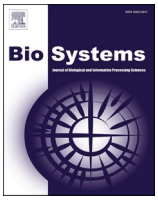
BioSystems 257 (2025) 105604

Contents lists available at ScienceDirect

BioSystems

journal homepage: www.elsevier.com/locate/biosystems



Beyond the genome: A multi-scale, agent-based taxonomy of biological codes and energetic constraints

Cris Micheli [a,b,*], Robert Prinz [c], Pier Luigi Gentili [d]

[a] NeuroTechX Chapter, Berlin, 12161, Germany
[b] Department of Management, Faculty of Management Sciences, Veritas University Abuja, Bwari Area Council, Federal Capital Territory, Nigeria
[c] Rechenkraft.net e.V., Marburg, 35037, Germany
[d] Department of Chemistry, Biology, and Biotechnology, University of Perugia, Perugia, 06123, Italy

A B S T R A C T

This work critically examines the organizational principles governing living systems and introduces emerging rules that pave the way for a computational approach to understanding life. It challenges the conventional assumption of the modern synthesis, which claims that the code of life resides solely in DNA, genetic networks, and epigenetics. Instead, we argue that the information essential for sustaining life is distributed across multiple scales. Drawing from diverse frameworks such as cybernetics and machine learning, we propose a fresh perspective on this fundamental question.

We begin by exploring the complexity of life and propose a thoughtfully constructed preliminary taxonomy of biological codes, while recognizing the potential for alternative frameworks. This interpretation integrates speculative ideas from concepts like constraint closure and agent-based modeling, framing the hierarchy of life as governed by a dynamic tension between stability and exploration. Building on this foundation, we analyze the compositional rules and properties of biological codes, uncovering their hierarchical and causal relationships across scales. We emphasize the cell's role as a fundamental cybernetic agent and discuss how this framework contributes to understanding natural phenomena such as cellular differentiation and collaboration.

The theoretical implications of this perspective highlight the importance of emergence and top-down interactions in fostering complexity. We argue that information distributed across multiple scales is necessary but not sufficient for sustaining life because living systems are open and dynamic, relying fundamentally on environmental interactions, subjected to entropy, mass, and energy exchanges. Additionally, any form of life functions as a physical information-processing system, further emphasizing the intricate interplay between structure and environment. We propose that available energy not only sustains autopoiesis in biological systems, but a fraction of it also drives their adaptation and evolution in an exploratory fashion.

Finally, we present a practical example and outline future directions for this approach. Specifically, we illustrate how our framework advances the understanding of protein folding agents, particularly in deciphering their regulatory dynamics and interactions with chaperones and organelles. By bridging theoretical concepts with practical examples, this work seeks to provide a framework for analyzing and manipulating complex biological systems, with potential implications for fields such as systems biology and Artificial Intelligence.

1. Introduction

For much of modern biology, the complexity of life has been attributed primarily to genetic determinism, with *DNA* regarded as the central script from which all biological functions emerge. While this perspective has yielded powerful insights, it increasingly appears insufficient to account for the full organizational and adaptive capacity of living systems. Recent research suggests that biological information is not confined to the genome but is distributed across multiple scales—emerging from energetic flows, cellular interactions, and system-level dynamics (Shapiro, 2021; Robin et al., 2021). This broader view invites a fundamental rethinking of biology's organizing principles and foundational dogmas, suggesting that life's coherence, complexity, and adaptability arise from an integrated energetic and informational framework spanning structure, function, and environment.

In response to this challenge, we propose a preliminary taxonomy of biological codes—not merely as a classificatory scheme, but as a heuristic tool for discovery. This taxonomy aims to illuminate the hidden architectures and regulatory motifs that drive phenomena from molecular and cellular patterns to collective behavior. Our goal is to develop a principled lens through which the mechanisms of coordination, adaptation, and decision-making in biological systems can be more clearly understood.

Central to this effort is the concept of the *biological agent* (see detailed

This article is part of a special issue entitled: Code Biology V published in BioSystems.

* Corresponding author. NeuroTechX Chapter, Berlin, 12161, Germany.
E-mail address: cris.micheli@neurotechx.com (C. Micheli).

https://doi.org/10.1016/j.biosystems.2025.105604
Received 7 July 2025; Received in revised form 24 September 2025; Accepted 25 September 2025
Available online 4 October 2025

definition in the Appendix), a hierarchical autonomous unit (Ashby, 1956) capable of sensing, processing, and responding under energetic constraints. Such agents, from molecules to organisms, face a fundamental trade-off between exploration and exploitation.

We propose that *biological codes* (e.g., genetic, epigenetic, bioelectric) coordinate agent behaviors by regulating energy allocation across two mutually exclusive modes: exploration—activities such as foraging, signaling, or morphogenesis that generate novelty—and exploitation—activities such as metabolic maintenance, repair, or stasis that sustain stability. These modes represent distinct (and mutually exclusive in time) strategic orientations that shape agent behavior over time.

This framing raises a fundamental question: What qualifies as a biological agent? According to the TAME framework, agency is scale-independent and emerges wherever goal-directed action occurs (Levin, 2022). This supports our hypothesis that agents are irreducible elements (Prigogine and Nicolis, 1971; de Castro and McShea, 2022) governed by codes that encode their goals and logic.

On this basis, we undertake the task of mapping each biological code—drawing on the catalog developed by Prinz (2024)—to its corresponding level of biological agency and to the energetic behavior it mediates (explorative, exploitative, or both). This mapping allows us to construct a functional taxonomy of biological codes based on agent behavior and energy dynamics. By incorporating cybernetic principles such as feedback, homeostasis, and energy efficiency, we aim to uncover generalizable patterns that govern life across scales. The following sections present our classification methodology, outline our key findings, and illustrate the utility of the energetic framework through a practical case study, demonstrating how the proposed taxonomy elucidates the dynamics of the well-known biological phenomenon of protein folding.

2. Towards a taxonomy of biological codes

2.1. Background

The concept that life requires not only energy and information but also meaning was first articulated in Marcello Barbieri's 'The Semantic Theory of Evolution' (1985), which introduced the idea that the existence of codes is what reveals meaning in nature. This perspective, later termed code biology, shifted the focus from a single genetic code to the recognition that many distinct organic codes underpin major evolutionary transitions and the organization of living systems. The systematic study of biological codes began with the discovery of the genetic code, which demonstrated that life is governed by rule-based conventions connecting distinct molecular domains, not just by chemical copying (Barbieri, 1985, 1998).

Building on Barbieri's pioneering work, Robert Prinz compiled and expanded a comprehensive database of biological codes, now maintained by the International Society of Code Biology. This resource has grown substantially: from 22 codes in 2012, when the Society was founded, to 237 codes by 2022, and 261 in the current version used in the manuscript (as of April 2024, see Suppl. Mat.), all of which are supported by peer-reviewed evidence. This rapid increase highlights that the diversity of biological codes in living systems is far greater than previously recognized, and that the field has not yet reached its empirical limits. The database serves as a foundational resource for empirical and quantitative research, classification, and theoretical development in code biology (see also, in this Issue, Prinz et al., 2025).

The methodology for collecting and verifying biological codes employs a rigorous, multistep approach. Systematic searches are conducted across scientific databases (e.g., PubMed, Google Scholar) and personal archives, encompassing peer-reviewed articles, preprints, conference abstracts, and online resources, including institutional websites. Some codes were initially reported by other researchers, prompting the first systematic collection and the creation of the database. Candidate codes are validated by selecting publications that explicitly use the term 'code' in the title, ensuring intentional engagement with the concept as defined by the authors. Whether these definitions align with Barbieri's or other criteria is evaluated case by case (Prinz, 2024; Barbieri, 2025). Only codes demonstrating clear, reproducible evidence of rule-based mapping should be retained. Community-driven standards, developed through iterative working group consensus, ensure consistent integration and version-controlled updates, with ongoing prioritization of entries requiring further empirical validation via experimental or computational methods.

The recognition of hundreds of biological codes, each grounded in experimental evidence, presents both theoretical and practical challenges. Theoretically, it suggests that biological codes are not limited to major macro-evolutionary events but also account for many other characteristics of living systems. Practically, it raises ongoing issues regarding the precise definition, classification, and differentiation of biological codes. This expanding landscape underscores the need for continued systematic research, robust classification frameworks, and deeper exploration of the role of codes in the evolution and organization of life.

Guided by those general criteria and Barbieri's specific criteria, the semantic search strategies used in this publication instantiate prompts based on terms like "biological code" and code-defining features (see definitions in Appendix).

2.2. Methodology

We classified biological codes through an agent-based modeling (ABM, Pleyer and Fleck, 2023; Zhang and DeAngelis, 2020; Stephan et al., 2024) perspective, which assumes that life is organized as interacting agents—ranging from molecules to populations—operating under energetic constraints. Agents adopt two main strategies: exploitation (optimizing stability and resource use) and exploration (generating novelty and adaptive flexibility). Each code was positioned along two axes: biological level (molecular to ecosystem) and behavioral type (exploitative, explorative, or mixed).

2.2.1. Classification approach

Our analysis (see a sample classification in Table 1) used the April 2024 version of the Code Biology Database. Each entry was assigned two labels—biological level and agentic behavior—based on references in the dataset. To assign labels we made use of a Large Language Model (LLM, Perplexity Pro, with Research, Web and Academic options

Table 1
Sample classification results from the Code Biology Database (April 2024 version). We show a few representative codes together with the assigned biological level, agent behavior category, and biological kingdom, following the methodology described in Section 2.2. For brevity, the justifications are only included in the Supplementary Materials.

Code Number	Code Name	Biological Level	Agent Behavior Type	Kingdom Involved
1	Acoustic	Population/ Ecosystem	Mixed	Animalia
7	Cancer	Cell	Explorative	Eukaryota
44	Cooperation	Population/ Ecosystem	Mixed	Animalia
60	Ecological	Population/ Ecosystem	Mixed	All
70	Genetic[a]	Cell	Exploitative	All
131	Mitochondrial genetic	Organelle	Exploitative	Eukaryota
140	Mutualism	Population/ Ecosystem	Exploitative	Plantae/ Animalia
165	Organogenesis	Organ (multiple)	Exploitative	Animalia
187	Protein folding	Molecular	Exploitative	Eukaryota
245	Toxin resistance	Organism	Explorative	Animalia

[a] The genetic code referred to here is the historical code, other 5 genetic/genomic codes are listed in the database.

enabled), an AI tool that provided the systematic labeling of biological codes and the post-processing of classification consistency. The procedure included (see Suppl. Mat.):

1. Code Selection: Extract codes from the database (input to the LLM).
2. Reference Review: Examine supporting evidence for each entry (LLM prompt 1).
3. Label Assignment: Classify codes by biological level and behavior, with literature-based justification (LLM prompt 1).
4. Documentation: Record results in a structured table (LLM output).
5. Post-Processing: Validate assignments to ensure that codes reflect the *smallest autonomous agent* responsible for goal-directed deployment, not merely the physical site of operation (LLM prompt 2, and new output).

2.2.2. Prompt development

Two large-language-model (LLM) prompts (see Suppl. Mat.) supported classification. The first assigned labels to codes, guided by predefined definitions of biological agent (autonomous, adaptive system) and biological code (rule-based mapping between distinct domains). The second refined results, verifying that level labels corresponded to teleological agency rather than mechanistic context. Agents' behavioral categories were defined as follows (see Table 1, column 4):

- Exploitative: stability and optimization of existing resources.
- Explorative: novelty generation and adaptive search.
- Mixed: dual or context-dependent roles.

Biological levels (Table 1, column 3) were categorized as molecular, organelle, cell, tissue, organ, organism, and population/ecosystem. Labels reflected the *operational agent employing the code*, not simply the physical scale at which it appeared (McShea, 2016).

2.3. Quantitative overview of the new taxonomy

The analysis of the biological code dataset reveals clear patterns in their distribution across biological levels, agent behaviors, and taxonomic kingdoms, offering insights into the state of current research (Fig. 1).

Our analysis shows that most biological codes operate at the cellular level. Of the 261 codes, 130 (49.8 %) are cellular, making this the dominant level. Organism-level codes follow with 73 (28.0 %), then organ-level (20, 7.7 %) and molecular-level (19, 7.3 %). Tissue and population/ecosystem levels each account for 7 codes (2.7 %), 2 codes (1 %) are undetermined, and just 1 code (0.4 %) is organelle-level.

Agent behavior distributions show exploitative strategies dominate with 145 codes (55.6 %), followed by mixed (107, 41.0 %) and exploratory (8, 3.1 %). At the cellular level, 74 codes are exploitative, 52 mixed, and 4 explorative. Similarly, at the organism level, 36 are exploitative, 35 mixed, and 2 explorative.

A striking pattern emerges when examining exploratory behaviors across biological levels. Exploratory codes are rare overall (8), absent in organs, tissues, and population/ecosystem levels, where only mixed codes have been found (e.g., CKD code, growth code, ecological code). This absence points to a lack of conceptual and empirical frameworks for purely exploratory codes. In section 3 we argue, following Barbieri (2012) and Prinz (2023a), that exploratory codes may evolve from exploitative or mixed ones. Paradoxically, the more complex the organization, the greater one might expect an exploratory demand—yet such codes remain undetected.

The taxonomic distribution reveals bias toward higher eukaryotes. Together, Eukaryota (109 codes, 39.5 %) and Animalia (125 codes, 45.3 %) account for 84.8 % of the dataset, underscoring a research focus on animals and eukaryotes. In contrast, Plantae contributes only 8 codes (2.9 %), Bacteria 13 (4.7 %), and "All kingdoms" 10 (3.6 %). Viral codes are nearly absent (1), while none are specifically linked to Prokariota. A few (15) apply across multiple kingdoms, such as the "Cell Wall Code" (Plantae, Fungi, Bacteria) and "Mutualism code" (Plantae, Animalia), representing conserved or convergent processes.

Exploitative and exploratory codes occur in all recorded kingdoms (Animalia, Eukaryota, Bacteria, Fungi, Plantae, "All kingdoms"). Exploitative codes appear in every kingdom, while exploratory ones, though rare, are still present across several. Cells emerge as central agents orchestrating processes such as differentiation, migration, and tissue formation, reflected in the dominance of cellular codes. Yet, codes at the organelle (0.4 %) and population/ecosystem (2.7 %) levels are largely absent, likely due to methodological challenges rather than biological irrelevance.

The overwhelming representation of Eukaryota and Animalia demonstrates a strong taxonomic bias. Current research emphasizes exploitation and exploration as multiscale strategies but remains heavily skewed toward animal and eukaryotic systems. Less-studied groups like Fungi, Bacteria, and viruses together form only a small share of the dataset, while Prokaryota beyond Bacteria, such as Archaea, are entirely

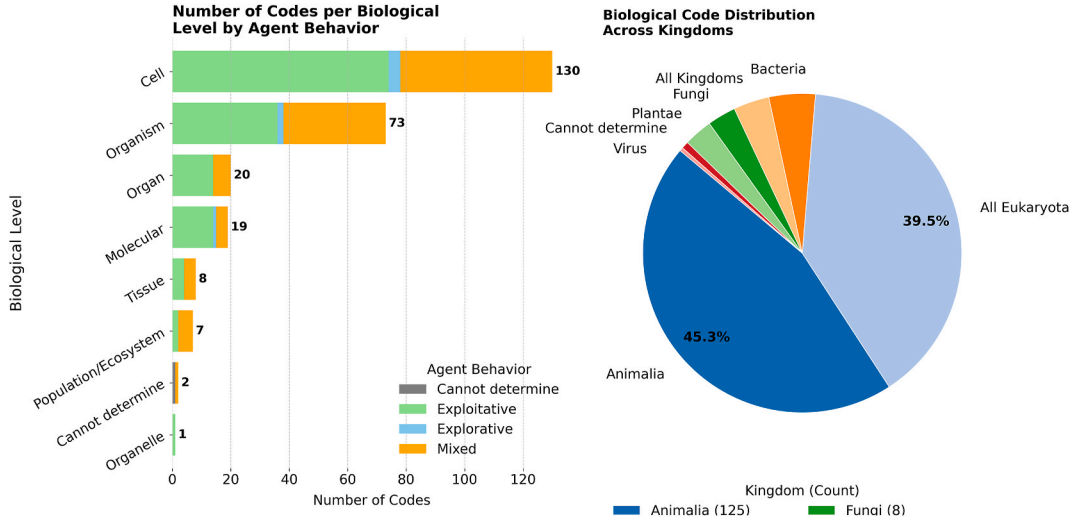


Fig. 1. Distribution of biological codes across biological levels, agent behaviors, and taxonomic kingdoms. Left: stacked histogram of codes across organizational levels with exploitative, mixed, and exploratory behaviors indicated per level. Right: chart showing taxonomic distribution, highlighting the predominance of Animalia and Eukaryota and underrepresentation of other kingdoms.

missing, representing clear opportunities for future research. An overview of references contextualizes how extensively codes are studied (Fig. 2).

Reference distribution highlights uneven research depth. A large fraction (114 codes, 43.7 %) are cited once, suggesting either highly specific mappings or under-citation in some areas. Codes cited 2–5 times account for 31.0 % (81 codes), 6–10 times 13.0 % (34), and 11–20 times 8.4 % (22). Only 10 codes (3.8 %) are cited more than 20 times, indicating foundational theories, cross-cutting concepts, or contested themes.

In summary, the citation landscape shows great breadth but uneven depth of research across categories, emphasizing the need for wider literature integration and deeper exploration of under-referenced codes.

2.4. *Emergent patterns in the new taxonomic framework of biological codes*

The data suggests a scale-dependent relationship between biological level and behavior complexity, with cellular and organismal levels showing a more diverse range of behaviors, including exploitative, explorative and mixed types. This methodology establishes a heuristic taxonomy by mapping each code within a rigorously defined two-dimensional space: agent behavior type (exploitative, explorative, or mixed) and the primary biological level of the agent utilizing the code (from molecular to ecosystem). This structured matrix not only organizes the diversity of biological codes, but crucially, provides a framework for identifying and exploring subcodes that may reflect distinct energetic and thermodynamical properties as discussed in section 4.

As will be shown below, it is the underlying energy and entropy landscapes that shape agents' behavioral states, driving temporal transitions between exploitative and explorative strategies (see also Fig. 3). Agents' behaviors are dynamic expressions of goal-directed action encoded in biological codes; the internal logic of these codes determines the pursuit and switching of behaviors that best support their objectives. Importantly, agents are not confined to a single behavioral state over time—as long as the goal encoded by the biological code is achieved, transitions between states are permissible and often advantageous (see Fig. 4).

The taxonomy matrix supports subclassification of codes according to how they govern energy expenditure, distinguishing between highly efficient exploitative processes and energetically costly but information-rich explorative processes (see section 3.3). By situating codes in this multidimensional parameter space, the taxonomy becomes a discovery tool for detecting patterns, clusters, and outliers, and for hypothesizing new subclasses that capture the thermodynamic nuances of biological information processing.

Ensemble agent behaviors, compared to single agent behaviors, are often more efficient because they enable fuzzy logic, collective problem solving, and adaptability in the face of fluctuating environments. Mixed-behavior codes, specifically, are those that enable agents to switch between at least two behavioral approaches in achieving goals. This behavioral flexibility is directly linked to the energetic landscape and temporal context of the system.

Consequently, future taxonomic subclassifications should account for the prevalence and distribution of distinct behavioral states, considering scale-free dynamics that reflect multi-agent interactions across scales. As further explored below, integrating energetic principles with taxonomy not only enriches our understanding of behavioral diversity and switching in living systems, but also highlights adaptive organization as a product of code-driven multi-state agency across biological hierarchies.

The next section presents a practical example of this approach: an analysis of agent behavior at both the molecular and cellular levels for code 187 (the protein folding code). This case illustrates how a single molecular code can mediate energetic, informational, and functional goals across organizational scales.

3. A paradigm of molecular swarm intelligence – an example from protein folding

To illustrate the explanatory power of our taxonomy, we present a case study of the *Protein Folding Code* (cd187_dbApr2024, Joo et al., 2015; Wallace, 2011, full list in Suppl. Mat.). This example shows how, when viewed through an energetic-agentic lens, a biological code can be systematically mapped to specific behavioral and organizational outcomes at the molecular scale. Based on our classification protocol, code entry 187 was determined to function at the *molecular level* and to support *exploitative agentic behavior*.

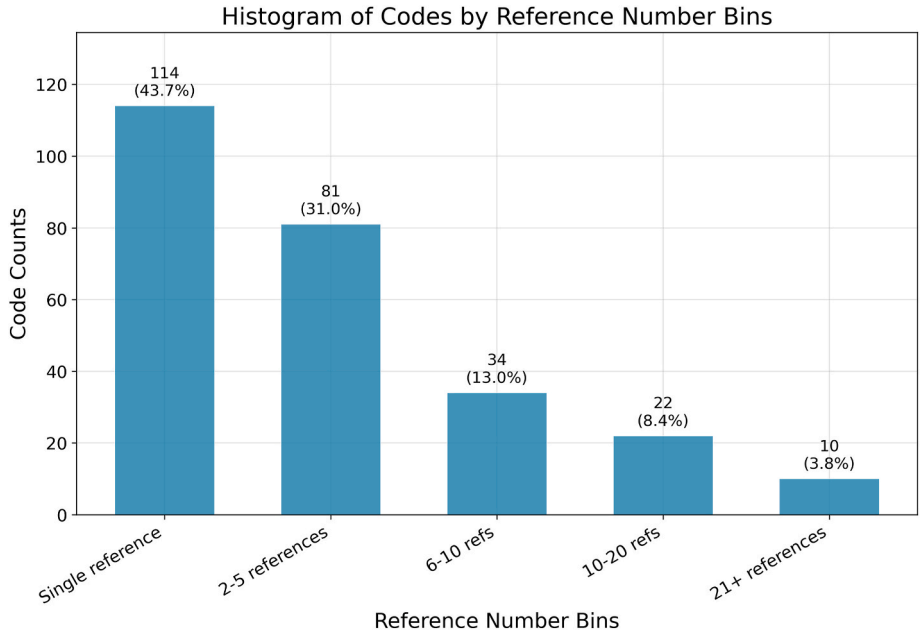


Fig. 2. Distribution of code counts by reference frequency. Most codes appear once, while a small proportion are cited more than 20 times. Percentages indicate each bin's share of the 261 codes.

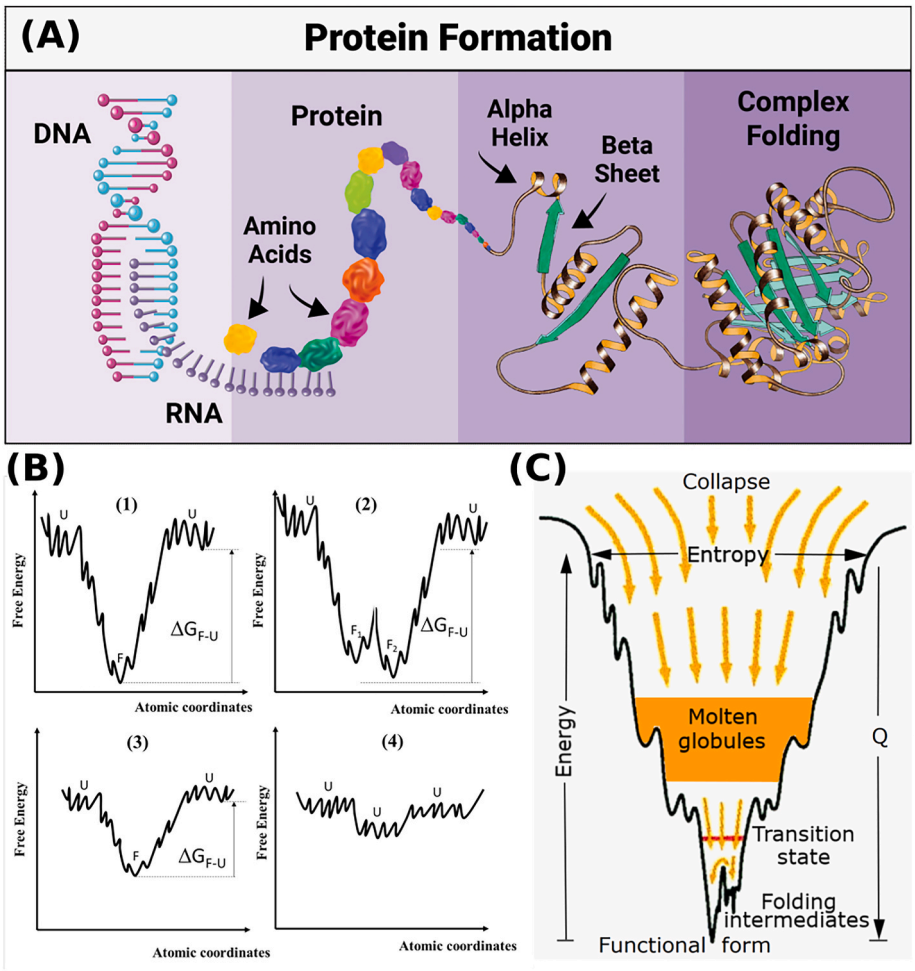


Fig. 3. Multi-regime dynamics in protein folding
A. Illustration of protein folding (adapted from NIMGS,<https://nigms.nih.gov/image-gallery/6603> licensed under CC BY-NC-SA 3.0<https://creativecommons.org/licenses/by-nc-sa/3.0/>). B. Free energy landscape of the Protein Folding Code, depicting four dynamic agent states (U: unfolded, F: folded): B.1: External Control State: Folding directed primarily by external constraints (e.g., chaperones), with minimal internal autonomy. B.2: Adaptive Transition State: metamorphic proteins with dual conformations and moderately constrained networks. B.3: Plastic Reorganization State: marginally stable proteins actively restructuring constraints, with high structural heterogeneity. B.4: Autonomous Optimization State: intrinsically disordered proteins or stable folded conformations with closed internal networks and efficient energy dissipation. Proteins may transition between these states depending on environmental and energetic conditions, illustrating how biological codes mediate multi-regime molecular agency. (Panels adapted from Biomimetics, 2024, 9, 121 under CC By 4.0). C. Thermodynamic funnel: proteins move from high macroscopic entropy (ΔSmacro) in unfolded states toward low-entropy native ensembles, navigating landscapes shaped by local interactions and environmental factors. The functional state represents a dynamic ensemble of conformations rather than a single static structure (adapted from Rumbley et al., 2001<https://www.pnas.org/doi/pdf/10.1073/pnas.98.1.105>, with permission).

The Protein Folding Code refers to a rule-based mapping between a protein's linear amino acid sequence and its three-dimensional conformation (Fig. 3). These mappings are governed by physicochemical constraints between amino acids (e.g. hydrophobic effects, hydrogen bonds, electrostatic interactions), encoded evolutionary solutions, and, in many cases, molecular chaperone support (Hill et al., 2011; Bard and Drummond, 2024). The code operates through a decentralized, bottom-up process akin to *swarm intelligence*—a distributed system where local interactions among amino acids and molecular forces produce globally ordered structures. This mechanistic view is strongly supported by classic and recent experimental studies on folding pathways, energy landscapes, and chaperone-assisted folding, which together demonstrate how physicochemical constraints and evolutionary solutions jointly govern the folding process (Jewett and Shea, 2010; Waudby et al., 2019).

This qualifies as a *biological code* in Barbieri's sense: a functional, non-deterministic correspondence between informational and structural domains mediated by physical but semi-arbitrary constraints. Although the process is strongly shaped by evolutionary pressures and energetics, the rules governing folding are not reducible to direct top-down (e.g. chemical) causation alone—they involve modular and reusable informational motifs (e.g., domains, secondary structure propensities) acting as coding elements.

3.1. *Agent identity and behavioral mode in the protein folding code*

The protein folding code offers a clear molecular paradigm for understanding biological agency as defined in our taxonomy. Here, the agent is the *entire protein molecule*, whose identity emerges from the collective, coordinated interactions among its amino acids. This swarm-like behavior enables the protein to self-organize and pursue a goal-directed trajectory: folding into its native, functional conformation. Such behavior fulfills the criteria for agency in our framework, as the protein acts as a transient, self-organizing system that minimizes free energy and achieves structural fitness through dynamic equilibrium between energetic and entropic constraints.

Within this context, the behavioral mode of the protein agent is classified as *exploitative*. The folding process is not aimed at generating

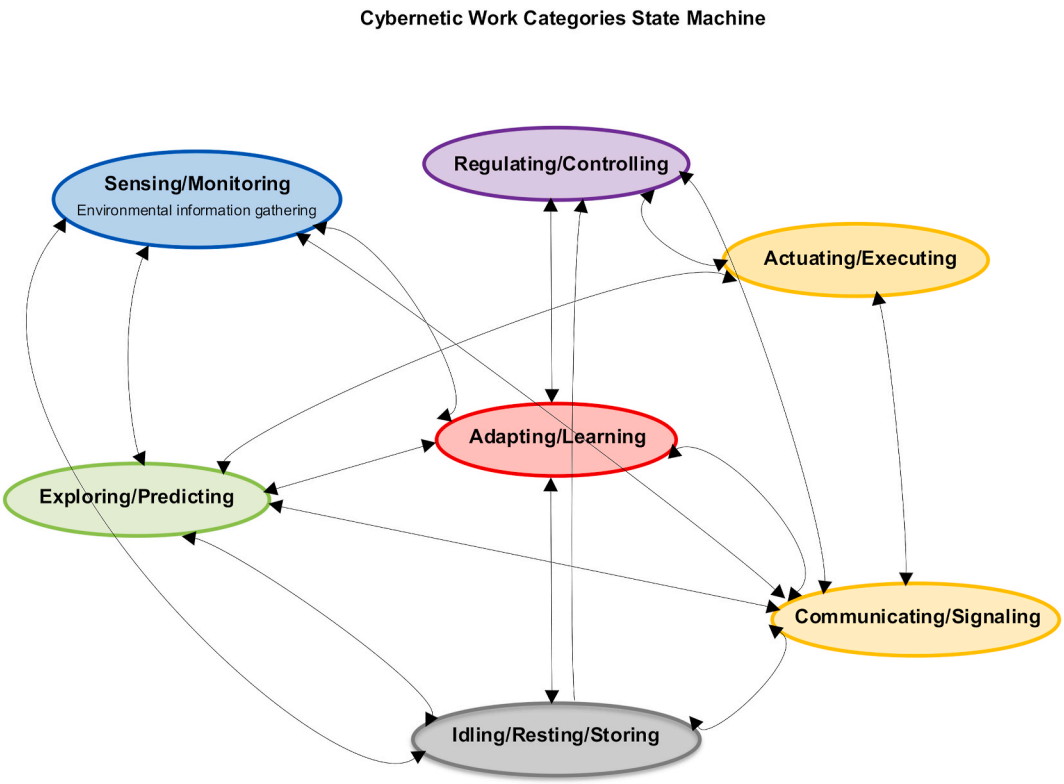


Fig. 4. Cybernetic work factorization in biological agents
This figure illustrates how biological agents partition available energy into distinct categories of cybernetic work, such as sensing, actuating, regulating, exploring, adapting, idling, and communicating. Each work category corresponds to a modular energy conversion process governed by the agent's underlying biological codes, which allocate energy between efficient, exploitative modes and more dissipative, explorative strategies. The framework highlights that living systems dynamically balance stability, adaptability, and complexity by orchestrating multiple codes and work categories, actively regulating energy dissipation and recovery to achieve their functional goals.
This is an arbitrary partition, and reflects the view of the authors that biological living entities could be treated as automata with classifiable dynamic and recurring states, to enhance predictability of the living system (aka agents states have a finite taxonomy of behaviors, which is skewed towards e.g. adapting or exploring for explorative behaviors and e.g. executing and controlling for exploitative ones).

novelty or expanding the system's functional repertoire; rather, it is directed toward realizing an optimal, pre-encoded structure essential for downstream cellular functions. Misfolding and subsequent degradation are not examples of exploration, but responses to failure in achieving the encoded goal. This reinforces our broader classification: biological codes can govern agent behavior at the *molecular level*, and exploitative strategies—focused on optimization and stability—are both detectable and classifiable even below the cellular threshold.

Importantly, this case demonstrates that molecular systems like protein folding, though lacking the top-down regulation seen in multicellular or neural systems, are nevertheless interpretable as coded, functional processes. The protein acts not as a passive machine, but as a rule-following, goal-directed actor within the biosphere. This perspective validates the scalability of our agent-based, energetically informed taxonomy and shows how meaningful structure and agency can be surfaced even in foundational layers of life.

3.2. *Multi-regime dynamics: agent states and transitions*

Protein folding stands as a quintessential example of how a biological code can govern agentic behavior across multiple regimes, revealing both the organizational logic and energetic foundations of molecular agency. The Protein Folding Code operates as a rule-based mapping from a protein's amino acid sequence to its three-dimensional structure, mediated by local interactions and distributed constraints that together enable the emergence of a stable, functional conformation (Fig. 3A).

The folding process is not monolithic; instead, it traverses a spectrum of agent states, each defined by distinct patterns of energetic and informational control. Recent experimental and computational work has shown that proteins, especially intrinsically disordered proteins (IDPs) and marginally stable domains, exhibit a continuum of conformational states—supporting the view that folding intermediates and end-states are best described as fuzzy sets with varying degrees of structural heterogeneity (Sharma et al., 2015; Fuxreiter, 2022; Sacquin-Mora and Prévost, 2021; Gianni et al., 2021).

Hence, each agent state (Fig. 3B, Gentili, 2024a) corresponds to a fuzzy set (Gentili, 2018) characterized by unique entropy profiles and structural heterogeneity:

- External Control State (Panel B.1): Folding dominated by strong external constraints, such as chaperone-mediated environments, where informational control is high, but internal autonomy is minimal. It exhibits low fuzzy entropy (Gentili and Perez-Mercader, 2022) forming a *crisp set* with minimal conformational diversity. For example, chaperones such as GroEL and Hsp70 have been shown to actively reshape the folding landscape, enabling proteins to reach their native states under otherwise non-permissive conditions by iterative annealing and ATP-dependent cycles (Todd et al., 1996; Deuerling and Bukau, 2004).
- Adaptive Transition State (Panel B.2): Systems with moderate constraint networks and partial closure (Montevil and Mossio, 2015), as seen in metamorphic proteins that can adopt dual conformations (intermediate fuzzy entropy, it enables partial structural plasticity with subdomains exhibiting semi-independent folding).

- Plastic Reorganization State (Panel B.3): Marginally stable proteins that actively restructure constraints and energy flows in response to environmental changes (Montevil and Mossio, 2015). This reflects hybrid or plastic behaviors, and it actively restructures constraints under stress (high fuzzy entropy, where heterogeneous conformational ensembles facilitate kinetic escape from energy minima).
- Autonomous Optimization State (Panel B.4): Internally closed, high-fidelity regimes—typified by intrinsically disordered proteins (IDPs, a significant portion of eukaryote proteins) i.e. interconverting native folds—where constraint networks efficiently channel energy dissipation into ordered, functional outcomes (maximal fuzzy entropy, Varadi et al., 2015; Aspromonte et al., 2024).

Depending on context, a single unfolded protein can dynamically transition between these macroscopic states though a multitude of microscopic states (Gentili, 2024), influenced by environmental pressures, energetic resources, and regulatory interactions with other agents such as chaperones or organelles. The evolutionary and functional importance of this conformational fuzziness has been highlighted in recent contributions (Gianni et al., 2021; Fuxreiter, 2022; Sacquin-Mora and Prévost, 2021; Fuxreiter, 2018), which show that alternative, sub-optimal contacts and dynamic disorder can encode functional specificity and robustness, expanding the repertoire of biological activities available to proteins. This multi-regime perspective thus highlights the protein as a molecular agent capable of switching between operational modes, each characterized by unique balances of entropy, energy dissipation, and informational closure.

3.3. Energetic and informational principles

Folding is fundamentally an energetic optimization process, minimizing Gibbs free energy while resolving the informational constraints encoded in the sequence. The thermodynamic funnel model (Fig. 3C–Rumbley et al., 2001; Wang et al., 2012) visually captures this process: the protein navigates an energy landscape from a high-entropy, unfolded ensemble toward a low-entropy, native conformation. The path is not a simple gradient descent but is shaped by the interplay of micro-level conformational diversity (ΔS_{micro}) and the macro-level assembly (ΔS_{macro}), balancing exploratory flexibility with the drive toward structural stability.

Quantitative studies have measured the information content required to specify a protein fold and the efficiency of energy-to-information conversion, finding values that closely match theoretical predictions and support the energetic-informational framing adopted here. This relationship can be formalized (Sánchez et al., 2022, Benítez and Jiménez, 2025; Sorokina et al., 2022; Dill, 1990) for a single domain (globular) protein as:

$$\Delta G_{fold} = \Delta H_{\text{fold}} - T\Delta S_{\text{micro}} + kT\ln(2)\cdot I_{fold} \quad \text{(eq. 1)}$$

where

ΔG_{fold}: Gibbs free energy change upon folding of a single domain protein (units in Joules)
ΔH_{fold}: Net energy change from bonds/contacts formation in Joules (e.g., hydrophobic burial, H-bonds).
ΔS_{micro}: Micro-level conformational entropy or cost of conformational diversity enabling exploration (Joules/Kelvin)
I_{fold}: Information gained through folding (in bits, inversely related to folding entropy decrease $\Delta S_{macro} = -kTln(2)\cdot I_{fold}$) in bits.
$kTln(2)$: Converts information (bits) into energy units (Joules), (k = Boltzmann constant, T = temperature in Kelvin).

This equation quantitatively links protein folding enthalpy change (ΔH_{fold}) to local conformational entropy (ΔS_{micro}) and folding information (I_{fold} *or* ΔS_{macro}) and is valid under the conditions: 1. the protein and its solvent environment are treated as a thermodynamic closed system capable of energy exchange (typically heat), but not matter; 2. the folding process is directed toward achieving a stable native conformation, consistent with goal-oriented behavior; 3. enthalpy, entropy, and temperature remain effectively constant or vary slowly during folding (quasi-equilibrium); and 4. physical constants and environmental parameters are stable throughout the process.

The equation is most accurate for spontaneous or chaperone-assisted folding under quasi-equilibrium or steady-state conditions, while for strictly non-equilibrium, energy-driven scenarios (such as continuous ATP-driven chaperone action), additional terms may be needed to capture ongoing energy input and dissipation. For example, for proteins with many conformers, instead of using binary information (bits in the I_{fold} *term)*, we might use fuzzy information (Gentili and Perez-Mercader, 2022). The Sánchez paper finds that the information content needed to specify a protein fold is about 2.2 ± 0.3 bits per folded amino acid, closely matching the information encoded in evolved protein sequences. They also report that the energy-to-information conversion efficiency during folding is around 50 %—lower than the theoretical maximum of 70 %, but much higher than in human-made machines (Sánchez et al., 2022).

The folding process can be viewed as a *state machine* where transitions between states are guided by an energy–information cost function (see Fig. 4). For multidomain proteins and intrinsically disordered proteins (IDPs), equation (1) can be generalized by adding terms that penalize domain interactions, entropic effects, and structural disorder. This allows the folding dynamics of complex proteins to be modeled as stochastic transitions between states, capturing how environmental factors and internal constraints influence folding pathways and stability.

In summary, the Sánchez equation (Sánchez et al., 2022) enables us to define the transition probabilities of a generalized folding state machine, which cycles through four agent states: External Control, where folding is dominated by chaperones or external constraints; Adaptive Transition, marked by partial constraint closure (Montevil and Mossio, 2015) and flexible domain behavior; Plastic Reorganization, characterized by high structural heterogeneity and environmental responsiveness; and Autonomous Optimization, where the protein achieves efficient, internally driven folding, as seen in many IDPs. The protein dynamically transitions between these states depending on its sequence, environment, and energetic context, allowing it to balance stability, adaptability, and functional order.

These agent states correspond directly to the thermodynamic constraint closure regimes illustrated by Gentili in Fig. 3B, linking the molecular folding code's energetic and informational logic to broader organizational principles. This alignment shows how the same underlying agent-based framework captures both the behavioral dynamics of protein folding and the general strategies by which biological systems manage energy, entropy, and information.

This generalization to all types of proteins is fully compatible with Jeremy England's theory of dissipative adaptation (England, 2015), which demonstrates that non-equilibrium systems —such as biological agents—naturally evolve toward configurations that maximize energy absorption and dissipation (in line with the principle of maximum entropy production, Volk and Pauluis, 2010), providing a statistical physics foundation for the emergence and regulation of life-like order in biological codes.

The integration of thermodynamic and informational constraints into agent behavior reflects a *systems biology* approach, which emphasizes the holistic modeling of biological networks and the dynamic interplay between system components. Having established how the protein folding code exemplifies agentic, energetic, and informational dynamics at the molecular level, we now broaden our perspective to generalize these principles across biological scales and code types. In the following section (3.4 Generalization and Implications), we synthesize the agent-based approach and energetic framework developed above to articulate a unified model for how biological codes operate in diverse

contexts, from molecules to ecosystems.

3.4. Generalization and Implications

The agent-based taxonomy developed here, grounded in energetic and informational dynamics, allows us to generalize beyond specific case studies like the protein folding code and to articulate a unified framework for understanding how biological codes operate across scales and contexts. This section synthesizes the manuscript's multi-scale agent-based approach with the broader system-theoretic and cybernetic concepts highlighted in classic theoretical contributions (Wiener, 1961; Ashby, 1956; England, 2013).

The literature supports viewing biological codes as the rule-based goals that guide agent behavior, shaping how energy is converted and dissipated within living systems. This perspective aligns with thermodynamic principles and agent-based models across scales. Below is a generalized equation capturing these dynamics, informed by the literature's energy dissipation frameworks (Yang et al., 2021; Kim, 2023; Schweitzer, 2019; Gusev and Martyushev, 2021; Otsuka,; Mikhailovsky, 2024).

$$\frac{dE_{total}}{dt} = \underbrace{\eta \cdot P_{env}(t)}_{\text{Energy input}} - \underbrace{[Wwork(t) + Q_{diss}(t)]}_{\text{Energy output}} + \underbrace{\alpha \cdot Q_{rec}(t)}_{\text{Waste heat recovery}} \quad \text{(eq. 2)}$$

where

$E_{total}(t)$: Total energy available to the agent (in Joules - J)
$P_{env}(t)$: Power input from environment (e.g. ATP hydrolysis, sunlight) in Watts – W
h: Energy conversion efficiency ($0 < h < 1$)
$W_{work}(t)$: Rate of useful work (e.g. mechanical, chemical, signaling) (W)
$Q_{diss}(t)$: Dissipation rate (irreversible entropy production) (W)
$Q_{rec}(t)$: Recoverable waste heat (e.g. during anaerobic digestion (Van Doren et al., 2017; Brownstein et al., 2022), or microbial growth (von Stockar and Liu, 1999)) (W)
α: Recovery efficiency ($0 < \alpha < 1$)

The equation is valid under the conditions that 1. the agent operates as an open thermodynamic system (not an isolated system) exchanging energy with its environment, 2. the agent exhibits autonomous goal-directed behavior, 3. energy terms (P_{env}, W_{work}, Q_{diss}, Q_{rec}) vary slowly relative to the agent's decision-making cycles, 4. work, dissipation, and recovery operate in parallel pathways with no interactions between terms, 5. energy conversion and recovery coefficients are time-independent or slowly varying, and 6. the open system is maintained by continuous energy input (non-equilibrium steady state). The equation does not apply, for example, to disordered systems with unpredictable energy partitioning or to isolated systems ($P_{env} = Q_{rec} = 0$).

Importantly, dissipation Q_{diss} in biological systems is not merely a passive loss but is actively regulated by the agent, which exerts ***top-down control*** to tune energy dissipation according to its goals and environmental demands. This regulatory capacity ensures that Q_{diss} reflects the agent's code-mediated strategy, rather than being a fixed or incidental byproduct (Friston, 2013; England, 2015). Thus, actively regulated dissipation (Q_{diss}) is a strategic tool: by modulating dissipation, agents use energy expenditure to selectively refine, reorganize, or encode information within their systems, ensuring that thermodynamic processes are harnessed in support of informational and behavioral adaptation. This regulatory capacity enables biological agents to exert control over their surrounding elements, with this control embodied by agential codes that operate through diverse hardware and software implementations—including cell walls, chemical pathways, signaling networks, and abstract regulatory codes.

Critically, this controlled dissipation represents one of the fundamental distinctions between living and non-living matter: while non-living systems exhibit passive energy dissipation governed solely by thermodynamic gradients, living systems instantiate intermediate sub-assemblies—e.g. molecular machines, cellular organelles, and hierarchical control networks—that actively modulate entropy production according to their functional requirements. These intermediate sub-assemblies serve as the physical embodiment of autonomous biological agents, enabling biocodes to channel energy flows in service of specific goals rather than allowing uncontrolled dissipation to equilibrium.

3.4.1. Opinions on energetic function of codes

We propose three main opinions regarding the energetic function of codes.

3.4.1.1. Opinion 1. Biological codes fundamentally act as energy routers, dynamically allocating energy between exploitative (high efficiency, low dissipation) and explorative (low efficiency, high dissipation) operational modes.

In short, biological codes act as energy routers that:

1. Allocate agents' energy flows between exploitative (high efficiency, low dissipation) and explorative (low efficiency, high dissipation) modes (Addicott et al., 2017; Klump et al., 2020, Brownstein et al., 2022).
2. Implement thermodynamic fuzzy logic to switch between dissipation regimes (Gentili and Perez-Mercader, 2022).
3. Optimize recovery strategies for waste heat, as seen in metabolic cycles (van Doren et al., 2017; Atkinson, 1968)

For example, in highly efficient, **exploitative codes** regimes, agents channel energy into maintaining stable, low-entropy states (see Fig. 3c), minimizing dissipation and maximizing the fidelity of functional outcomes—exemplified by the precise folding of proteins into their native conformations.

The efficiency–patterning trade-off observed in protein folding is a general principle that recurs across biological systems governed by codes. However, when agents are tasked with generating or maintaining complex spatial patterns, as seen in processes like Turing pattern formation during tissue morphogenesis (cd134_ dbApr2024, Turing, 1952; Aydin et al., 2022) or the self-organization of bacterial colonies, the energy cost rises substantially (Zhang et al., 2023). Here, a greater fraction of ***available energy*** (i.e. the sum of all accessible energy sources to the agent for work, originating from environmental input and waste heat recovery: hP_{env} *and* αQ_{rec}) is dissipated to support the robustness and adaptability of spatial organization, often at the expense of overall energetic efficiency. This increased dissipation reflects the informational cost of sensing, interpreting, and responding to environmental cues necessary for precise pattern formation (England, 2013), potentially requiring the agent to switch through several different states such as monitoring, regulating, and acting (Ashby, 1956).

Available energy and dissipated energy differ fundamentally in their **entropy levels** and, consequently, in their ability to contribute to work. Available energy—such as environmental input or recoverable waste heat—is **low-entropy**, structured energy that can be directed toward performing biological or mechanical work. In contrast, dissipated energy is **high-entropy**, typically released as heat, and represents energy that has been degraded and is no longer usable for work (Nelson and Cox, 2021). This distinction is central to understanding energy flow and efficiency in biological systems (Schrödinger, 1944).

This entropy balance, mediated by the underlying biological code, reflects a fundamental design constraint: systems that prioritize spatial or functional complexity, because of their spatio-temporal energy demands and complex modular organization, must accept higher energetic costs, while those that maximize efficiency tend to produce simpler, more static outcomes.

In contrast, **exploratory codes** sacrifice efficiency to enable adaptive spatial patterning and system-level coordination. These regimes exhibit:

- Lower energy conversion efficiency ($\eta \sim 0.3-0.5$) (Ge and Qian, 2010; Santillán et al., 1997; Skinner and Dunkel, 2021)
- Higher dissipation ($Q_{diss} \propto \sigma_{EPR}$, where σ_{EPR} is the entropy production rate (Ge and Qian, 2010))
- Increased waste recovery ($\alpha > 0$) in cyclical processes.

Exploratory codes facilitate adaptive search, novelty generation, and long-term flexibility, often at the expense of immediate conversion efficiency. For instance, codes involved in developmental patterning (plants, fungi or animals, cd028_dbApr2024) or immune system diversification (cd098_dbApr2024) are more explorative, allowing agents to adapt to new challenges or environments by incurring higher energetic costs and increased dissipation.

The dataset reveals that purely exploratory codes are extremely rare, while exploitative and mixed codes are much more common across all biological levels. This distribution suggests that most biological systems initially favor stability and resource optimization (exploitative or mixed strategies), and only under certain conditions or evolutionary pressures do more specialized, exploratory codes emerge. Thus, the rarity and patterning of exploratory codes in the taxonomy suggests that they might arise as specializations of more general, exploitative or mixed codes, reflecting an evolutionary and functional continuum rather than discrete, unrelated categories.

Finally, **mixed codes** which combine both exploitative and explorative features, may serve as transitional forms, providing the structural and energetic foundation from which specialized exploratory codes can evolve. These codes could be better modified or adapted, to generate new codes (Prinz, 2023a).

3.4.1.2. Opinion 2. In general, biological codes orchestrate a spectrum of cybernetic work categories, each corresponding to distinct agent activities that collectively define the energetic landscape of living systems. These categories—sensing/monitoring, actuating/executing, regulating/controlling, exploring/predicting, adapting/learning, idling/resting/storing, and communicating/signaling—are not merely functional labels but represent modular components within the generalized work equation introduced above (Ashby, 1956; Maturana and Varela, 1980; Gahrn-Andersen and Prinz, 2021).

Each work component (see in eq. (1), $W_{work} = W_{sensing} + W_{actuating} + \ldots + W_{idling}$) in the energetic equation can thus be decomposed into these cybernetic subcategories, with distinct efficiency (η) and recovery (α) parameters, reflecting the diversity of biological strategies. The different kinds of work described here can be understood as distinct **degrees of freedom** (Farnsworth, 2017, 2023) available to biological agents, each representing a specific modality through which agents interact with and transform their environment, rather than as subcategories of codes.

Decomposing W_{work} into sub-activity terms implies that each corresponds to the time derivative of a distinct Gibbs energy (or free energy) component (G_i). This framework reveals how biological agents achieve efficiency through modular energy partitioning, coupled processes, and hierarchical control. However, non-additive interactions between components necessitate integrated regulation to maintain thermodynamic stability. The approach underscores that biological work is not monolithic but a coordinated ensemble of specialized energy conversions.

3.4.1.3. Opinion 3. A fundamental requirement for biological agents is to increase and maintain internal information in order to persist and adapt within their environment. According to Friston (2013) this is realized through the formation of Markov blankets—statistical boundaries that separate internal from external states, enabling agents to actively regulate their interactions and maintain homeostasis. This organizational closure is not arbitrary, but physically constrained by thermodynamic limitations: specifically, by the balance between micro-level conformational entropy (ΔS_{micro}) and macro-level informational order (ΔS_{macro}), as formalized in equation (1) of the manuscript. To accumulate more information (i.e., to increase I_{fold} or reduce ΔS_{macro}), agents must dissipate energy and reduce accessible microstates, which is only possible by creating new boundaries—compartmentalization (Prinz, 2022a; Prinz, 2022b; Gahrn-Andersen and Prinz, 2021)—that partition the system into autonomous units.

Compartmentalization thus serves as the physical mechanism for generating Markov blankets, enabling the emergence and stabilization of new degrees of freedom, and allowing agents to accumulate, store, and process information beyond the limits set by their unconstrained thermodynamic state. This process is central to the evolution of biological complexity, as it allows for the hierarchical nesting of autonomous agents, each defined by its own Markov blanket and informational closure (Kirchhoff et al., 2018). Critically, this framework highlights that biological agents may deploy multiple codes and work categories in succession or combination, dynamically partitioning energy to balance stability, adaptability, and complexity across organizational levels.

Our taxonomy aligns with the principles of *systems biology*, which seeks to integrate data and models across scales to understand the emergent behavior of complex biological systems (Chen and Wu, 2013). By providing an agent-based, energetically grounded classification, this framework serves as a *systems biology* discovery tool for mapping the organizational logic of living matter.

This paradigm is fundamentally descriptive, providing a robust and structured lens for analyzing how biological codes orchestrate agent behavior and energy allocation. While the current taxonomy and energetic framework offer valuable insights into system organization, they do not yet enable quantitative prediction of agent behavior or system outcomes. Bridging this gap—from descriptive classification to predictive modeling—will require the integration of computational tools, empirical validation, and theoretical advances, as outlined in the following section on future developments.

4. Discussion and future vistas

The proposed agent-based taxonomy developed in this work represents just the beginning of a more comprehensive classification system for biological codes. As the field continues to expand, we anticipate the emergence of numerous new taxonomic subdivisions, and associated biological codes' metadata, that will refine our understanding of code-mediated biological processes. The current framework's ability to classify codes across biological levels and behavioral strategies provides a foundation for discovering previously unrecognized patterns and relationships within living systems. Further subdivision could include the functional or structural role of codes (Prinz, 2023b, 2025), their cybernetic agential relevance or their cybernetic feedback loops across biological levels.

Because of its flexibility and multidisciplinary approach, our discovery methodology, grounded in agent-based modeling and energetic constraints, offers unprecedented opportunities to elucidate therapeutic intervention strategies across multiple biological scales. The systematic mapping of codes to specific agent behaviors creates a powerful lens through which researchers can identify critical nodes for therapeutic targeting.

4.1. Limitations of the framework

We acknowledge the methodological limitations associated with using language models (LLMs) for the classification of biological codes. The classification process, while systematic and transparent, relies heavily on prompts developed for the language model to assign behavioral and biological level labels to each code. This approach introduces potential bias, as the model's outputs are influenced by both the training

data and the specific wording of the prompts.

We attempted to mitigate these limitations by employing a post-processing step, where a secondary prompt is used to verify and refine the initial labels, ensuring that the classification reflects the teleological agency of the biological entities involved rather than merely their physical location. However, the process still depends on the interpretive capabilities of the language model and the quality of the input definitions. In this sense, our use of AI tools highlights not only its utility in managing large and heterogeneous biological datasets but also the methodological challenges of bias, reproducibility, and the need for continuous validation, which must be explicitly addressed in future applications. Specifically, we might make use in future of external validation by means of e.g. independent benchmarks or human-in-the-loop approaches to assess the accuracy and consistency of the model's classifications. Additionally, the iterative, trial-and-error nature of prompt development suggests that our methodology may be sensitive to subtle changes in prompt phrasing, potentially affecting reproducibility. These factors highlight the need for further methodological work, such as incorporating human-in-the-loop validation and systematic bias assessment, to strengthen the reliability of language model-based classification in future research.

Additionally, a key limitation of the current framework is that it is fundamentally descriptive rather than predictive. While the taxonomy and energetic-agentic perspective provide a structured lens for classifying and analyzing how biological codes orchestrate agent behavior and energy allocation, they do not yet enable quantitative prediction of system dynamics or future outcomes.

To move toward a predictive framework, several strategies can be pursued. A promising strategy is ***Chemical AI*** (Gentili, 2025). Chemical AI is trying to mimic biological intelligence competencies using inanimate chemical systems in wetware, in liquid solutions, which is the characteristic phase of life. In Chemical AI, molecules and macromolecules operate as if they were the hardware, whereas phenomena such as chemical reactions, diffusion, advection, migration, convection, and chemical waves operate as if they were the software. In other words, the hardware and software of traditional AI systems are merged in the wetware of reactive chemical solutions. The development of Chemical AI could enable the implementation of the cybernetic work categories shown in Fig. 4 (Gentili and Stano, 2024). Alternatively, integrating ***multi-agent artificial intelligence*** approaches would allow simulation of interacting agents governed by code-mediated rules and energetic constraints, making it possible to forecast emergent system-level behaviors (Soheilypour and Mofrad, 2018). Similarly, developing ***nested or programmable state machines***—such as those modeled on recombinase-based DNA logic (Roquet et al., 2016)—could capture the order-dependent transitions and feedback central to biological processes, enabling quantitative predictions of agent state dynamics. Advances in ***large-scale AI and machine learning***, as demonstrated by recent models predicting gene activity from genomic context (Fu et al., 2025), also suggest that data-driven approaches can uncover underlying rules and provide prospective forecasts for complex biological systems.

4.2. Final remarks and potentials of the framework

While this framework provides a structured, energetically grounded taxonomy for understanding biological codes and agent behavior, it is important to acknowledge its current limitations and future potentials. Notably, we have not yet addressed the full richness of agent–agent interactions—such as those modeled by Turing pattern formation, diffusion-driven instabilities, or game-theoretic approaches—which are central to many biological processes but can still be conceptualized within this paradigm as interacting state machines (Ruan et al., 2024). Phenomena like cellular differentiation and collaboration exemplify agents (cells) not only switching states but also monitoring and acting upon themselves and other agents, illustrating the versatility of the state machine framework for capturing both autonomous and collective behaviors.

The evolutionary dimension is equally significant: the agential state machine paradigm, especially when combined with swarm intelligence, allows the description of complex multicellular phenomena such as metamorphosis, growth, and embryogenesis as multi-level interactions among agents with distinct or even competing goals. When these goals are in conflict, systems experience internal stress, necessitating a rebalancing of entropy and information across organizational levels. This often drives the emergence of new codes or compartments, or the recruitment and mixing of existing codes, fueling evolutionary innovation and the accumulation of biological information—a process reminiscent of practopoiesis (Nikolić, 2015) and the emergence of new phenotypes through code recombination (Kirschner and Gerhart, 2005). As evolution proceeds and organisms become more complex, the lower organizational layers tend to become increasingly machinified, likely as a consequence of resource allocation and energetic constraints (Tlusty and Libchaber, 2025). In fact, as agential information is distributed throughout the biological levels of an organism, it does not remain constant in the different evolved subparts or subagents (Levin, 2022); rather, it dynamically shifts in quantity and form as it is processed and integrated at multiple levels of biological organization (de Castro and McShea, 2022).

Interestingly, claims that the genome is a landing point rather than a starting point (West-Eberhard, 2003) are validated within this context, whereby new codes, structures, and interactions that emerge during evolution can become consolidated and stored in the genome through a circular evolutionary pattern—a circularity that arises from ongoing interactions with the environment and other agents, driving reciprocal eco-evolutionary feedback and adaptation.

Moreover, the agent-based, energetically grounded taxonomy of biological codes offers valuable empirical and theoretical insights that artificial intelligence systems can emulate. By learning from the multi-scale, goal-directed behaviors and energy allocation strategies observed in biological agents, AI architectures—particularly those incorporating agent-based and neuromorphic principles—can enhance their adaptability, efficiency, and robustness. This framework thus provides a promising blueprint for developing AI models that better capture the dynamic balance between exploration and exploitation inherent in living systems, potentially advancing AI frameworks toward more life-like, context-aware intelligence. It is important to emphasize, however, that in the present study AI tools served primarily a methodological and practical role in data classification, rather than a speculative conceptual framing.

Finally, this framework points toward a general principle: cybernetic non-equilibrium open systems—including but not limited to living organisms—tend toward efficiency, stability, and the accumulation of information (Zeng and Wang, 2024). Identifying the rules and constraints that govern these tendencies remains a key challenge for both biology and systems science. By extending the agent-based, energetically informed taxonomy to include agent–agent interactions, evolutionary code mixing, and the dynamics of open non-equilibrium systems, this framework offers a promising foundation for future predictive and integrative models of biological complexity.

CRediT authorship contribution statement

Cris Micheli: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Project administration, Methodology, Formal analysis, Data curation, Conceptualization. **Robert Prinz:** Writing – review & editing, Writing – original draft, Supervision, Resources. **Pier Luigi Gentili:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Methodology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

P. L. Gentili acknowledges the European Union—NextGenerationEU under the Italian Ministry of University and Research (MUR) National Innovation Ecosystem grant ECS00000041—VITALITY. He is grateful to Università degli Studi di Perugia and MUR for support within the project Vitality. P. L. Gentili also acknowledges MUR for support through the project PRIN2022-LUNARLIGHT - Prot. 2022NHLX2M.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.biosystems.2025.105604.

Data availability

The extended classification table is available on request. The codes database is available as Supplementary Material

References

Addicott, M.A., Pearson, J.M., Sweitzer, M.M., Barack, D.L., Platt, M.L., 2017. A primer on foraging and the explore/exploit trade-off for psychiatry research. Neuropsychopharmacology 42 (10), 1931–1939. https://doi.org/10.1038/npp.2017.108.

Ashby, W.R., 1956. An Introduction to Cybernetics. Chapman & Hall.

Aspromonte, M.C., Nugnes, M.V., Quaglia, F., Bouharoua, A., DisProt, Consortium, Tosatto, S.C.E., Piovesan, D., 2024. DisProt in 2024: improving function annotation of intrinsically disordered proteins. Nucleic Acids Res. 52 (D1), D434–D441. https://doi.org/10.1093/nar/gkad928. Erratum in: Nucleic Acids Res. 2025 Feb 27;53(5): gkaf228. doi: 10.1093/nar/gkaf228.

Atkinson, D.E., 1968. The energy charge of the adenylate pool as a regulatory parameter: interaction with feedback modifiers. Biochemistry 7 (11), 4030–4034.

Aydin, O., Passaro, A.P., Raman, R., Spellicy, S.E., Weinberg, R.P., Kamm, R.D., Sample, M., Truskey, G.A., Zartman, J., Dar, R.D., Palacios, S., Wang, J., Tordoff, J., Montserrat, N., Bashir, R., Saif, M.T.A., Weiss, R., 2022. Principles for the design of multicellular engineered living systems. APL Bioeng. 6 (1), 010903. https://doi.org/10.1063/5.0076635.

Barbieri, M., 1985. The semantic theory of evolution. Riv. Biol. 78 (1), 21–48.

Barbieri, M., 1998. The organic codes. The basic mechanism of macroevolution. Riv. Biol. 91 (3), 481–513.

Barbieri, M., 2012. Codepoiesis – the deep logic of life. Biosemiotics 5 (3), 297–299. https://doi.org/10.1007/s12304-012-9162-4.

Barbieri, M., 2025. The concepts of code biology. Biosystems 248, 105400. https://doi.org/10.1016/j.biosystems.2025.105400.

Bard, J.A.M., Drummond, D.A., 2024. Chaperone regulation of biomolecular condensates. Front. Biophys. 2, 1342506. https://doi.org/10.3389/frbis.2024.1342506.

Benítez, M.J., Jiménez, J.S., 2025. Gibbs free energy and enthalpy–entropy compensation in protein folding. Biophysica 5 (1), 2.

Brownstein, A.J., Veliova, M., Acin-Perez, R., Liesa, M., Shirihai, O.S., 2022. ATP-consuming futile cycles as energy dissipating mechanisms to counteract obesity. Rev. Endocr. Metab. Disord. 23 (1), 121–131. https://doi.org/10.1007/s11154-021-09690-w.

Chen, B.S., Wu, C.C., 2013. Systems biology as an integrated platform for bioinformatics, systems synthetic biology, and systems metabolic engineering. Cells 2 (4), 635–688. https://doi.org/10.3390/cells2040635.

de Castro, C., McShea, D.W., 2022. Applying the Prigogine view of dissipative systems to the major transitions in evolution. Paleobiology 48 (4), 711–728. https://doi.org/10.1017/pab.2022.7.

Deuerling, E., Bukau, B., 2004. Chaperone-assisted folding of newly synthesized proteins in the cytosol. Crit. Rev. Biochem. Mol. Biol. 39 (5–6), 261–277. https://doi.org/10.1080/10409230490892496.

Dill, K.A., 1990. Dominant forces in protein folding. Biochemistry 29 (31), 7133–7155. https://doi.org/10.1021/bi00483a001.

England, J.L., 2013. Statistical physics of self-replication. J. Chem. Phys. 139 (12), 121923. https://doi.org/10.1063/1.4818538.

England, J.L., 2015. Dissipative adaptation in driven self-assembly. Nat. Nanotechnol. 10 (11), 919–923. https://doi.org/10.1038/nnano.2015.250.

Farnsworth, K.D., 2017. Can a robot have free will? Entropy 19 (5), 237. https://doi.org/10.3390/e19050237.

Farnsworth, K.D., 2023. How biological codes break causal chains to enable autonomy for organisms. Biosystems 232, 105013. https://doi.org/10.1016/j.biosystems.2023.105013.

Friston, K., 2013. Life as we know it. J. R. Soc. Interface 10 (86), 20130475. https://doi.org/10.1098/rsif.2013.0475.

Fu, X., Mo, S., Buendia, A., Laurent, A.P., Shao, A., Alvarez-Torres, M.D.M., Yu, T., Tan, J., Su, J., Sagatelian, R., Ferrando, A.A., Ciccia, A., Lan, Y., Owens, D.M., Palomero, T., Xing, E.P., Rabadan, R., 2025. A foundation model of transcription across human cell types. Nature 637 (8047), 965–973. https://doi.org/10.1038/s41586-024-08391-z.

Fuxreiter, M., 2018. Fuzziness in protein Interactions-A historical perspective. J. Mol. Biol. 430 (16), 2278–2287. https://doi.org/10.1016/j.jmb.2018.02.015.

Fuxreiter, M., 2022. Electrostatics tunes protein interactions to context. Proc. Natl. Acad. Sci. U. S. A. 119 (31), e2209201119. https://doi.org/10.1073/pnas.2209201119.

Gahrn-Andersen, R., Prinz, R., 2021. How cyborgs transcend Maturana's concept of languaging: a (bio)engineering perspective on information processing and embodied cognition. Riv Ital Filosof Linguaggio 15 (2), 104–120. https://doi.org/10.4396/2021204.

Ge, H., Qian, H., 2010. Physical origins of entropy production, free energy dissipation, and their mathematical representations. Phys. Rev. E - Stat. Nonlinear Soft Matter Phys. 81 (5 Pt 1), 051133. https://doi.org/10.1103/PhysRevE.81.051133.

Gentili, P.L., 2018. The fuzziness of the molecular world and its perspectives. Molecules 23 (8), 2074. https://doi.org/10.3390/molecules23082074.

Gentili, P.L., Perez-Mercader, J., 2022. Quantitative estimation of chemical microheterogeneity through the determination of fuzzy entropy. Front. Chem. 10, 950769. https://doi.org/10.3389/fchem.2022.950769.

Gentili, P.L., 2024. The conformational contribution to molecular complexity and its implications for information processing in living beings and chemical artificial intelligence. Biomimetics (Basel). 9 (2), 121. https://doi.org/10.3390/biomimetics9020121.

Gentili, P.L., Stano, P., 2024. Living cells and biological mechanisms as prototypes for developing chemical artificial intelligence. Biochem. Biophys. Res. Commun. 720, 150060. https://doi.org/10.1016/j.bbrc.2024.150060.

Gentili, P.L., 2025. Chemical AI in the limelight: the contribution of photochromic materials and oscillatory chemical reactions. Adv. Opt. Mater. 13 (15), 2200016. https://doi.org/10.1002/adom.202500016.

Gianni, S., Freiberger, M.I., Jemth, P., Ferreiro, D.U., Wolynes, P.G., Fuxreiter, M., 2021. Fuzziness and frustration in the energy landscape of protein folding, function, and assembly. Acc. Chem. Res. 54 (5), 1251–1259. https://doi.org/10.1021/acs.accounts.0c00813.

Gusev, A.O., Martyushev, L.M., 2021. An evolution based on various energy strategies. Entropy (Basel) 23 (3), 317. https://doi.org/10.3390/e23030317.

Hill, A.F., Barnham, K.J., Bottomley, S.P., Cappai, R. (Eds.), 2011. Protein Folding, Misfolding, and Disease: Methods and Protocols. Methods Mol Biol. Humana Press, Totowa (NJ), p. 249. https://doi.org/10.1007/978-1-60327-223-0. ISBN: 978-1-60327-223-0.

Jewett, A.I., Shea, J.E., 2010. Reconciling theories of chaperonin accelerated folding with experimental evidence. Cell. Mol. Life Sci. 67 (2), 255–276. https://doi.org/10.1007/s00018-009-0164-6.

Joo, H., Chavan, A.G., Fraga, K.J., Tsai, J., 2015. An amino acid code for irregular and mixed protein packing. Proteins 83 (12), 2147–2161. https://doi.org/10.1002/prot.24929.

Kim, C.S., 2023. Free energy and inference in living systems. Interface Focus 13 (3), 20220041. https://doi.org/10.1098/rsfs.2022.0041.

Kirchhoff, M., Parr, T., Palacios, E., Friston, K., Kiverstein, J., 2018. The markov blankets of life: autonomy, active inference and the free energy principle. J. R. Soc. Interface 15 (138), 20170792. https://doi.org/10.1098/rsif.2017.0792.

Kirschner, M.W., Gerhart, J.C., 2005. The Plausibility of Life: Resolving Darwin's Dilemma. Yale University Press.

Klump, H.H., Völker, J., Breslauer, K.J., 2020. Energy mapping of the genetic code and genomic domains: implications for code evolution and molecular Darwinism. Q. Rev. Biophys. 53, e11. https://doi.org/10.1017/S0033583520000098.

Levin, M., 2022. Technological approach to mind everywhere: an experimentally-grounded framework for understanding diverse bodies and minds. Front. Syst. Neurosci. 16, 768201. https://doi.org/10.3389/fnsys.2022.768201.

Maturana, H.R., Varela, F.J., 1980. Autopoiesis and cognition: the realization of the living. Dordrecht (NL): springer. Boston Stud. Philosophy and History of Sci. 42. https://doi.org/10.1007/978-94-009-8947-4. ISBN: 978-90-277-1016-1.

McShea, D.W., 2016. Hierarchy: the source of teleology in evolution. In: Eldredge, N., Pievani, T., Serrelli, E. (Eds.), Evolutionary Theory: a Hierarchical Perspective. University of Chicago Press, Chicago (IL), pp. 86–102. https://doi.org/10.7208/9780226426198-003-0006.

Mikhailovsky, G.E., 2024. Life, its definition, origin, evolution, and four-dimensional hierarchical structure. Biosystems 237, 105158. https://doi.org/10.1016/j.biosystems.2024.105158.

Montevil, M., Mossio, M., 2015. Biological organization as closure of constraints. J. Theor. Biol. 372, 179–191. https://doi.org/10.1016/j.jtbi.2015.02.029.

Nelson, D.L., Cox, M.M., 2021. Lehninger's Principles of Biochemistry, eighth ed. W. H. Freeman, New York.

Nikolić, D., 2015. Practopoiesis: or how life fosters a mind. J. Theor. Biol. 373, 40–61. https://doi.org/10.1016/j.jtbi.2015.03.003.

Otsuka J. The negative entropy in organisms; its maintenance and extension. J. Mod. Phys., 9, 2156-2169. doi: 10.4236/jmp.2018.912136.

Pleyer, J., Fleck, C., 2023. Agent-based models in cellular systems. Front. Physiol. 10, 968409. https://doi.org/10.3389/fphy.2022.968409.

Prigogine, I., Nicolis, G., 1971. Biological order, structure and instabilities. Q. Rev. Biophys. 4 (2), 107–148. https://doi.org/10.1017/S0033583500000615.
Prinz, R., 2022a. A simple measure for biocomplexity. Biosystems 217, 104670. https://doi.org/10.1016/j.biosystems.2022.104670.
Prinz, R., 2022b. The modularity codes. Biosystems 219, 104735. https://doi.org/10.1016/j.biosystems.2022.104735.
Prinz, R., 2023a. Nothing in evolution makes sense except in the light of code biology. Biosystems 229, 104907. https://doi.org/10.1016/j.biosystems.2023.104907.
Prinz, R., 2023b. Biological codes: a field guide for code hunters. Biol Theory 19 (2), 120–136. https://doi.org/10.1007/s13752-023-00444-2.
Prinz, R., 2024. Code biology database [Internet]. Int. Soc. Code Biol. Available from: https://www.codebiology.org/.
Prinz, R., 2025. Functional Classification of Organic Codes. https://doi.org/10.13140/RG.2.2.12439.89767.
Prinz, R., Bucher, P., Kun, Á., Paredes, O., Aragno, A., Shelby, C., Gumbel, M., Fimmel, E., Strüngmann, L., 2025. Codes across (life)sciences. Biosystems 254, 105515. https://doi.org/10.1016/j.biosystems.2025.105515. Epub ahead of print.
Robin, A.N., Denton, K.K., Horna Lowell, E.S., Dulay, T., Ebrahimi, S., Johnson, G.C., Mai, D., O'Fallon, S., Philson, C.S., Speck, H.P., Zhang, X.P., Nonacs, P., 2021. Major evolutionary transitions and the roles of facilitation and information in ecosystem transformations. Front Ecol. Evol. 9, 711556. https://doi.org/10.3389/fevo.2021.711556.
Roquet, N., Soleimany, A.P., Ferris, A.C., Aaronson, S., Lu, T.K., 2016. Synthetic recombinase-based state machines in living cells. Science 353 (6297), aad8559. https://doi.org/10.1126/science.aad8559.
Ruan, S., Ma, Y., Xiang, C., 2024. Real-time coordinated control strategy for the hybrid electric propulsion system of a flying car with model adaptation and game theory. Proc. Inst. Mech. Eng. Part I J Syst Control Eng. 238 (1), 3–18. https://doi.org/10.1177/09544070241248560.
Rumbley, J., Hoang, L., Mayne, L., Englander, S.W., 2001. An amino acid code for protein folding. Proc. Natl. Acad. Sci. USA 98 (1), 105–112. https://doi.org/10.1073/pnas.98.1.105.
Sánchez, I.E., Galpern, E.A., Garibaldi, M.M., Ferreiro, D.U., 2022. Molecular information theory meets protein folding. J. Phys. Chem. B 126 (43), 8655–8668. https://doi.org/10.1021/acs.jpcb.2c04532.
Santillán, M., Arias-Hernández, L.A., Angulo-Brown, F., 1997. Some optimization criteria for biological systems in linear irreversible thermodynamics. Il Nuovo Cimento D 19 (1), 99.
Sacquin-Mora, S., Prévost, C., 2021. When order meets disorder: modeling and function of the protein interface in fuzzy complexes. Biomolecules 11 (10), 1529. https://doi.org/10.3390/biom11101529.
Schrödinger, E., 1944. What is Life? the Physical Aspect of the Living Cell. Cambridge University Press, Cambridge (UK), p. 91. ISBN: 978-0-521-42708-8.
Schweitzer, F., 2019. An agent-based framework of active matter with applications in biological and social systems. Eur. J. Phys. 40 (1), 14003. https://doi.org/10.1088/1361-6404/aaeb63.
Shapiro, J.A., 2021. Why the third way of evolution is necessary. Theor. Biol. Forum 114 (2), 13–26. https://doi.org/10.19272/202111402002.
Sharma, R., Raduly, Z., Miskei, M., Fuxreiter, M., 2015. Fuzzy complexes: specific binding without complete folding. FEBS Lett. 589 (19 Pt A), 2533–2542. https://doi.org/10.1016/j.febslet.2015.07.022.
Skinner, D.J., Dunkel, J., 2021. Improved bounds on entropy production in living systems. Proc. Natl. Acad. Sci. U. S. A. 118 (18), e2024300118. https://doi.org/10.1073/pnas.2024300118.
Soheilypour, M., Mofrad, M.R.K., 2018. Agent-based modeling in molecular systems biology. Bioessays 40 (8), e1800020. https://doi.org/10.1002/bies.201800020.
Sorokina, I., Mushegian, A.R., Koonin, E.V., 2022. Is protein folding a thermodynamically unfavorable, active, energy-dependent process? Int. J. Mol. Sci. 23 (1), 521. https://doi.org/10.3390/ijms23010521.
Stephan, S., Galland, S., Labbani Narsis, O., Shoji, K., Vachenc, S., Gerart, S., Nicolle, C., 2024. Agent-based approaches for biological modeling in oncology: a literature review. Artif. Intell. Med. 152, 102884. https://doi.org/10.1016/j.artmed.2024.102884.
Tlusty, T., Libchaber, A., 2025. Life sets off a Cascade of machines. Proc. Natl. Acad. Sci. U. S. A. 122 (4), e2418000122. https://doi.org/10.1073/pnas.2418000122.
Todd, M.J., Lorimer, G.H., Thirumalai, D., 1996. Chaperonin-facilitated protein folding: optimization of rate and yield by an iterative annealing mechanism. Proc. Natl. Acad. Sci. U. S. A. 93 (25), 14502–14507. https://doi.org/10.1073/pnas.93.25.14502.
Turing, A.M., 1952. The chemical basis of morphogenesis. Philos. Trans. R. Soc. Lond. B Biol. Sci. 237 (641), 37–72. https://doi.org/10.1098/rstb.1952.0012.
Van Doren, L.G., Posmanik, R., Bicalho, F.A., Tester, J.W., Sills, D.L., 2017. Prospects for energy recovery during hydrothermal and biological processing of waste biomass. Bioresour. Technol. 225, 67–74. https://doi.org/10.1016/j.biortech.2016.11.030.
Varadi, M., Vranken, W., Guharoy, M., Tompa, P., 2015. Computational approaches for inferring the functions of intrinsically disordered proteins. Front. Mol. Biosci. 2, 45. https://doi.org/10.3389/fmolb.2015.00045.
Volk, T., Pauluis, O., 2010. It is not the entropy you produce, rather, how you produce it. Philos. Trans. R. Soc. Lond. B Biol. Sci. 365 (1545), 1317–1322. https://doi.org/10.1098/rstb.2010.0019.
von Stockar, U., Liu, J., 1999. Does microbial life always feed on negative entropy? Thermodynamic analysis of microbial growth. Biochim. Biophys. Acta 1412 (3), 191–211. https://doi.org/10.1016/s0005-2728(99)00065-1.
Wallace, R., 2011. Structure and dynamics of the 'protein folding code' inferred using Tlusty's topological rate distortion approach. Biosystems 103 (1), 18–26. https://doi.org/10.1016/j.biosystems.2010.09.007.
Wang, J., Oliveira, R.J., Chu, X., Whitford, P.C., Chahine, J., Han, W., Wang, E., Onuchic, J.N., Leite, V.B., 2012. Topography of funneled landscapes determines the thermodynamics and kinetics of protein folding. Proc. Natl. Acad. Sci. USA 109 (39), 15763–15768. https://doi.org/10.1073/pnas.1212842109.
Waudby, C.A., Dobson, C.M., Christodoulou, J., 2019. Nature and regulation of protein folding on the ribosome. Trends Biochem. Sci. 44 (11), 914–926. https://doi.org/10.1016/j.tibs.2019.06.008.
West-Eberhard, M.J., 2003. Developmental Plasticity and Evolution. Oxford University Press.
Wiener, N., 1961. Cybernetics: or Control and Communication in the Animal and the Machine, second ed. MIT Press.
Yang, X., Heinemann, M., Howard, J., Huber, G., Iyer-Biswas, S., Le Treut, G., Lynch, M., Montooth, K.L., Needleman, D.J., Pigolotti, S., Rodenfels, J., Ronceray, P., Shankar, S., Tavassoly, I., Thutupalli, S., Titov, D.V., Wang, J., Foster, P.J., 2021. Physical bioenergetics: energy fluxes, budgets, and constraints in cells. Proc. Natl. Acad. Sci. U. S. A. 118 (26), e2026786118. https://doi.org/10.1073/pnas.2026786118.
Zeng, Q., Wang, J., 2024. Non-equilibrium enhancement of classical information transmission. Entropy (Basel). 26 (7), 581. https://doi.org/10.3390/e26070581.
Zhang, B., DeAngelis, D.L., 2020. An overview of agent-based models in plant biology and ecology. Ann. Bot. 126 (4), 539–557. https://doi.org/10.1093/aob/mcaa043.
Zhang, D., Zhang, C., Ouyang, Q., Tu, Y., 2023. Free energy dissipation enhances spatial accuracy and robustness of self-positioned turing pattern in small biochemical systems. J. R. Soc. Interface 20 (204), 20230276. https://doi.org/10.1098/rsif.2023.0276.