, no. 10, p. 793, 2018, https://doi.org/10.1098/rsfs.2018.0041

[17] B. Ruzzante, L. Del Moro, M. Magarini, and P. Stano, "Synthetic Cells Extract Semantic Information From Their Environment," IEEE Transactions on Molecular, Biological and Multi-Scale Communications, vol. 9, no. 1, pp. 23–36, 2023, doi:10.1109/TMBMC.2023.3244399.

[18] S. Bartlett et al., "Physics of Life: Exploring Information as a Distinctive Feature of Living Systems," PRX Life, vol. 3, 037003, 2025, doi:10.1103/rsx4-8x5f

From Physical Difference to Meaning: A Constructor-Theoretic Framework for Prebiotic Information in Casimir-Lifshitz-Coupled Protocell Clusters

Michael Massoth
Department of Computer Science, Hochschule Darmstadt (h_da)
University of Applied Sciences Darmstadt, member of European University of Technology (EUt+)
Darmstadt, Germany
e-mail: michael.massoth@h-da.de

***Abstract*- This paper develops a physical framework for the prebiotic emergence of information and meaning. Building on Constructor Theory, we define information as a reproducible physical difference and meaning as a difference with stable functional consequences. Casimir–Lifshitz–coupled protocell clusters serve as a minimal model that exhibits reproducible attractors, ordered transitions, and autonomous task structures. We show that such clusters carry both informational states (e.g., distances, geometries, gradients) and meaningful states that regulate prebiotic tasks such as approach, exchange, or stabilization. This approach integrates physical mechanisms, computational mechanics, and early proto-semantic functions into a coherent account of information formation before biology.**

Keywords-constructor theory; prebiotic information; Casimir–Lifshitz forces; protocell clusters; ε-machines and attractor memory; proto-semantic function.

I. INTRODUCTION

This is the third of seven papers in the series: "A Constructivist Proto-Bio-Information Theory: A Physically Grounded Nano-Systems Architecture for Prebiotic Emergence, Information, Proto-Semantic Function, and Sustainability of Protocell Aggregation and Cluster Formation".

Massoth [1] shows that Casimir–Lifshitz forces generate robust attraction and stable protocell clusters across 5–200 nm under realistic prebiotic conditions. These physically enforced mesoscale assemblies form the structural substrate in which this paper grounds the emergence of information and meaning.

Massoth [2] demonstrates that such clusters form reproducible mesoscale attractors with autonomous ε-machine dynamics. This attractor-based organization provides the computational and informational foundation that this paper develops into a physical account of proto-information and proto-semantics.

This work applies the Constructor Theory of Information proposed by Deutsch and Marletto [4] to a computational-mechanics framework of ε-machines and attractor-based memory developed by Rosas *et al.* [3], establishing a unified physical and informational architecture for protocell-cluster dynamics.

The structure of the paper is as follows: In Section II, we introduce the motivation, conceptual gap, and guiding research questions, framing prebiotic information and meaning within a constructor-theoretic perspective. Section III reviews related work on semantic information, protocells, and Constructor Theory, positioning our approach relative to existing biological and computational models.

Section IV outlines the core principles of Constructor Theory, with a focus on tasks, constructors, and the physical definition of information. In Section V, we formalize information as a reproducible physical difference and demonstrate how protocell cluster attractors satisfy constructor-theoretic information criteria.

Section VI develops the central concept of meaning as a functionally exploited difference, showing how specific protocell states control downstream task realizations. Section VII applies this framework to Casimir–Lifshitz–coupled protocell clusters, identifying concrete tasks and constructor-like behaviors.

Finally, Section VIII summarizes the conceptual contributions, answers the guiding research questions, relates the framework to Computational Mechanics, and discusses implications, limitations, and future research directions.

II. MOTIVATION, RESEARCH QUESTION AND RELEVANCE

A. Motivation: Information and Meaning in Prebiotic Systems

How prebiotic systems generated information, function, and eventually meaning before genes, enzymes, or metabolism emerged remains a central challenge in origins-of-life research. Classical models describe self-organization yet offer little insight into when a physical pattern qualifies as information or under which conditions differences become functionally consequential. The distinction between information formation and meaning typically appears only once biological codes exist.

Here we propose an alternative approach. Constructor Theory defines information through physically possible tasks: information is a reproducible physical difference. Meaning arises when such a difference produces stable functional consequences that influence processes such as persistence or reproduction. Prebiotic information thus becomes a physical question: Which tasks were allowed by early Earth

conditions, which substrates could support them, and when did differences begin to be used rather than merely produced?

B. Approach, Gap, and Guiding Research Questions

We apply Constructor Theory to prebiotic protocells that interact through Casimir–Lifshitz forces and electrolyte-mediated fields. In contrast to the accompanying Computational Mechanics paper, which analyzes ε-machines, we focus here on meaning and function. Our framework proceeds in three steps: core principles of Constructor Theory; information as a reproducible physical difference; and meaning as a difference functionally exploited within prebiotic task networks.

Current models of chemical self-organization rarely specify which informational tasks were physically feasible, when differences are merely structural patterns, or when they become functionally active. It also remains unclear at what point protocells qualify as constructors capable of executing repeatable tasks.

This motivates three guiding research questions:

RQ#1: Under which physical conditions can protocell clusters act as constructors that repeatedly perform tasks such as approach, exchange, or coupling?

RQ#2: How can pure informational differences be distinguished from meaningful states with distinct functional outcomes?

RQ#3: Which classes of tasks in Casimir-coupled clusters can support proto-semantic stability, and how does this relate to concepts such as informational and causal closure?

C. Relevance for Biology and Computer Science

Our approach is deliberately interdisciplinary. In biology, protocells appear not only as reaction compartments but as early functional units capable of performing tasks without genes or enzymes. In computer science, Constructor Theory broadens the notion of a program to networks of physical tasks, offering new perspectives on natural information processing.

III. RELATED WORK

Theories of semantic information define meaning via viability or functional relevance, yet often operate at the level of abstract agents or single synthetic cells. Experimental work such as Ruzzante et al., “Synthetic Cells Extract Semantic Information From Their Environment” (2023) [24] demonstrates semantic information processing, but does not address its prebiotic physical realization [23] [25]. Constructor Theory of Information from D. Deutsch and C. Marletto, “Constructor Theory of Information” [4] provides a physical definition of information as a reproducible difference, but has rarely been applied to origins-of-life systems.

This paper integrates these strands by applying Constructor Theory to Casimir–Lifshitz–coupled protocell clusters whose ε-machine organization was established in [2]. Information is defined as a reproducible physical difference; meaning arises when such differences systematically enable or disable tasks. Protocell clusters thus act as partial constructors, in which cluster states regulate prebiotic tasks such as approach, exchange, or stabilization. Meaning is thereby grounded in physical task structure rather than symbolic representation.

IV. CORE PRINCIPLES OF CONSTRUCTOR THEORY

Constructor Theory (CT) of information [4] describes physical processes in terms of tasks—transformations that are possible, impossible, or only conditionally realizable. Instead of trajectories, CT focuses on transitions that can, in principle, be reproduced indefinitely. This makes CT well suited for prebiotic protocells, where information and early meaning arise not through symbols but through physically achievable transformations.

A. Tasks and Constructors

A task $A=\{x_i \rightarrow y_i\}$ maps input attributes of a substrate to output attributes. These attributes may be geometric, chemical, or energetic, such as distances, field configurations, or cluster shapes. A constructor is a system that performs a task repeatedly and reliably without losing its ability to do so. Perfect stability is unnecessary; what matters is the theoretical possibility of arbitrarily increasing accuracy. Enzymes, technical protocols, and stable protocell clusters all function as constructors because they transform states reproducibly without being consumed. CT thus generalizes concepts across chemistry, biology, and computer science. In a prebiotic context, protocells can be viewed as natural constructors that translate environmental differences into stable internal or collective states.

B. Possible versus Impossible Tasks

CT expresses physical laws as constraints on which tasks are permitted. A task is possible if physics allows a constructor to perform it with arbitrarily high accuracy. It is impossible if it would violate fundamental principles such as energy conservation or the quantum no-cloning theorem. CT therefore studies the space of what can occur, not detailed time evolutions. This is valuable for protocell clusters: their microscopic dynamics are complex, but mesoscale processes—approach, exchange, stabilization—form well-defined task classes.

Casimir–Lifshitz–coupled protocells thus support a set of physically allowed tasks arising from interaction forces, geometry, and thermal fluctuations. CT provides a unified vocabulary without requiring a full microphysical derivation.

C. Constructor Theory of Information

In CT, a substrate carries information when its attributes are distinguishable and copiable through physically possible tasks. Distinguishability requires a task that maps two attributes to different, readable outputs while the constructor remains functional. Copyability requires a task that transfers a state onto a second, initially blank substrate.

Many physical systems—such as general quantum states—do not satisfy this copying requirement. Information in CT is therefore a reproducible physical difference, not a symbolic one. Protocell clusters exhibit such differences through stable distances, geometries, and field states. A classical bit is defined not by symbols "0/1" but by the existence of a task family capable of generating, copying, and transforming those differences.

V. Information as a Reproducible Physical Difference

A. From Symbolic to Physical Notions of Information

The Shannon framework describes information in terms of probabilities, encoding, and channel transmission, assuming senders, receivers, and symbolic structures—features absent in prebiotic environments. Constructor Theory replaces this symbolic view with a physical one: information is a reproducible difference in the world, independent of interpretation or statistics.

A difference between two states counts as information only if physically possible tasks exist that can generate, copy, transmit, or read it with arbitrarily increasing accuracy. Information is therefore a property of substrates and their space of possible transformations, not merely of descriptions. The key question shifts from "How many bits?" to "Which differences can be reliably realized in this world?"

This perspective aligns with results from [2]: mesoscale attractors in protocell clusters form reproducible macro-states with stable, recurrent transitions, robust to microscopic noise [10]. These attractors satisfy the CT criteria for informational states.

B. Reproducible Differences as Tasks

In Constructor Theory, reproducible differences can be described directly as tasks. The most elementary task is *distinguishing*: : ($x_A \rightarrow y_A$), ($x_B \rightarrow y_B$) with $y_A \neq y_B$.

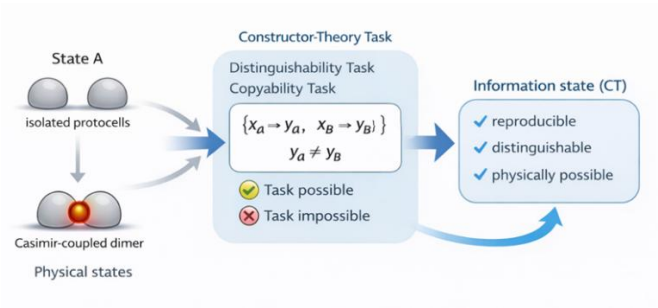


Figure 1. From physical difference to information in Constructor Theory.

In Figure 1, two physical protocell states are shown, isolated vesicles and a Casimir-coupled dimer. Only when physically possible distinguishability and copyability tasks exist does a mere physical difference qualify as an information state in Constructor Theory, emphasizing reproducibility, distinguishability, and physical realizability.

If such a task is physically possible, then x_A and x_B are distinguishable, and the difference $\Delta(x_A,x_B)$ carries information. A difference counts as information only if the laws of nature permit its reproducible realization. Random, unstable, or purely formal differences do not constitute information unless they can be reliably instantiated [13].

The ε-machine analysis from Paper 2 supports this criterion. Protocell clusters satisfy *informational closure*: macrostates of cluster configurations predict future states better than the full microscopic dynamics. Differences between attractors are therefore not only reproducible but also causally effective - strong evidence for information in the CT sense [11].

C. Protocellular Examples of Reproducible Differences

Protocell cluster systems present several classes of prebiotic differences that satisfy CT information criteria and are further supported by results from [2][15].

Isolated vs. Coupled Protocells:

The transition from isolated protocells to dimers is robustly generated by Casimir–Lifshitz forces, reproducible, and forms a well-defined macro-state. The associated distance signature (e.g., L ≈ 5–20 nm) can be read out through physically possible tasks, such as changes in local field structures or membrane deformations. In the ε-machine model, this difference corresponds to distinct causal states E_i, and the transition follows defined paths T_{ij} - precisely the structure expected for a CT informational difference.

(2) Distinct Cluster Geometries (tetrahedral, octahedral, icosahedral):

These attractors differ clearly in contact graphs, symmetry, energetic minima, and return dynamics. Identified as stable attractors in Paper 2, they meet CT requirements for distinguishable and reproducibly realizable states.

The attractor-based memory- formally $\|Z_{t+1} - A\| < \|Z_t - A\|$, - provides the stability CT demands for reproducible differences.

(3) Ion gradients and electrochemical differences:

Protocells capable of maintaining consistently high versus low ion gradients realize a difference that can be copied (via gradient transport) or distinguished (via membrane responses). Such differences define an informational variable, provided that ($x_{high} \rightarrow y_{high}$), ($x_{low} \rightarrow y_{low}$) are physically possible tasks.

(4) ε-machine states as informational differences:

A major result from [2] is that macro-states Z_t differ not only geometrically but through their future-state distributions. Two states belong to different ε-state classes E_i and E_j when $P(Z_{t+1} \mid Z_t \in E_i) \neq P(Z_{t+1} \mid Z_t \in E_j)$.

From a CT perspective, this implies that tasks exist that can not only generate these differences but also exploit them functionally, because different transitions lead to different consequences.

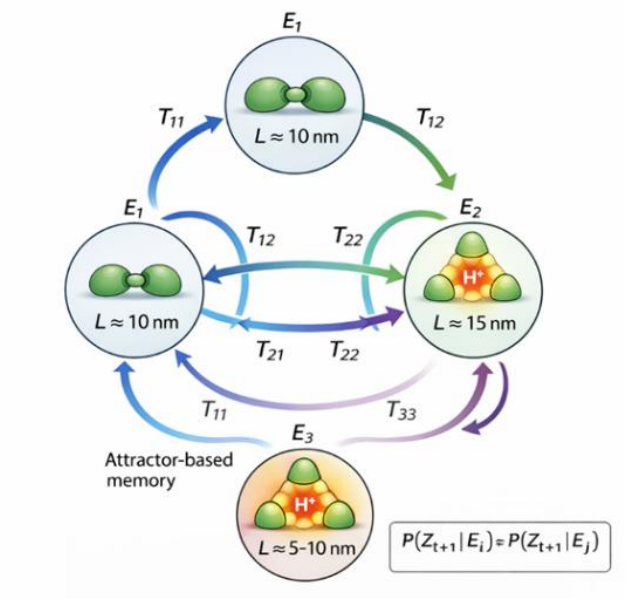


Figure 2. ε-machine attractors with geometry-dependent proton-gradient states.

In Figure 2, an ε-machine network illustrates how distinct protocell-cluster attractors encode reproducible informational states. Dimers and tetrahedral clusters form macro-states with characteristic Casimir separations L and transition pathways. Central proton (H^+) gradients in tetrahedral states indicate geometry-dependent quantum-field-induced energy reservoirs underlying predictive state differentiation.

Interim Conclusion: A difference qualifies as information in Constructor Theory only when it is physically reproducible. Results from [2] show that protocell clusters possess precisely such reproducible differences: stable attractors, robust transitions, attractor-based memory, and structured ε-machines. They therefore satisfy the necessary criteria to serve as prebiotic information carriers in the CT framework.

VI. Meaning as a Functionally Exploited Difference

A. From Information to Meaning: Additional Conditions

In Constructor Theory, information is a reproducible physical difference [4][16] that can be generated, copied, or read through tasks. This alone does not constitute meaning. The crucial addition is functional: meaning is information that plays a stable role within a network of tasks.

An informational difference (x_A, x_B) gains meaning when it systematically determines which tasks a system performs and what consequences follow [23].

The substrate then acts not only as a variable but as a controller of possible transformations. A difference has meaning when constructors exist whose tasks are realizable only for specific informational values:

$x_A \Rightarrow T_A$, $x_B \Rightarrow T_B$, $T_A \neq T_B$.

In this sense, information (x_A or x_B) acts as a control variable. Paper [2] provides empirical support: the ε-machines of prebiotic protocell clusters exhibit distinct causal states with different transition patterns and functional outcomes. This mapping from state to transformation is exactly how CT defines meaning.

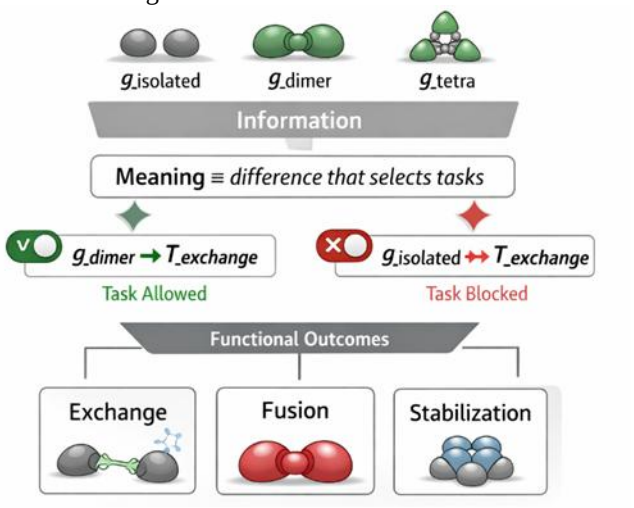


Figure 3. From information to meaning via task-gating in protocell clusters

Figure 3, Schematic illustrating the transition from information to meaning via task-gating in Casimir–Lifshitz–coupled protocell clusters. Cluster states act as informational variables, while meaning emerges when specific states selectively enable or block downstream tasks. Meaning is thus defined as a physical difference that controls functional transformations.

B. Task Networks and Functional Roles

Physical systems consist of chained task networks, where the output of one task becomes the input of another. Within such networks, functional roles emerge: some states act as triggers initiating tasks; others act as switches selecting pathways; still others serve as signals guiding subsystem responses.

Formally, a task is possible only if the relevant informational variable has the appropriate value. A state gains meaning when it enables or restricts the realizability of other tasks.

For biologists, a meaningful state modulates adaptive behavior or responses. For computer scientists, meaning parallels the role of a bit in program flow, such as a conditional branch. For sustainability science, meaning resembles the role of a state within a regulatory system. Massoth [2] offers a prebiotic realization of these roles: the ε-machine of the cluster specifies which transitions are possible and thus which functional consequences arise.

C. Prebiotic Emergence of Meaning in Protocell Clusters

The central question is whether simple Casimir–Lifshitz–coupled protocells can already realize functional roles and thereby generate proto-semantic structures [12]. The answer

is yes, provided that differences in cluster states systematically enable different tasks.

(1) Coupled vs. Isolated Protocells as a Functional Difference:
The distinction between isolated protocells and dimers (or higher configurations such as tetrahedra or octahedra) is established in Paper 2 as a stable informational state with geometry variable $G=\{g_{isolated}, g_{dimer}, g_{tetra}, \ldots\}$. This difference acquires meaning when it modulates downstream tasks—for example:

(a) increased likelihood of fusion,
(b) enhanced molecular exchange due to reduced separation,
(c) altered collective membrane modes or field configurations.

If a task such as an exchange process $T_{exchange}$ is possible only in the dimer state, Dimer $\Rightarrow T_{exchange}$, isolated $\not\Rightarrow T_{exchange}$, then the state "dimer'' carries meaning. This is precisely the constructor-theoretic condition.
The ε-machine analysis in Paper 2 confirms this: dimer states form distinct causal classes E_i and evolve differently from isolated states, giving rise to different functional consequences—i.e., meaning.

(2) Casimir-shaped Distances and Field Configurations as Meaningful Differences:
Casimir–Lifshitz forces generate discrete, energetically preferred separations. Distances and orientations thus act as informational variables, such as distance variable $L=\{L_{near}, L_{far}\}$ or gradient variable $D=\{\Delta_{low}, \Delta_{high}\}$. These differences acquire meaning when they enable different tasks,
for example:

• higher reaction rates at L_{near},
• more efficient energy transfer through collective field modes,
• stronger synchronization of ionic gradients.

If the set of possible tasks depends on the realized distance- $L_{near} \Rightarrow T_{fusion}$, $L_{far} \Rightarrow T_{diffusion}$,
then the distance difference is functional, and therefore meaningful.
Massoth [2] shows that these separations are attractor states stabilized by attractor-based memory and that they induce distinct transitions T_{ij}. These differentiated paths constitute the functional consequences that define meaning in Constructor Theory (CT).

Interim Conclusion: Meaning arises when physical differences are not only reproducible but also functionally effective. Casimir-stabilized protocell clusters satisfy both conditions. They carry informational states through stable attractors and ε-machine–structured transitions, and meaningful states through systematic modulation of prebiotic tasks such as approach, exchange, fusion, stabilization, or gradient formation. Such clusters therefore represent a minimal physical context in which proto-semantic structures can emerge.

VII. Application: Casimir–Lifshitz–Coupled Protocell-clusters as Constructors

We now apply the constructor-theoretic framework of information to prebiotic protocell systems that form mesoscale clusters through Casimir–Lifshitz attraction and electrolyte-mediated coupling. Paper [2] shows that these clusters implement an ε-machine with stable attractors and structured transition pathways, generating an autonomous macro-level with attractor-coded memory. These properties make protocell clusters natural candidates for constructor-theoretic tasks.

A. Casimir–Lifshitz–Coupled Protocell-Clusters as Physical Substrates

Protocells with radii of $R \approx 100$–1000 nm interact in saline prebiotic environments via Casimir–Lifshitz forces. The resulting potential landscapes contain stable and metastable minima, giving rise to long-lived configurations such as dimers, tetrahedral clusters, or 13-membered icosahedra. Massoth [2] demonstrates that these clusters form robust attractors with characteristic binding energies and structured transition dynamics. They satisfy informational closure, as future states are better predicted from macro-states than from full microscopic data.
In Constructor Theory, this implies that protocell clusters provide physical substrates carrying stable attributes—distances, geometries, contact graphs, and gradients—and reproducible transformation pathways. These attributes and transitions supply the material basis for possible tasks.

B. Elementary Tasks in Protocell Clusters

Under prebiotic conditions, three fundamental tasks can be identified that Casimir-coupled clusters can perform repeatedly and robustly [14]. The ε-machine analysis from Paper 2 confirms their structural stability.
Task A: Approach and binding of two protocells:
$(x_{isolated} \rightarrow x_{dimer})$ with $A=\{g_{isolated} \rightarrow g_{dimer}\}$
Two vesicles with suitable orientation, located in a region where the Casimir–Lifshitz force is net attractive, reliably relax into a stable dimer separation L (surface-to-surface distance). The ε-machine encodes this transition as a strictly causal edge $E_{isolated} \rightarrow E_{dimer}$.
CT interpretation: If this transition can be realized with arbitrarily increasing accuracy in principle, Task A is possible (A✓).

Task B: Exchange and modification of gradients within the dimer: $(x_{gradient} \rightarrow x_{gradient'})$ with $C=\{(\Delta_{low} \rightarrow \Delta'_{high})\}$
Once a dimer is formed, contact regions and field couplings enable repeated exchange of ions or molecules. Massoth [2] shows that dimers persist long enough to develop attractor-

based memory, making these exchange patterns reproducible.
CT interpretation: The dimer functions as a constructor that transforms gradients.
Task C: Stabilization of specific cluster configurations:
$(X \rightarrow X)$ over Δt
A cluster maintains its geometry despite thermal fluctuations by relaxing back to its attractor along characteristic return trajectories.
CT interpretation: This structural persistence constitutes a maintenance task enabled by autonomous attractor memory [2], defining a maintenance constructor.

C. Protocell Clusters as (Partial) Constructors

A system qualifies as a constructor if it can perform a task repeatedly without losing its ability to do so. Protocell clusters meet this criterion. Casimir–Lifshitz attraction stabilizes repeated approach events; long lifetimes support robust exchange processes; and the ε-machine defines reproducible transitions and return pathways [24]. The cluster thus transforms input states into output states while retaining its capacity for task execution.
Protocells are not intentional agents; their constructor capacity emerges from physical law. They may be considered partial constructors, since some resources—such as temperature or ion supply—derive from the environment. Constructor Theory permits such side effects as long as the constructor determines the structure of the transformation.

D. Information vs. Meaning States in Protocell Clusters

Constructor Theory provides a natural classification of the cluster states identified in [2]. Information states are physical differences for which tasks of generation, copying, or discrimination exist. These include monomers, dimers, and higher-order clusters, distinct ion gradients, defined distance and field configurations, and ε-machine attractors. Information states satisfy: $(x_A \neq x_B) \wedge A\checkmark, B\checkmark$.
Meaning states are those information states that trigger different downstream tasks. A dimer, for example, can enable further approach or gradient modification. Smaller separations can activate transport or fusion tasks that are impossible at larger separations. Distinct attractors lead to different transition pathways. A meaning state satisfies the constructor-theoretic condition that different information values realize systematically different tasks: $x_A \Rightarrow T_i$, $x_B \Rightarrow$ Tj, $T_i \neq T_j$.

Interim conclusion: Casimir–Lifshitz–coupled protocell clusters are both physical aggregates and natural constructors that implement reproducible tasks such as approach, exchange, and stabilization. They carry information states as reproducible differences and meaning states as differences with functional consequences in the task network. They thus realize the constructor-theoretic transition from information to meaning in a concrete prebiotic setting and complement the mechanisms of informational closure and attractor-based memory described in [2].

VIII. CONCLUSION AND FUTURE WORK

A. Summary of the Conceptual Contribution

This work links Constructor Theory to a realistic prebiotic scenario. In this framework, information is a reproducible physical difference, and meaning arises when such a difference has stable functional consequences within a network of tasks. Casimir–Lifshitz–coupled protocell clusters serve as a minimal model: their stable attractors form reproducible information states, and their transition and return pathways enable basic tasks such as approach, exchange, and stabilization. Physical differences thus become functional categories.

Our contribution is twofold: information is the capacity of a substrate to realize specific tasks and meaning emerges when these differences systematically guide transformations.

B. Answering the research questions

RQ#1: Under which physical conditions can protocell clusters act as constructors that repeatedly perform tasks such as approach, exchange, or coupling?
Protocell clusters function as repeatable constructors when nanoscale separations, dielectric contrasts and moderate ionic strengths jointly create Casimir–Lifshitz–dominated interaction landscapes with discrete, thermally stable attractors. Under these conditions, vesicles can repeatedly approach, bind and reconfigure without structural degradation, and the resulting attractor network provides an autonomous mesoscale level capable of supporting reproducible tasks such as coupling, exchange and geometric stabilization.

RQ#2: How can pure informational differences be distinguished from meaningful states with distinct functional outcomes?
Pure informational differences correspond to reproducible physical distinctions—cluster geometries, separation states or ionic gradients—that can, in principle, be generated, discriminated or copied. These differences become meaningful when they systematically change which downstream tasks are possible and thus produce distinct functional consequences. In Casimir-coupled protocells, meaning arises when particular spatial or gradient states act as control variables that gate transformations such as fusion, exchange or attractor maintenance.

RQ#3: Which classes of tasks in Casimir-coupled clusters can support proto-semantic stability, and how does this relate to concepts such as informational and causal closure?
Proto-semantic stability is supported by three mutually reinforcing task classes: approach and coupling tasks that establish persistent relational structures; exchange and transformation tasks that reorganize internal gradients in a state-dependent manner; and maintenance tasks that restore perturbed configurations to their characteristic attractors. Together, these tasks form a mesoscale dynamic that is

informationally and causally closed, meaning that future transitions are best predicted from macro-states rather than microstates. Within this closed task network, certain informational states reliably select—and are stabilized by—specific transformations, constituting a physically grounded precursor to semantic organization in prebiotic protocell systems.

C. Relation to Computational Mechanics

This approach complements paper [2], which describes protocell clusters as ε-machines whose causal states encode attractors and their mesoscale switching pathways. This macro-level is informationally and causally closed and displays a proto-software-like dynamic.

Here, we ask which of these states carry information and which acquire meaning. Computational Mechanics characterizes structure and predictability; Constructor Theory characterizes tasks and functional roles. Together, they show how ε-machines identify stable differences while CT determines which become functionally effective.

D. Relevance for Biology and Computer Science

The framework offers a prebiotic definition of function independent of genes or enzymes. Protocells appear as active physical units that stabilize distances, generate gradients, and modulate exchange processes. Function arises when such differences reliably trigger or regulate tasks.

In computer science, the approach broadens the notion of a program: programs become networks of physical tasks guided by state variables. Protocell clusters thus implement a natural, substrate-bound computation, where information states act as switches and meaning states as control structures.

E. Limitations and Future Work

The contribution is conceptual. The proposed information and meaning states require experimental and numerical validation. Models of Casimir–Lifshitz potential landscapes and mesoscale dynamics could test the robustness of approach, exchange, and stabilization tasks under realistic conditions. Experiments with protocell models—vesicles or coacervates—may probe analogous attractors and their functional consequences.

A future integrated theory should unify Computational Mechanics and Constructor Theory to determine which macrostates function as constructors and how meaning states emerge from ε-machine organization. This work represents a first step toward a physical explanation of how stabilized differences gave rise to functional categories and early proto-semantic structures.

REFERENCES

[1] M. Massoth, “Attractive Casimir-Lifshitz Forces as a Universal Driver of Prebiotic Protocell Aggregation and Cluster Formation,” *BIOTECHNO*, 2026.

[2] M. Massoth, “Emergent Information Formation in Prebiotic Protocell Clusters: A Computational Mechanics Framework of ε-Machines and Attractor Memory,” *BIOTECHNO*, 2026.

[3] F. E. Rosas, *et al.*, “Software in the natural world: A computational approach to hierarchical emergence,” *arXiv* preprint, 2024, doi:10.48550/arXiv.2402.09090.

[4] M. Deutsch and C. Marletto, “Construction theory of information,” *Proc. R. Soc. A*, vol. 471, art. no. 20140540, 2015, doi:10.1098/rspa.2014.0540.

[5] H. B. G. Casimir, “On the attraction between two perfectly conducting plates,” *Proc. K. Ned. Akad. Wet.*, vol. 51, pp. 793–795, 1948.

[6] E. M. Lifshitz, “The theory of molecular attractive forces between solids,” *Sov. Phys. JETP*, vol. 2, pp. 73–83, 1956.

[7] J. N. Israelachvili, *Intermolecular and Surface Forces*, 3rd ed., Academic Press, London, 2011.

[8] P. C. W. Davies, “Does quantum mechanics play a non-trivial role in life?” *Biosystems*, vol. 78, pp. 69–79, 2004.

[9] I. Gözen, *et al.*, “Protocells: Milestones and recent advances,” *Small*, vol. 18, art. no. 2106624, 2022.

[10] J. P. Crutchfield, “Inferring statistical complexity,” *Phys. Rev. Lett.*, vol. 63, no. 2, pp. 105–108, 1989, doi:10.1103/PhysRevLett.63.105.

[11] E. P. Hoel, L. Albantakis, and G. Tononi, “Quantifying causal emergence,” *Proc. Natl. Acad. Sci. U.S.A.*, vol. 110, no. 49, pp. 19790–19795, 2013, doi:10.1073/pnas.1314922110.

[12] S. A. Kauffman, *The Origins of Order: Self-Organization and Selection in Evolution*, Oxford Univ. Press, Oxford, 1993.

[13] C. R. Shalizi and J. P. Crutchfield, “Computational mechanics: Pattern and prediction, structure and simplicity,” *J. Stat. Phys.*, vol. 104, nos. 3–4, pp. 817–879, 2001, doi:10.1023/A:1010388907793.

[14] G. Nicolis and I. Prigogine, *Self-Organization in Nonequilibrium Systems*, Wiley, New York, 1977.

[15] M. M. Hanczyc, S. M. Fujikawa, and J. W. Szostak, “Experimental models of primitive cellular compartments,” *Science*, vol. 302, no. 5645, pp. 618–622, 2003, doi:10.1126/science.1089904.

[16] D. Deamer, *Assembling Life: How Can Life Begin on Earth and Other Habitable Planets?*, Oxford Univ. Press, Oxford, 2017.

[17] S. Pei, Z. Zhang, T. Li, *et al.*, “Direct measurement of the van der Waals force between a pair of microspheres based on photonic force microscopy,” *Opt. Eng.*, vol. 60, art. no. 084101, 2021.

[18] B. Liu, B. Poolman, and A. J. Boersma, “Ionic strength sensing in living cells,” *ACS Chem. Biol.*, vol. 12, pp. 2510–2514, 2017.

[19] A. J. Boersma, I. S. Zuhorn, and B. Poolman, “A sensor for quantification of macromolecular crowding in living cells,” *Nat. Methods*, vol. 12, pp. 227–229, 2015.

[20] M. Moazzami-Gudarzi, G. Trefalt, I. Szilagyi, P. Maroni, and M. Borkovec, “Nanometer-ranged attraction induced by multivalent ions between similar and dissimilar surfaces probed using an atomic force microscope,” *Phys. Chem. Chem. Phys.*, vol. 18, pp. 8739–8751, 2016.

[21] V. Valmacco, G. Trefalt, P. Maroni, and M. Borkovec, “Forces between silica particles in the presence of multivalent cations,” *J. Colloid Interface Sci.*, vol. 472, pp. 108–115, 2016.

[22] H.-J. Butt, “Measuring electrostatic, van der Waals, and hydration forces in electrolyte solutions with an atomic force microscope,” *Biophys. J.*, vol. 60, pp. 1438–1444, 1991.

[23] A. Kolchinsky and D. H. Wolpert, "Semantic information, autonomous agency and non-equilibrium statistical physics," Interface Focus, vol. 8, no. 6, art. no. 20180041, 2018, doi:10.1098/rsfs.2018.0041.

[24] B. Ruzzante, L. Del Moro, M. Magarini, and P. Stano, "Synthetic Cells Extract Semantic Information From Their Environment," IEEE Transactions on Molecular, Biological and Multi-Scale Communications, vol. 9, no. 1, pp. 23–36, 2023, doi:10.1109/TMBMC.2023.3244399.

[25] S. Bartlett et al., "Physics of Life: Exploring Information as a Distinctive Feature of Living Systems," PRX Life, vol. 3, 037003, 2025, doi:10.1103/rsx4-8x5f

[26] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 79, no. 8, pp. 2554–2558, 1982.

Physical Origin of Proto-Information: Syntax, Semantics and Pragmatic Emergence in Prebiotic Protocell Clusters

Michael Massoth
Department of Computer Science, Hochschule Darmstadt (h_da)
University of Applied Sciences Darmstadt, member of European University of Technology (EUt+)
Darmstadt, Germany
e-mail: michael.massoth@h-da.de

***Abstract*- The emergence of information and meaning before genes or enzymes existed is a key open problem in prebiotic research. This paper develops a physics-based framework for the rise of early information in protocell systems. We show that Casimir–Lifshitz interactions, acting under plausible primordial-soup conditions, generate stable dimers and ordered clusters in which nanometer-scale resonant gaps create structured fluctuation spectra and proton-rich reaction zones. On this basis, Computational Mechanics is used to describe the syntactic layer via ε-machines, functional semantics via lumping operations, and the pragmatic layer via emergent macro-ε-machines that execute cluster-level tasks in the sense of Construction Theory. Together these elements yield a coherent model of proto-biological meaning: one that explains prebiotic information structures without genes, enzymes, or biochemical networks and identifies a physical precursor to functional organization.**

***Keywords-Casimir*–Lifshitz coupling; computational mechanics; ε-machines; prebiotic information; construction theory; protocell clusters.**

I. INTRODUCTION

This is the fourth of seven papers in the series: "A Constructivist Proto-Bio-Information Theory: A Physically Grounded Nano-Systems Architecture for Prebiotic Emergence, Information, Proto-Semantic Function, and Sustainability of Protocell Aggregation and Cluster Formation".

Massoth [1] shows that Casimir–Lifshitz forces generate robust attraction and stable protocell clusters across 5–200 nm under realistic prebiotic conditions. These physically enforced mesoscale assemblies form the structural substrate in which this paper grounds the emergence of information and meaning.

Massoth [2] demonstrates that such clusters form reproducible mesoscale attractors with autonomous ε-machine dynamics. This attractor-based organization provides the computational and informational foundation that this paper develops into a physical account of proto-information and proto-semantics.

Massoth [3] shows that protocell clusters generate reproducible differences and functional meaning states through constructor-theoretic tasks. These proto-functional structures prepare the ground for this paper, which unifies syntactic, semantic, and pragmatic dimensions into a coherent framework of proto-biological meaning.

In addition, we integrate the 'Software in the Natural World' framework of Rosas et al. [4] with the Constructor Theory of Information [5] developed by Deutsch and Marletto, applying both to the hierarchical and task-based organization of protocell cluster dynamics.

The structure of the paper is as follows: In Section II, we review existing approaches to semantic information, Computational Mechanics, and Constructor Theory, highlighting the absence of a unified, physically grounded framework for prebiotic meaning. Section III introduces the physical and conceptual background of prebiotic protocell systems and formulates the central research questions concerning the emergence of syntax, semantics, and pragmatics prior to genetics. Section IV develops the physical and information-theoretic foundations of structured protocell dynamics, combining Casimir–Lifshitz coupling with Computational Mechanics, functional lumping, and constructor-theoretic task concepts. In Section V, we analyze the emergence of syntactic structure in protocell clusters, showing how reproducible fluctuation patterns are captured by ε-machines. Section VI addresses the transition from syntax to semantics, demonstrating how functional equivalence relations give rise to proto-semantic categories grounded in physical invariance. Section VII develops the pragmatic level, where entire protocell clusters form macro-ε-machines that perform reproducible tasks and qualify as prebiotic partial constructors. Section VIII answers the guiding research questions by synthesizing the syntactic, semantic, and pragmatic results within a unified physical framework. Finally, Section IX summarizes the conceptual contributions, situates the results within origins-of-life research, and outlines experimental and theoretical directions for future work.

II. RELATED WORK

Recent theoretical approaches to semantic information—most notably the work of Kolchinsky and Wolpert [15], Ruzzante et al. [16], and S. Bartlett et al. [17] - demonstrate that information is not intrinsically tied to genes or replication.

Existing literature treats syntax (Computational Mechanics), semantics (semantic information), and function (Constructor Theory) largely in isolation. Even integrative discussions such as Rosas et al. [4] stop short of identifying a shared physical origin. A unified, prebiotically plausible framework linking all three levels remains absent.

This paper closes this gap by showing that syntax, semantics, and pragmatics emerge from the same Casimir–Lifshitz–shaped attractor landscape of protocell clusters. ε-machines capture syntactic structure, functional lumping defines semantic classes, and macro-ε-machines implement cluster-level tasks in the sense of Constructor Theory [5]. The result is a fully physical account of proto-biological meaning, in which information precedes replication and function precedes biochemistry.

III. Prebiotic Emergence of Information and Meaning

How information, functional structure, and early meaning relations arose before genes, enzymes, or metabolic networks existed remains a major open challenge in origins-of-life research. The earliest protocells must already have exploited reproducible differences, produced functionally relevant states, and stabilized collective patterns; otherwise, the transition to biochemical evolution could not have occurred. Yet a mechanistic, physics-based model that explains how such proto-informational processes emerged is still largely missing.

Recent results show that Casimir–Lifshitz interactions can generate robust coupling between protocells under prebiotic conditions, forming stable dimers and ordered clusters. These clusters create nanometer-scale resonant gaps that structure the electromagnetic mode spectrum and produce characteristic energetic signatures—structured fluctuation spectra. The same gaps can host quantum-field-driven proton reservoirs, offering a physical precursor to chemiosmotic gradients.

These structured energy patterns may serve as substrates for proto-informational processes. To analyze them systematically, we use Computational Mechanics: ε-machines capture minimal predictive structure (syntax), lumping procedures define functional macrostates (semantics), and stable cluster geometries generate collective attractors (pragmatics).

Against this background, the paper addresses three central research questions:

RQ#1 Syntax: How does Casimir–Lifshitz coupling generate recurrent, predictable patterns describable by ε-machines?
This concerns the causally minimal states that arise from structured fluctuation profiles, gap geometries, and field modulations.

RQ#2 Semantics: Under which physical conditions do these syntactic patterns become functional categories with stable consequences for proton distribution, permeability, or adsorption?
Here we examine how lumping operations define semantic classes whose energetic signatures produce distinct physico-chemical effects.

RQ#3 Pragmatics: How do entire protocell clusters form emergent macro-ε-machines that perform reproducible tasks in the sense of Construction Theory?
This concerns proto-functional stability: when a dimer or tetrahedron becomes a coherent unit capable of repeated tasks such as stabilization, proton focusing, or exchange.

Together, these questions establish a physics-based framework for the emergence of prebiotic meaning—a transition from fluctuations to syntax, from functional consequences to semantics, and from cluster coherence to proto-pragmatic organization. This approach explains information formation before genetics and links concepts across biology, physics, and computer science.

IV. Physical and Information-Theoretic Foundations of Structured Protocell Dynamics

The framework developed here is built on the coupling of two physical layers: (1) Casimir–Lifshitz–driven formation of structured fluctuation spectra between protocells, and (2) their representation as syntactic, semantic, and pragmatic information structures within Computational Mechanics and constructor-theoretic information.

A. Casimir–Lifshitz coupling as a generator of structured fluctuations

Between two protocells with radii R_1, R_2 and a gap L, Casimir–Lifshitz theory produces a mode-selective energy spectrum that structures local electromagnetic fluctuations. In the biologically relevant regime, 5 nm ≤ L ≤ 100 nm, the algebraic high-temperature term dominates:

$$F_{CL}(L) \approx -(A_{eff}/6) \cdot (R_{eff} / L^2), \text{ with } R_{eff} = (R_1R_2)/(R_1 + R_2),$$

and A_{eff} an effective Hamaker constant integrating the spectral dielectric response of the membrane-water-system.

The corresponding potential $U_{CL}(L) \propto -(A_{eff}/6) \cdot (R_{eff} / L)$ creates strong and reproducible nanometer-scale configurations that generate characteristic energetic signatures—structured fluctuation spectra. These signatures define proto-syntactic patterns because they restrict the system to a limited set of reproducible configuration classes.

In highly ordered protocell clusters (e.g., tetrahedral arrangements), the superposition of modes forms a topologically nontrivial attractor landscape. At the cluster center, this landscape produces quantum-field–induced proton enrichment zones—a physical mechanism for early

functional gradients (Paper 5). Experimental protocell systems demonstrate that membrane-bounded compartments can indeed support such early functional gradients and stable microchemical domains [14].

B. Computational Mechanics and the mathematical structure of syntax

Computational Mechanics characterizes a system through the causally minimal, prediction-relevant states *S* derived from past observations x(t:–∞,0):

$$\varepsilon(x(t:-\infty,0)) = \{x' : P(x(t:0,+\infty) \mid x(t:-\infty,0)) = P(x(t:0,+\infty) \mid x'(t:-\infty,0))\}.$$

An ε-machine is thus a minimal state-based representation that captures all reproducible patterns required to predict a system's future behavior, without invoking symbolic codes or external control. It provides a mathematically well-defined notion of syntax as causal structure rather than symbolic description.

In this work, Computational Mechanics is not used as a generic modeling framework, but as a diagnostic tool to identify which physical degrees of freedom in protocell clusters give rise to reproducible, prediction-relevant macrostates. The observable process X_t corresponds to physically accessible quantities—such as inter-protocell distances, local field intensities, proton densities, or adsorption states—sampled over time in Casimir–Lifshitz–coupled assemblies. The resulting ε-states therefore do not represent abstract symbols, but equivalence classes of physical configurations that share the same conditional future behavior under field-mediated coupling.

The full set S of ε-states defines the syntactic level of the system. For protocell clusters, the ε-machine identifies reproducible fluctuation patterns and groups them into causally relevant states. Their temporal evolution is described by the stochastic transition matrix $T_{ij} = P(S_{t+1} = j \mid S_t = i)$.

The entropy rate $h_\mu = H[X_t \mid S_t]$ measures residual unpredictability within syntactically ordered patterns and quantifies the interaction complexity of protocell coupling.

This framework allows a principled distinction between mere structural order and genuine informational structure: only ε-states associated with distinct future consequences qualify as informational variables in the constructor-theoretic sense.

C. From syntax to semantics: lumping as functional category formation

Semantics arises when several syntactic states produce equivalent functional effects and can be grouped into macrostates Z_k via a mapping A:

$$Z_k = A(S_i, S_j, \ldots) \text{ if } \Phi(S_i) \approx \Phi(S_j),$$

where Φ(S) denotes a physical effect such as changes in proton distribution, adsorption profiles, or membrane permeability. A lumping is semantically valid if:

$$P(\text{Outcome} \mid S_i) \approx P(\text{Outcome} \mid S_j).$$

Semantics thus becomes a functional equivalence relation over energetic signatures. Examples:

- Protocell dimer:
  different fluctuation patterns that yield the same proton-focusing profile constitute a semantic microstate.
- Tetrahedral cluster:
  multiple syntactic field configurations producing the same central proton funnel.

D. Pragmatics: Emergence and Construction Theory

Construction Theory models functional possibilities as tasks: $A = \{x \rightarrow y\}$. A protocell cluster acts as a constructor when it performs a task repeatedly without losing functionality:

$$A(t: 0, +\infty): (x, C) \rightarrow (y, C), \text{ where } C \text{ is the cluster state.}$$

A task is physically possible if such a constructor can exist in principle.

For protocell clusters, we define proto-pragmatic tasks such as stable proton focusing, enhanced adsorption within the gap, and geometry-regulated exchange channels.

These tasks are described by macro-ε-machines, whose states correspond not to individual protocells but to entire cluster geometries (e.g., dimers or tetrahedra).

At the pragmatic level, this causal state structure extends naturally to macro-ε-machines that govern cluster-level tasks. Casimir–Lifshitz–coupled protocell clusters possess a finite set of causally relevant macrostates with distinct future consequences, thereby providing a physically grounded notion of syntax.

V. SYNTAX: PATTERN FORMATION AND E-MACHINES IN PROTOCELL CLUSTERS

In prebiotic protocell assemblies, syntax denotes the level of structured pattern formation that emerges from initially disordered fluctuations and becomes stabilized through physical coupling.

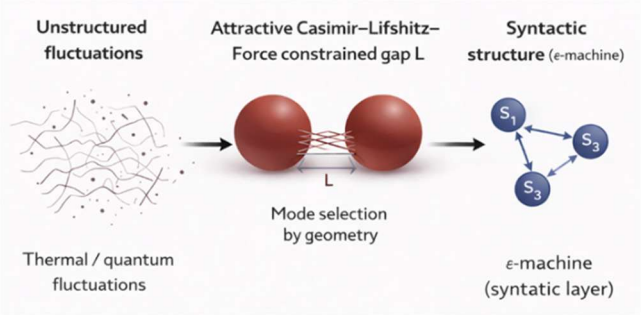


Figure 1. Physical origin of syntactic structure in protocell clusters.

In Figure 1, initially unstructured thermal and quantum fluctuations become geometrically constrained by attractive Casimir–Lifshitz coupling in a protocell dimer. Mode selection within the nanoscale gap reduces fluctuation diversity and gives rise to a finite set of reproducible causal states, represented as an ε-machine that constitutes the syntactic layer of prebiotic information.

Analysis of attractive Casimir–Lifshitz forces in single- and multi-cell geometries shows that protocells in saline environments generate characteristic regional fluctuation spectra that are neither random nor locally isolated. Casimir–Lifshitz interactions are fluctuation-induced forces that arise when electromagnetic vacuum fields are confined by material boundaries, here provided by adjacent protocell membranes, leading to reproducible, geometry-dependent coupling at nanometer scales.

Nanometer-scale gaps filter and modulate these electromagnetic modes so that specific field profiles recur. These energy patterns remain dynamic, yet they form spatially and temporally correlated sequences that exceed thermal noise. Such reproducible configurations constitute the syntactic basis of the system. Computational Mechanics provides a rigorous tool for reconstructing this structure.

The observable sequence $X(t:-\infty,+\infty)$ includes the relevant physical variables—distance dynamics, proton density, field strengths, and local permittivities. The ε-machine extracts the minimal state classes needed for prediction. Two histories are equivalent when they yield identical future distributions:

$$\varepsilon(x(t:-\infty,0)) = \{x' : P(x(t:0,+\infty) \mid x) = P(x(t:0,+\infty) \mid x')\}.$$

Each equivalence class defines a causal state S_t. The full set of these states forms the system's syntactic structure. While the underlying dynamics is continuous and high-dimensional, the ε-machine offers a discrete, reduced representation that preserves all information relevant for prediction.

In protocell clusters, recurring fluctuation motifs act as attractors of syntactic states. A dimer with a stable Casimir–Lifshitz gap generates characteristic sequences of distance-dependent field profiles; a tetrahedral cluster produces corresponding superposed near-field patterns. These ordered sequences are encoded in the transition matrix

$T_{ij} = P(S_{t+1} = j \mid S_t = i)$.

The emergence of such reproducible fluctuation sequences corresponds to the behavior of dissipative structures formed in far-from-equilibrium regimes [13].

Geometric fluctuations shift transition probabilities but remain confined to a limited set of robust patterns. Syntax thus arises from deterministic physical constraints combined with stochastic microscopic variability, which together permit only a finite number of causally meaningful patterns.

Syntactic order is quantified by the entropy rate

$h_\mu = H[X_t \mid S_t]$, where X_t denotes the next observable event (e.g., proton-density change or field-profile shift) and S_t the current syntactic state. The entropy rate measures the remaining unpredictability given the internal cluster state. Low values indicate that future behavior is strongly determined by syntax. Protocell clusters exhibit markedly lower entropy rates than isolated vesicles: Casimir–Lifshitz attraction channels the fluctuation spectrum and confines state dynamics to a few stable attractors.

Syntax is therefore not an abstract category but a direct physical consequence of cluster geometry, material-dependent mode structure, and feedback between proton distribution and electromagnetic fields. The ε-machine reveals this order and shows how protocell assemblies can exhibit structured "responses"—not through control or intention, but through reproducible organization within the fluctuation space.

VI. Semantics: Functional Category Formation in Protocell Clusters

In prebiotic protocell assemblies, semantics marks the transition from mere pattern formation to functional meaning. Syntax captures recurring fluctuation sequences; semantics emerges when such patterns produce systematically different physical effects. Proto-semantic structure is therefore defined as a relation between an energetic signature and its functional consequence.

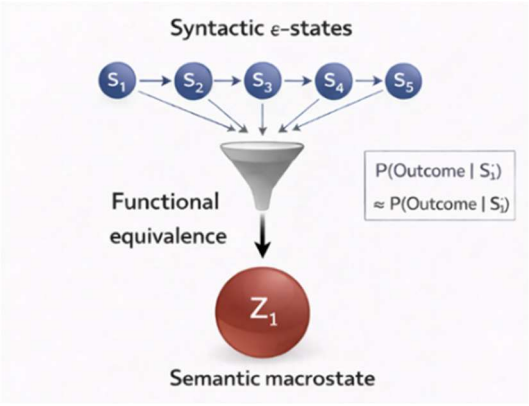


Figure 2. Semantic macrostates emerge by functional lumping of syntax.

This corresponds to the principle of causal emergence, in which macro-level differences can exert greater causal specificity than their microstate constituents [11]. Meaning arises not through symbolic representation but through stable couplings between fluctuation patterns and measurable changes in proton distribution, membrane permeability, or adsorption dynamics.

In Figure 2, multiple syntactic ε-states with distinct transition structures are grouped through a lumping operation based on functional equivalence. When different syntactic realizations produce indistinguishable physical outcomes, they collapse into a single semantic macrostate. Semantics thus arises as an invariance of physical effects across syntactic diversity, not as symbolic encoding.

The mathematical basis is the lumping procedure of Computational Mechanics. Starting from syntactic ε-states S_t, semantics appears when several such states have equivalent effects. A semantic class Z_k is formed via a mapping A when $\Phi(S_i) \approx \Phi(S_j)$, where Φ denotes observable outcomes such as proton-density shifts or field-energy changes. Two syntactic states belong to the same semantic category if they yield the same macroscopic response. The condition

$$P(\text{Outcome} \mid S_i) \approx P(\text{Outcome} \mid S_j)$$

defines this functional equivalence: meaning emerges as invariance of physical effects across different syntactic realizations. The lumping criterion used here is mathematically equivalent to strong lumpability in Markov chains (Kemeny and Snell), but is interpreted physically in terms of functional equivalence of outcomes rather than state aggregation alone.

In dimers, distinct field-fluctuation patterns often lead to the same functional consequence, such as directed proton focusing in the gap. Quantum-field–induced mode selection creates resonance zones whose impact on proton dynamics remains similar despite syntactic variability. This yields a semantic class that groups all patterns producing the same proton-based effect.

Tetrahedral clusters generate a richer semantics. The four vesicles jointly form a three-dimensional energetic minimum at the cluster center, concentrating protons and acting as a proto-chemiosmotic reservoir. Different syntactic sequences—arising from minor distance fluctuations or altered mode superposition—produce the same functional macrostate: a stable proton-funnel geometry. Here, semantics appears as geometric invariance of functional effects. The semantic state is defined not by a single pattern but by all patterns that generate the same proton-reservoir dynamics.

In Computational Mechanics, such functional stability corresponds to a reduction of syntactic diversity into a small set of macrocategories. These lumped classes form a proto-semantic alphabet whose elements are defined by their effects. Meaning thus arises not from external coding but from the physical robustness of functional outcomes. Semantics becomes a direct consequence of energetic invariance, not a layer added on top of physics.

These invariances are experimentally accessible. Measurements of proton density in gap regions, membrane permeability, or field-intensity profiles offer observables whose outcome distributions can be compared under different syntactic excitations. If these distributions coincide, a semantic lumping class is identified. Prebiotic semantics thus becomes an empirically accessible, mathematically precise, and physically grounded layer. It marks the transition from structural order to functional efficacy and provides the foundation for later pragmatic and emergent organization in entire protocell clusters.

VII. Pragmatics and Emergence in Protocell Systems

The pragmatic layer of prebiotic protocell assemblies describes the processes through which syntactic patterns and semantic categories combine into collective functional units. Syntax captures the ordering of fluctuations, and semantics captures their functional effects.

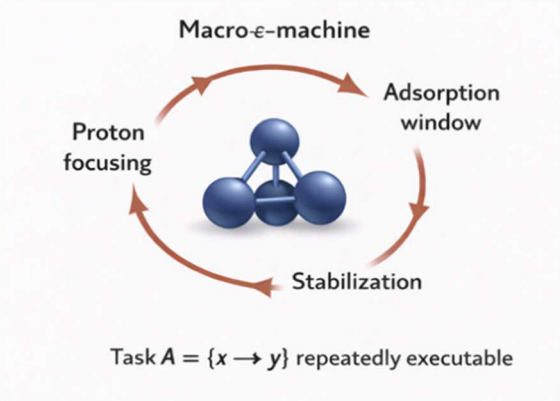


Figure 3. Macro-ε-machines implement repeatable prebiotic functional tasks.

In Figure 3, a tetrahedral protocell cluster is represented as a macro-ε-machine whose collective state enables the repeated execution of physical tasks. Proton focusing, adsorption windows, and stabilization form a closed task loop that persists despite microscopic fluctuations. The cluster thus functions as a prebiotic constructor in the sense of Construction Theory.

Pragmatics emerges only when whole protocell aggregates perform reproducible transformations that exceed the capabilities of individual vesicles. Such emergent task-capabilities mirror classical self-organizing processes in which new functional roles arise spontaneously from collective system dynamics [12]. At this level, protocell clusters act as physical constructors in the sense of Construction Theory: they carry out tasks repeatedly, robustly, and without losing their functional core. These emergent tasks arise not from biological programs but from

the attractor dynamics of macro-ε-machines, whose states represent complete cluster geometries.

Dimers and tetrahedral clusters stabilized by Casimir–Lifshitz coupling form such macro-states. Their evolution is described by transition probabilities between geometrically defined macro-ε states. In dimers, coupling distance is the central state variable from which a reduced set of functional categories emerges. Tetrahedra possess an additional internal resonance zone, created by overlapping near fields, which forms a localized energetic subsystem. Both cluster types exhibit attractors with stable statistical signatures that persist despite microscopic noise. These macro-attractors carry proto-pragmatic meaning because they implement recurrent, functionally relevant tasks.

A Proto-Pragmatic-Functional (PPF) meaning category is: A stable, pre-biological macrostructure in protocell clusters that reliably performs a specific physical task and therefore has functional significance — without genes, enzymes, or chemical codes.

PPF-Meaning Category:
High Lipid Packing Density → Activation Mode

Regions of high lipid packing density reduce local permeability, stabilize nonpolar molecules, and form robust membrane patches. Despite microscopic variability, they produce recurrent patterns of compressed membrane modes that the ε-machine groups into one syntactic class. Functionally, this corresponds to an activation mode: these areas act as preferred sites for adsorption and structural coupling. Pragmatically, protocell clusters recreate such activation zones, supporting early precursor chemistry.
Syntax: Recurrent patterns of compressed membrane modes stabilized by Casimir–Lifshitz-induced pressure gradients.
Semantics: "High packing density means activation mode" - consistent adsorption and stabilization effects.
Functional value: Enhanced adsorption, structural stability, and energetic coupling; early hubs for precursor chemistry.

PPF-Meaning Category:
Resonant Modes → Preferred Adsorption

Resonant modes arise in dimers when specific distances and orientations amplify electromagnetic frequencies. Functionally, they lead to reliably increased adsorption of small molecules. Pragmatically, the dimer acts as a constructor that forms, maintains, and regenerates these adsorption windows under fluctuation. In the macro-ε-machine, the system returns to the same adsorption attractor despite varying microstates.
Syntax: Recurrent patterns of enhanced resonant modes in the fluctuation spectrum.
Semantics: "Resonant modes mean increased adsorption" - stable correlation between mode amplification and molecular attachment.
Functional value: More efficient adsorption, improved precursor chemistry, and faster local reaction kinetics.

PPF-Meaning Category:
Field Gradient → Preferred Proton Transfer

Field gradients created by mode selection concentrate protons in the gap and are a key functional feature. They arise through quantum-field-driven proton focusing and rest on stable syntactic patterns of damping and mode selection. Semantically, they represent facilitated proton transfer. Pragmatically, they create a proto-chemiosmotic space in which proton flow is more ordered than in isolated protocells. Such gradient zones act as energetic constructors because they repeatedly shape and stabilize reaction profiles.
Syntax: Recurrent patterns of stable field gradients and proton-focusing mode selection.
Semantics: "Field gradient means efficient proton transfer" - functional equivalence across syntactic variants.
Functional value: Improved energetics of simple redox and pH-driven reactions; proto-chemiosmotic precursors.

PPF-Meaning Category:
Energetic wells → Transfer Windows

Energetic wells form when dimers or tetrahedra generate local potential minima. These regions open transfer windows that ease passage or accumulation of specific molecules. Their syntactic patterns vary, but their functional impact remains stable. Pragmatically, these minima act as recurring resonance and transport zones and represent early proto-molecular channels in both pair and cluster geometries.
Syntax: Recurrent profiles of local potential minima in the near field of dimers or tetrahedra.
Semantics: "Energetic wells mean transfer windows" — consistently increased exchange and permeation probabilities.
Functional value: Earliest form of a proto-molecular channel; facilitated material transport between protocells.

PPF-Meaning Category:
Synchronization Mode → Cluster Stability

In larger clusters, pragmatics appears in collective attractors with new functional roles. When coupling strength integrates several protocells, synchronization modes arise that generate correlated oscillations across the assembly. These syntactically coherent patterns correspond semantically to increased cluster stability. Pragmatically, they reduce dissipation because synchronized membrane dynamics enable more efficient exchange processes. The macro-ε-machine thus captures a functional model that exceeds the sum of individual vesicles.
Syntax: Correlated oscillatory field modes that reproduce across multiple protocells.

Semantics: "Synchronous mode means cluster stability" - robust convergence to coherent dynamics.
Functional value: Reduced dissipation, more stable material cycles, coordinated adaptation to external fluctuations.

PPF-Meaning Category:
Interference Patch → Reaction-Enhancing Zone

Local interference patches ("hotspots") amplify selected electromagnetic modes and create energetic concentrations that promote reactions. Syntactically, they appear as stable condensation patterns; semantically, as zones of increased reactivity. Pragmatically, they function as recurrent reaction nodes: the cluster acts as a constructor that generates, stabilizes, and renews favored chemical microenvironments.
Syntax: Localized hotspots of energetic concentration stabilized by overlapping near fields.
Semantics: "Hotspot means reaction-enhancing zone" - strong coupling to increased adsorption and molecular retention.
Functional value: Local heating, higher reaction turnover, efficient chemical microdomains.

The proto-pragmatic-functional architecture of protocell systems thus emerges from the ability of dimers and tetrahedra to repeatedly occupy stable functional states. Macro-ε-machines formalize these emergent patterns, while Construction Theory clarifies their role as physical task realizers. At this level, the transition from physical fluctuation to proto-biological functionality begins: clusters become carriers of energetic, material, and informational tasks — early precursors of biological action systems.

VIII. ANSWERING THE RESEARCH QUESTIONS

RQ#1 Syntax: How does Casimir–Lifshitz coupling generate recurrent, predictable patterns describable by ε-machines?

Casimir–Lifshitz coupling generates syntax by constraining initially broadband thermal fluctuations into a restricted set of recurrent field and gap configurations that are well captured by ε-machines. In dimers, tetrahedra, and higher-order clusters, nanometre-scale gaps filter the mode spectrum and produce characteristic fluctuation and proton-density profiles that reappear with high frequency, so that Computational Mechanics reconstructs a small set of causal states with structured transition probabilities rather than a featureless noise process. These ε-states and their transitions define the syntactic layer: the minimal predictive patterns imposed by geometry, material parameters and field-mediated coupling.

RQ#2 Semantics: Under which physical conditions do these syntactic patterns become functional categories with stable consequences for proton distribution, permeability, or adsorption?

Semantics arises when subsets of these syntactic states are lumped into macrostates that share the same physico-chemical consequences. Under prebiotic saline conditions, different microscopic fluctuation patterns can produce indistinguishable proton-focusing profiles, adsorption propensities, or permeability shifts in the gap; such ε-states form semantic classes because they lead to the same outcome distributions for observables such as proton density or local reaction rates. In this view, meaning corresponds to energetic invariants: a semantic category is defined not by symbolic labels but by the robustness of its functional effects across underlying syntactic variability.

RQ#3 Pragmatics: How do entire protocell clusters form emergent macro-ε-machines that perform reproducible tasks in the sense of Construction Theory?

At the pragmatic level, entire dimers and tetrahedral clusters behave as macro-ε-machines whose states are cluster geometries and collective field configurations, and whose transitions implement reproducible tasks in the sense of Constructor Theory. Stable Casimir–Lifshitz attractors support repeated execution of proton focusing, adsorption enhancement, or exchange-channel formation without loss of cluster functionality, so that these aggregates qualify as prebiotic partial constructors. In this way, protocell clusters realize a proto-pragmatic architecture: from constrained fluctuations to syntax, from functional invariance to semantics, and from cluster-level attractors to emergent task spaces that constitute an early, fully physical notion of meaning before genetics.

IX. CONCLUSION AND FUTURE WORK

This final section summarizes the conceptual contributions of the framework, situates the results within current origins-of-life research, and outlines experimentally and theoretically accessible directions for further investigation.

A. The Physical Emergence of Early Proto-Biological Information and Meaning

This work shows that prebiotic protocell assemblies can develop a structured, information-bearing organization even in the absence of genes, enzymes, or metabolic networks. The present study is conceptual and does not aim at a full numerical reconstruction of ε-machines; instead, it identifies the physical conditions under which syntactic, semantic, and pragmatic structures necessarily emerge.

Low-entropy structural order alone, such as crystalline lattices (e.g., NaCl), does not constitute information or meaning in the present sense. Information requires multiple reproducible macrostates with distinct future consequences and meaning requires that such differences gate physical

tasks; crystalline systems realize a single attractor state and therefore lack informational and functional differentiation.

Unlike agent-based or viability-centered notions of semantic information [15], the present approach grounds meaning entirely in physical task realizability.

Casimir–Lifshitz coupling does more than merely stabilize protocell aggregates: it shapes the fluctuation landscape in nanometer-scale gaps, producing recurrent patterns, functional invariances, and collective attractors. From these physically constrained dynamics, syntactic, semantic, and pragmatic layers of proto-biological organization emerge in a unified manner.

Within this framework, syntax arises from ordered fluctuation patterns captured by ε-machines, semantics from functional invariance under syntactic variability, and pragmatics from cluster-level tasks reproducibly executed by macro-ε-machines. Meaning is therefore not introduced symbolically or biologically, but emerges as a physical property of structured, field-mediated interactions. Information precedes replication, and function precedes biochemistry.

This perspective complements existing origins-of-life models [15] that emphasize chemical or metabolic closure by identifying a prior stage of organization: closure at the level of physically stabilized task spaces. Protocell clusters act as prebiotic partial constructors in the sense of Constructor Theory, repeatedly realizing tasks such as proton focusing, adsorption enhancement, and exchange-channel formation without molecular machinery. Functional organization thus originates in field-structured geometries before genetic encoding becomes possible. By identifying field-structured task spaces as a precursor to biochemical organization, this framework offers a physically testable bridge between prebiotic physics and the emergence of biological function.

B. Future work

The framework developed here is experimentally and computationally accessible. ε-machines of syntactic fluctuation patterns can be reconstructed using near-field measurements, optical tweezers, pH-sensitive microsensors, and finite-element simulations of Casimir–Lifshitz interactions. Microfluidic platforms enable controlled assembly of protocell dimers and clusters, allowing systematic mapping of attractor landscapes and functional task stability.

At a theoretical level, future work should further integrate Computational Mechanics and Constructor Theory to characterize which macrostates qualify as constructors and how proto-pragmatic task spaces scale with cluster size and environmental complexity. Such integration may clarify the transition from field-mediated task organization to chemically driven metabolic networks.

More broadly, this approach suggests a revised narrative for the emergence of proto-liveliness: information is instantiated as reproducible field structure in confined gaps, while meaning and function emerge only at the cluster level—from field ordering imposed by geometric embedding and constrained by Casimir–Lifshitz boundary conditions across the coupled network of intermembrane spaces.

REFERENCES

[1] M. Massoth, “Attractive Casimir-Lifshitz Forces as a Universal Driver of Prebiotic Protocell Aggregation and Cluster Formation,” *BIOTECHNO*, 2026.

[2] M. Massoth, “Emergent Information Formation in Prebiotic Protocell Clusters: A Computational Mechanics Framework of ε-Machines and Attractor Memory,” *BIOTECHNO*, 2026.

[3] M. Massoth, “From Physical Difference to Meaning: A Constructor-Theoretic Framework for Prebiotic Information in Casimir–Lifshitz-Coupled Protocell Clusters,” *BIOTECHNO*, 2026.

[4] F. E. Rosas, *et al.*, “Software in the natural world: A computational approach to hierarchical emergence,” *arXiv* preprint, 2024, doi:10.48550/arXiv.2402.09090.

[5] M. Deutsch and C. Marletto, “Construction theory of information,” *Proc. R. Soc. A*, vol. 471, art. no. 20140540, 2015, doi:10.1098/rspa.2014.0540.

[6] H. B. G. Casimir, “On the attraction between two perfectly conducting plates,” *Proc. K. Ned. Akad. Wet.*, vol. 51, pp. 793–795, 1948.

[7] E. M. Lifshitz, “The theory of molecular attractive forces between solids,” *Sov. Phys. JETP*, vol. 2, pp. 73–83, 1956.

[8] J. N. Israelachvili, *Intermolecular and Surface Forces*, 3rd ed., Academic Press, London, 2011.

[9] P. C. W. Davies, “Does quantum mechanics play a non-trivial role in life?” *Biosystems*, vol. 78, pp. 69–79, 2004.

[10] I. Gözen, *et al.*, “Protocells: Milestones and recent advances,” *Small*, vol. 18, art. no. 2106624, 2022.

[11] E. P. Hoel, L. Albantakis, and G. Tononi, “Quantifying causal emergence,” *Proc. Natl. Acad. Sci. U.S.A.*, vol. 110, no. 49, pp. 19790–19795, 2013, doi:10.1073/pnas.1314922110.

[12] S. A. Kauffman, *The Origins of Order: Self-Organization and Selection in Evolution*, Oxford Univ. Press, Oxford, 1993.

[13] G. Nicolis and I. Prigogine, *Self-Organization in Nonequilibrium Systems*, Wiley, New York, 1977.

[14] M. M. Hanczyc, S. M. Fujikawa, and J. W. Szostak, “Experimental models of primitive cellular compartments,” *Science*, vol. 302, no. 5645, pp. 618–622, 2003, doi:10.1126/science.1089904.

[15] A. Kolchinsky and D. H. Wolpert, “Semantic Information, Autonomous Agency and Non-Equilibrium Statistical Physics,” Entropy, vol. 20, no. 10, p. 793, 2018, https://doi.org/10.1098/rsfs.2018.0041

[16] B. Ruzzante, L. Del Moro, M. Magarini, and P. Stano, “Synthetic Cells Extract Semantic Information From Their Environment,” IEEE Transactions on Molecular, Biological and Multi-Scale Communications, vol. 9, no. 1, pp. 23–36, 2023, doi:10.1109/TMBMC.2023.3244399.

[17] S. Bartlett et al., “Physics of Life: Exploring Information as a Distinctive Feature of Living Systems,” PRX Life, vol. 3, 037003, 2025, doi:10.1103/rsx4-8x5f

[18] D. Deamer, *Assembling Life: How Can Life Begin on Earth and Other Habitable Planets?*, Oxford Univ. Press, Oxford, 2017.

Quantum Field-induced Proton H^+ Gradients in Prebiotic Protocell Clusters

Michael Massoth
Department of Computer Science, Hochschule Darmstadt (h_da)
University of Applied Sciences Darmstadt, member of European University of Technology (EUt+)
Darmstadt, Germany
e-mail: michael.massoth@h-da.de

***Abstract*- The origin of the first hydrated proton (H^+) gradients before the rise of biological proton pumps or membrane-bound enzyme complexes remains a central unresolved problem in prebiotic research. Here we show that protocell clusters, formed under saline primordial-soup conditions, are not only stabilized by Casimir–Lifshitz attraction but also generate selective modulation of dynamic Matsubara modes within their nanometer-scale gaps. Because the static mode (n = 0) is fully suppressed by Debye screening, the remaining dynamic modes create geometry-defined resonance zones with lowered vacuum energy. These zones exponentially enrich hydrated protons (H^+) through Casimir–Polder coupling. Dimers act as one-dimensional proton (H^+) channels, while tetrahedral clusters produce three-dimensional proton funnels. The resulting quantum-field-induced proton reservoirs offer a physically plausible precursor to chemiosmotic (H^+) gradients and may have provided the first energetic coupling mechanism between early protocells.**

Keywords-*Casimir–Lifshitz forces; protocell clusters; prebiotic proton gradients; dynamic Matsubara modes; Casimir–Polder potentials; chemiosmotic precursors.*

I. INTRODUCTION

This is the fifth of seven papers on the series: "A Constructivist Proto-Bio-Information Theory: A Physically Grounded Nano-Systems Architecture for Prebiotic Emergence, Information, Proto-Semantic Function, and Sustainability of Protocell Aggregation and Cluster Formation".

Massoth [1] showed that attractive Casimir–Lifshitz forces formed an unavoidable aggregation mechanism under the saline and thermally active conditions of the early Earth. Classical DLVO interactions collapse in the relevant 5–200 nm range [12], while Casimir–Lifshitz attraction remains algebraic, scaling roughly as $1/L^2$, and therefore dominates mesoscale interactions.

Massoth [2] built on this result and demonstrated how these forces generate stable mesoscale protocell clusters, especially dimers and tetrahedra with six loose contact points and typical gap widths of $L \approx 10$ nm. In this nanoregime, Casimir–Lifshitz attraction exceeds thermal energy and suppresses Brownian motion. The outcome is long-lived, geometry-defined protocell clusters that can support both prebiotic energy organization and early, physically implemented information states.

Massoth [3] shows that protocell clusters can generate reproducible differences and functional meaning states. These proto-functional states form the conceptual basis for this paper, where proton gradients appear as early functional signals emerging from physical constraints.

Massoth [4] reveals that Casimir-Lifshitz-shaped resonance zones and ε-machines create syntactic, semantic, and pragmatic layers of early information. This multi-layered framework supports this paper, which interprets proton reservoirs as energetically encoded proto-semantic structures within nanoscale reaction zones.

Structure of the paper: Section II introduces the physical framework of Casimir–Lifshitz–driven protocell clustering and defines the relevant geometric configurations. Section III develops the quantum-field mechanism underlying proton accumulation in nanoscale resonance zones and provides order-of-magnitude estimates. Section IV discusses the prebiotic energetic, informational, and proto-biological implications of these proton reservoirs. Section V concludes with experimental perspectives and directions for future work.

II. QUANTUM-FIELD ORIGINS OF PREBIOTIC PROTON GRADIENTS AND RESEARCH QUESTION

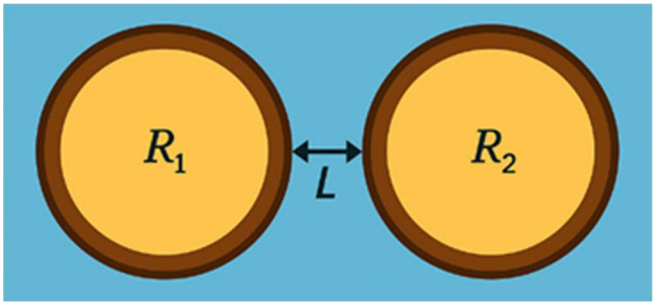


Figure 1. Geometry of a prebiotic protocell dimer

In Figure 1, schematic representation of two spherical protocells with radii R_1 and R_2 separated by a minimal surface-to-surface distance L in saline aqueous solution (primordial soup). This sphere–sphere geometry defines the fundamental configuration used to model Casimir–Lifshitz interactions between protocell membranes under prebiotic conditions.

Energy landscapes were fundamental to the origin of life, yet the emergence of the first proton gradients—long before

biological proton pumps, membrane complexes, or metabolic networks existed—remains a central unresolved question in prebiotic research. Modern chemiosmotic systems rely on complex molecular architectures.

This raises a key research question: could simple protocell assemblies generate proton (H^+) differences [10] from their physical environment and create local energy potentials solely through geometry and the properties of the quantized electromagnetic field?

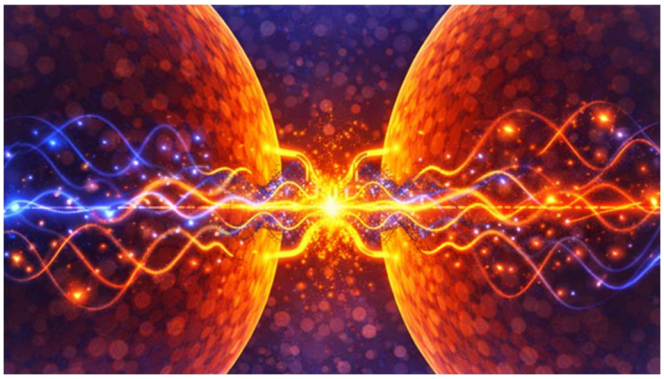


Figure 2. Thermal and quantum vacuum field fluctuations within a nanoscale gap of the prebiotic protocell dimer membranes

This paper explores exactly that scenario [20]. We show that protocell clusters, which can arise spontaneously under prebiotic conditions through Casimir-Lifshitz attraction, induce selective modulation of electromagnetic fluctuations in their nanometer-sized gap regions. Under saline "primordial soup" conditions with ionic strengths of 50–200 mMol, the static Matsubara mode is completely suppressed by Debye screening. As a result, only the fast dynamic modes of the thermal quantum field govern the interaction between cells. These modes create geometry-dependent resonance zones where local vacuum energy is reduced, enabling hydrated protons to accumulate due to their high polarizability.

In Figure 2, two spherical protocells form a narrow intermembrane gap (L= 2-50 nm) in which thermal and quantum electromagnetic fluctuations are geometrically confined. The boundary conditions imposed by the opposing membranes selectively reduce the density of dynamic electromagnetic vacuum modes within the gap. This mode suppression creates a localized minimum of vacuum energy, which gives rise to an attractive Casimir-Lifshitz force that stabilizes the dimer configuration.

The aim of this paper is to develop this quantum-field mechanism systematically and to show that protocell assemblies can generate proton reservoirs solely through their nanoscale structure. These reservoirs serve as plausible precursors to chemiosmotic gradients. This reveals a potentially universal physical process that could have supported early energy provision and functional coupling among protocells, offering a new foundation for prebiotic self-organization and early bioenergetic evolution.

III. The Role of Prebiotic Energetics for Protocells

Bioenergetics is one of life's most universal foundations. All cells use ion gradients—especially proton gradients—to perform work, drive metabolism, and store energy in chemical form. The combined pH gradient (ΔpH) and membrane potential (ΔΨ) form the proton-motive force, which powers ATP synthase and many other processes. This principle is conserved across all domains of life, from bacteria to mitochondria. Proton gradients are therefore a central energetic currency of biology.

This raises a key question in origins-of-life research: how could the first ion gradients arise before specialized membrane proteins, transport channels, or metabolic networks existed? Conventional models often invoke external sources such as geochemical gradients at hydrothermal vents or mineral interfaces. These scenarios provide macroscopic gradients but do not explain how nanoscale, locally stabilized gradients could emerge directly at or between protocells—nor how such gradients could persist long enough to enable early energy conversion rather than dissipating immediately.

Recent proto-physiological work shows that primitive protocells can generate stable ion asymmetries without membrane proteins or transporters. Matveev [24], for example, demonstrated that proteinoid-based model cells can concentrate K^+ ions up to 1600-fold relative to the medium—solely through sorption and phase organization, without pumps or channels. Protocells must therefore be viewed as non-membranous biophysical phases capable of generating electrostatic potentials, selective ion distributions, and energetic nonequilibrium states through their internal structure.

This perspective supports the central thesis of this paper: proton gradients in protocell clusters may not require biological origin. They could arise from the physical structure of vesicle assemblies and their quantum-geometric interactions with electromagnetic fluctuation modes. Protocells could thus have generated physically driven, locally stabilized proton reservoirs long before true proton pumps evolved.

The present paper addresses precisely this research gap. We investigate how quantum- and thermally induced fluctuation forces—specifically Casimir-Lifshitz interactions in saline prebiotic environments—could shape the spatial organization and energetic dynamics of protocells. These forces act at nanometer separations and can determine both stabilization and relative arrangement. In such cluster geometries, the electromagnetic mode spectrum is modified in ways that generate proton fluxes, nanofluidic transport, and local potential minima.

This leads to an alternative view of prebiotic bioenergetics: proton gradients may not be purely biological innovations but emergent products of physical coupling, fluctuation dynamics, and dissipative organization in protocell clusters. For biologists, this suggests that essential elements of chemiosmotic energy conversion may have roots in the physical behavior of simple vesicles. For computer scientists, such gradients illustrate how early protocell assemblies could stabilize and couple states through field structures. For sustainability science, ion gradients appear as early examples of nonequilibrium-driven self-stabilization—a process reflecting core features of prebiotic resilience and energy efficiency.

IV. THE SPECIAL GEOMETRY OF PROTOCELL CLUSTERS

Massoth [2] showed that protocells in prebiotic environments likely did not exist as isolated compartments. Theoretical models and recent experimental analogues indicate that nanoscale vesicles and amphiphile aggregates can spontaneously assemble into clusters in saline solutions through Casimir–Lifshitz attraction. This attraction arises from the reduction of electromagnetic fluctuations in the gap between two membranes and makes certain geometries energetically favored. Two simples, but structurally central clusters emerge: the protocell dimer and the protocell tetrahedron.

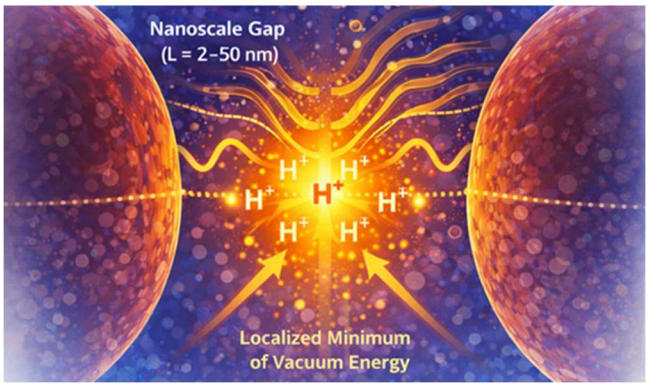


Figure 3. Localized minimum of quantum vacuum energy at the nanoscal gap of the prebiotic protocell dimer spheres with membranes

A dimer forms when two protocells approach within a few nanometers [11]. This produces a well-defined gap region, typically L=2–50 nm [22] wide depending on vesicle size, membrane stiffness, and ionic conditions. This distance lies squarely in the regime where quantum-field mode selection is strongest and dynamic Matsubara modes dominate the electromagnetic field. The intermembrane gap thus acts as a nanoscale resonance cavity that shapes field fluctuations and can generate proton potentials along a one-dimensional axis. Because of their frequency and mechanical stability, dimers represent the simplest and most common form of protocellular cooperation.

In Figure 3, resonance zones emerge in which relevant QED vacuum fluctuations are strongly suppressed. This creates local minima of vacuum energy. Hydrated protons drift passively into these minima and accumulate in a stable manner. Casimir-Lifshitz coupling drives nanoscale quantum-geometric confinement and proton (H^+) enrichment. Attraction between adjacent protocell membranes (L = 2–50 nm) concentrates protons within the intermembrane gap and generates localized electrochemical potentials.

A more structurally influential motif is the protocell tetrahedron—a three-dimensional arrangement of four vesicles whose centers form an almost regular tetrahedron. Such configurations arise when several protocells interact simultaneously through Casimir-Lifshitz attraction. The pairwise gaps again fall within the critical 5–50 nm range, while the center of the tetrahedron becomes a three-dimensional resonance space enclosed by four interfaces. This geometry suppresses selected electromagnetic modes far more strongly than in a dimer, creating a pronounced energetic minimum in the tetrahedral core.

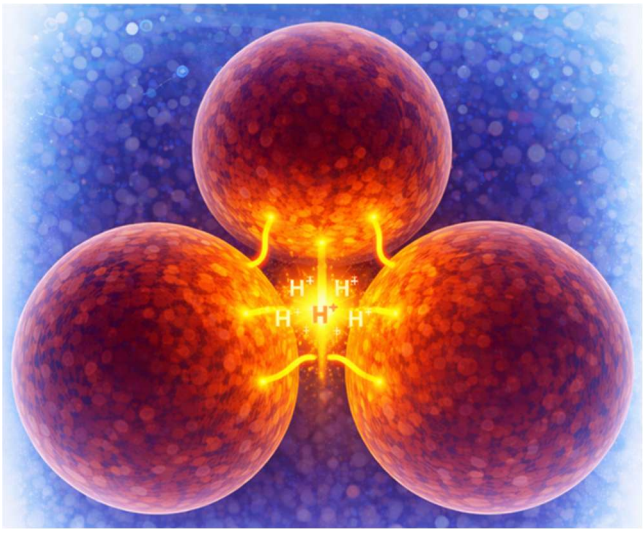


Figure 4. Resonance zone within a protocell tetrahedron

In Figure 4, four protocell spheres (one not visible) with membranes form a tetrahedral geometry. This multi-interface confinement suppresses dynamic vacuum modes and produces a central energy minimum. A three-dimensional resonance zone forms at the geometrical center. The resulting minimum of vacuum energy stabilizes proton (H^+) accumulation within the tetrahedral core. Generating local electrochemical potentials within the resonance zone.

Both geometries—dimer and tetrahedron—are therefore not only energetically preferred but also physically significant. They define the nanoregions in which the electromagnetic field of the primordial environment is most strongly filtered, modulated, and focused. The resulting resonance zones form the foundation for the quantum-field-induced proton reservoirs examined in the following sections.

V. METHODS, NUMERICAL CALCULATIONS AND ORDER-OF-MAGNITUDE VALIDATION

Model Assumptions and Limitations:
These assumptions define a minimal physical model aimed at isolating the contribution of quantum-geometric effects. Protocell radii of 500–1000 nm, intermembrane separation gap L of 2–100 nm, temperatures of 20–80 °C, and ionic strengths of 50–200 mMol define a physically plausible prebiotic parameter window in which Casimir–Lifshitz interactions can contribute to non-chemical mesoscale aggregation. Protocell membranes are treated as smooth dielectric interfaces. Simple fatty-acid membranes exhibit low effective permittivities ($\varepsilon_{membrane} \approx 2–4$), well below that of saline water ($\varepsilon_{water} \approx 75–78$), satisfying the sign condition for attractive Casimir–Lifshitz forces. Protein–membrane-based compartments (PMBCs), with higher effective permittivities ($\varepsilon_{membrane} \approx 3–8$) and increased mechanical stability, are expected to exhibit particularly pronounced and experimentally accessible interactions. Hydrated protons are represented by an effective polarizability, and the electrolyte is described within the linear Debye–Hückel regime. All numerical calculations are obtained using an effective Hamaker constant $A_{eff} = 5 \times 10^{-21}$ Joule, consistent with reported values for membrane–water systems. No active transport, chemical buffering, or metabolic processes are included, and cluster geometries are assumed static on proton diffusion timescales.

Derjaguin proximity-force approximation (PFA):
In the biologically relevant regime $L \ll R$ (here L=2–100 nm and R=500–1000 nm), the exact Casimir–Lifshitz description for a sphere–medium–sphere system can be reduced, via the Derjaguin proximity-force approximation (PFA), to a simple scaling form:

$$F_{CL}(L) \propto A_{eff} R_{eff}/L^2. \quad (1)$$

Here, the effective curvature radius R_{eff} acts as the dominant amplification factor for biological cluster stability [13]. Larger protocells therefore generate systematically stronger coupling at identical material parameters. The following applies:

$$F_{CL}(L) \approx -(A_{eff}/6) * (R_{eff}/L^2) \quad (2)$$
$$\text{with } R_{eff} = (R_1 * R_2)/(R_1 + R_2).$$

Here, A_{eff} denotes an effective Hamaker constant that integrates the spectral dielectric response of the membrane–water system. The force scales linearly with the effective curvature radius R_{eff} and decays algebraically as $1/L^2$.
For the parameter ranges considered below, the resulting potential wells reach several k_BT, making them relevant for mesoscale cluster stabilization.

The corresponding effective binding potential $U_{CL}(L)$ follows directly from integration:

$$U_{CL}(L) \propto -(A_{eff}/6) * (R_{eff}/L), \quad (3)$$

with A_{eff} the membrane–water Hamaker constant, R_{eff} the reduced curvature of the two spheres, and L the minimal surface-to-surface distance. The approximation holds for $L \ll R_i$, L larger than the membrane thickness, and smooth, non-adhesively functionalized interfaces.

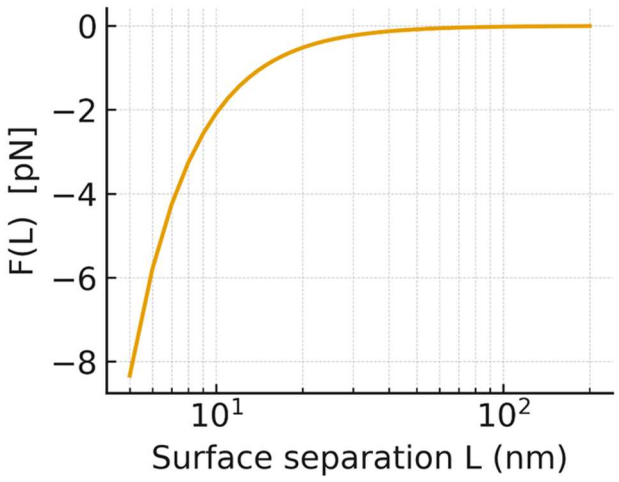


Figure 5. Attractive Casimir–Lifshitz force $F_{CL}(L)$ between two PMBC-like protocells (R_1=R_2=500 nm) as a function of separation L

Figure 5 shows the attractive Casimir‑Lifshitz force $F_{CL}(L)$ between two PMBC-like protocells (R_1=R_2=500 nm) as a function of surface separation L (logarithmic x-axis). The algebraic decay $F_{CL}(L) \propto 1/L^2$ produces significant mesoscale attraction over L=5‑100 nm.

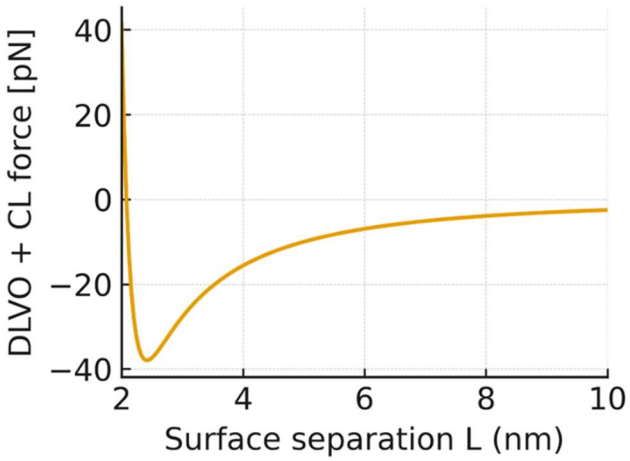


Figure 6. Resulting total force from F_{DLVO} and F_{CL} contributions over L=2–10 nm

Comparison with classical interaction mechanisms: For protocell radii of ~500–1000 nm and separations of ~2–200 nm, a regime emerges in which Casimir–Lifshitz forces become comparable to, or exceed, residual Derjaguin–Landau–Verwey–Overbeek (DLVO) contributions. At prebiotically realistic distances ($L \gtrsim 2$–100 nm), the Casimir–Lifshitz component provides a non-DLVO-compatible, algebraically decaying attraction that persists beyond screened electrostatic interactions. The combined force landscape yields detectable attractive wells (up to ~−38 pN) that stabilize bound protocell configurations against Brownian motion, as calculated and shown in Figure 6. These results indicate that, under saline early-Earth conditions, Casimir–Lifshitz interactions constitute a robust physical aggregation mechanism without requiring universal dominance over classical forces.

VI. Quantum-Physical Derivation of Proton (H^+) Gradients

Proton (H^+) Behavior between two Casimir Plates:
As a first step, consider two parallel Casimir plates immersed in a saline aqueous environment that models a prebiotic soup with ionic strengths of 50–200 mMol. These plates represent protocell membranes or simplified dielectric boundaries. Their separation lies in the nanoscale range (L = 2–100 nm, typically 2–50 nm), where both classical electrostatics and quantum-field-driven Casimir–Lifshitz forces act.
In free vacuum, quantum electromagnetic fluctuations are isotropic. Between two Casimir plates, this symmetry breaks. Only virtual photon modes compatible with the boundary conditions remain allowed. The result is a thinned and shifted spectrum of dynamical vacuum fluctuations.
In an electrolyte, Debye screening adds a second effect. Slow, quasi-static fields are suppressed over 1–2 nm [9]. In Casimir–Lifshitz terms, this eliminates the static thermal Matsubara mode (n = 0) for separations $L \gg 1$–2 nm. Thus only dynamical Matsubara modes ($n \geq 1$) define the effective vacuum field. Their imaginary frequencies are:

$$\xi_n = 2\pi n k_B T/\hbar \text{ with } n \geq 1, \quad (4)$$

where k_B is Boltzmann's constant and T is the temperature.
These modes lie in the far infrared and terahertz range and generate the characteristic quantum dynamic structure in the gap. The resulting field-induced energy density [21] can be written as a sum over the local mode densities $f_n(z)$ at position z:

$$U_{field}(z) \approx k_B T \sum_{n=1}^{\infty} f_n(z), \quad (5)$$

where $f_n(z)$ quantifies the presence of the nth Matsubara mode. Boundary conditions, material properties, and plate separation determine where modes are suppressed. Minima of this potential define quantum-induced resonance zones—regions of reduced vacuum energy and shortened fluctuation spectra.
The dynamical ($n \geq 1$) modes respond sensitively to nanoscale geometry. Because they are not screened in electrolytes, they shape the entire local quantum spectrum. Mode selection acts as a geometric filter that suppresses some field modes while amplifying others. Regions of maximal mode suppression form stable resonance minima with lowered vacuum energy. These minima are robust against thermal noise and produce defined proton reservoirs whose concentrations rise exponentially with the depth of the mode-suppression well.

The key question is how protons (H^+) behave in this quantum-shaped landscape. In water, protons exist as hydrated, highly polarizable clusters [23]. They couple strongly to the fast ($n \geq 1$) Matsubara modes. Their interaction with the modified vacuum field is captured by a Casimir–Polder expression:

$$U_{CP}(z) \approx \frac{1}{2} k_B T \sum_{n=1}^{\infty} \alpha(i\xi n) fn(z), \quad (6)$$

where $\alpha(i\xi_n)$ is the frequency-dependent polarizability of the hydrated proton cluster and $f_n(z)$ the local mode density.
Because $\alpha(i\xi_n) > 0$, protons (H^+) prefer regions where the sum $\sum_{n=1}^{\infty} \alpha(i\xi n) fn(z)$ is minimal [23]: $U_{CP}(z_{min}) = \min_z U_{CP}(z)$.

In symmetric plate–medium–plate systems, this minimum typically lies at the center of the gap, forming a "proton reservoir" without any membrane potential or protein channels.
The stationary proton distribution follows a Boltzmann profile [23]:

$$P(z) \propto \exp[-U_{CP}(z)/k_B T]. \quad (7)$$

Even modest potential shifts of a few $k_B T$ yield exponential proton (H^+) enrichment. For separations of 5–50 nm, the resonance zone accumulates markedly elevated proton concentrations—a quantum-geometric precursor of chemiosmotic gradients.

Interim conclusion: In saline prebiotic media, Debye screening removes the static (n = 0) Matsubara mode. Only dynamical ($n \geq 1$) Matsubara modes contribute to Casimir–Lifshitz forces. These modes shape a spatially varying quantum potential in which protons (H^+) drift toward resonance minima through Casimir–Polder coupling. The result is a purely physical mechanism that generates nanoscale proton (H^+) reservoirs—an early form of prebiotic energy organization.

Proton (H^+) behavior in the reaction zones of protocell dimers and tetrahedra:

Extending the plate model to realistic protocell clusters shows that dimer and tetrahedral configurations strongly structure the accessible electromagnetic fluctuations. Because the n = 0 Matsubara mode is absent under saline conditions, reaction zones are shaped exclusively by the dynamical (n ≥ 1) Matsubara modes. These far infrared and terahertz modes define the local quantum-field landscape for proton (H^+) motion.

As before, the effective proton (H^+) potential is governed by the reduced local mode density $f_n(r)$:

$$U_{CP}(z) \approx \frac{1}{2} k_B T \sum_{n=1}^{\infty} \alpha(i\xi n) fn(z). \quad (8)$$

Resonance zones occur where the mode-suppression term reaches its minimum. Hydrated protons (H^+) respond to these nanoscale potentials according to the Casimir–Polder interaction. Because $\alpha(i\xi_n) > 0$, protons accumulate where fluctuating modes are maximally damped.

Protocell Dimer: One-dimensional Proton (H^+) Guidance

In a dimer, two opposing membranes separated by L form an ordered nanoscale slit. A linear resonance zone develops along the symmetry axis. Dynamical modes (n ≥ 1) are strongly suppressed near the midline, producing a pronounced minimum in $U_{CP}(z)$. The steady-state proton (H^+) distribution obeys:

$$P(z) \propto \exp[-U_{CP}(z)/k_B T]. \quad (9)$$

This creates a one-dimensional proton conduit—a quantum-geometric "proto-proton (H^+) channel" that forms without proteins.

Protocell Tetrahedron: Three-dimensional Resonance Well

In a tetrahedron, modes from four interfaces interfere. Multiple scattering produces deep suppression of specific modes at the center. The resulting potential shows a strong three-dimensional minimum:

$$U_{CP}(\mathbf{r}_{Tetrahedron\text{-}center}) = \min_r U_{CP}(\mathbf{r}). \quad (10)$$

This forms a three-dimensional proton funnel. Proton (H^+) density in the tetrahedral core can exceed the bulk medium outside the tetrahedron cluster by several orders of magnitude:

$$P_{center} / P_{bulk} \propto \exp[\Delta U / k_B T], \quad (11)$$

where ΔU is the depth of the resonance well. Larger ΔU produces exponentially greater proton (H^+) enrichment.

Prebiotic physical implications:

The two geometries generate distinct proton (H^+) landscapes:

- Dimer: directed one-dimensional proton flow
- Tetrahedron: a three-dimensional proton hub

Both structures produce proton (H^+) gradients through geometry and quantum fluctuations alone—a possible physical origin of chemiosmotic organization, predating biological proton pumps.

The total energy of an N-protocell cluster in close contact scales with the number of membrane–membrane bonds N_{bonds} (N):

$$\begin{aligned} E_N(L) &\approx -N_{bonds}(N) * I U_{CL}(L) I \\ &= -N_{bonds}(N) * (A_{eff}/6) * (R_{eff}/L) \\ &= -N_{bonds}(N) * \Lambda(R,L) * k_B T. \end{aligned} \quad (12)$$

For R = 500 nm, L = 10 nm, and $A_{eff} = 5\times10^{-21}$ J, a dimer already binds with ≈ −5.1 $k_B T$, much higher than the thermal noise fluctuations of +1.0 $k_B T$ at ~25 °C.

TABLE I. TOTAL CASIMIR–LIFSHITZ INTERACTION ENERGIES

N	Protocell Structure	N_{bonds}	$E_N/k_B T \approx$
2	Dimer	1	−5.1
3	Triangularer Trimer	3	−15.2
4	Tetrahedron	6	−30.5
6	Oktahedron	12	−61.0
7	Pentagonal Bipyramid	15	−76.2
13	Icosahedral 13-Cluster	42	−213.4

Table I shows the total Casimir–Lifshitz interaction energies $E_N/k_B T$ for representative N-protocell cluster geometries, illustrating how the number of pairwise bonds Nbonds drives mesoscale stabilization.

A potential well of -5 $k_B T$ yields roughly 150-fold proton (H^+) enrichment in a tetrahedral reaction zone. Even such modest depths allow protocells to stabilize reaction sites, catalyze proton-driven processes, support local energy potentials, generate pH differences of 2–3 units, and initiate primitive metabolism-like behavior.

Order-of-magnitude estimate:

The depth of the geometry-induced Casimir–Polder potential well is expected to lie in the range ΔU ≈ 1–5 $k_B T$ for protocell dimers and ΔU ≈ 3–8 $k_B T$ for tetrahedral clusters, depending on the intermembrane separation L and cluster geometry. For narrow gaps (L ≈ 2–10 nm), stronger mode suppression yields ΔU ≈ 3–5 $k_B T$ in dimers and ΔU ≈ 5–8 $k_B T$ in tetrahedra, whereas wider gaps (L ≈ 20–50 nm) correspond to ΔU ≈ 1–3 $k_B T$ and ΔU ≈ 3–5 $k_B T$, respectively. The resulting proton enrichment follows a Boltzmann factor $P_{local}/P_{bulk} \approx \exp(\Delta U/k_B T)$, implying concentration ratios from order-unity up to ~10^2 for dimers and up to ~10^3 for tetrahedra.

These enrichments correspond to local pH shifts of approximately ΔpH ≈ 0.5–2 for dimers and ΔpH ≈ 1–3 for

tetrahedral resonance zones, indicating that nanoscale protocell clusters can sustain biologically relevant proton gradients without active transport mechanisms.

VII. PREBIOTIC SIGNIFICANCE AND IMPLICATIONS

The quantum-field-induced proton reservoirs that form in protocell clusters offer a new perspective on prebiotic energetics. Stable proton gradients can arise solely from geometry, polarizability, and thermal quantum fluctuations. Energetic states therefore did not require early biological machines but may have been embedded directly in the physical structure of protocell assemblies. Proton traps in dimers and proton funnels in tetrahedra create local pH centers that promote and stabilize reactions, providing energetic advantages without pumps or membrane proteins.

These mechanisms extend beyond energy supply. In modern cells, proton gradients act not only as energy sources but as key regulatory variables. Quantum-field-based reservoirs may therefore have enabled early self-organization, coordination, and coupling within protocell groups. Dimers would serve as directed energetic channels. Tetrahedra would act as three-dimensional reaction hubs supporting rudimentary metabolic steps. Protocells appear not as isolated compartments but as cooperative micro-assemblies sharing energy and state information through quantum-modulated nanoscale gaps.

The informational implications are similar. Proton gradients in resonance zones can function as early physical memory or state variables, analogous to analog nodes in distributed systems. Their stability and geometric coupling support proto-informational states without polymers or genetic coding. In this sense, energy gradients themselves may represent one of the earliest forms of functional information, consistent with constructor-theoretic ideas.

From a sustainability perspective, these gradients illustrate that core principles of efficient energy systems—generation, storage, controlled dissipation, and resilience—may already have been realized at prebiotic nanoscales. Protocell clusters behave as open dissipative systems that capture energy, structure it, and use it locally. Quantum-field-driven gradient formation thus represents a proto-sustainable organizational mode: a minimal mechanism generating energetic stability in a fluctuating environment without biological infrastructure.

The emergence of quantum-field-induced proton reservoirs transforms protocell clusters from passive aggregates into active energetic systems. By stabilizing localized proton gradients through geometry alone, protocell assemblies acquire functional, non-equilibrium states that persist over biologically relevant timescales. These states represent a physically grounded precursor to chemiosmotic organization and mark a critical step toward proto-living behavior, in which energy, structure, and function become intrinsically coupled.

Overall, these results suggest that early steps toward life were strongly shaped by the capacity of simple protocell assemblies to build quantum-geometrically defined energy structures. Proton reservoirs in the nanoscale gaps of dimers and tetrahedra may have marked the transition from chemistry to proto-biology—a regime in which energy, geometry, and quantum-field structure first merged into functional self-organization.

VIII. CONCLUSION AND FUTURE WORK

The results presented here fit into a growing line of research that places physical mechanisms at the center of life's emergence. Classical models of prebiotic energetics often invoke geochemical or molecular gradients. Our analysis shows instead that protocell clusters can become active energetic landscapes themselves. Quantum-field-induced proton reservoirs provide a plausible route toward early chemiosmotic precursors, long before biological proton pumps existed. This offers a physics-based explanation for how early protocells could stabilize their energy budgets and couple to one another.

The underlying mechanisms are experimentally accessible today. Liposome-based model systems can generate nanoscale gaps of L=2–50 nm, where Casimir–Lifshitz forces become measurable. Nanofluidic platforms can resolve proton distributions and pH profiles inside these gaps. Terahertz and infrared spectroscopy can characterize the mode selection of dynamic fluctuations. Together, these tools allow direct tests of the predicted resonance zones and proton traps. Empirical validation of quantum-field-bound energy structures is therefore within reach.

This mechanism does not imply sustained transmembrane chemiosmosis but local, geometry-bound proton reservoirs.

The implications extend beyond origin-of-life studies. In synthetic biology, artificial protocell systems could be engineered to exploit quantum-geometric proton gradients for energy storage, catalytic enhancement, or stabilization of functional states. Such "synthetic protocells" would not rely on biological machines but draw usable energy directly from their physical arrangement. Likewise, the findings point to sustainable nanoscale energy concepts based on dissipative self-organization without external pumping mechanisms.

This work thus highlights a broad research frontier where physics, biology, nanotechnology, and information science converge: the exploration of quantum-geometric energy structures as foundational elements of both prebiotic and synthetic life.

REFERENCES

[1] M. Massoth, "Attractive Casimir-Lifshitz Forces as a Universal Driver of Prebiotic Protocell Aggregation and Cluster Formation," *BIOTECHNO*, 2026.

[2] M. Massoth, "Emergent Information Formation in Prebiotic Protocell Clusters: A Computational Mechanics Framework of ε-Machines and Attractor Memory," *BIOTECHNO*, 2026.

[3] M. Massoth, "From Physical Difference to Meaning: A Constructor-Theoretic Framework for Prebiotic Information in Casimir–Lifshitz-Coupled Protocell Clusters," *BIOTECHNO*, 2026.

[4] M. Massoth, “Physical Origin of Proto-Information: Syntax, Semantics and Pragmatic Emergence in Prebiotic Protocell Clusters,” *BIOTECHNO*, 2026.

[5] F. E. Rosas, *et al.*, “Software in the natural world: A computational approach to hierarchical emergence,” *arXiv* preprint, 2024, doi:10.48550/arXiv.2402.09090.

[6] M. Deutsch and C. Marletto, “Construction theory of information,” *Proc. R. Soc. A*, vol. 471, art. no. 20140540, 2015, doi:10.1098/rspa.2014.0540.

[7] H. B. G. Casimir, “On the attraction between two perfectly conducting plates,” *Proc. K. Ned. Akad. Wet.*, vol. 51, pp. 793–795, 1948.

[8] E. M. Lifshitz, “The theory of molecular attractive forces between solids,” *Sov. Phys. JETP*, vol. 2, pp. 73–83, 1956.

[9] J. N. Israelachvili, *Intermolecular and Surface Forces*, 3rd ed., Academic Press, London, 2011.

[10] P. C. W. Davies, “Does quantum mechanics play a non-trivial role in life?” *Biosystems*, vol. 78, pp. 69–79, 2004.

[11] I. Gözen, *et al.*, “Protocells: Milestones and recent advances,” *Small*, vol. 18, art. no. 2106624, 2022.

[12] S. K. Lamoreaux, “Demonstration of the Casimir force in the 0.6–6 μm range,” *Phys. Rev. Lett.*, vol. 78, pp. 5–8, 1997.

[13] J. L. Garrett, D. A. T. Somers, and J. N. Munday, “Measurement of the Casimir force between two spheres,” *Phys. Rev. Lett.*, vol. 120, art. no. 040401, 2018.

[14] P. Rodriguez-Lopez, “Casimir energy and entropy in the sphere–sphere geometry,” *Phys. Rev. B*, vol. 84, art. no. 075431, 2011.

[15] S. J. Rahi, T. Emig, N. Graham, R. L. Jaffe, and M. Kardar, “Scattering theory approach to electrodynamic Casimir forces,” *Phys. Rev. D*, vol. 80, art. no. 085021, 2009.

[16] G. Bimonte and T. Emig, “Exact results for classical Casimir interactions: Dirichlet and Drude model in the sphere–sphere and sphere–plane geometry,” *Phys. Rev. Lett.*, vol. 109, art. no. 160403, 2012.

[17] A. Canaguier-Durand, G.-L. Ingold, M.-T. Jaekel, A. Lambrecht, P. A. Maia Neto, and S. Reynaud, “Classical Casimir interaction in the plane–sphere geometry,” *Phys. Rev. A*, vol. 85, art. no. 052501, 2012.

[18] B. B. Machta, S. L. Veatch, and J. P. Sethna, “Critical Casimir forces in cellular membranes,” *Phys. Rev. Lett.*, vol. 109, art. no. 138101, 2012.

[19] D. S. Ether, I. S. Nogueira, and H. M. Nussenzveig, “Probing the Casimir force with optical tweezers,” *Europhys. Lett.*, vol. 112, art. no. 44001, 2015.

[20] S. Pei, Z. Zhang, T. Li, *et al.*, “Direct measurement of the van der Waals force between a pair of microspheres based on photonic force microscopy,” *Opt. Eng.*, vol. 60, art. no. 084101, 2021.

[21] M. Boström and B. E. Sernelius, “Thermal effects on the Casimir force in the 0.1–5 μm range,” *Phys. Rev. Lett.*, vol. 84, pp. 4757–4760, 2000.

[22] H.-J. Butt, “Measuring electrostatic, van der Waals, and hydration forces in electrolyte solutions with an atomic force microscope,” *Biophys. J.*, vol. 60, pp. 1438–1444, 1991.

[23] S. Cukierman, “Et tu, Grotthuss! And other unfinished stories,” *Biochim. Biophys. Acta*, 2006, doi:10.1016/j.bbabio.2005.12.001.

[24] V. V. Matveev, “Comparison of fundamental physical properties of the model cells (protocells) and living cells reveals the need for protophysiology,” *Int. J. Astrobiol.*, vol. 16, no. 1, pp. 97–106, 2017, doi:10.1017/S1473550415000476.

Prebiotic Sustainability in the Nanoworld: A Physical Framework for Early Protocell Cluster Stability

Michael Massoth
Department of Computer Science, Hochschule Darmstadt (h_da)
University of Applied Sciences Darmstadt, member of European University of Technology (EUt+)
Darmstadt, Germany
e-mail: michael.massoth@h-da.de

***Abstract*- This paper develops a physically grounded concept of prebiotic sustainability in the nanoworld. It shows how early protocell clusters were able to form stable, persistent, and regenerative states of order through fluctuation-induced forces, nanofluidic flows, and proton-based gradients—before genes or enzymes. The model describes sustainability as an emergent property of open non-equilibrium systems and integrates three dimensions: structural, energetic, and informational sustainability. The resulting framework enables testable hypotheses about stability, gradient duration, and regeneration processes. Thus, the approach opens up new experimental avenues for protocellular self-preservation and provides criteria for evaluating early prebiotic scenarios.**

Keywords-Prebiotic sustainability; protocells; Casimir–Lifshitz forces; ion gradients; dissipative structures; origins of life.

I. INTRODUCTION

This work is the sixth paper in a seven-paper series: "A Constructivist Proto-Bio-Information Theory: A Physically Grounded Nano-Systems Architecture for Prebiotic Emergence, Information, Proto-Semantic Function, and Sustainability of Protocell Aggregation and Cluster Formation".

Massoth [1] shows that Casimir–Lifshitz forces generate robust attraction and stable protocell clusters at 2–100 nm under prebiotic conditions. This coupling provides the structural basis on which this paper formulates prebiotic sustainability and long-term cluster stability.

Massoth [2] demonstrates that such clusters form reproducible mesoscale attractors and autonomous ε-machine dynamics. These attractor-based macrostates supply the informational substrate that this paper [2] extends into a dimension of prebiotic sustainability.

Massoth [3] shows that protocell clusters generate reproducible differences and functional meaning states. These functional states anchor the informational sustainability developed in this paper [3] as a physically stabilizable proto-function.

Massoth [4] reveals that structured resonance zones and ε-machines produce syntactic, semantic, and pragmatic layers of proto-biological information. This tri-level architecture underpins the present work, which frames sustainability as the emergent stabilization of such multilayered information processes.

Massoth [5] identifies Matsubara-mode selection and Casimir–Polder coupling as sources of stable proton (H^+) gradients in dimers and tetrahedra. These quantum-field-driven energy structures form the energetic core that this paper develops into "energetic sustainability" and links to experimental tests.

The structure of the paper is as follows: Section I presents the motivation, state of the art, and introduces prebiotic sustainability as a physical non-equilibrium principle. Section II develops the nanophysical baseline model and derives fluctuation-induced coupling, flows, and gradient persistence conditions. Section III defines prebiotic sustainability formally and decomposes it into structural, energetic, and informational dimensions. Section IV introduces a three-level sustainability framework and maps it to reliability-engineering concepts and metrics. Section V operationalizes sustainability via a normalized structural distance and experimentally accessible order parameters. Section VI formulates testable hypotheses and an experimental roadmap for validating prebiotic sustainability. Section VII summarizes implications, limitations, and future directions for origins-of-life and synthetic protocell research.

II. MOTIVATION AND RELEVANCE

The origin of life is usually described biochemically—through prebiotic syntheses, reaction networks, and early compartments. Yet these models explain only partially how nanoscale protocells stayed stable in a fluctuating, salt-rich environment, or how early energetic and functional patterns emerged. A clear physical framework is still missing, one that specifies the conditions under which such systems could maintain structures and gradients long before genes or enzymes existed.

This paper introduces prebiotic sustainability in the nanoworld: the ability of an open nanoscale protocell system with typical radii of R=200-1000 nm to preserve its structural and functional identity despite fluctuations, dissipation, and continuous energy flow. Sustainability is used here not normatively, but as an emergent non-equilibrium principle

that links biophysics, origins-of-life research, and sustainability science.

From an engineering perspective [21], the proposed notion of prebiotic sustainability can be read as reliability under non-equilibrium constraints: stability corresponds to low failure probability under perturbations, persistence to extended state retention times, and regenerability to recovery after disruption. This interpretation aligns with reliability engineering views that treat complex systems as coupled structures of hardware, environment, and stochastic errors, and motivates measurable *failure* and *recovery* metrics for protocell clusters.

The approach integrates key physical mechanisms established earlier: fluctuation-induced coupling such as Casimir–Lifshitz forces [8][9] that drive robust aggregation at 2–100 nm [1]; mesoscale attractors that form stable, information-bearing macrostates [2]; and quantum-field-induced proton reservoirs [3] that create early bioenergetic gradients. Together, these results show that protocell assemblies were neither chemically isolated nor energetically disordered, but could form stable patterns through fluctuations, flows, and field modulation [6].

This raises a central question: How can stability in the prebiotic nanoworld be defined in a way that is measurable and captures physical, energetic, and informational processes at once? Classical models describe local structure but not its persistence in open, noise-dominated environments. Yet without sustained stabilization of membranes, clusters, and gradients, early selection and functionality would have been unlikely. The physical meaning of prebiotic sustainability as an emergent non-equilibrium property is summarized schematically in Figure 1.

Figure 1. The Concept of Prebiotic Sustainability

In Figure 1, a conceptual illustration of prebiotic sustainability in the nanoworld is shown. A protocell dimer forms the minimal stable unit within an open, non-equilibrium environment. Fluctuation-induced forces drive nanofluidic flows and proton gradients, generating persistent patterns despite continuous dissipation. Sustainability emerges as the coupled stabilization of structural integrity, energetic gradients, and informational pattern recurrence under sustained inflow and entropy outflow, without genetic or enzymatic control.

Here, sustainability is defined as a three-layered property: structural sustainability (persistence of membranes and clusters), energetic sustainability (lifetimes of proton-driven gradients), and informational sustainability (robustness of recurrent state patterns without genes).

This perspective provides a physical foundation for protocell self-maintenance. Stability emerges as a natural outcome of energy flow, dissipative self-organization, and nanoscale coupling. It also points toward "synthetic sustainable protocells," engineered systems that reproduce stable structures and gradient profiles.

Prebiotic sustainability thus appears as a general physical ordering principle—active long before biological complexity and potentially critical for the transition from simple self-organization to early bioenergetic and informational functions.

III. State of the Art

This section provides a formal definition of prebiotic sustainability and introduces a taxonomy that separates structural, energetic, and informational dimensions as experimentally accessible criteria.

A. Origins-of-Life Approaches

Origins-of-life research is commonly divided into three major lines: chemical evolution, compartment-based models, and bottom-up synthetic biology. These form the conceptual background within which a physical notion of sustainability for protocells can be developed.

Chemical evolution studies how organics arise from simple precursors; Miller–Urey showed that amino acids form under strong energy fluxes. Amphiphiles later became central because they self-assemble into membranes and create isolated reaction spaces.

Compartment models view protocells [10] as simple physicochemical interfaces. Fatty-acid vesicles form, grow, and encapsulate solutes; coacervates show strong concentration effects; polymer- and protein-based PMBCs offer more durable systems for stability studies. Bottom-up synthetic biology builds "minimal cells" that reproduce basic functions such as primitive metabolism or gradient formation, demonstrating that simple compartments can organize energy flow.

All three approaches focus on chemistry and reaction spaces, but nanoscale physical mechanisms—fluctuation forces [10], field modulation, nanofluidic flows, gradient-stabilizing couplings—remain understudied. This paper addresses this gap by integrating these effects into a physical model of prebiotic sustainability.

B. Dissipative Structures according to Prigogine

A second foundation comes from non-equilibrium thermodynamics and Prigogine's dissipative structures. Systems far from equilibrium can form stable order when

energy and matter continuously flow; classic examples include Bénard convection, chemical oscillations, and autocatalytic networks. These structures persist only while energy is supplied.

Applied to protocells, this suggests that life began as a persistent non-equilibrium state driven by geochemical gradients, light, and thermal fluctuations. Dissipation is essential, not a defect.

Prebiotic sustainability describes this regime: energy flows maintain membranes, clusters, and gradients despite constant internal motion. The system changes but stays recognizably the same. Early protocell clusters likely behaved this way—open, fluctuation-driven systems that produced stable patterns across diffusive timescales. Dissipative structures thus provide the thermodynamic basis for a testable concept of prebiotic sustainability.

C. Reliability Engineering

Reliability engineering [21] provides a mature language for quantifying failure, recovery, and robustness in noisy, safety-critical open systems. In particular, recent surveys emphasize the need for formal definitions, taxonomies of errors, and measurable reliability metrics, which directly motivates the present operationalization of sustainability via stability, persistence, and regenerability.

IV. A Physical Model for Prebiotic Sustainability in the Nanoworld

The model of prebiotic sustainability assumes that early protocells existed in an ionic, strongly fluctuating environment—without enzymes, genes, or active pumps. The baseline scenario considers two vesicles of different radii (R_1, R_2), typically 200–1000 nm in size. They reside in water at separations of L=2–100 nm, exactly the range where Casimir and Lifshitz forces [9] act while classical DLVO interactions are strongly screened at ionic strengths typical of plausible prebiotic brines (Debye length ~1 nm). Ions diffuse freely. The entire system is open and dissipative. Any persistent structure must therefore arise from coupled fluctuations, forces, flows, and ion gradients.

Under these conditions, a characteristic process chain forms. Fluctuations modulate electromagnetic modes between the membranes and generate weak but sustained Casimir–Lifshitz forces [8][9]. These forces act nonspecifically, even in salt-rich media. They pull protocells together and promote cluster formation. As separations shrink, water is expelled from the gap. This generates nanofluidic flows that advect ions. Asymmetries in size or membrane properties [19] amplify these flows. Local ion concentration differences emerge—especially for protons, which are highly mobile and respond strongly to nanoscale geometry. Early electrochemical potentials form and can persist over diffusive timescales. Such gradients reshape the microenvironment between vesicles and extend the lifetime of emerging structures. This marks an initial step toward proto-functional stability.

The following applies to the strength of the attractive Casimir-Lifshitz force between two protocells:

$$F_{CL}(L) \approx -(A_{eff} / 6) * (R_{eff} / L^2) \quad (1)$$
$$\text{with } R_{eff} = (R_1 * R_2)/(R_1 + R_2).$$

Here, A_{eff} denotes an effective Hamaker constant [10] that integrates the spectral dielectric response of the membrane–water system. For numerical calculations we use an effective Hamaker constant $A_{eff} = 5 \times 10^{-21}$ Joule, representative of membrane-water-membrane systems. The force scales linearly with the effective curvature radius R_{eff} and decays algebraically as $1/L^2$.

In the L=2–100 nm range, these Casimir-Lifshitz-forces can exceed thermal contributions [13][14][15][16][17][18][20]. Proton diffusion [11] times scale as $\tau_{diff} \approx L^2/D$. Because protons diffuse rapidly, these times are very short. For stable gradients, the persistence time $\tau_{gradient}$ must therefore match or exceed τ_{diff}. Only then do energy flows stabilize the structure rather than disrupt it. Prebiotic sustainability emerges from the interplay of energy flux, geometry, and gradient lifetime.

The model leads to a clear conclusion: the earliest proto-functional stability did not arise *despite* fluctuations, but *because of* them. Fluctuations generate forces. Forces generate flows. Flows generate gradients. And gradients generate stable patterns that can serve as precursors of biological function. Sustainable protocell clusters thus make active use of fluctuations. They convert noise into persistent micro-energy potentials. Sustainability becomes an emergent non-equilibrium phenomenon and a necessary precondition for early prebiotic self-organization.

Physical Pipeline of Prebiotic Sustainability:

For readers with an engineering or systems background, the core mechanism of the proposed framework can be summarized as a five-step physical pipeline:

- Fluctuations in confined nanoscale gaps generate effective coupling forces between adjacent protocell membranes.
- Coupling stabilizes proximity, driving cluster formation and inducing nanofluidic expulsion and directed microflows in the intermembrane gap.
- Flows combined with geometric or material asymmetry create local ion and proton concentration biases without pumps or enzymes.
- Resulting gradients extend state lifetimes, enabling persistent energetic and structural configurations beyond diffusive relaxation times.
- Attractor-like return dynamics allow disrupted clusters to reassemble into functionally equivalent states, providing regenerability without replication.

Together, these steps explain how stability, persistence, and recovery emerge as purely physical properties of open, fluctuation-driven protocell systems.

Massoth [2] showed that protocells in prebiotic environments likely did not exist as isolated compartments. Theoretical models and recent experimental analogues indicate that nanoscale vesicles and amphiphile aggregates can spontaneously assemble into clusters in saline solutions through Casimir–Lifshitz attraction. This attraction arises from the reduction of electromagnetic fluctuations in the gap between two membranes and makes certain geometries energetically favored. Two simple but structurally central clusters emerge: the protocell dimer and the protocell tetrahedron.

V. Definition and Taxonomy of Prebiotic Sustainability

The physical model shows that early protocell clusters could form stable mesoscale structures because they integrated fluctuations, flows, and ion gradients. Sustainability here does not arise from biological control but from non-equilibrium physics. The goal is a precise definition that remains accessible to biologists, computer scientists, and sustainability researchers, while integrating key results from earlier papers—Casimir–Lifshitz forces, attractor-based informational patterns, and quantum-field-induced proton reservoirs.

A. Definition of Prebiotic Sustainability

Prebiotic sustainability denotes the ability of an open, nano-compartmentalized protocell system to preserve its structural and functional identity over time, despite strong fluctuations and continuous energy flow. This occurs without enzymes, genes, or pumps. The system remains recognizably the same entity even as it produces entropy and undergoes constant internal dynamics.

Physically, this implies *controlled openness*: the system is far from equilibrium, produces entropy (dS/dt > 0), and uses fluctuations actively to generate forces, flows, and gradients. For biologists, this is proto-bioenergetic self-maintenance; for computer scientists, the stabilization of pattern states; for sustainability science, the preservation of identity under change.

Formally, sustainability can be expressed using an order-parameter vector θ(t) describing cluster geometry, contacts, or gradient profiles. A system is sustainable when the normalized structural deviation

$$D(\theta(t+\Delta t), \theta(t)) < \delta \quad (2)$$

while energy flow persists. Sustainability is not equilibrium but the persistence of macroscopic signatures in a fluctuating, dissipative system.

Specification of Prebiotic Sustainability in the Nanoworld:

A protocell system (cluster) is sustainable if the entropy production rate σ(t) > 0 in steady operation and for S(Total) with (Total = system + environment), the following applies

$$\frac{dS(Total)}{dt} > 0 \text{ and } D(\text{Structure}, t+\Delta t \mid t) = \frac{||\,\theta(t+\Delta t)-\theta(t)\,||}{||\theta(t)||} < \delta$$

consistent with open, dissipative operation, under continuous energy input and dissipation, with θ(t) representing experimentally accessible order parameters (cluster geometry, ion gradients), and δ a small, system-dependent tolerance for structural drift.

Here, ‖·‖ is the Euclidean norm, and δ may correspond to "<3% structural change per Δt".

Interpretation:

- $dS_{system}/dt > 0$: The system produces entropy → it is open and dissipative, not equilibrated.
- $D(Structure, t+\Delta t \mid t) < \delta$: Despite flow and dissipation, the structure remains sufficiently similar → identity under change.
- *Continuous energy input:* Classic Prigogine setting; without energy input, every dissipative structure collapses.

The key lies in *D*: what counts as "structure" in a protocell cluster?

D measures the normalized distance between relevant order parameters at two times t and t+Δt. Framework prebiotic sustainability in the nanoworld means that this macroscopic signature drifts only weakly, even though energy is continuously dissipated. In practice the order parameter θ(t) could be, for example: mean nearest neighbor distance, width of distance distribution, mean number of vesicles per cluster, mean number of contacts per vesicle, proportion of vesicles in clusters, mean and variance of pH differences in the intermembrane gaps, and so on.

B. Three Dimensions of Prebiotic Sustainability

These three dimensions jointly describe early self-maintenance in protocellular systems and can be experimentally assessed.

(i) Structural Sustainability:

Structural sustainability is the ability of a protocell cluster to maintain its spatial organization. It includes the integrity of individual membranes and the persistence of collective arrangements such as dimers or larger supramolecular patterns. Thermal noise, mechanical perturbations, and diffusion may distort the architecture only moderately. Key metrics include lifetime, coalescence and fission rates, and return to stable attractor shapes. This directly connects to [1], which showed that Casimir–Lifshitz forces provide robust mesoscale binding in the 2–100 nm regime.

(ii) Energetic Sustainability:

Energetic sustainability refers to the stability of electrochemical potentials—especially proton and ion gradients within or between protocells. The critical question is whether the persistence time of a gradient ($\tau_{gradient}$) matches or exceeds the diffusive relaxation time (τ_{diff}). This dimension links the fluctuation-driven model of the present paper with the quantum-field-induced proton reservoirs of [5], where narrow gaps act as resonant proton traps. Energetic sustainability means that the system organizes energy flows into long-lived, functionally relevant potential landscapes.

(iii) Informational Sustainability:

Informational sustainability describes the stability and recurrence of physical patterns [2][7] on mesoscale levels. These include stationary concentration profiles, characteristic cluster spacings, or recurring flow cycles that act as elementary informational states without genes. Such patterns must remain reproducible despite fluctuations and function as attractor-based proto-memory structures. This dimension connects to [2], where reproducible differences were defined as physical information, and to the attractor architectures observed in protocell clusters.

Summary:

- *Structural sustainability:* the "vessel" (compartment or protocell) remains intact.
- *Energetic sustainability:* "energy in the form of gradients" remains available long enough to matter.
- *Informational sustainability:* "patterns" and "states" remain recognizable over time.

Together, these three dimensions form a coherent physical taxonomy of sustainability in the prebiotic nanoworld. They place protocellular persistence into a unified framework that integrates structure, energy, and information.

The following sections translate this taxonomy into a three-level sustainability framework—stability, persistence, and reproducibility/regenerability—and derive concrete experimental hypotheses.

VI. Conceptual Sustainability Framework (Three-Level Model)

The definition of prebiotic sustainability leads to a three-level framework: stability, persistence, and reproducibility/ regenerability. Each level builds on the previous one. Together, they connect Casimir–Lifshitz forces, early proton landscapes, and attractor-based patterns into a coherent model of protocellular self-maintenance.

The definition of prebiotic sustainability leads naturally to a three-level framework, summarized in Figure 2.

In Figure 2, Schematic three-level framework of prebiotic sustainability. Structural stability arises from Casimir–Lifshitz coupling between protocells. Persistence is achieved when proton-gradient lifetimes exceed diffusive relaxation times ($\tau_{gradient} \geq \tau_{diff}$).

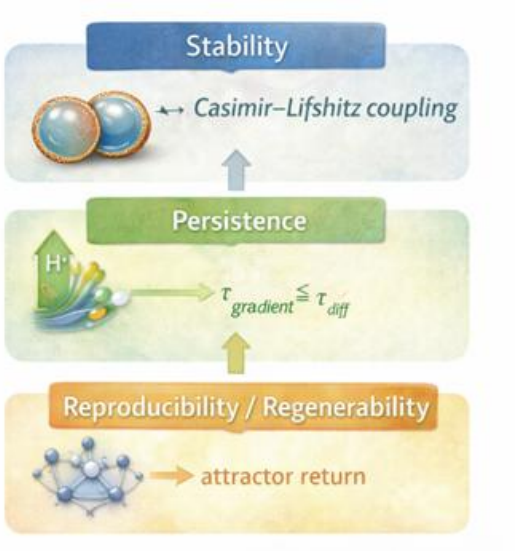


Figure 2. Three-Level Sustainability Framework

Reproducibility and regenerability emerge from attractor-based return dynamics, enabling recovery of functionally equivalent states after perturbation in the absence of genetic or enzymatic control.

TABLE I. MAPPING OF PREBIOTIC SUSTAINABILITY LEVELS TO RELIABILITY ENGINEERING CONCEPTS

Prebiotic Sustainability Level	Reliability Engineering Analogon	Representative Measurement / Metric
Stability	Low hazard rate / high Mean Time To Failure (MTTF)	Cluster break-up rate; mean lifetime of dimer or cluster structures
Persistence	State retention time / operational availability	Ratio τgradient / τdiff; duration of stable ion or proton gradients
Regenerability	Mean Time To Recovery (MTTR) / recoverability	Reassembly time after perturbation; return probability to functional attractor state

Table I relates the physical notion of prebiotic sustainability to established reliability-engineering metrics, enabling quantitative assessment of protocell cluster stability, persistence, and recovery under non-equilibrium conditions.

This mapping enables importing mature reliability engineering metrics for quantifying sustainability in protocell clusters.

A. Stability

Stability describes the ability of a protocell cluster to withstand perturbations without losing its basic structure. Membranes should not rupture immediately, and fusion or

fission events should not cause collapse. Despite thermal noise, characteristic distances and geometries must survive.

The key drivers are fluctuation-induced forces. Casimir–Lifshitz attractions in the 10–200 nm range prevent uncontrolled separation and create weak but continuous coupling. These forces act without specific chemistry and keep compartments within valid structural regimes. Membrane composition—flexible fatty acids, lipids, or polymer shells—reinforces this effect.

Stability therefore defines a baseline order space in which protocells remain identifiable despite noise. For computer scientists, it corresponds to remaining inside the valid state space. For biologists, it is the preservation of compartments. For sustainability research, it represents the minimal condition for long-term viability.

B. Persistence

Persistence extends stability through time. A state must last long enough for gradients, reaction pathways, or pattern formation to operate. The critical factor is the lifetime of ion gradients, especially in nanoscale gaps.

The relevant timescale is $\tau_{diffusion}$, while $\tau_{gradient}$ denotes the actual gradient lifetime. Persistence occurs when $\tau_{gradient} \geq \tau_{diffusion}$. Narrow geometries, flows, and fluctuations can slow relaxation and create proto-bioenergetic plateaus.

Persistence means that energy flows do not immediately destroy structures but temporarily stabilize them. For biologists, this resembles an early energetic steady state without ATP synthase. For computer scientists, it corresponds to the retention time of a state. For sustainability science, it represents the physical form of long-term viability.

C. Reproducibility and Regenerability

The third level describes the ability to return to a functionally similar state after disruption—without genes, enzymes, or complex replication mechanisms. This is not biological reproduction but physical reproducibility of structures and gradients.

Regenerability arises from the same mechanisms that drive cluster formation: fusions replace damaged compartments; fission and re-assembly create new ones; flows can rebuild collapsed proton potentials. In the language of earlier papers, this corresponds to attractor-based return dynamics. The restored state is not identical but functionally equivalent.

For biologists, this resembles rudimentary autopoiesis. For computer scientists, it is a form of physical error correction. For sustainability research, it represents the earliest regenerative capability.

Summary:
This framework defines a robust physical notion of sustainability:

- Stability prevents immediate collapse.
- Persistence supports long-lived energetic and structural patterns.
- Reproducibility/Regenerability enables return to functional states after disturbance.

A prebiotic protocell cluster that satisfies all three levels can be considered sustainable—long before biological evolution or molecular control. The next section translates this framework into concrete experimental hypotheses.

VII. OPERATIONAL DEFINITION OF THE STRUCTURAL DISTANCE D

To quantify prebiotic sustainability in protocell clusters, we define a structural distance D that measures the normalized drift of the mesoscale system signature over time. Let $\theta(t)$ denote a vector of experimentally accessible order parameters capturing cluster geometry, energetic gradients, and recurrent dynamical patterns.

$$\theta(t) = (\theta_1(t), \theta_2(t), \ldots, \theta_n(t)) \quad (3)$$

The structural distance D between two times t and $t + \Delta t$ is defined as a weighted, normalized root-mean-square drift:

$$D(t, \Delta t) = \sqrt{\left(\Sigma_{i=1}^{n} w_i \cdot [\, (\theta_i(t + \Delta t) - \theta_i(t)) / \theta_{i,ref} \,]^2 \right)} \quad (4)$$

where:

- $\theta_i(t)$ = i-th order parameter at time t (dimensionful or dimensionless).
- $\theta_{i,ref}$ = reference scale used for normalization (e.g., baseline mean, initial value, or experimentally defined characteristic scale).
- $w_i \geq 0$ = weights (often set to $w_i = 1$ if no prior weighting is justified).
- n = number of included order parameters.

Here w_i are dimensionless weights ($\Sigma\, w_i = 1$) and $\theta_{i,ref}$ are reference scales used for normalization. In practice, $\theta_{i,ref}$ are chosen as robust characteristic values measured in an initial calibration window $[t_0, t_0 + T_0]$, or, where appropriate, as physically motivated scales.

Operational sustainability criterion:
A protocell cluster is operationally "sustainable" over Δt if it remains dissipative and its mesoscopic signature drifts only weakly:

$$dS\_system/dt > 0 \text{ and } D(t,\Delta t) < \delta \quad (5)$$

with δ a small tolerance (experiment-dependent; e.g., δ = 0.03–0.05 for "3–5% normalized drift per Δt").
Sustainability thus corresponds to persistent identity under non-equilibrium conditions, rather than equilibrium stability.

Practical guidance for choosing θ(t):

θ(t) should combine a minimal set of geometric, energetic, and (optionally) pattern/attractor descriptors that are directly measurable (microscopy + pH/ion imaging), e.g. nearest-neighbor distance statistics, contact number, cluster size, mean/variance of ΔpH in gaps, and estimated $\tau_{gradient}$.

Operational θ-vector for microscopy, pH-mapping, and microfluidics:

For experimental realizations based on time-resolved microscopy, fluorescence pH maps, and microfluidic perturbations, we propose the following minimal and sufficient parameter set (n = 9):

Structural (geometry):

$\theta_1(t)$ Mean nearest-neighbor distance $\langle L_{nn} \rangle$

$\theta_{1,ref}$ = median$\langle L_{nn} \rangle$ in $[t_0, t_0 + T_0]$

$\theta_2(t)$ Standard deviation of nearest-neighbor distances $\sigma(L_{nn})$

$\theta_{2,ref}$ = median $\sigma(L_{nn})$

$\theta_3(t)$ Cluster size (number of vesicles in the largest connected cluster)

$\theta_{3,ref} = N_{cluster}(t_0)$

$\theta_4(t)$ Mean contact number $\langle k \rangle$ (graph-based adjacency)

$\theta_{4,ref}$ = median $\langle k \rangle$

$\theta_5(t)$ Cluster fraction $f_{cluster}$ (vesicles in clusters / total vesicles)

$\theta_{5,ref} = f_{cluster}(t_0)$

Energetic (pH maps, gradients):

$\theta_6(t)$ Mean proton gradient $\langle \Delta pH_{gap} \rangle$ (gap vs. bulk)

$\theta_{6,ref} = |\langle \Delta pH_{gap} \rangle|$ at t_0

$\theta_7(t)$ Variance of gap gradients $\sigma(\Delta pH_{gap})$

$\theta_{7,ref}$ = median $\sigma(\Delta pH_{gap})$

Dynamical / sustainability metrics:

$\theta_8(t)$ Gradient persistence time $\tau_{gradient}$ (autocorrelation decay time)

$\theta_{8,ref} = \tau_{diff} = L_{gap}^2 / D_{H^+}$. Here L_{gap} is the median intermembrane separation extracted from microscopy over $[t_0, t_0 + T_0]$, and T_{obs} is the total recording duration.

$\theta_9(t)$ Return (recovery) time T_{return} after perturbation

$\theta_{9,ref} = T_{obs}$ (total observation window)

We set the sustainability threshold δ to 0.05 as a conservative default (≈5% normalized drift per Δt), and in practice calibrate δ from control recordings by choosing $\delta = P95(D_{control})$ (95th percentile, equivalently, $\delta \approx 3\sigma_D$) to account for platform-specific measurement noise

Interpretation:

Small values of D indicate that a protocell cluster remains within the same effective attractor basin despite continuous entropy production and microscopic rearrangements. Large D signals structural drift, loss of gradient coherence, or transition to a different dynamical regime.

In this formulation, D provides a compact, experimentally accessible measure of identity preservation under flow and fluctuation, enabling direct comparison of protocell systems in terms of structural, energetic, and informational sustainability.

VIII. Testable Hypotheses and Experimental Roadmap

Prebiotic sustainability must be experimentally verifiable. To achieve this, the key mechanisms of the model—structural coupling, gradient persistence, and simple regeneration—are translated into measurable quantities.

The roadmap defines four hypotheses that map onto the three levels of the framework (stability, persistence, regenerability) and can be tested in minimal protocell models. The goal is to demonstrate physical sustainability phenomena without genetic or enzymatic complexity. The following hypotheses translate the physical sustainability framework into experimentally testable regimes (Fig. 1).

Figure 1 summarizes the experimental roadmap (H1–H4) for probing prebiotic sustainability in protocell systems. It contrasts lifetimes of isolated and clustered protocells, visualizes proton-gradient formation in nanoscale gaps, relates gradient persistence to heat dissipation, and maps sustainability regimes in parameter space defined by effective curvature, gradient lifetime, and temperature.

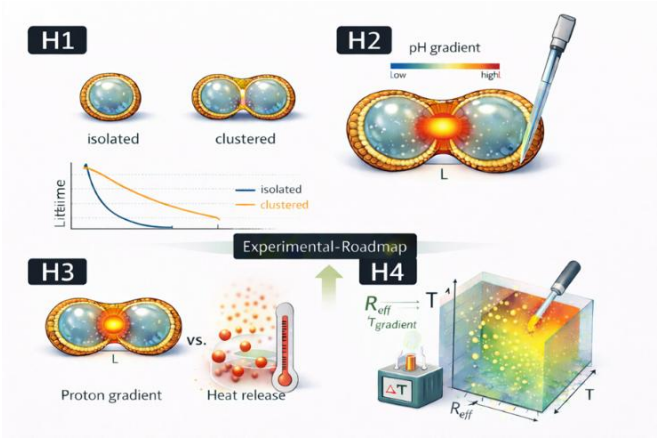


Figure 3. Experimental roadmap for testing prebiotic sustainability mechanisms

A. Hypothesis H1 – Structural Sustainability:

Protocell clusters outlive isolated vesicles because fluctuation-induced forces stabilize their collective structure.

Casimir–Lifshitz attractions couple vesicles at distances of 2–100 nm [8][9][13][14]. Single vesicles in saline environments often degrade quickly. Clusters, by contrast, form mesoscale networks that protect and stabilize compartments.

Experiment H1E:

Time-resolved microscopy (confocal or holographic). Compare two conditions: dilute solutions (isolated vesicles) and higher concentrations (cluster formation). Measure lifetime, decay rates, structural drift, and spontaneous fusion events. Longer lifetimes in clusters would directly demonstrate structural sustainability.

B. *Hypothesis H2 – Asymmetry and Ion Gradients:*

Asymmetric vesicle pairs generate measurable proton and ion gradients in the gap region through geometry, fluctuations, and flow alone.

The process chain is: fluctuation → force → flow → gradient. Differences in radii or membrane stiffness produce asymmetric gaps. Displaced fluid creates directed nanofluidic flows and local concentration shifts.

Experiment H2E:

Fluorescence-based pH microscopy or micro-pH electrodes in microfluidic chambers. Observe vesicle pairs with defined size ratios ($R_1 > R_2$). Expect a stable pH difference in the gap and proton drift toward the asymmetric region. Evidence of this would show that early, protein-free gradient dynamics could arise purely from physical principles.

C. *Hypothesis H3 – Energetic Sustainability and Entropy Production:*

Protocell configurations with long-lived ion gradients exhibit lower entropy production per time and volume than configurations with short-lived gradients.

Long-lived gradients indicate efficient use of energy flows. Short-lived gradients dissipate rapidly and produce more heat.

Experiment H3E:

Combine gradient measurements (Δc, $\Delta \mu$, $\tau_{gradient}$) with thermal measurements (isothermal microcalorimetry or μ-thermography). If stable gradients correlate with reduced local heat release, energetic sustainability is experimentally supported.

D. *Hypothesis H4 – Sustainability and the Three-Layer Framework:*

Observable dynamics can be statistically grouped into the three sustainability levels—stability, persistence, and regenerability. Cluster dynamics can be classified via low/medium/high D regimes.

The result is a multidimensional dataset of lifetime, gradient persistence, reassembly rates, and other parameters. The question is whether these values form robust classes.

Experiment H4E:

Vary vesicle parameters (size, stiffness, salt concentration, temperature, pH). Analyze time series using clustering, Principal Component Analysis (PCA), Density-Based Spatial Clustering of Applications with Noise (DBSCAN), t-distributed Stochastic Neighbor Embedding (t-SNE), or spectral methods.

Three classes are expected:

(1) Stable but not persistent
(2) Stable and persistent but weakly regenerable
(3) Fully sustainable

Such a result would empirically validate the three-level sustainability framework.

IX. CONCLUSION AND FUTURE WORK

The concept of prebiotic sustainability shifts the focus of origins-of-life research. Instead of emphasizing individual reaction pathways, it highlights a physical understanding of stability, persistence, and regenerability in nanoscale protocell clusters. Our analysis shows that fluctuation-induced forces, nanofluidic flows, and proton-based gradients can generate long-lived ordered states even in simple, non-genetic compartments. These states persist over diffusive timescales and can return to functionally similar configurations after disturbances. Prebiotic sustainability thus appears as an emergent property of open, fluctuation-coupled systems—not as a late product of biological optimization.

The proposed triad of structural, energetic, and informational sustainability provides a scale-invariant framework connecting physics, prebiotic chemistry, and early information. Combined with the three levels of stability, persistence, and regenerability, this framework enables direct comparison of different protocell types—from fatty-acid vesicles to PMBC systems. Hypotheses H1–H4 and the experimental roadmap show that these concepts can be tested using current protocell technologies, microfluidics, pH microscopy, calorimetry, and data-driven analysis.

In the long term, the approach opens two major directions. For origins-of-life research, it offers a way to compare scenarios by their sustainability profiles: Which protocell types maximize structural, energetic, and informational sustainability—and could this have served as a key selection criterion in prebiotic evolution? For synthetic biology and sustainability science, the framework points toward "synthetic sustainable protocells," engineered systems that replicate physical sustainability and could be applied in biohybrid or technological nanosystems.

Prebiotic sustainability thus emerges as a foundational ordering principle that may have enabled the transition from pure physical self-organization to early bioenergetic and informational functions. The central challenge now is to reveal this three-dimensional sustainability signature in the laboratory—to make measurable the traces of a world in which sustainability was already at work long before life existed in its modern form.

REFERENCES

[1] M. Massoth, "Attractive Casimir-Lifshitz Forces as a Universal Driver of Prebiotic Protocell Aggregation and Cluster Formation," *BIOTECHNO*, 2026.

[2] M. Massoth, "Emergent Information Formation in Prebiotic Protocell Clusters: A Computational Mechanics Framework of ε-Machines and Attractor Memory," *BIOTECHNO*, 2026.

[3] M. Massoth, "From Physical Difference to Meaning: A Constructor-Theoretic Framework for Prebiotic Information in Casimir–Lifshitz-Coupled Protocell Clusters," *BIOTECHNO*, 2026.

[4] M. Massoth, "Physical Origin of Proto-Information: Syntax, Semantics and Pragmatic Emergence in Prebiotic Protocell Clusters," *BIOTECHNO*, 2026.

[5] M. Massoth, "Quantum Field-Induced Proton H^+ Gradients in Prebiotic Protocell Clusters," *BIOTECHNO*, 2026.

[6] F. E. Rosas, *et al.*, "Software in the natural world: A computational approach to hierarchical emergence," *arXiv* preprint, 2024, doi:10.48550/arXiv.2402.09090.

[7] M. Deutsch and C. Marletto, "Construction theory of information," *Proc. R. Soc. A*, vol. 471, art. no. 20140540, 2015, doi:10.1098/rspa.2014.0540.

[8] H. B. G. Casimir, "On the attraction between two perfectly conducting plates," *Proc. K. Ned. Akad. Wet.*, vol. 51, pp. 793–795, 1948.

[9] E. M. Lifshitz, "The theory of molecular attractive forces between solids," *Sov. Phys. JETP*, vol. 2, pp. 73–83, 1956.

[10] J. N. Israelachvili, *Intermolecular and Surface Forces*, 3rd ed., Academic Press, London, 2011.

[11] P. C. W. Davies, "Does quantum mechanics play a non-trivial role in life?" *Biosystems*, vol. 78, pp. 69–79, 2004.

[12] I. Gözen, *et al.*, "Protocells: Milestones and recent advances," *Small*, vol. 18, art. no. 2106624, 2022.

[13] S. K. Lamoreaux, "Demonstration of the Casimir force in the 0.6–6 μm range," *Phys. Rev. Lett.*, vol. 78, pp. 5–8, 1997.

[14] J. L. Garrett, D. A. T. Somers, and J. N. Munday, "Measurement of the Casimir force between two spheres," *Phys. Rev. Lett.*, vol. 120, art. no. 040401, 2018.

[15] P. Rodriguez-Lopez, "Casimir energy and entropy in the sphere–sphere geometry," *Phys. Rev. B*, vol. 84, art. no. 075431, 2011.

[16] S. J. Rahi, T. Emig, N. Graham, R. L. Jaffe, and M. Kardar, "Scattering theory approach to electrodynamic Casimir forces," *Phys. Rev. D*, vol. 80, art. no. 085021, 2009.

[17] G. Bimonte and T. Emig, "Exact results for classical Casimir interactions: Dirichlet and Drude model in the sphere–sphere and sphere–plane geometry," *Phys. Rev. Lett.*, vol. 109, art. no. 160403, 2012.

[18] A. Canaguier-Durand, G.-L. Ingold, M.-T. Jaekel, A. Lambrecht, P. A. Maia Neto, and S. Reynaud, "Classical Casimir interaction in the plane–sphere geometry," *Phys. Rev. A*, vol. 85, art. no. 052501, 2012.

[19] B. B. Machta, S. L. Veatch, and J. P. Sethna, "Critical Casimir forces in cellular membranes," *Phys. Rev. Lett.*, vol. 109, art. no. 138101, 2012.

[20] M. Boström and B. E. Sernelius, "Thermal effects on the Casimir force in the 0.1–5 μm range," *Phys. Rev. Lett.*, vol. 84, pp. 4757–4760, 2000.

[21] E. Zaitseva and V. Levashenko, "Reliability engineering in healthcare: Opportunities and challenges", Reliability Engineering & System Safety, Volume 267, Part B, 2026, https://doi.org/10.1016/j.ress.2025.111933.

[22] H.-J. Butt, "Measuring electrostatic, van der Waals, and hydration forces in electrolyte solutions with an atomic force microscope," *Biophys. J.*, vol. 60, pp. 1438–1444, 1991.

Network-Based Analysis of Gut Microbiome Profiles in Autism Spectrum Disorder

Shado Badr Basem
Software engineering, The British University in Egypt
The British University in Egypt
Cairo, Egypt
Email: shado.Badr@bue.edu.eg

Hesham Ali
University of Nebraska at Omaha
Nebraska, USA
Email: hali@unomaha.edu

Andreas Pester
The British University in Egypt
Cairo, Egypt
Email: Andreas.Pester@bue.edu.eg

***Abstract*—The last two decades have witnessed significant increase in research activities associated with Autism Spectrum Disorders (ASD). Many research studies have been conducted to provide better understanding of ASD and attempt to find biological or behavioral signals to help in classification and early detection efforts. In addition, a recent report from the Center of Disease Control (CDC) reported a substantial increase in the percentage of children diagnosed with Autism as compared to the percentage reported a decade earlier. In this study, we explore the possibility of obtaining a biological signal that connects Autism with Gut Microbiome profiles. We employ the concept of similarity networks to analyze micobiome datasets collected in two ASD studies. The key bacterial genera present across two independent datasets are Ruminococcaceae, Fastidiosipila, Firmicutes, Lachnospiraceae, which are consistently associated with variations in microbial community structure. These findings demonstrate the utility of combining similarity networks using cosine similarity and Bray–Curtis dissimilarity to try and identify a unique microbial structure for ASD and controls , which can help us diagnose ASD early on, this study has shown different microbial structure that although are not yet definitive in this preliminary stage, can later be further analyzed to serve as potential biomarkers for early ASD detection and support the development of microbiome-based therapeutic strategies to alleviate symptom severity.**

Keywords-Health Informatics; Autism Spectrum Disorder; Similarity Networks; Biological Markers; Gut Microbiomes Profiles.

I. INTRODUCTION

Concerns over the rising prevalence of individuals being diagnosed with Autism Spectrum Disorder (ASD) have escalated over recent decades. Statistics indicate that almost 1 in every 68 individuals worldwide is diagnosed with ASD [1]. ASD is identified as a persistent deficit in social communication or behavior that can usually be repetitive or restrictive for affected individuals. Diagnostic recategorization and improved screening criteria may be the cause for this surge in numbers, where almost 62.2 million individuals live with ASD, as estimated in 2016 [2]. Numerous studies have highlighted a possible relationship with multiple biomarkers that can aid in detecting ASD early in an individual's life; some of these biomarkers include movement, brain activity, and gut microbiome.

The human gut microbiome comprises a complex population of tens of thousands of microbes. This population includes a variety of organisms, such as bacteria, fungi, archaea, protozoa and viruses [3]. The number of bacteria can vary throughout the gastrointestinal tract, with the colon harboring the highest abundance and most diverse species when compared to the stomach and the small intestine. Hence, the gut microbiomes represent a fundamental role in the host's homeostasis, which is nutrient, immune development, metabolism and defense against pathogens [4]. The gut's microbial balance has been directly linked to human health, where numerous studies have shown that the gut bacteria are directly involved in the fundamental biological processes in humans. Furthermore, the gut microbiome composition can be influenced by the host genetics, diet, age, or external factors, such as geographical location [5]. This study aims to form a relationship between ASD and microbiomes by exploring potential correlations that may act as a diagnostic biomarker.

This paper is comprised of a related work in Section II highlighting the most recent and important findings in the ASD microbiome field related to this paper, followed by a methodology in Section III where the datasets used will be discussed in more details, the metrics used, along with the proposed pipeline and the methodology steps implemented. The results for the implementation will then be presented in Section IV and discussed Section V to highlight the most important findings. Finally, a conclusion for this paper and the future direction of the research will be presented in Section VI.

II. RELATED WORK

In this section a discussion of the most important publications related to the ASD microbiome research are mentioned along with their most important findings.

A review paper by O'Donnell et al. [6] discussed that all species, including animals and humans, have a co-evolving microbial community that colonizes them. However, a short-term diet can have a temporary effect on the gut microbiota; long-term dietary habits can affect the abundance of taxa and have a subsequent impact on the host's health. The most prominent changes can be observed when an individual follows a westernized, fiber-poor diet or a more plant-based, fiber-rich diet, such as the Mediterranean diet. These diets can be associated with higher microbial diversity in the gut.

A paper by Hrnciarova et al. [7] proposed a randomized, double-blinded, placebo-controlled pilot study to evaluate the efficacy of biological response modifier juvenil, which is a mixture of peptides, nucleotides, amino acids, and other components derived from animals, in modulating the microbiome of children with ASD, specifically whether juvenil can help in alleviating the symptoms of ASD. There were 20 ASD and 12 Neurotypical children involved in this pilot study of Czech nationality between the ages of 3 and 7. The samples were collected from each of these children for the duration of 3 months and one group was administered Juvenil, while the other group was administered a placebo.

The results of this pilot study showed that the Juvinel achieved its purpose by making the microbiota of the ASD children quite similar to that of the Neurotypical children. However, it did not result in a complete restoration of the microbiota composition. Furthermore, the behavior of all the ASD children involved in this study was evaluated using CARS2-ST. The group of children who were administered the Juvinel showed a 12.4% reduction in their autism symptoms, which was double that (6.6%) of the placebo group.

Another research by Peralta-Marzal et al. [8] used machine learning (recursive ensemble feature selection) to determine how gut health affects individuals who are diagnosed with ASD, where the primary focus was trying to choose a set of bacterial taxa that defines ASD classification. This was done using sibling-controlled dataset.

The results were obtained by applying RFFS, a method used to discover biomarkers. RFFS is composed of 8 classifiers. This study was successful in identifying 26 bacterial taxa that consistently distinguished ASD from the controls across three independent cohorts, including lifestyle, which influences both ASD and gut microbiome studies. This paper helped prove that there exists a signature for the microbiomes that could identify individuals with ASD from the controls and help reinforce gut health as a valid biomarker.

In a paper published by Taniya et al. [9], several factors were discussed regarding the connection between ASD and the gutbrain axis, mainly the effects of microbiota early dysbiosis in the gestation period, mode of delivery, uncontrolled usage of antibiotics, and stress. These factors affect the gut microbiome and lead to dysbiosis, which then impacts the Central Nervous System function by the production of neurotoxins. The effect of Clostridium found in the colon of children can be used as an indication of ASD development. This paper also discussed the Food and Drug Administration's recognition of microbial transplant therapy as a fast-track treatment for ASD.

This study by Fourquier et al. [10] focused on the gut microbiome of children with ASD in Arizona and Colorado. The results showed differences in microbial structure based on the children's location, suggesting the importance of geographical location when discussing microbiomes. A longitudinal analysis performed linked microbiome shifts with the severity of the displayed ASD behaviors, such as lethargy and inappropriate speech. These findings stress the importance of longitudinal tracking of the ASD subjects, as much more information can still be found from such research.

In a review by Sivamaruthi et al. [11], which again stressed the importance of the gut-brain axis in ASD, we can see how the diet and microbial composition have influenced the symptoms of ASD. This further supports that probiotic supplement can have a powerful potential in improving gut balance, which in turn would help in reducing some of the behavioral symptoms caused by ASD. Though the results still need further clinical trials. Some of the main Risk factors that were stressed in this paper were maternal diet, lifestyle, and early microbial colonization as risk factors that may contribute to ASD development.

From the reviewed studies, it is clear that there is a definitive connection between the gut microbiome and the ASD symptoms which might even allow for the use of the gut microbiome as a biomarker for diagnosis or treatment. It has been shown that clinical interventions, such as dietary modulation, use of biological response modifiers, such as juvenil, and probiotic supplements have shown promise in shifting the microbial profile of ASD individuals with ASD to one that more closely resembles that of a neurotypical individual. Moreover, longitudinal studies have shown that the microbial composition affects not only the diagnosis but also the severity of symptoms of ASD.

In this research, the main aim is to apply a network-based similarity and clustering approach to ASD microbiome datasets. The focus is mainly on the structural properties of microbial communities, not solely on abundance. This can help us make the most of the datasets we can now access. Having access to such datasets will help us build a possible microbial profile that can help in the early identification of ASD in younger individuals.

Improving the overall treatment plan and achieving much better symptom control for individuals.

III. METHODOLOGY

The methodology for the paper will now be presented where the datasets used will be discussed in details, along with the metrics, proposed pipeline and the methodology.

A. Dataset

Multiple 16S rRNA gene sequencing datasets were utilized. These datasets include the taxonomic profiles that were derived from the amplification and sequencing of the phylogenetic marker for the characterization and profiling of microbial communities.

In the Dataset collected by Kang et al. [12], fecal samples were obtained from children between the ages of 4 and 17 years with ASD and neurotypicals to act as controls. The dataset contains 21 neurotypical children and 23 children with ASD. The children did not have a first-degree relationship. Furthermore, the children with ASD were assessed with the Autism Treatment Evaluation Checklist (ATEC) and the Pervasive Development Disorder Behavior Inventory (PDDBI).

The genomic DNA was extracted using a PowerSoil DNA extraction kit, the sequencing analysis was performed using the Quantitative Insights into Microbial Ecology (QIIME). An Operational Taxonomic Unit (OTU) was then obtained in table form by clustering the sequence similarity at 97% where the OTU represents the abundance of each detected bacterium in each sample.

A second dataset (GSE113690) [13] that was utilized in this paper comprises 143 subjects who are clinically diagnosed as ASD, where the sex was male: female 130:13, and the average age was 4.937 ± 0.155. While the Control group had an average age of 5.189 ± 0.170 and a male-to-female ratio of 127:16, 16S rRNA gene sequencing was used in this dataset to evaluate the microbial composition of the ASD and the controls.

Moving on to preprocessing the datasets, a similar structure was followed, where the low-abundance OTUs were removed, meaning the taxa that had near-zero values across most samples were removed. This was done to avoid any bias in the network and to ensure that only the bacteria with any influence remained. For the second dataset, a low-abundance filter had to be applied to ensure that any noise was prevented, which meant eliminating any OUT with counts ≤ 1 in more than 50% of samples.

B. Measures

The measures applied will be the Cosine Similarity and the Bray-Curtis dissimilarity.

1) Cosine similarity

The cosine similarity is a metric that is commonly used to perform tasks, such as information retrieval and data mining.

$$cos(\theta) = \frac{\sum_{i=1}^{d} x_i \cdot x'_i}{\sqrt{\sum_{i=1}^{d} x_i^2} \cdot \sqrt{\sum_{i=1}^{d} x'^2_i}} \quad (1)$$

Cosine similarity focuses more on the direction, where it measures the similarity as the cosine of the angle between two vectors, where two vectors that are considered similar will have a small angle between them [14].
possible link between the microbial composition of individuals with ASD and the individuals who act as controls.

2) Bray-Curtis

The Bray-Curtis index qualifies the dissimilarity between two different taxa over multiple samples. Moreover, BrayCurtis dissimilarity is used to complement the Cosine Similarity; it is used to quantify the compositional difference in the microbiome profile between the ASD and the control Subjects. It focuses on the absolute dissimilarity in the microbiome studies.

$$BC_{ij} = 1 - \frac{2C_{ij}}{S_i + S_j} \quad (2)$$

Where is the number of samples in which taxa i and taxa j are commonly occurring. and are the number of times that taxa i or taxa j occurred, respectively. If two taxa have a strong relationship, their index will be closer to zero, while if two taxa are entirely independent, they will have an index of 1 [15].

C. Proposed pipeline

This is a proposed general pipeline applied to the two datasets we are currently using in this paper. The pipeline aims to identify any standard microbial profiles that ASD and control cases may share, using these profiles as a biomarker for diagnosing ASD.

D. Methodology steps

In this paper, an attempt is made to construct a similarity structure using the cosine similarity. Cosine similarity assesses the similarity degree between an individual's microbial composition and tries to connect individuals with a similar composition.

1) Data loading and preprocessing

Importing the abundance matrix and its corresponding metadata file allowed for the preprocessing step to occur by replacing any null values with zeros and normalizing the

samples to relative abundance by dividing each taxon count by the total count per sample.

2) Cosine similarity computation

Compute the pairwise cosine similarity between all the sample vectors based on the normalized taxonomic profile. This allowed for a cosine similarity matrix to be generated, where represents the cosine similarity between samples i and j.

3) Graph construction

Building an undirected graph where each node represents a sample. Only the samples with a threshold exceeding 0.75 similarity were included in the graph.

4) Clustering and community detection

Communities were then detected in the constructed graph, identified by exhibiting high microbial similarity and their average relative abundance.

5) *Taxonomy profiling of clusters*

For these clusters, computations were performed to calculate the top-contributing taxa within each cluster using their average relative abundance.

6) Visualization using Bray-Curtis dissimilarity

The Bray-Curtis dissimilarity was then computed to visualize the community differences. The graph was color-coded according to the metadata to differentiate the ASD from the control samples clearly. A heatmap was then generated to illustrate the dissimilarity of the samples.

7) *Analysis of the obtained results*

The analysis of the obtained results was then performed to attempt and identify a microbial profile for ASD individuals and controls.

E. Pseudo code

1) Read the abundance file where rows = sample, cols = taxa.
2) Null values, 0
3) For each sample, divide by row sum so it becomes relative abundance.
4) S = cosine similarity for relative abundance
5) For i<j : if s[i,j] ≥ 0.75 , add weighted edge (I,j,s[I,j]) to graph G.
6) Read metadata
7) Map sample to sample, autistic = red, neurotypical = blue.
8) Draw a graph coloured by group and label nodes with the sample ID.
9) Cluster connected components.
10) For each component k, write the subset of the counts table for its samples to a CSV. 11) BC_Index (relative-abundance, metric is BrayCurtis).
12) Heat-map color by diagnosis.

For dataset loading, the abundance files were the first step, which was then followed by filling all the null values with 0.

Furthermore, all the samples were converted to relative abundance, so all the rows sum to 1. It is then followed by computing the cosine similarity between every pair of samples, which is modeled as an undirected weighted graph, with an edge of similarity only if the pair similarity exceeds 0.75. This threshold allows only the most important relationships to be shown and avoids those that may result in noise.

The metadata is read from its respective file. Each sample is mapped to its metadata to obtain its diagnosis, with autistic samples mapped to the color red and neurotypical samples to the color blue. These samples are then displayed in a weighted graph to visualize the relationship between them. Then each connected component is treated as a cluster, and its data is saved in a CSV file. This is a topology-defined clustering, where if all the members are mutually reachable via over 0.75 edges, then they belong together. The Bray-Curtis index is then used to visualize the data as a heatmap.

IV. RESULTS

Dataset 1, collected by Kang et al. shows several clusters, these clusters seem to have a mixture of both ASD and controls. where the blue nodes signify the controls and the red nodes signify the ASD samples.

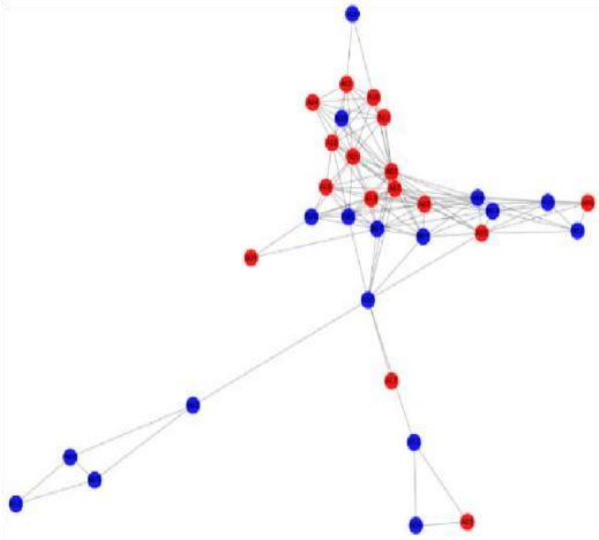


Figure 1. A graph showing the Cosine Similarity undirected graph visualization of samples whose similarity exceeds 0.75 for the first (Kang) dataset.

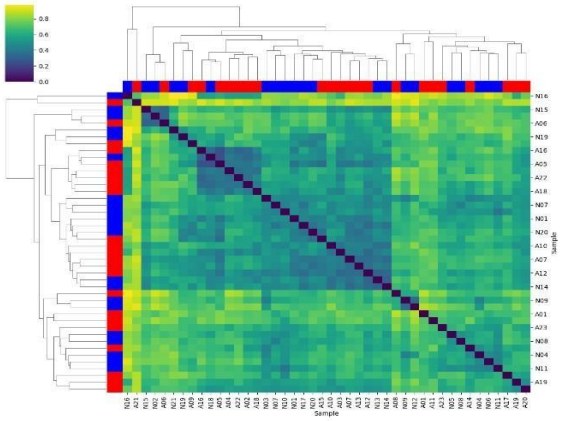


Figure 2. Bray-Curtis dissimilarity heatmap of the first (Kang) dataset.

This dataset provides extensive metadata that can aid in the enrichment analysis. The ATEC and the PDD-BI, which were collected exclusively for individuals with ASD, can help us infer where on the autism scale each individual lies. This will provide clearer information about the microbial structure.

In Figure 1, we can note that there are some dendrogram branches that are dominated by ASD samples (red), while others are dominated by neurotypical samples (blue). Some others have a mixture of both blue and red.

Figure 2 shows the Bray-Curtis dissimilarity of the Kang dataset, where the red indicates the ASD samples and the Blue indicates the neurotypical samples. This was done to help visualize the diagnosis of the samples. For the BrayCurtis dissimilarity, the lower values and darker colors mean that these samples share a similar microbial composition, while the higher values with more yellowish hues indicate greater differences. While the second dataset, as shown in Figure 3, does not show a clear separation, but we can still infer some information from its microbial structure, since this dataset doesn't provide much metadata according to the severity of the ASD.

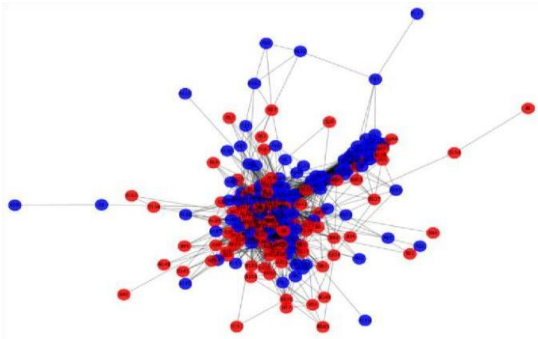


Figure 3. shows the Cosine Similarity undirected graph visualization of samples whose similarity exceeds 0.75 for the second dataset (GSE113690).

Furthermore, using the Bayes-Curtis dissimilarity shown in Figure 4, we can see that the heat map provides us with a more understandable view of the top 100 samples. Most values in this heatmap have an intermediate range.

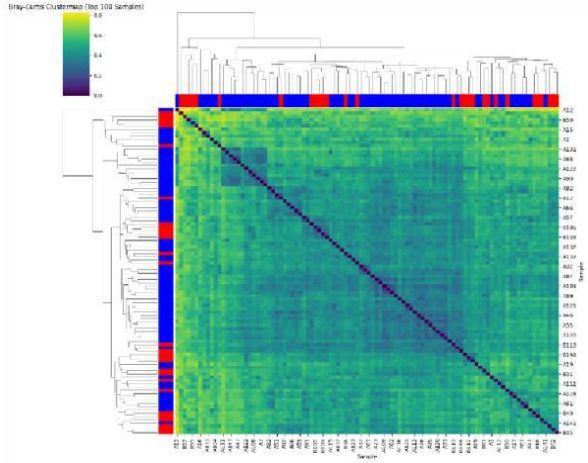


Figure 4. Bray-Curtis dissimilarity heatmap for the second (GSE113690) dataset.

The colors indicate a mixed but uneven distribution, with red samples representing ASD individuals and blue samples representing neurotypical individuals

V. DISCUSSIONS

The analysis of Figure 1 shows some promising microbial structures detected, which show us what a typical gut microbiome diversity looks like. We had 3 clusters where the first cluster had an almost even distribution of ASD and control samples, with 21 control samples and 21 ASD samples, where the most prominent bacteria profile with a mean relative abundance of 0.158, where the most common were Bacteroidota, Bacteroidia, Bacteroidales, Bacteroidaceae, Phocaeicola.

Following through with the second cluster had only one ASD, the sample had a mean relative abundance of 0.298, with the most common bacteria being Firmicutes, Clostridia, Lachnospirales, Lachnospiraceae, Hungatella. Finally, the third cluster has one ASD sample with its mean relative abundance being 0.669 and the most prominate bacteria being Verrucomicrobiota, Verrucomicrobiae, Verrucomicrobiales, Akkermansiaceae, Akkermansia.

Taking a deeper look into these results we can see that in the first cluster there is equal samples of control and ASD, where the most prominent bacteria found can be an indication of a balanced gut barrier and a strong gut barrier when they are present in moderation, but if there is an imbalance or over abundance, this can show various health issues. While for the second cluster the most prominent microbial profile can indicate a healthy and diverse gut eco system, yet Hungatella ‘s abundance may have a potential role in diseases. Moving on to the third cluster where the

presence of this microbial profile can indicate a reduction in the gut inflammation.

Using a genus-level cosine network on the (GSE113690) balanced dataset, we can identify 6 main clusters, where for the first cluster had 43 ASD samples and 47 controls and a mean relative abundance of 0.125, which the most common bacteria being Firmicutes, Clostridia, Clostridiales, Ruminococcaceae, which can have an effect on cognitive function, Fastidiosipila, Clostridiales_bacterium.

The second cluster had 34 ASD samples and 31 controls with a mean relative abundance of 0.104 where the most common bacteria were Proteobacteria, Alphaproteobacteria, Rickettsiales.

The third cluster is comprised of 12 ASD samples and 35 Controls with a mean relative abundance of 0.202 with the most common bacteria being Firmicutes, Clostridia, Clostridiales, Lachnospiraceae, Tyzzerella.

The fourth cluster has 14 ASD samples and 16 Control samples, with a mean relative abundance of 0.079 with the most common bacteria being Firmicutes, Clostridia, Clostridiales, Ruminococcaceae, Fastidiosipila.

The fifth cluster had an equal amount of ASD and Controls with 2 samples each, a mean relative abundance of 0.103 and the most common bacteria being Firmicutes, Clostridia, Clostridiales, Ruminococcaceae, Fastidiosipila. Finally, the sixth cluster has 2 control samples with a relative abundance of 0.116 and the most common bacteria being Proteobacteria, Gammaproteobacteria, Pseudomonadales, Moraxellaceae, Perlucidibaca.

For the first cluster, the bacterial profile is essential to maintain a healthy gut and strengthening the gut barrier. While for the second cluster, has a Rickettsiales , which is considered a pathogen that is responsible for multiple diseases, such as Rocky Mountain spotted fever or typhus [16]. For the third cluster we can see that that the controls are much more abundant where Tyzzerella can act as a biomarker for certain conditions, such as Clostridioides difficile infection. For the fourth cluster and fifth cluster, we can see a similar microbial profile, where we can see that both almost has an equal sample of ASD and controls inside them. where Firmicutes is one of the most abundant bacteria in the human gut. Finally for the sixth cluster, it has Perlucidibaca, which is not typically considered a resident of human gut microbiome.

To conclude this discussion, a gut microbial profile for either, the ASD or the Controls, cannot be deduced at this point, but some promising microbial structures have been detected, which show us what a typical gut microbiome diversity looks like. These results are not definitive yet at this state to act as a biomarker, but they are definitely a step in the right direction to help build a separate microbial profile for controls and ASDs.

VI. CONCLUSION AND FUTURE WORK

The results are still in the early stages and not clear microbial profile has been identified at this point in this ongoing research. Different profiles have been identified, which can benefit from further research to be able to pinpoint how each bacterium in a certain profile interact together, which can show us how healthy the gut microbiome really is, contributing to being able to build a solid microbial profile for ASD individuals. The current datasets were not able to provide us with severity indication, and have shown very similar structures for both the ASD and the controls, proving that both profiles do share a certain degree of similarity in their microbial profile, but at this point the exact difference between such profiles cannot be determined.

Most of the profiles detected are normal occurrences in a healthy gut microbial structure, so further studies will need to be implemented to be able to extract distinct gut microbial profiles that will help us in Diagnosing ASD early on.

From childbirth, a steady interaction is formed between the human body and its environmental microbial structure [18]. To pinpoint the exact effect of different environments on the gut microbiota, a larger sample of individuals will need to be studied over a long period to determine the environmental impact on the gut microbiome of any individual. This will be able to show the effect of the environment, the impact of the diet, and the exposure to medication and antibiotics [19]. It has been proven that the gut microbiota can be affected by antibiotic intake, where antibiotic prescriptions can be a viable solution to treat some more prominent ASD symptoms, by using these prescriptions to alter their microbiota to match that of more neurotypical individuals [20].

Furthermore, using a multi-cohort dataset can prove to be a beneficial way to try and gather a larger dataset. In addition, combining multiple biomarkers can help in detecting ASD symptoms earlier on and more efficiently, as it will help detect ASD at an earlier age, which in turn will help alleviate some of the severe symptoms that can be treated if detected early on.

Finally, more metadata should be collected about the samples, since ASD is not only related to the physical symptoms, but the attitude and the behavior of the diagnosed individual can act as an indicator that can help diagnose ASD earlier. A more refined dataset with not only microbiome data, but also combined with behavior and attitude of the affected individuals and controls can provide for useful for more extensive research in ASD biomarkers. These Datasets can help us understand and classify the severity of the ASD diagnosis and treat the patient accordingly.

ACKNOWLEDGMENT

I would like to give my deepest appreciation and gratitude to my dear supervisors for their constant support,

effort and advisement. Furthermore, I would like to thank my family and friends for all their support and encouragement.

REFERENCES

[1] M. A. Taniya et al., "Role of Gut Microbiome in Autism Spectrum Disorder and Its Therapeutic Regulation," Front. Cell. Infect. Microbiol., vol. 12, p. 915701, July 2022, doi: 10.3389/fcimb.2022.915701.

[2] D. Johnson, V. Letchumanan, S. Thurairajasingam, and L.H. Lee, "A Revolutionizing Approach to Autism Spectrum Disorder Using the Microbiome," Nutrients, vol. 12, no. 7, p. 1983, July 2020, doi: 10.3390/nu12071983.

[3] S. A. Ahmed et al., "Study of the gut Microbiome Profile in Children with Autism Spectrum Disorder: a Single Tertiary Hospital Experience," J. Mol. Neurosci., vol. 70, no. 6, pp. 887–896, June 2020, doi: 10.1007/s12031-020-01500-3.

[4] K. L. Glassner, B. P. Abraham, and E. M. M. Quigley, "The microbiome and inflammatory bowel disease," J. Allergy Clin. Immunol., vol. 145, no. 1, pp. 16–27, Jan. 2020, doi: 10.1016/j.jaci.2019.11.003.

[5] K. T. Mohamed, S. Shabayek, N. F. Mahmoud, M. M. Tawfick, and A. M. S. Hanora, "Investigation of bacterial gut microbiome in diverse Egyptian populations 'pilotstudy,'" Cell. Mol. Biol., vol. 71, no. 1, pp. 75–87, Feb. 2025, doi: 10.14715/cmb/2025.70.1.8.

[6] J. A. O'Donnell, T. Zheng, G. Meric, and F. Z. Marques, "The gut microbiome and hypertension," Nat. Rev. Nephrol., vol. 19, no. 3, pp. 153–167, Mar. 2023, doi: 10.1038/s4158102200654-0.

[7] J. Hrnciarova et al., "Modulation of Gut Microbiome and Autism Symptoms of ASD Children Supplemented with Biological Response Modifier: A Randomized, DoubleBlinded, Placebo-Controlled Pilot Study," Nutrients, vol. 16, no. 13, p. 1988, June 2024, doi: 10.3390/nu16131988.

[8] L. N. Peralta-Marzal et al., "A robust microbiome signature for autism spectrum disorder across different studies using machine learning," Sci. Rep., vol. 14, no. 1, p. 814, Jan. 2024, doi: 10.1038/s41598-023-50601-7.

[9] M. A. Taniya et al., "Role of Gut Microbiome in Autism Spectrum Disorder and Its Therapeutic Regulation," Front. Cell. Infect. Microbiol., vol. 12, p. 915701, July 2022, doi: 10.3389/fcimb.2022.915701.

[10] J. Fouquier et al., "The Gut Microbiome in Autism: StudySite Effects and Longitudinal Analysis of Behavior Change," mSystems, vol. 6, no. 2, p. 10.1128/msystems.00848-20, Apr. 2021, doi: 10.1128/msystems.00848-20.

[11] B. S. Sivamaruthi, N. Suganthy, P. Kesika, and C. Chaiyasut, "The Role of Microbiome, Dietary Supplements, and Probiotics in Autism Spectrum Disorder," Int. J. Environ. Res. Public. Health, vol. 17, no. 8, p. 2647, Apr. 2020, doi: 10.3390/ijerph17082647.

[12] D.-W. Kang et al., "Differences in fecal microbial metabolites and microbiota of children with autism spectrum disorders," Anaerobe, vol. 49, pp. 121–131, Feb. 2018, doi: 10.1016/j.anaerobe.2017.12.007.

[13] Z. Dan et al., "Altered gut microbial profile is associated with abnormal metabolism activity of Autism Spectrum Disorder," Gut Microbes, vol. 11, no. 5, pp. 1246–1267, Sept. 2020, doi: 10.1080/19490976.2020.1747329.

[14] P. Xia, L. Zhang, and F. Li, "Learning similarity with cosine similarity ensemble," Inf. Sci., vol. 307, pp. 39–52, June 2015, doi: 10.1016/j.ins.2015.02.024.

[15] Z. Liu, A. Ma, E. Mathé, M. Merling, Q. Ma, and B. Liu, "Network analyses in microbiome based on high-throughput multi-omics data," Brief. Bioinform., vol. 22, no. 2, pp. 1639–1655, Mar. 2021, doi: 10.1093/bib/bbaa005.

[16] C. D. Putnam, "Evolution of the methyl directed mismatch repair system in Escherichia coli," DNA Repair, vol. 38, pp. 32–41, Feb. 2016, doi: 10.1016/j.dnarep.2015.11.016.

[17] S. Y. Geerlings, I. Kostopoulos, W. M. De Vos, and C. Belzer, "Akkermansia muciniphila in the Human Gastrointestinal Tract: When, Where, and How?," Microorganisms, vol. 6, no. 3, p. 75, July 2018, doi: 10.3390/microorganisms6030075.

[18] G. A. Ogunrinola, J. O. Oyewale, O. O. Oshamika, and G. I. Olasehinde, "The Human Microbiome and Its Impacts on Health," Int. J. Microbiol., vol. 2020, pp. 1–7, June 2020, doi: 10.1155/2020/8045646.

[19] R. K. Weersma, A. Zhernakova, and J. Fu, "Interaction between drugs and the gut microbiome," Gut, vol. 69, no. 8, pp. 1510–1519, Aug. 2020, doi: 10.1136/gutjnl-2019320204.

[20] M. T. Sorbara and E. G. Pamer, "Microbiome-based therapeutics," Nat. Rev. Microbiol., vol. 20, no. 6, pp. 365– 380, June 2022, doi: 10.1038/s41579-021-00667-9.

From Prebiotic Sustainability to Proto-Liveliness: A Biocomputational Framework with Structural, Energetic and Informational Metrics for Protocell Cluster Evolution

Michael Massoth
Department of Computer Science, Hochschule Darmstadt (h_da)
University of Applied Sciences Darmstadt, member of European University of Technology (EUt+)
Darmstadt, Germany
e-mail: michael.massoth@h-da.de

***Abstract*- This paper presents a quantitative, operational framework for classifying protocell clusters along a continuum from passive stability to proto-biological liveliness. We define a three-dimensional Prebiotic Sustainability Space with measurable structural (S), energetic (E), and informational (I) metrics, and introduce a Life–Death Threshold ($L > L_{crit}$) that integrates these metrics into a single rule distinguishing stable, dying, and proto-living systems. Proto-liveliness is formalized as the capacity of a protocell ensemble to maintain structural connectivity, support persistent gradients, and regenerate attractor-based functional states. We further distinguish material from organizational quantities and formulate a Bridge Principle in which coupled S–E–I feedback loops render recurrent physical patterns persistence-relevant (proto-semantic) and dynamically effective (proto-pragmatic). Microfluidic protocell arrays and AI-based simulations are proposed as complementary routes for empirical classification and validation within this unified physical–informational framework.**

Keywords-Prebiotic sustainability; protocell cluster; proto-liveliness; dissipative structures; origins of life.

I. INTRODUCTION

This seventh paper concludes the series "A Constructivist Proto-Bio-Information Theory: A Physically Grounded Nano-Systems Architecture for Prebiotic Emergence, Information, Proto-Semantic Function, and Sustainability of Protocell Aggregation and Cluster Formation."

Massoth [1] shows that Casimir–Lifshitz forces produce robust attraction and stable 5–200 nm protocell clusters, providing the structural basis for the present formulation of prebiotic sustainability.

Massoth [2] demonstrates that these clusters yield reproducible mesoscale attractors and autonomous ε-machine dynamics, supplying the informational substrate extended here into a sustainability dimension.

Massoth [3] establishes that cluster dynamics generate reproducible differences and functional meaning states, grounding the informational sustainability developed in this work, while Massoth [4] reveals that resonance zones and ε-machines generate syntactic, semantic, and pragmatic information layers, whose stabilization frames sustainability as an emergent multilayer process.

Massoth [5] identifies Matsubara-mode selection and Casimir–Polder coupling as sources of persistent proton gradients, forming the energetic core that this paper extends into energetic sustainability and connects to experimental tests, while Massoth [6] develops a physically grounded concept of prebiotic sustainability in the nanoworld.

The structure of the paper is as follows: Section II motivates the transition problem from fluctuation-stabilized protocell clustering to prebiotic information and proto-agency and defines the key retention/recovery and dissipative-structure metrics. Section III introduces the Prebiotic Sustainability Space (S, E, I) and operationalizes structural, energetic, and informational sustainability with measurable scores. Section IV formulates quantitative criteria for proto-liveliness via coupled thresholds and a Life–Death function in S–E–I space and illustrates exemplar classifications. Section V distinguishes material from organizational quantities and formalizes four levels (statistical–syntactic–semantic–pragmatic) linking dynamics to function. Section VI states the Bridge Principle and derives coupled structural, energetic, and informational feedback loops as the minimal mechanism generating proto-semantic and proto-pragmatic organization. Section VII relates the framework to autopoiesis, work-cycle theories, and information-theoretic accounts and clarifies the novel contributions. Section VIII proposes experimental and simulation probes (microfluidic arrays and AI-based models) to measure S, E, I and test the Life–Death threshold. Section IX summarizes implications, limitations, and future directions for empirical validation and design of synthetic proto-biological systems.

II. FROM PHYSICAL STABILITY TO PREBIOTIC INFORMATION AND AGENCY

Protocellular systems serve as key models of early chemical evolution [12]. They represent minimal, physically stabilized compartments capable of taking up energy, generating gradients, and producing spatial or temporal patterns. Theory and experiment show that vesicles can couple into clusters through short-range, fluctuation-induced forces [11], in particular Casimir and Casimir-like nanoscale attractions [9][10][13][14][15][16]. These interactions impose spatial order and support collective dynamics such as

 ISBN: 978-1-68558-361-3

reinforced gradients, synchronized fluctuations, and shared energy flows.

Recent work has enabled increasingly precise characterization of the structural and energetic stability of such assemblies. Spectral descriptors of cluster geometry quantify robustness; characteristic timescales of gradient formation and decay capture energetic performance. Together, these approaches demonstrate that protocell clusters can form persistent, non-equilibrium structures maintained under continuous dissipation [19].

What remains unresolved, however, is how such *passive persistence* transitions into *functional liveliness*—that is, into a regime where a system actively reproduces and restores the conditions of its own persistence.

This motivates the central research question (RQ): *Which physical conditions mark the transition from mere stability to self-regenerative, proto-biological agency in protocell clusters?*

A. Physical Origin of Prebiotic Information in Protocell Clusters

Prebiotic information in protocell clusters arises when fluctuation-stabilized mesoscale geometries form reproducible macrostates whose future dynamics are better predicted from the macrostate than from the full microscopic configuration. Casimir–Lifshitz coupling restricts the accessible configuration space to a small set of energetically favored cluster geometries, giving rise to discrete attractors that persist despite thermal noise.

These attractor macrostates function as a physical memory substrate: past configurations bias future evolution, enabling pattern retention and recovery without polymers, genetic encoding, or biochemical control. Information, in this prebiotic sense, is not symbolic but structural and dynamical, grounded in geometry, interaction topology, and reproducible transition pathways.

B. Operational Measures of Pattern Retention and Recovery

Pattern retention is quantified by the persistence of predictive state structure in a reconstructed ε-machine, for example by the statistical complexity C_μ, which measures the memory stored in causal states, and by the excess entropy E which captures the predictable information shared between past and future observations. Pattern recovery is quantified by perturb-and-return statistics, including the regeneration time T_{regen} required to return to a dominant attractor basin after disturbance, and the return probability $P_{return}(T)$ that the system re-enters this basin within time T.

From an information-theoretic perspective, pattern retention and recovery are quantified by informational closure and attractor-based memory. Coarse-grained macrostates $Z_t=f(X_t)$ satisfy a strong reduction in conditional entropy, $H(Z_{t+1} \mid Z_t) \ll H(Z_{t+1} \mid X_t)$, indicating that cluster geometry retains predictive information about future configurations. Recovery after perturbations follows deterministic return trajectories into the same attractor basin, expressed as

$$\|Z_{t+}1-A_i\| < \|Z_t-A_i\|.$$

Mutual information and excess entropy quantify long-range temporal correlations encoded in ε-machine causal states.
Long-lived dissipative structures are identified by sustained entropy production $dS/dt > 0$, low normalized structural drift $D(t,\Delta t) < \delta$, high return probability $P(Z_{t+1} \in A_i \mid Z_t \in A_i) \approx 1$, and persistence of energetic patterns with $\tau_{gradient} \geq \tau_{diff}$. Together, these metrics demonstrate metastable, information-bearing dissipative protocell clusters.

In this paper, a *long-lived dissipative structure* is operationally indicated when structural cohesion (e.g., positive algebraic connectivity λ_2), energetic persistence ($\tau_{gradient} / \tau_{diff} > 1$), and informational recovery (high C_μ, short T_{regen}, and large $P_{return}(T)$) are simultaneously observed.

A related challenge is to describe functional self-organization, system identity, and the emergence of semantic and pragmatic structure within a purely physical formalism, without invoking modern cellular complexity or genetic storage.

C. From Prebiotic Information to Sustainability and Proto-Agency

Prebiotic sustainability denotes the maintenance of structural, energetic, and informational order. Prebiotic life emerges only when a sustainable system acquires autonomous capacities for self-restoration, attractor reconstruction, and functional expansion. To formalize this transition, we introduce a Prebiotic Sustainability Space, capturing structural (S), energetic (E), and informational (I) sustainability as core parameters of proto-biological persistence. Building on this, a Life–Death Threshold expressed through metric inequalities classifies cluster states along a continuum from stable to proto-liveliness.

We further distinguish material quantities—fields, forces, gradients, energy flows—from organizational ones that encode functional patterns, attractors, and proto-semantic structure. A Bridge Principle characterizes the recursive feedback loops through which physical patterns acquire functional significance and generate early forms of agency. Here, proto-agency denotes purely physical capacities for self-restoration, state reconstruction, and constraint preservation, without cognition, intention, or genetic control. A bio-computational framework demonstrates that all quantities are tractable and allow sharp classification of system states.

III. Prebiotic Sustainability Space (S, E, I)

This section defines the three-dimensional sustainability space and specifies measurable structural (S), energetic (E), and informational (I) scores for placing protocell clusters on a common quantitative scale.

A. Motivation

"Prebiotic" is used here in the origins-of-life sense (pre-genetic, pre-cellular organization). Prebiotic protocells could form higher-order units only if three fundamental domains remained jointly stable: structural integrity (S), persistence of dissipative energy flows (E), and reproduction of internal patterns and functional relations (I). These dimensions define a sustainability space in which proto-biological systems can persist, interact, and generate selectable variation. The transition from a merely stable aggregate to a functionally autonomous unit becomes intelligible only when structural, energetic, and informational conditions are considered together.

Definition: Prebiotic sustainability denotes the ability of an open protocell ensemble to maintain its structural, energetic, and informational identity while continuously dissipating energy. A system is sustainable if it preserves its mesoscale architecture, supports long-lived nonequilibrium gradients, and regenerates recurrent physical patterns despite fluctuations and molecular turnover.

B. Structural Sustainability S

Structural sustainability captures a cluster's capacity to maintain physical organization despite defects, thermal noise, or transient forces. It can be quantified using graph-theoretic measures. A key parameter is the second eigenvalue λ_2 of the Laplacian of the contact graph, which measures algebraic connectivity. For $\lambda_2 > 0$ the cluster remains connected; larger values indicate higher robustness. Structural redundancy R_{struct} measures the number of alternative load-bearing paths. Together they define $S = f(\lambda_2, R_{struct})$,

which increases with network connectivity and path diversity. S thus characterizes the physical basis for maintaining spatial organization over relevant timescales.

Operationalization of Structural Sustainability S:

Structural sustainability captures a cluster's capacity to maintain physical organization despite defects, thermal noise, and transient forces. We represent a protocell cluster as a *contact graph* G(t), where nodes are vesicles and edges indicate sustained proximity/contact (e.g., gap $L \leq L_{contact}L$ for longer than a dwell time $\tau_{contact}$).

A key robustness descriptor is the *algebraic connectivity* of the graph, defined by the second eigenvalue of the (combinatorial) Laplacian $L_G(t)$:

$$\lambda_2(t) = \text{second\_smallest\_eigenvalue}\,(L_G(t))$$

For $\lambda_2(t) > 0$ the contact graph is connected; larger λ_2 indicates stronger global cohesion and lower fragmentation risk under random edge/node loss. Because λ_2 depends on the whole topology, it serves as a mesoscale stability indicator rather than a purely local measure.

We complement λ_2 by a redundancy term that captures alternative load-bearing paths. One simple operational choice is an edge-redundancy score (normalized to [0,1]):

$$R_{struct}(t) = 1 - (B(t) / E(t))$$

where B(t) is the number of bridge edges (edges whose removal disconnects the graph) and E(t) is the number of edges. High R_{struct} implies multiple alternative paths and lower single-edge failure sensitivity.

We define a dimensionless structural sustainability score S(t) by mapping the two measures into [0,1]:

$$S(t) = \sigma(\,(\lambda_2(t) - \lambda_{2,ref}) / s_\lambda\,) \cdot \sigma(\,(R_{struct}(t) - R_{ref}) / s_R\,)$$

where $\sigma(x) = 1 / (1 + e^{(-x)})$ is a squashing function, $\lambda_{2,ref}$ and R_{ref} are reference values (e.g., medians in a calibration window), and s_λ, s_R set sensitivity.

Interpretation. S increases when clusters remain connected ($\lambda_2 > 0$) and structurally redundant (high R_{struct}), providing the physical substrate for persistent gradients and recurrent patterns.

C. Energetic Sustainability E

Energetic sustainability measures the ability of a cluster to generate, stabilize, and maintain gradients as dissipative flows [19]. Central is the ratio $\tau_{gradient} / \tau_{diff}$, which determines whether gradients build up faster than diffusion erodes them. The efficiency η quantifies the fraction of absorbed energy converted into structure-stabilizing dynamics. These parameters define

$$E = g(\tau_{gradient} / \tau_{diff}, \eta),$$

which attains high values when energy-driven patterns emerge reproducibly and withstand perturbations. Energetic sustainability forms the physical foundation of dissipative structure formation and early proto-metabolic feedback.

Operationalization of Energetic Sustainability E:

Energetic sustainability measures the ability of a cluster to generate, stabilize, and maintain gradients as dissipative energy flows. The core energetic object is the *electrochemical potential difference* across a relevant interface (e.g., gap vs bulk, or inside vs outside for vesicles). For protons: $\Delta\mu_{H+}(t) = k_B T \ln(\, c_{in}(t) / c_{out}(t)\,) + e\,\Delta\psi(t)$

Equivalently in pH form:

$$\Delta\mu_{H+}(t) = 2.303\, k_B T \cdot \Delta pH(t) + e\,\Delta\psi(t)$$

The free energy stored in a proton gradient (over an effective number of charge carriers N_{H+}) is:

$$E_{grad}(t) = N_{H+}(t) \cdot \Delta\mu_{H+}(t).$$

In minimal protocell systems where $\Delta\psi$ is small or not measured, the pH term provides a conservative lower-bound proxy. A necessary persistence condition is that gradients survive diffusion. We therefore use the ratio of gradient lifetime to diffusive relaxation time:

$$\kappa(t) = \tau_{gradient}(t) / \tau_{diff}(t)$$

with $\tau_{diff}(t) \approx L_{gap}(t)^2 / D_{H+}$ in a gap of characteristic width L_{gap} and proton diffusion coefficient D_{H+}. Energetic sustainability requires $\kappa > 1$, meaning dissipative processes build/maintain gradients faster than diffusion erodes them.

We further introduce an energetic allocation/efficiency term $\eta(t)$ that quantifies what fraction of absorbed/available energy supports stabilization (gradients, structural cohesion, pattern retention) rather than immediate dissipation.

Operationally (experiment-compatible), η can be defined from measurable proxies:

$$\eta(t) = P_{stab}(t) / (P_{in}(t) + \varepsilon)$$

where P_{stab} is the estimated power associated with maintaining gradients/structure (e.g., gradient energy change rate plus stabilization-relevant work proxies) and P_in is the net incoming power (thermal, chemical, radiative, or imposed microfluidic work), with ε preventing division by zero. In practice, η is typically estimated comparatively across conditions rather than absolutely.

We define energetic sustainability E(t) as a bounded score:

$$E(t) = \sigma((\kappa(t) - 1) / s_\kappa) \cdot \sigma((\eta(t) - \eta_{ref}) / s_\eta)$$

Threshold note (η). Values such as η>0.4 are not universal constants; they are *operational working thresholds* indicating that stabilization is not a minor side effect but a dominant energetic role. In experiments/simulations, η_{ref} and s_η should be calibrated from control recordings (e.g., isolated vesicles vs clusters).

D. Informational Sustainability I

Informational sustainability denotes a system's capacity to regenerate internal functional patterns and their causal relations. It captures organizational competence beyond structure and energetics. The number of stable attractors N_{attr} reflects functional resolution, while statistical complexity C_μ quantifies the memory required to reproduce system dynamics [18]. The regeneration time T_{regen} measures how rapidly perturbations are resolved by returning to a stable attractor. Together they define $I = h(N_{attr}, C_\mu, T_{regen})$, whose higher values indicate systems capable of reliably reproducing their own patterns and developing proto-semantic structure.

Operationalization of informational Sustainability I:

Informational sustainability denotes a system's capacity to regenerate internal functional patterns and their causal relations—organizational competence beyond geometry and energetics. We operationalize I using three measurable ingredients: (i) the number of functionally distinct, stable macrostates (*attractors*), (ii) the memory/structure required to reproduce observed dynamics (statistical complexity), and (iii) the speed of recovery after perturbations.

Let $N_{attr}(t)$ be the number of stable attractors identified in a state-space reconstruction (e.g., clustering of trajectories, recurrence plots, or ε-machine state structure). Let $C_\mu(t)$ denote *statistical complexity*, i.e., the Shannon entropy of causal states in an ε-machine reconstruction:

$$C_\mu(t) = H[S_t]$$

Let T_{regen} be the characteristic regeneration time after a standardized perturbation (e.g., flow pulse, salt jump, temperature step), and T_{break} a characteristic decay/break time (e.g., mean time to fragmentation or gradient collapse under the same regime).

We define a complete, bounded informational sustainability score:

$$I(t) = (\ln(N_{attr}(t)) / \ln(N_{max})) \cdot (C_\mu(t) / (C_\mu(t) + C_{\mu,ref})) \cdot (T_{break} / (T_{break} + T_{regen}))$$

where N_{max} is an operational maximum used for normalization (e.g., maximum observed in a dataset), and $C_{\mu,ref}$ sets the scale at which complexity contributes substantially.

Threshold note ($N_{attr} \geq 2$). The condition $N_{attr} \geq 2$ is an operational criterion for *functional resolution*: at least two distinct stable macrostates must exist for state-dependent behavior (e.g., "adsorption-favoring" vs "proton-focusing" regimes) and for meaningful regeneration to be non-trivial. This criterion is logically independent of energetic allocation η: a system may expend substantial energy on stabilization (high η) yet still collapse into a single dominant attractor (N_attr≈1). Proto-liveliness requires both energy-supported stability and non-trivial attractor structure.

E. Sustainability Space Representation

To visualize how structural, energetic, and informational sustainability jointly delimit the transition from passive persistence to proto-liveliness, we introduce a three-dimensional sustainability space in which protocell cluster states can be positioned relative to a life–death threshold (Figure 1).

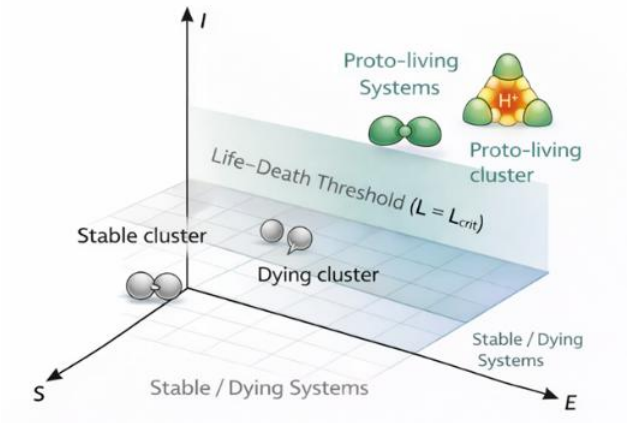


Figure 1. Prebiotic sustainability space and life–death threshold.

In Figure 1, protocell clusters are positioned within a three-dimensional sustainability space defined by structural (S), energetic (E), and informational (I) axes. A life–death threshold separates stable or dying assemblies from proto-living systems. Dimers, isolated protocells, and tetrahedral clusters illustrate how increasing coupled sustainability enables the transition to proto-liveliness.

The dimensions S, E, and I jointly span a Prebiotic Sustainability Space in which protocellular organizations can be placed. Systems with high values in all axes possess both physical robustness and functional autonomy, occupying regions of potentially proto-liveliness behavior. If any axis drops below a critical threshold, the system enters a decay zone where structural coherence, energy flows, or informational patterns cannot be sustained.

This three-dimensional space provides the conceptual basis for defining a numerical Life–Death Threshold.

The three scores define a point in a bounded sustainability space: $(S(t), E(t), I(t)) \in [0,1]^3$.

This representation enables direct comparison of protocell organizations and supports empirical phase-style classification: clusters with high S but low E exhibit cohesive structure without persistent gradients; clusters with high E but low I show sustained gradients without reliable pattern regeneration; and clusters high in all three dimensions are candidates for proto-liveliness under the criteria introduced next.

IV. From Sustainability to Proto-Liveliness: Quantitative Criteria

This section introduces a coupled threshold logic and derives a quantitative Life–Death criterion that classifies protocell clusters as stable, dying, or proto-living within S–E–I space.

A. Need for a quantitative threshold

Structural, energetic, and informational sustainability are necessary conditions for persistence, yet none alone marks the transition to proto-liveliness. The transition is best treated as a coupled threshold phenomenon in the three-dimensional S–E–I space: only when structure, gradients, and recoverable patterns mutually reinforce each other does a protocell cluster move beyond passive persistence toward autonomous self-restoration and functional renewal.

Only when all three dimensions exceed system-critical values does a state emerge in which a protocell cluster actively reproduces its own conditions of persistence. Below these thresholds, a cluster remains either merely physically stable—through mechanical cohesion or dissipative gradients—or disintegrates before functional autonomy can arise. The challenge is to formulate criteria that quantify this transition and allow experimental and theoretical testing.

Definition (Proto-liveliness). Prebiotic liveliness is defined as a sustainable protocell ensemble that autonomously regenerates functional patterns, restores dominant attractor states, and reorganizes energy flows, and thereby stabilizes its identity beyond passive physical persistence.

B. Life–Death Threshold

To operationalize this view, each sustainability dimension is assigned a minimal value above which it contributes to self-supporting organization:

$$S > S_{min},\ E > E_{min},\ I > I_{min}.$$

These conditions are necessary but not sufficient, since functional autonomy requires coordinated interplay among all three components. To capture this interdependence, a proto-liveliness function is introduced:

$$L = w_S\, S + w_E\, E + w_I\, I,$$

where w_S, w_E, and w_I are weighting factors derived from models, empirical data, or heuristic reasoning. Proto-liveliness manifests when the combined value exceeds a critical threshold: $L > L_{crit}$.

L quantifies the emergent capacity of a system to integrate structural integrity, energetic flows, and informational patterns into coherent, reciprocally coupled behavior.

Definition:

A protocell cluster is considered living when it actively maintains structural connectivity, organizes energy flows as stable gradients over diffusive timescales, and generates attractor-based functional states that can regenerate after perturbation. A dying system loses these capacities, and a dead system lacks gradients, connectivity, and attractor reconstruction.

L_{crit} marks the point at which a cluster no longer merely reacts but exhibits active self-regeneration and primitive agency.

C. Example classification

Applying these criteria enables a differentiated classification of protocellular organizations.

A *living* cluster exhibits strong structural connectivity, typically with a positive second Laplacian eigenvalue λ_2, generates stable gradients such that $\tau_{gradient} / \tau_{diff} > 1$, and allocates a significant share of energy to stabilizing processes, e.g., $\eta > 0.4$. $\eta > 0.4$ is tied to energy allocation (a "substantial fraction" supporting stabilization rather than dissipation). $N_{attr} \geq 2$ is tied to functional resolution (at least two distinct stable macrostates).

It possesses multiple functional attractors ($N_{attr} \geq 2$) with substantial statistical complexity C_μ, and regenerates patterns faster than they decay ($T_{regen} < T_{break}$). Such systems robustly satisfy $L > L_{crit}$.

A *dying* cluster shows moderate structural stability, unstable or collapsing gradients, and regeneration times approaching decay intervals ($T_{regen} \approx T_{break}$), reducing both I and L below autonomy-supporting values.

A *dead* cluster is structurally fragmented ($\lambda_2 \to 0$), lacks stable attractors ($N_{attr} \to 0$), loses internal structure ($C_\mu \to 0$), and exhibits diverging regeneration times ($T_{regen} \to \infty$). These systems fall below all sustainability thresholds and remain non-living.

This classification provides the basis for quantitative tables, phase diagrams, and visualizations used later to characterize proto-liveliness transitions with precision.

Clarifying note. The thresholds (e.g., $\kappa > 1$, $N_{attr} \geq 2$, $T_{regen} < T_{break}$, $\eta > \eta_{ref}$) are intended as *operational*, dataset-calibrated criteria. They define falsifiable inequalities that can be measured in microfluidic arrays or simulated ensembles and refined as empirical evidence accumulates.

V. MATERIAL VS. ORGANIZATIONAL QUANTITIES AND THE FOUR LEVELS

This section separates material from organizational variables and uses a four-level hierarchy to show where persistence-relevant meaning and function first emerge in protocell-cluster dynamics.

A. Why distinguish these domains?

The physical description of prebiotic systems comprises two categorically distinct classes of quantities. Material quantities refer to directly measurable parameters—forces, distances, diffusion coefficients, field strengths, energy densities—that determine the mechanical and energetic constraints under which protocell clusters exist. Organizational quantities instead capture emergent patterns and functional relations that cannot be reduced to individual molecules. They describe how a system structures states, transitions, and responses such that certain processes are preferentially produced, regenerated, or stabilized.

The distinction is essential because proto-liveliness is not determined by material parameters alone but by their embedding in self-consistent organizational forms. Material quantities define the level of physical realization; organizational quantities define functional order. Only their interplay generates the proto-liveliness phenomena analyzed in the remainder of the model.

B. The four levels of information: statistical, syntactic, semantic, and pragmatic

To clarify how functional organization emerges from purely material dynamics, we distinguish four hierarchical levels of system description, separated by a critical transition at which functional meaning first appears (Figure 2).

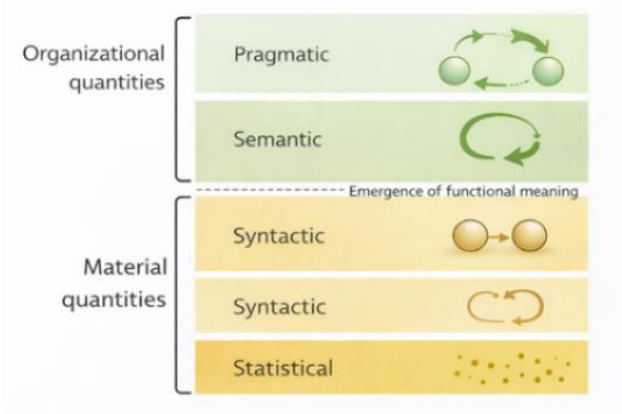


Figure 2. From material dynamics to organizational function.

In Figure 2, four hierarchical levels—statistical, syntactic, semantic, and pragmatic—organize system dynamics from material quantities to organizational structure. Functional meaning emerges at the boundary between syntactic transitions and semantic attractors, marking the shift from purely physical processes to persistence-relevant, functionally organized behavior.

(1) *Statistical level:*

At the statistical level, raw data, fluctuations, and ensemble distributions are described. Concentrations, densities, and thermal noise appear as stochastic variables. Patterns have no identity; they are deviations or aggregates without functional relevance.

(2) *Syntactic level:*

The syntactic level comprises lawful transitions between material states. It is captured by dissipative dynamics, reaction equations, and coupled differential equations. The system state x(t) follows

$$x(t + \Delta t) = F(x(t)),$$

entirely determined by physical mechanisms. This level encodes chemical and mechanical rules without semantic or functional interpretation.

(3) *Semantic level:*

At the semantic level, patterns acquire meaning because they contribute to system persistence. A state x has semantic value when its reappearance after perturbation enhances survival. Meaning arises through recursive stabilization: patterns that support persistence are selectively regenerated. Formally, this is reflected in an attractor A whose maintenance maximizes survival probability. For a relevant state x:

$$P(x(t + \Delta t) \in A \mid \text{disturbance}) \rightarrow 1,$$

indicating that the system functionally orients itself toward attractor restoration.

(4) *Pragmatic level:*

The pragmatic level involves patterns that causally modulate future dynamics and act to preserve the system. Such patterns alter transition probabilities, regulate energy flows, or adjust structural couplings. The dynamics take a bidirectional form:

$$x(t + \Delta t) = F(x(t), y(t)) \text{ and } y(t + \Delta t) = G(y(t), x(t)),$$

where y(t) denotes an organizational variable. At this level, an early form of system-level agency ("agency light") emerges, enabling active influence on conditions of persistence.

C. Core definitions

An organizational quantity structures system dynamic at a higher level without being tied to specific physical carriers. Examples include attractor architectures, regeneration pathways, and functional coupling relations. These quantities emerge from material dynamics yet are not reducible to them, as they represent invariants over trajectories and become systemically causal.

The functional meaning of a state x is defined as its contribution to increasing a system's probability of persistence: $P(\text{Persistence} \mid x) > P(\text{Persistence})$.

This relation quantifies meaning within a consistent physical–functional framework [8]. It links materially grounded dynamics to organizational structure and prepares the transition to the feedback mechanisms discussed in the next chapter, where semantic patterns give rise to pragmatically effective ones.

TABLE I. SUMMARY EQUATION PANEL

Level	Formal Expression	Interpretation
Statistical	$p(x)$; $\langle x^n \rangle$; $S(\omega)$	Ensemble distributions, moments, spectra, and fluctuation statistics
Syntactic	$x(t + \Delta t) = F(x(t))$	Material transition laws governing physical dynamics
Semantic	$P(x(t + \Delta t) \in A \mid$ disturbance $) \rightarrow 1$	Persistence-relevant attractor A defines functional meaning
Pragmatic	$x(t + \Delta t) = F(x(t), y(t))$ $y(t + \Delta t) = G(y(t), x(t))$	Organizational feedback via y(t) modulating future dynamics

Table I, transition from statistical description to pragmatic, function-modulating dynamics through recurrent stabilization and organizational feedback. Semantic meaning arises from persistence-oriented attractors, while pragmatic function emerges when organizational variables causally modulate future system dynamics.

VI. THE BRIDGE PRINCIPLE: FUNCTIONAL FEEDBACK LOOPS

This section states the Bridge Principle and formalizes the minimal feedback architecture by which structural, energetic, and informational loops jointly generate proto-semantic and proto-pragmatic organization.

A. *Central Claim*

The Bridge Principle describes the transition from purely physical patterns to functional organization. This transition begins when recurrent patterns modulate system dynamics and thereby maintain the conditions of their own realization. Physical patterns thus become carriers of functional meaning and seeds of proto-liveliness autonomy.

The principle shows how local structures, energy flows, and dynamic patterns become recursively coupled so that a system not only persists but actively stabilizes the conditions of its reproduction. Early forms of meaning, functionality, and autonomy arise without genetic information, enzymatic networks, or complex metabolic architectures. The Bridge Principle marks the minimal organizational threshold at which physical dynamics shift into proto-semantic and proto-pragmatic behavior, generating an operational precursor of biological agency.

The transition rests on stable feedback cycles in which recurrent patterns stabilize and regenerate the very conditions enabling them. Patterns gain semantic status when they promote persistence and pragmatic force when they causally modulate system dynamics. The shift from material to organizational quantities occurs when such patterns achieve functional stabilization. Semantic properties emerge from recurrence; pragmatic properties from causal influence. The three loops operate as coupled mechanisms of functional organization.

B. *The Three Feedback Loops: structural, energetic, and informational*

To make explicit how functional autonomy arises from recursive physical organization, we represent proto-liveliness as the mutual coupling of structural, energetic, and informational feedback loops (Figure 3).

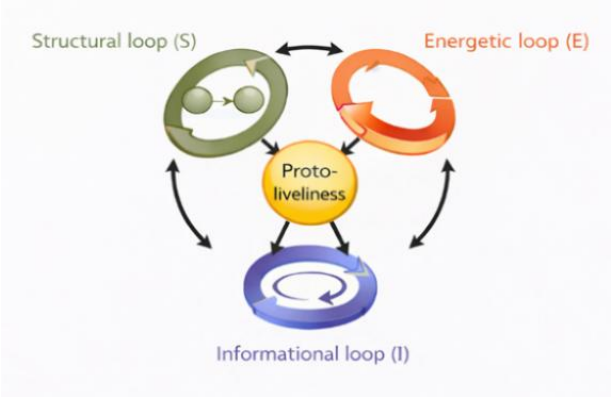


Figure 3. Coupled structural, energetic, and informational feedback loops.

In Figure 3, three mutually coupled feedback loops—structural (S), energetic (E), and informational (I)—form a recursive triad whose interaction stabilizes proto-liveliness. None of the loops is hierarchical; proto-liveliness emerges only from their continuous bidirectional coupling, which jointly sustains structure, gradients, and attractor-based functional patterns.

(1) *Structural loop:*

The structural loop captures how spatial organization increases the likelihood of stable force distributions and thereby reinforces its own persistence. Dense, redundant connectivity—mediated, for example, by Casimir-like nanoscale attractions—increases the Laplacian eigenvalue λ_2. Higher λ_2 lowers the probability of fragmentation, yielding the recursive relation: $\lambda_2 \uparrow \Rightarrow S \uparrow \Rightarrow \lambda_2$ stabilizes.

Spatial structure thus generates the conditions for its own renewal and acts as a dimensionless organizational operator at the mesoscale.

(2) *Energetic loop:*

Energetic feedback arises when energy flows stabilize gradients that in turn facilitate renewed energy uptake. The key parameter is $\tau_{gradient} / \tau_{diff}$, which determines whether dissipative patterns build faster than diffusion erodes them. For $\tau_{gradient} / \tau_{diff} > 1$ a stable gradient cycle can exist. Persistent gradients increase energetic efficiency η, allowing more of the incoming energy to support structure- and

pattern-forming processes. This yields a reinforcing cycle: ($\tau_{gradient}$ / τ_{diff} ↑), η ↑ ⇒ E ↑ ⇒ gradients stabilized.

(3) *Informational loop:*

The informational loop appears when internal attractors generate dynamic patterns that also stabilize those attractors. An attractor with statistical complexity C_μ [18] generates robust regeneration pathways and shortens the regeneration time T_{regen}. Shorter T_{regen} increases the likelihood of returning to the same functional state after perturbation: C_μ ↑, T_{regen} ↓ ⇒ I ↑ ⇒ C_μ stabilizes.

Here proto-semantic meaning becomes visible: a pattern is "meaningful" because its recurrence enhances system persistence, even in the absence of encoded information. All three loops interact tightly. Structure influences energy flows; energy flows stabilize patterns; patterns stabilize structure. Together they produce states with both semantic (pattern-based stability) and pragmatic (dynamic modulation) properties. The Bridge Principle therefore describes the mechanism through which physical dynamics are transformed into functional organization.

VII. Discussion and Relation to Existing Theories

This section situates the proposed metrics and thresholds within established theories of living organization and highlights what the present framework adds in terms of operational measurability and physical grounding.

A. Autopoiesis (Maturana & Varela)

The autopoietic framework characterizes living systems as networks of processes that generate the components sustaining the network itself [17]. The present theory shares this focus but provides a physically grounded, metrically defined account of prebiotic units. While Autopoiesis specifies structural closure qualitatively, the Sustainability Space enables quantitative assessment of the conditions under which protocell clusters achieve structural coherence, energetic recursion, and pattern regeneration: $S > S_{min}$, $E > E_{min}$, $I > I_{min}$.

Incorporating nanoscale coupling forces—especially Casimir-like interactions—adds an explicit physical mechanism for mesoscale coherence. These forces stabilize contact geometries assumed but not mechanistically explained in classical autopoietic models. The framework thus supplies a dynamically quantifiable foundation for autopoietic organization.

B. Kauffman's Work Cycles

Kauffman's notion of work cycles identifies energy flows as drivers of chemical organization [20]. This idea is formalized in $k = \tau_{gradient} / \tau_{diff} > 1$, which ensures that dissipative structures arise faster than diffusion erodes them. What is new here is the integration of structural and informational dimensions. Even strong energy input does not yield proto-liveliness organization when structural coherence is absent ($S \approx 0$) or regenerative patterns are lacking ($I \approx 0$). Only the joint action of S, E, and I determines whether work is translated into functional organization. Energetic sustainability is necessary but not sufficient.

C. Information-theoretic Proto-Liveliness Definitions

Information-theoretic proposals - such as Rosas et al. [7] -define proto-liveliness as the ability to preserve or regenerate informational patterns. This view aligns with informational sustainability, $I = h(N_{attr}, C_\mu, T_{regen})$.

N_{attr} reflects functional resolution, C_μ the memory required for dynamical reproduction, and T_{regen} the rate of recovery. Unlike abstract formulations, the present framework provides a physical implementation of these informational structures. Patterns arise from concrete couplings- geometry, osmotic gradients, Casimir-like forces, and dissipative energy networks. Information processing is not symbolic manipulation but an emergent consequence of physical dynamics enabling attractor formation and stability.

D. What is genuinely new

The key innovation is the introduction of a quantitative Life–Death Threshold within the S–E–I Sustainability Space.

The condition $L = w_S S + w_E E + w_I I > L_{crit}$ provides the first integrated metric for distinguishing stable, dying, and proto-living systems. The Bridge Principle further offers a physically plausible mechanism by which patterns acquire functional significance without genetic coding or enzymatic regulation. Structural, energetic, and informational feedback loops generate proto-pragmatic efficacy, enabling a system to actively stabilize the conditions that support its persistence.

This approach unites physical principles with biological and information-theoretic perspectives, offering an empirically testable account of proto-liveliness organization. Sustainability denotes passive persistence of structural, energetic, and informational order; liveliness entails active self-regeneration and functional renewal under continuous energy flow.

VIII. Outlook: Simulation and Experimental Probes

This section proposes complementary experimental and computational strategies for measuring S, E, and I and for testing the Life–Death threshold under controlled perturbations.

A. Microfluidic protocell arrays

Microfluidic platforms provide controlled environments for probing the sustainability parameters S, E, and I. Vesicle clusters can be reproducibly generated and geometrically confined. Structural sustainability S is obtained via high-resolution imaging and graph-theoretic reconstruction of contact matrices. Algebraic connectivity λ_2 and structural redundancy yield $S = f(\lambda_2, R_{struct})$.

Energetic processes are monitored through fluorescence-based pH, redox, or ion indicators that report gradient formation and lifetime, enabling measurement of

$\tau_{gradient}$ / τ_{diff} and the energetic efficiency η. Informational metrics such as T_{regen} follow from applying defined perturbations and tracking return to dominant patterns. Attractor dynamics, N_{attr}, and C_μ can be extracted from reconstructed trajectories.
Such arrays offer a direct experimental route to evaluating the Life–Death Threshold $L > L_{crit}$.

B. AI-based simulations

Simulations access parameter regimes difficult to probe experimentally. AI-driven optimization—particularly reinforcement learning—adjusts λ_2, η, and N_{attr} to maximize the life function $L = w_S S + w_E E + w_I I$.

This identifies organizational forms representing especially robust proto-liveliness states in the Sustainability Space. Agent-based models with large vesicle populations allow analysis of mesoscale pattern formation, emergent attractor landscapes, and functional robustness under stochastic conditions. They also reveal how the Bridge Principle operates at the population level and how structural, energetic, and informational feedback loops interact. The models generate hypotheses about required combinations of S, E, and I and highlight parameter regions most relevant for experiments.

C. Potential impact

Combining microfluidic experiments with AI-based simulations enables empirical validation of a quantitative physics of prebiotic liveliness. The quantities λ_2, $\tau_{gradient}$ / τ_{diff}, η, N_{attr}, C_μ, and T_{regen} form a coherent set of measurable indicators for functional autonomy, allowing an experimental distinction between stable, dying, and proto-living systems.

The framework also has technological relevance. It may guide the design of artificial proto-biological systems in which dissipative patterns carry functional information. Such proto-biocomputing devices would couple physical self-organization with information-processing dynamics, opening new possibilities in biotechnology, materials science, and unconventional computation. The model's criteria thus illuminate early evolutionary processes while offering an architecture for future bio-inspired technologies.

IX. CONCLUSION AND FUTURE WORK

This work establishes a quantitative framework for the emergence of proto-liveliness in Casimir-coupled protocell clusters. By defining structural (S), energetic (E), and informational (I) sustainability and integrating them into a Life–Death Threshold, we identify precise physical conditions under which a merely persistent assembly becomes functionally autonomous. The Bridge Principle reveals how recurrent physical patterns gain semantic and pragmatic roles through coupled structural, energetic, and informational feedback loops, providing a minimal mechanism for early agency.

The framework shows that proto-liveliness is not the product of any single domain but arises from the coordinated reinforcement of connectivity, dissipative gradients, and attractor-based pattern regeneration. This approach connects autopoietic, energetic, and information-theoretic theories while grounding them in measurable physical quantities.

Microfluidic protocell arrays and AI-based simulations offer concrete routes to testing the model, enabling empirical discrimination between stable, dying, and proto-living systems. As a broader implication, the theory outlines principles for designing artificial proto-biological systems in which dissipative patterns carry functional information, suggesting potential pathways toward bio-inspired materials, chemical computing, and synthetic life.

Take-home message: Life begins when matter organizes itself such that structure, energy, and information cooperatively stabilize the conditions of their own continuation.

REFERENCES

[1] M. Massoth, "Attractive Casimir-Lifshitz Forces as a Universal Driver of Prebiotic Protocell Aggregation and Cluster Formation," *BIOTECHNO*, 2026.

[2] M. Massoth, "Emergent Information Formation in Prebiotic Protocell Clusters: A Computational Mechanics Framework of ε-Machines and Attractor Memory," *BIOTECHNO*, 2026.

[3] M. Massoth, "From Physical Difference to Meaning: A Constructor-Theoretic Framework for Prebiotic Information in Casimir–Lifshitz-Coupled Protocell Clusters," *BIOTECHNO*, 2026.

[4] M. Massoth, "Physical Origin of Proto-Information: Syntax, Semantics and Pragmatic Emergence in Prebiotic Protocell Clusters," *BIOTECHNO*, 2026.

[5] M. Massoth, "Quantum Field-Induced Proton H^+ Gradients in Prebiotic Protocell Clusters," *BIOTECHNO*, 2026.

[6] M. Massoth, "Prebiotic Sustainability in the Nanoworld: A Physical Framework for Early Protocell Cluster Stability," *BIOTECHNO*, 2026.

[7] F. E. Rosas, *et al.*, "Software in the natural world: A computational approach to hierarchical emergence," *arXiv* preprint, 2024, doi:10.48550/arXiv.2402.09090.

[8] M. Deutsch and C. Marletto, "Construction theory of information," *Proc. R. Soc. A*, vol. 471, art. no. 20140540, 2015, doi:10.1098/rspa.2014.0540.

[9] H. B. G. Casimir, "On the attraction between two perfectly conducting plates," *Proc. K. Ned. Akad. Wet.*, vol. 51, pp. 793–795, 1948.

[10] E. M. Lifshitz, "The theory of molecular attractive forces between solids," *Sov. Phys. JETP*, vol. 2, pp. 73–83, 1956.

[11] J. N. Israelachvili, *Intermolecular and Surface Forces*, 3rd ed., Academic Press, London, 2011.

[12] I. Gözen, *et al.*, "Protocells: Milestones and recent advances," *Small*, vol. 18, art. no. 2106624, 2022.

[13] S. K. Lamoreaux, "Demonstration of the Casimir force in the 0.6–6 μm range," *Phys. Rev. Lett.*, vol. 78, pp. 5–8, 1997.

[14] J. L. Garrett, D. A. T. Somers, and J. N. Munday, "Measurement of the Casimir force between two spheres," *Phys. Rev. Lett.*, vol. 120, art. no. 040401, 2018.

[15] S. J. Rahi, T. Emig, N. Graham, R. L. Jaffe, and M. Kardar, "Scattering theory approach to electrodynamic Casimir forces," *Phys. Rev. D*, vol. 80, art. no. 085021, 2009.

[16] A. Canaguier-Durand, G.-L. Ingold, M.-T. Jaekel, A. Lambrecht, P. A. Maia Neto, and S. Reynaud, "Classical Casimir interaction in the plane–sphere geometry," *Phys. Rev. A*, vol. 85, art. no. 052501, 2012.

[17] H. R. Maturana and F. J. Varela, *Autopoiesis and Cognition: The Realization of the Living*, D. Reidel, Dordrecht, 1980, doi:10.1007/978-94-009-8947-4.

[18] J. P. Crutchfield and K. Young, "Inferring statistical complexity," *Phys. Rev. Lett.*, vol. 63, pp. 105–108, 1989, doi:10.1103/PhysRevLett.63.105.

[19] G. Nicolis and I. Prigogine, *Self-Organization in Nonequilibrium Systems*, Wiley-Interscience, New York, 1977.

[20] S. A. Kauffman, "The sciences of complexity and 'origins of order'," *PSA: Proc. Biennial Meet. Philos. Sci. Assoc.*, vol. 1990, no. 2, pp. 299–322, 1990, doi:10.1086/psaprocbienmeetp.1990.2.193076.

Computational Discovery of Inhibitors Targeting Alphavirus nsP2 Proteases

José Augusto Leoncio Gomide, Gabriel de Souza Segadães, Victor Rodrigues Terra, Yasmim Nunes Mesquita, Nilson Nicolau Junior
Instituto de Biotecnologia
Universidade Federal de Uberlândia
Rua Acre, 1004, 38405-319, Uberlândia, Brasil
e-mail: jose.gomide@ufu.br, gabriel.segadaes@ufu.br, victor.terra@ufu.br, yasmim.mesquita@ufu.br, nicolaujr@ufu.br

***Abstract*— The paper tackles the absence of approved drugs for debilitating alphavirus infections by targeting the nonstructural protein 2 protease, a key enzyme for viral replication. A pharmacophore-guided virtual screening workflow combined with molecular docking and all-atom molecular dynamics simulations was employed to investigate potential inhibitors. Within the scope of computational biology and biotechnological applications, this work reports new results from a pharmacophore-guided screening and docking of the NuBBE database and subsequent structural dynamics simulations. Molecular dynamics analysis demonstrated that the ligand nubbe417 exhibits a dual-inhibition potential against both Chikungunya and Mayaro viruses by structurally locking the catalytic loop in a rigid, closed state. The paper concludes that integrating *in silico* screening with dynamic structural analysis is a viable strategy to identify natural compounds, laying the groundwork for future *in vitro* validation of nubbe417 and expanded screening of additional compound databases to develop effective therapies against emerging arboviruses.**

Keywords- Chikungunya virus; Mayaro virus; pharmacophore modeling; molecular dynamics.

I. INTRODUCTION

Alphaviruses are enveloped, single-stranded positive-sense RNA viruses belonging to the Togaviridae family, primarily transmitted by mosquitoes and divided into two clinical groups: arthritogenic and encephalitic [1]. Among the arthritogenic viruses, Chikungunya (CHIKV) and Mayaro (MAYV) stand out as major public health threats, causing acute fever accompanied by debilitating polyarthralgia and, frequently, chronic pain that can persist for months or even years [2]. CHIKV caused global outbreaks, while MAYV remains South American but shows expansion potential [1], [2]. These infections significantly reduce quality of life and productivity, particularly in resource-limited endemic areas [2].

Despite their global impact, CHIKV and MAYV remain neglected tropical diseases, with no specific therapies or widely available vaccines [3]. Current treatment is limited to analgesics and anti-inflammatory drugs, and although efforts in antiviral development exist, none have yet been approved [3]. This therapeutic gap highlights the urgent need for new drug discovery strategies.

Alphavirus replication depends on the expression and processing of nonstructural polyproteins, mediated by the nsP2 protein, which contains a C-terminal cysteine protease domain, whose proteolytic activity is essential for releasing components of the replicase complex, being indispensable for viral replication [4], [5]. Structural studies reveal conserved pockets across different alphavirus species, suggesting potential for broad-spectrum antivirals [6].

The druggability of nsP2 has already been demonstrated: covalent inhibitors with electrophilic groups and computational approaches such as virtual screening, docking, and repurposing of approved drugs have identified promising compounds [4], [5], [7]. Previous computational studies [4], [5] have successfully utilized virtual screening and docking to identify potential inhibitors for alphavirus nsP2 proteases. However, a common weakness in existing literature is the primary focus on single-virus targets and the reliance on static binding models, which often fail to account for the dynamic flexibility of the interdomain loop that regulates active-site access.

This raises a critical research question: can a single natural compound achieve broad-spectrum inhibition by effectively modulating the conformational dynamics of this conserved loop across different alphaviruses? The purpose of this article is to address this gap by identifying a dual-inhibitor for both Chikungunya (CHIKV) and Mayaro (MAYV) viruses. Unlike standard virtual screening workflows, this study introduces a novelty by proposing a 'loop-locking' inhibition mechanism, where the ligand stabilizes the catalytic loop in a rigid, closed state to prevent substrate processing. Although this approach is currently limited by its purely *in silico* nature and the necessity for future *in vitro* validation, it provides a mechanistic framework for developing broad-spectrum therapies targeting the structural dynamics of emerging arboviruses.

Despite the global burden of CHIKV and MAYV, effective treatments remain elusive. To address this gap, we target the highly conserved nsP2 protease, a critical component for viral replication. By utilizing bioinformatics tools for rational drug design, this study aims to identify potent inhibitors against nsP2, establishing a foundation for their subsequent pharmaceutical application.

The rest of the paper is structured as follows. Section 2 presents the Materials and Methods, Section 3 the Results, Section 4 the Discussion, and Section 5 the Conclusion and Future Work.

II. MATERIALS AND METHODS

To identify potential inhibitors, this study utilized an integrated computational pipeline combining structure-based pharmacophore modeling with virtual

screening, molecular docking, and all-atom molecular dynamics simulations. The computational resources and software protocols utilized in this study are detailed below.

A. Infrastructure

All preprocessing, pharmacophore construction, virtual screening, and docking were performed on PowerEdge T550 workstation (Dell Technologies, United States), equipped with an Intel® Xeon® Silver 4316 processor (32 physical cores, 2.3 GHz). These tasks were executed primarily within the Molecular Operating Environment (MOE, v. 2024.0601, Chemical Computing Group, Canada).

Molecular Dynamics (MD) simulations were performed with GROMACS (v. 2023.2, GROMACS Development Team, Sweden/Germany) on the Santos Dumont Supercomputer (Laboratório Nacional de Computação Científica — LNCC, Brazil). Simulations were executed on Intel Xeon Gold 6154 ("Skylake") processors per node (36 physical cores total, 2.1 GHz).

B. Structure Preparation

The crystal structure of the nsP2 protease from Chikungunya virus (CHIKV) (PDB ID: 3TRK) was retrieved from the Protein Data Bank (Protein Data Bank, RCSB, United States). The protease from Mayaro virus (MAYV), for which no experimental structure exists, was modeled using ColabFold/AlphaFold2 (v 1.5.5, ColabFold Project) based on its sequence retrieved from National Center for Biotechnology Information (NCBI) (NCBI Reference Sequence: NP_740688.1). MAYV model quality was assessed via SAVES v.6.0 (Ramachandran plot). Both nsP2 structures were prepared in MOE using QuickPrep.

C. Pharmacophore Modeling, Virtual Screening and Docking

Binding pockets were identified using the Site Finder module, and pharmacophoric interaction points of the receptor were mapped using the Pharmacophore Query Editor. Electrostatic potential surfaces and interaction-energy grids were analyzed to refine essential pharmacophoric features.

Because no crystallized ligands or experimentally confirmed inhibitors exist for MAYV nsP2, pharmacophore validation was performed only for the CHIKV nsP2 using known inhibitors reported in the literature. Validation utilized property-matched decoys generated via DUDE-Z. Docking scores and pharmacophore metrics were analyzed to calculate ROC curves and enrichment factors using a specific MOE script. (ROC Calculator for MOE Docking, v. XX; R. Staub, GitHub repository) [8]. The validated pharmacophore model was subsequently employed to screen the NUBBE Database (Núcleo de Bioensaios, Biossíntese e Ecofisiologia de Produtos Naturais, UNESP, Brazil). It should be noted that pharmacophore validation was performed exclusively for CHIKV nsP2 due to the absence of experimentally confirmed MAYV inhibitors. The application of this model to MAYV assumes structural and functional conservation of the catalytic site, which is supported by sequence and structural similarity but remains a limitation of the present study. Database compounds were curated and protonation-normalized using OpenEye Applications Software (v. 24.05.15.0, OpenEye Scientific, United States), including pKa adjustment at physiological pH (7.4), tautomer enumeration, and 3D conformer generation.

Filtered ligands were docked into both nsP2 proteases. The validated pharmacophore was used as a 3D constraint during ligand placement, followed by scoring with London dG and refinement via induced-fit optimization, generating 30 poses per ligand. A consensus selection strategy was adopted, in which only ligands appearing within the top five ranked sets for both receptors were advanced to MD simulations.

D. Molecular Dynamics

In order to perform MD, the ligand topologies were generated using SwissParam (Swiss Institute of Bioinformatics, Switzerland). The system was solvated and neutralized by adding counterions, and the ionic strength was adjusted to 0.15 M NaCl to mimic physiological conditions. Energy minimization was performed using the steepest-descent algorithm. Equilibration consisted of a 100-ps NVT phase followed by a 100-ps NPT phase. Temperature was maintained at 300 K using the V-rescale thermostat. Pressure was controlled at 1 atm using the Parrinello–Rahman barostat with isotropic coupling. Production MD runs (200 ns) were carried out using the TIP3P water model and the CHARMM27 force field. Analyses included protein–ligand hydrogen bonding, protein RMSD, RMSF and comparison between the first and last MD frames.

III. Results

The outcomes of our integrated computational pipeline are detailed in this section, covering the complete trajectory from initial target characterization to ligand identification. These findings provide a comprehensive overview of the structural insights gained through both static modeling and dynamic simulations of the protease complexes.

E. Protein Molecular Modeling

The crystal structure of CHIKV nsP2 (PDB ID: 3TRK; 2.40 Å resolution) was used as the reference model, while the MAYV nsP2 structure was obtained through AlphaFold2. Structural validation of the MAYV model demonstrated good stereochemistry, with 92.6% of residues in favored regions of the Ramachandran plot, supporting its suitability for subsequent analyses. In both proteins, the identified active site corresponds to the same catalytic amino-acid ensemble described in the literature for CHIKV, including the residues that form the flexible interdomain loop controlling access to the catalytic pocket. Despite differences in residue numbering relative to the reference study, structural equivalence was confirmed.

This loop plays a critical functional role, as it contains residues essential for catalysis and substrate recognition. As reported previously [8], mutation of Asn547 (equivalent to Asn79 in the present numbering) results in a threefold

increase in Km, underscoring its importance for proper substrate positioning. Thus, alterations in the dynamics of this loop directly affect the catalytic competence of the protease.

F. Pharmacophore Modeling, Virtual screening and Docking

Pharmacophore analysis of the catalytic site revealed a robust set of interaction points, which were subsequently refined using electrostatic and energy-interaction maps. Validation with known CHIKV inhibitors yielded satisfactory performance (AUC = 0.88), reinforcing the discriminatory power of the model and supporting its use in the virtual screening of the NuBBE database. Docking allowed the selection of the top five ranked compounds for each protease; among them, only nubbe417 appeared at the top for both, motivating its selection for molecular dynamics simulations.

G. Molecular Dynamics: Structural Stability from RMSD and Local Fluctuations (RMSF)

The global structural stability and local residue flexibility of the nsP2 proteases were evaluated through Root Mean Square Deviation (RMSD) and Root Mean Square Fluctuation (RMSF) analyses. The resulting profiles, which reveal distinct behaviors between the apo and ligand-bound forms for both CHIKV and MAYV, are comprehensively presented in Figure 1.

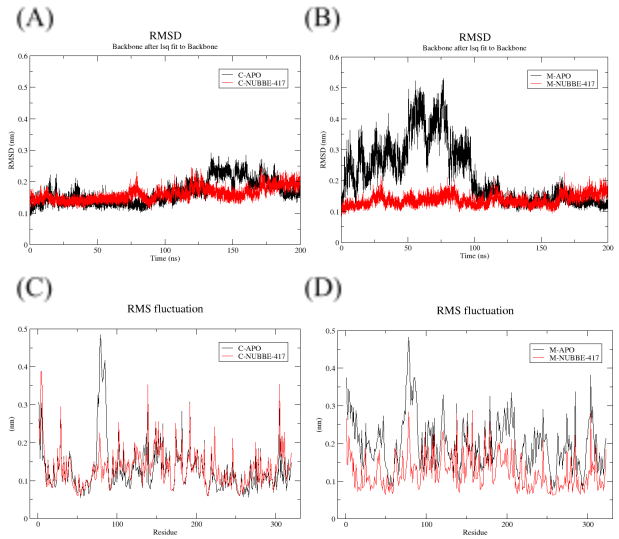


Figure 1. Structural stability and local fluctuations of CHIKV and MAYV nsP2 proteases. (A) RMSD profile for CHIKV nsP2 systems; (B) RMSD profile for MAYV nsP2 systems; (C) RMSF analysis of CHIKV nsP2 residues; (D) RMSF analysis of MAYV nsP2 residues. Black lines represent the apo systems, while red lines indicate the nubbe417-bound complexes.

The RMSD profiles obtained after the MD analysis revealed distinct behaviors between apo and complex forms. For CHIKV nsP2, both systems remained globally stable and showed similar fluctuations; 0.165 nm and 0.160 nm to apo and complex systems, respectively. In contrast, MAYV nsP2 displayed a pronounced difference: the apo form showed substantial initial instability, stabilizing only after approximately 100 ns with a mean of 0.219 nm, whereas the complex remained stable from the onset with a mean of 0.136 nm. Nubbe417 stabilized conformations, especially in the more flexible MAYV nsP2.

RMSF analysis further highlights the stabilizing effect of the ligand. In both proteases, the most mobile segment corresponds to the catalytic loop (residues 76–86 in the numbering used here). For CHIKV nsP2, the apo form displayed fluctuations of 0.3887 ± 0.0596 nm, while the complex reduced this mobility to 0.1456 ± 0.0349 nm. Similarly, in MAYV nsP2, the apo form fluctuated at 0.3798 ± 0.0554 nm, whereas the complex decreased mobility to 0.1649 ± 0.0547 nm. Reduced mobility shows the ligand pins the loop, promoting catalytic-site closure.

H. Molecular Dynamics: Dynamic Stability and Conformational Changes

The dynamic stability and conformational changes within the catalytic loop were investigated by comparing the first and last Molecular Dynamics (MD) frames (Figure 2). In this visualization, initial positions are represented by light colors—yellow for CHIKV, light blue for MAYV, and light gray for nubbe417—whereas the final states are depicted in dark colors—orange for CHIKV, dark blue for MAYV, and dark gray for the ligand. A distinct 'loop opening' effect was observed in both apo systems, leading to significant structural expansion of the catalytic pocket. Conversely, the presence of the inhibitor prevents this movement, inducing a prominent closure of the loop in CHIKV and promoting overall structural stabilization in MAYV. This 'loop-locking' mechanism is reinforced by persistent interactions with residues such as Asn79 and Leu202, which effectively maintain the protease in an inhibited, closed conformation .

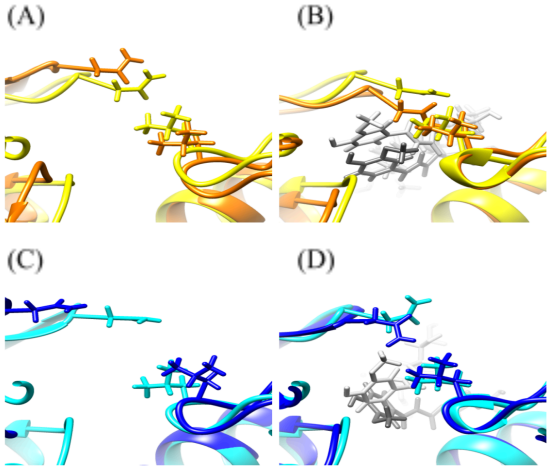


Figure 2. Conformational transitions of the nsP2 catalytic loop during MD simulations. Comparison between the initial (light colors) and final frames (dark colors) for the apo and complexed systems of both viruses. Protein is shown in ribbon representation, with catalytic residues and nubbe417 displayed in sticks.

Loop dynamics were quantified by comparing the distances between the α-carbons (Cα) of residues Asn79 and Leu202 across the initial and final conformations of the MD simulations1. As detailed in Table I, these distance measurements provide a numerical basis for assessing the ligand-induced closure of the catalytic pocket in both CHIKV and MAYV systems.

TABLE I. DISTANCE VARIATION BETWEEN Cα ATOMS OF ASN79 AND LEU202 IN APO AND LIGAND-BOUND NSP2 SYSTEMS.

System	Distance measurements (Cα - Asn79–Leu202)		
	First Frame (Å)	*Last Frame (Å)*	*Δ (Å)*
Apo CHIKV	7.96	10.83	+2.87
CHIKV-nubbe417	8.18	6.58	-1.60
Apo MAYV	9.83	14.19	+4.36
MAYV-nubbe417	9.39	8.74	-0.65

In the apo systems, the distance between Asn79 and Leu202 increased by 2.87 Å in CHIKV and expanded by a substantial 4.36 Å in MAYV relative to the initial structures. In sharp contrast, the inhibitor-bound complexes exhibited a narrowing of this distance. The CHIKV-nubbe417 complex tightened by 1.60 Å, while the MAYV-nubbe417 complex stabilized with a reduction of 0.65 Å. In these bound systems, the Asn79 side chain is specifically oriented toward the ligand and Leu202.

I. Molecular Dynamics: Hydrogen Bond Interactions

The stability of the protein-ligand complexes was further investigated by quantifying the number of hydrogen bonds formed throughout the 200-ns molecular dynamics simulations. As presented in Figure 3, the persistent nature of these interactions for both CHIKV and MAYV systems highlights the robust anchoring of nubbe417 within the catalytic pocket.

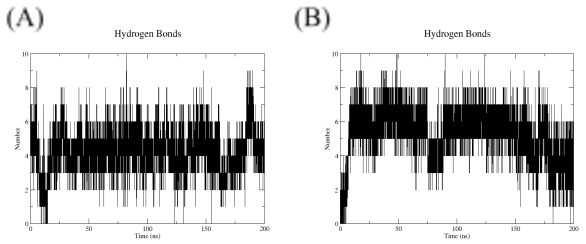


Figure 3. Time-dependent evolution of protein-ligand hydrogen bond interactions. The plots show the number of hydrogen bonds maintained between nubbe417 and the catalytic pocket for (A) CHIKV and (B) MAYV nsP2 systems throughout the 200-ns MD simulations.

The complexes maintained a consistent number of hydrogen bonds throughout the simulations, reinforcing the observed stability. In CHIKV–nubbe417, the average number of hydrogen bonds was 4.17 ± 1.24, while in MAYV–nubbe417 it was 5.21 ± 1.48. These interactions keep the ligand firmly anchored within the catalytic pocket, reduce local flexibility, and propagate stabilization to adjacent regions.

IV. DISCUSSION

The combined analyses of RMSD, RMSF, hydrogen-bond stability, and structural comparisons provide a coherent mechanistic interpretation of ligand-induced stabilization. The reduction in RMSF within the catalytic loop correlates strongly with the conformational transitions observed from initial to final frames: apo forms display a flexible, open loop, whereas ligand-bound complexes modulates the conformational state of the loop, stabilizing a closed architecture essential for protease inhibition. Persistent hydrogen bonds anchor nubbe417, dampening loop motion and enforcing inhibitory compactness. Taken together, these results are consistent with the regulatory 'gating' role described by Narwal et al. [8], where the loop modulates active site access. Nubbe417 exploits this by structurally locking the loop in a closed conformation, effectively barricading the catalytic dyad. Therefore, the stabilization induced by nubbe417 may potentially affect the protease functionality by reducing local flexibility, enforcing a closed active-site geometry, and maintaining interactions that reinforce this inhibited structural state.

The performance of nubbe417 in both proteases—structural stability, reduction of catalytic-loop mobility, and formation of persistent interactions—demonstrates that natural-product-derived molecules represent valuable sources of bioactive scaffolds with antiviral potential. Their ability to stabilize key functional regions suggests a mechanism of action capable of interfering with the proteolytic processing essential for viral replication. From a biomedical perspective, identifying compounds capable of modulating the catalytic loop is particularly significant, given its dual role in substrate access and positioning. Ligands that stabilize closed conformations may act as competitive or allosteric inhibitors. Additionally, the effective performance of nubbe417 against two distinct alphavirus proteases highlights its potential as a broad-spectrum antiviral lead.

These computational findings gain further relevance when contextualized with phytochemical and biological studies of *Chioccoca alba* (Rubiaceae). A previous work reported the isolation and structural characterization of iridoid and seco-iridoid glucosides from *C. alba*, including the compound identified in our study as nubbe417 [9]. The chemical framework of these metabolites confirmed the presence of functional groups capable of hydrogen bonding and stabilizing protein–ligand interactions, consistent with our simulation results. More recently, another study evaluated methanolic root extracts of *C. alba* against Chikungunya and Mayaro viruses, demonstrating inhibition levels above 70% at 60 µg/mL [10]. Notably, nubbe417 was originally isolated from *C. alba* roots , the exact plant organ exhibiting antiviral activity. Moreover, the methanolic extraction used in these biological assays is chemically suitable for solubilizing polar glycosides, supporting the physical presence of this iridoid in the active extract.

While Pires et al. [10] attributed the extract's potency primarily to flavonoids via docking, the use of crude extracts allows for the presence of other bioactive compounds. Thereby, our results suggest that iridoids like nubbe417—previously isolated from *C. alba* [9]—could be significant contributors to this antiviral activity, potentially acting synergistically or as potent specific inhibitors of nsP2. This hypothesis is further supported by the established bioactivity of the iridoid class, which has been previously documented to inhibit other cysteine proteases [11]. However, this hypothesis requires direct confirmation through *in vitro* testing of nubbe417 to establish its specific role as nsP2 protease inhibitor.

As with any structure-based virtual screening approach, potential limitations include docking score bias, force-field dependency, and the risk of false-positive predictions. Molecular dynamics simulations mitigate some of these limitations by incorporating protein flexibility and solvent effects; however, experimental validation remains essential to confirm binding affinity and inhibitory activity.

In summary, when compared with existing computational studies that primarily emphasize docking scores or covalent inhibition, the present work highlights a complementary mechanism based on dynamic stabilization of the catalytic loop. The dual-virus consistency observed for nubbe417, combined with its natural origin and previously reported antiviral activity at the extract level, distinguishes this study by proposing a mechanistically informed, cross-alphavirus inhibition strategy rather than a virus-specific solution.

V. Conclusion and Future Work

The present study demonstrates that an integrated computational workflow combining pharmacophore modeling, docking, and molecular dynamics simulations can identify natural product-derived compounds capable of modulating the structural dynamics of alphavirus nsP2 proteases. The ligand nubbe417 consistently stabilized the catalytic loop of both Chikungunya and Mayaro virus proteases, supporting a loop-locking inhibition mechanism with potential broad-spectrum relevance.

As future work, expanded virtual screening campaigns involving additional natural and synthetic compound libraries and *in vitro* protease inhibition assays, antiviral cell-based experiments are planned. These efforts will be essential to experimentally validate the proposed mechanism and advance nubbe417 or related scaffolds toward lead optimization.

Acknowledgment

The authors thank the Federal University of Uberlândia (UFU) and the Institute of Biotechnology (IBTEC/UFU) for institutional support. We also acknowledge the Santos Dumont supercomputing facility for providing the computational resources essential for this study. Financial support from FAPEMIG, CNPq, and CAPES is gratefully acknowledged.

References

[1] R. Abdelnabi, S. Jacobs, L. Delang and J. Neyts, ”Antiviral drug discovery against arthritogenic alphaviruses: tools and molecular targets”, Biochemical Pharmacology, vol. 174, p. 113777, 2020.

[2] D. S. Metibemu, et al., ”Inhibitors of the structural and nonstructural proteins of alphaviruses”, ACS Infectious Diseases, vol. 10, no. 8, pp. 2715–2734, 2024.

[3] D. F. P. Rivera, S. M. Anthony and B. Pahar, ”Alphaviruses: recent insights into pathogenesis, immunity, and approaches for treatment and prevention. In: Emerging and Re-emerging Infections”, pp. 479–526, 2024.

[4] X. Hu, et al., ”*In silico* screening of inhibitors of the Venezuelan equine encephalitis virus nonstructural protein 2 cysteine protease”, Viruses, vol. 15, no. 7, p. 1503, 2023.

[5] P. Kumar, D. Kumar and R. Giri, ”Targeting the nsp2 cysteine protease of Chikungunya virus using FDA approved library and selected cysteine protease inhibitors”, Pathogens, vol. 8, no. 3, p. 128, 2019.

[6] A. Zaid, et al., ”Arthritogenic alphaviruses: epidemiological and clinical perspective on emerging arboviruses” The Lancet Infectious Diseases, vol. 21, no. 5, pp. e123–e133, 2021.

[7] S. Wang, S. Mahalingam and A. Merits, ”Alphavirus nsP2: a multifunctional regulator of viral replication and promising target for anti-alphavirus therapies”, Reviews in Medical Virology, vol. 35, no. 2, pp. e70030, 2025.

[8] R. Staub, ”ROC Calculator for MOE Docking. GitHub repository”, Available online December, 2025: https://github.com/RubenStaub/ROC-calculator-for-MOE-docki.

[9] M. Narwal, et al., ”Crystal structure of Chikungunya virus nsP2 cysteine protease reveals a putative flexible loop blocking its active site”, International Journal of Biological Macromolecules, vol. 116, pp. 451–462, 2018.

[10] C. A. Carbonezi, et al., ”Iridoid and seco-iridoid glucosides from *Chioccoca alba (Rubiaceae)*”, Phytochemistry, vol. 51, no. 6, pp. 781–785, 1999.

[11] E. C. F. Pires, et al., ”Antiviral potential of *Chiococca alba (L.) Hitchc.* plant extracts against Chikungunya and Mayaro viruses”, International Journal of Molecular Sciences, vol. 25, no. 21, p. 11397, 2024.

[12] D. Y. Huang, et al., ”Anti-coronavirus and anti-pulmonary inflammation effects of iridoids, the common component from Chinese herbal medicines for the treatment of COVID-19”, vol. 78, pp. 1003-1012, 2024.