AMISHI_2.docx

by Amishi Roy

Submission date: 03-Jun-2026 03:19PM (UTC+0530)
Submission ID: 2975575251
File name: AMISHI_2.docx (790.63K)
Word count: 7220
Character count: 43169

ABSTRACT

The process of protein aggregation is a vital one which depends on time and necessitates models which transcend closed-system approaches to become models of dynamic kinetics and proteostasis. The generalized modelling approach connects the physical mechanism of aggregation with the biological system of aging. This physical model is composed of kinetic laws, the behaviours of the source terms (monomer generation/elimination), and thermos-reversible assembly. The model predicts the scaling behaviours of aggregation which depend on whether the process is limited by the formation of aggregates or the generation of monomers.

The biological aspect of the modelling framework takes into account the effect of the decline in the process of proteostasis in a systems biology model. Coupled differential equations form the basis of this biological model with consideration of the decline in physiological function, hormone regulation, genetic instability, and gender-related effects. Importantly, the non-linear dynamics of proteostasis with respect to age-related stresses and inefficient clearing must be accounted for since they can lead to a tipping point where a rapid build-up occurs.

Keywords: Protein Aggregation, Dynamic Kinetics, Proteostasis, Systems Biology

INTRODUCTION AND LITERATURE REVIEW

1.1 Introduction

Biophysics is an interdisciplinary science that combines physics and biology in an effort to understand the workings of biological systems, from their basic components like molecules to more complex components like living organisms. Using the concepts and techniques of physics, this field aims to discover the physical rules responsible for shaping biological structures, and molecule-to-molecule interactions necessary for survival.

In general, biophysics can be two types: Molecular Biophysics and Cellular Biophysics.

- Molecular Biophysics finds out the behaviour, structure formation, and dynamics of biomolecules such as proteins, nucleic acids, and membranes. By

analysing the interactions between these molecules, we can identify the factors that cause misfolding leading to illnesses such as Alzheimer's, Parkinson's, and sickle cell anaemia.

- Cellular Biophysics finds the functioning of cells from a physical perspective, especially concerning transport mechanisms and signalling. It connects molecular behaviour to cell function and to the physiology of the full organisms.

Biomolecules: The Building Blocks of Life

A molecule is a very small unit of a substance made up of atoms binded together. In living systems, these are known as biomolecules, which has a variety of organic compounds that carry out the chemical and structural tasks of life. Biomolecules can be:

- Macromolecules, such as proteins, carbohydrates, and nucleic acids, with large structures.
- Micromolecules, such as amino acids, lipids, hormones, and enzymes, which often act as smaller but are important regulators or metabolic intermediates.

1.2 POLYMER PHYSICS

Polymers are large molecules made up of many tiny, repeating units called **monomers**, joined together by covalent bonds. They are everywhere in daily life, from the natural polymers present in proteins and DNA to synthetic materials like plastics, rubbers, and fibers.

Structure and Formation

Polymers are **macromolecules** identified by long-chain structures, which can be **linear**, **branched**, or **cross-linked**. The molecular weight of a polymer can be represented as:

$$M_{\text{poly}} = N \times M_{\text{mono}} \qquad \rightarrow (1.1)$$

where N is the degree of polymerization, and M_{poly} and M_{mono} shows the molecular weights of the polymer and monomer, respectively.

Polymers are of two main types:

- **Homopolymers**: they are made from one type of monomer (e.g., polyethylene).

- **Copolymers (Heteropolymers)**: they are made from two or more monomer types (e.g., nylon, proteins).

Bonding and Growth

Polymerization can occurs mainly via two mechanisms:

- **Addition polymerization**, here monomers join without producing byproducts (e.g., polyethylene).
- **Condensation polymerization**, here the reaction produces small molecules like water as byproducts (e.g., proteins).

The Polywater Hypothesis

Polywater was thought to be consists of polymerized water molecules. The idea arose because each water molecule can form four hydrogen bonds, which gives it a functionality of $f = 4$. This theoretically allowed for an extensive, interconnected 3D hydrogen-bonded network. Since, hydrogen bonds are weak and thermal energy at room temperature is comparable to their bond energy, such a structure is unstable.

Polydispersity

Real polymer samples contain chains of different lengths. This molecular weight distribution is described by the Polydispersity Index (PDI):

$$\text{PDI} = \frac{M_w}{M_n} \qquad \rightarrow (1.2)$$

where M_w is the weight-average molecular weight and M_n the number average molecular weight. A PDI value of 1 shows uniform chains, on the other hand higher values suggest greater diversity in chain length. All natural polymers are polydisperse. Achieving uniform chain lengths requires post-synthesis filtration and fractionation.

1.3 LENGTH SCALES

Long-chain macromolecules such as polymers and proteins can accept a big number of orientations because of internal bond rotations and external factors like solvent quality, temperature, and pressure. This focuses on the equilibrium, static properties of single, linear polymer chains and focuses to relate their overall dimensions to the properties of their

constituent monomers using models like the *Freely Jointed Chain* and the *Random Flight Model*.

a. **Radius of gyration (R_g)**

The radius of gyration describes how far the monomer units are distributed from the polymer's centre of mass. Mathematically, it's expressed as:

$$\overrightarrow{R_g^2} = \frac{1}{N^2}\sum_{i,j=1}^{N}\langle(\overrightarrow{R_i} - \overrightarrow{R_j})^2\rangle \qquad \rightarrow (1.3)$$

Here,

N is the total number of monomer units,

R_i:-i is the position vector of the i-th monomer

R_g increases with chain length and flexibility and reflects how the polymer's mass is distributed in space.

b. Hydrodynamic radius (R_h).

It shows how the polymer moves through a fluid, taking account of the frictional drag and surrounding solvent layer. R_h is usually smaller than R_g, and their ratio R_g/R_h serves as a useful indicator of polymer conformation.

$$R_h = \frac{k_B T}{6\pi\eta D} \qquad \rightarrow (1.4)$$

here,

- k_B is Boltzmann's constant,
- T is absolute temperature
- η is the viscosity of the solvent at T,
- D is the centre of mass diffusion coefficient of the polymer

R_g is a structural property R_h reflects how a polymer interacts with the solvent during motion.

c. Contour length and dry radius.

The contour length is the maximum length of a fully stretched polymer. The **dry radius** represents the size of a polymer in the absence of solvent, depending on its mass and density.

If the chain has molecular weight M and density ρ, its dry radius R_d will be given by

$$R_d = \left(\frac{3M}{4\pi\rho N_A}\right)^{\frac{1}{3}} \quad \rightarrow (1.5)$$

The Freely Jointed Chain and Random Flight Models

In the *Freely Jointed Chain Model*, the polymer is shown as *N* rigid segments of fixed length l, each freely rotating in three-dimensional space. In each case, the direction of the bond does not depend on the other bonds, and this leads to a random walk path.

The Random Flight model builds on the Random Walk model by including relations amongst successive bonds, which makes it more of a real form. In averaging over many paths, it shows that a polymer obeys Gaussian statistics.

d. Persistence Length and Kuhn Segments

The stiffness of real polymers differs. The persistence length is the length over which the orientation of the polymer chain loses correlation. Flexible polymers has a short persistence length, whereas biopolymers such as DNA and actin have long persistence lengths because of their strength.

$$\langle cos\theta(s)\rangle = e^{-s/l_p} \quad \rightarrow (1.6)$$

Relating Kuhn Length and End-to-End Distance

The freely jointed chain gives:

$$R^2 = Nl^2 \quad \rightarrow (1.7)$$

This is called the mean-squared end-to-end distance for a chain of N segments of length l. To generalize this to real polymers, we have defined the **Kuhn length** b and write:

$$R^2 = N_K b^2 \quad \rightarrow (1.8)$$

Here:

- $N_K = \frac{Nl}{b}$ is the number of Kuhn segments
- K_b is the effective segment length for a real chain (includes stiffness),
- The total contour length remains L=Nl, so:

$$R^2 = Lb \qquad \rightarrow (1.9)$$

e. Scaling Laws and Chain Size

Scaling relations, in polymer science deals with how characteristic size measures (like the end-to-end distance or R_g) grow with the number of monomers N or molecular weight M. For an *ideal chain* (no excluded volume interactions), $R_g \propto N^{0.5} \propto M^{0.5}$. Conversely, for polymers in *good solvents*, where monomer–monomer repulsions predominate, and therefore scaling exponent increases to about 0.6.

f. Excluded Volume and Polymer Solvent Environment

Excluded volume refers to the property of polymer segments not being able to live in space. Hence, a polymer becomes expanded, and the Gaussian chain approximation becomes invalid. Thus, the distribution obtained will favour extended chains with higher values of R_g.

g. Gaussian Chain Approximation and Experimentally Observed Deviations

Under the Gaussian Chain approximation (in the continuum limit), where the number of segments is large, and their length is small, the end-to-end distance of the polymer becomes normally distributed. This Gaussian Chain model is the basis for experiments (light scattering or diffusion) and quantitative analysis of deviations from the ideal case in real-world polymers.

1.4 MOLECULAR BIOPHYSICS

Molecular biophysics is a science which combines fundamental concepts from physics, chemistry, biology, mathematics, and nanotechnology in order to describe quantitatively the processes taking place in living items.

a) Life and Non-Equilibrium Thermodynamics

Organisms create system in a random world. In order to have their organized, low-entropy condition, organisms:

1. Absorb ordered energy and matter from their surroundings, and

2. Export disorder by way of heat and by-products.

Organisms are open systems and continuously exchange energy and matter, allowing them to maintain a highly random and nonequilibrium state which helps them resist the natural flow of energy towards entropy. This process allows them to sustain metabolism, grow, and reproduce.

b) Biological Unity and Its Origin

Life itself emerged as a result of the self-organization of non-biological organic molecules in the form of cells structured using three major classes of biopolymers: nucleic acids, proteins, and polysaccharides. These molecules' monomers such as nucleotides, amino acids, and monosaccharides become building blocks that govern biological life forms.

Table1: Classification of biopolymer

Biopolymer	**Monomer Unit**	**Primary Function**	**Key Biophysical Features**
Nucleic Acids (DNA/RNA)	Nucleotides	Storage and transmission of genetic information	Chiral, highly charged polyelectrolytes
Proteins	Amino acids	Catalytic, structural, and motor functions	Chiral (L-form), polyampholyte with pH-dependent charge
Polysaccharides	Monosaccharides	Energy storage and structural support	Chiral (D-form), high charge density

c) Chirality in Biomolecules

Chirality is the quality of having a geometrical configuration where a molecule cannot be superimposed on its mirror image. This occurs in cases where a chiral carbon atom attaches to four different substituents and creates two mirror image structures, enantiomers.

In biological systems:

- Proteins only contain L-amino acids, and
- Carbohydrates only use D-sugars

This is called biological homochirality. Such preference of one kind of enantiomer over another is not thermodynamically favored, indicating the tendency of life forms towards orderliness. Optical activity is characteristic of these compounds, as they rotate plane polarized light in opposite directions.

d. Charged Macromolecules: Polyelectrolytes and Polyampholytes

There are two types of charged biomolecules, polyelectrolytes and polyampholytes, which play an important part in living organisms' structure and functions.

Polyelectrolytes

They possess a single charge, either positive or negative, irrespective of pH. They are made up of a balance between their electrostatic potential and configurational entropy, extending themselves in low salt concentrations where the charges do not screen one another.

Example: DNA and polysaccharides.

Polyampholytes

They comprise both acidic and basic groups; thus, having a net charge depending upon pH. They have variable conformations with changing pH and ion concentration.

Examples: Proteins

e. Electrostatic Interaction in Solution

The charged nature of macro-molecules (macro-ions) in aqueous media leads to electrostatic interactions among them. In solution, charged molecules tend to attract oppositely charged ions resulting in the following configuration:

The Stern layer, composed of tightly bound ions moving with the macroion, and

- The diffuse layer, containing mobile ions distributed in solution.

Together, these layers reduce the effective electric field—a phenomenon known as electrostatic screening—which allows macromolecules to interact transiently and specifically within crowded cellular environments.

The Debye-Hückel approach explains the way the above-mentioned potentials decay mathematically.

1. The core area, containing no movable ions.

2. The Stern layer with the potential satisfying the Laplace equation.

3. The diffuse area where the decay of the potential is governed by the Poisson-Boltzmann equation.

This theoretical approach takes into account how the electrostatic interactions become altered due to varying the following parameters: the ion concentration, the temperature, and the dielectric constant.

Key Debye–Hückel Parameters

i) Debye-Hückel Parameter (κ)

The Debye-Hückel parameter (κ) is defined as**:**

$$\kappa = \left(\frac{8\pi N\varepsilon^2}{D k_B T}\right)^{\frac{1}{2}} \qquad \rightarrow (1.10)$$

where:

- N = number of charges per unit volume,
- ε = charge on each mobile ion,
- D = dielectric constant of the medium,
- k_B = Boltzmann constant,
- T = temperature.

Quantifies the strength of electrostatic shielding; higher ion content means stronger screening.

ii) Debye Screening Length (κ^{-1})

Quantifies how far away an electrostatic field can influence a macroion.

iii) Ionic Strength (I)

$$I = \frac{1}{2}\sum_i c_i z_i^2 \qquad \rightarrow (1.11)$$

where:

- Z_i = charge on ion i.

Directly proportional to total ion concentration; controls the value of κ in $\kappa \propto \sqrt{I}$.

iv) Bjerrum Length (l_B)

The separation between two charged particles at which the electrostatic force is equal to the thermal energy.

$$l_B = \frac{e^2}{4\pi\varepsilon\varepsilon_0 k_B T} \qquad \rightarrow (1.12)$$

where:

- e = electronic charge,
- ε = dielectric constant of the medium.

1.5 STRUCTURE BIOPHYSICS

Biopolymers are macromolecules made up of monomeric building blocks. They form the basis of many biological processes in living cells.

a. Biopolymers (Nucleic acids)

Biopolymers act as stores and messengers of genetic material. Made up of nucleotide monomers, nucleic acids display structural intricacies enabling them to carry out their biological functions

.i) Nucleotide Structure

Each nucleotide consists of three components:

- A pentose sugar (D-deoxyribose in DNA, D-ribose in RNA),
- A nitrogenous base (purines: adenine-A, guanine-G; pyrimidines: cytosine-C, thymine-T in DNA, uracil-U in RNA), and
- One to three phosphate groups.

A base + sugar = nucleoside; nucleoside + phosphate = nucleotide. Beyond genetics, nucleotides power cellular metabolism (ATP, GTP).

ii) DNA: The Double Helix

ii) DNA (Deoxyribonucleic acid)

Deoxyribonucleic acids contain four deoxyribonucleotides (deoxyadenosine, deoxyguanosine, deoxycytidine, and deoxythymidine). The base complementarity follows the Chargaff rules:

- A-T pair: 2 hydrogen bonds
- G-C pair: 3 hydrogen bonds (greater stability)

Watson-Crick B-DNA molecule adopts a right-handed helical structure:

- The hydrophilic sugar-phosphate backbone is directed to the exterior while hydrophobic bases form internal stacks,
- There are 10 base pairs per helical turn (3.4 nm or 0.34 nm/bp).

Other forms are: A-form DNA (shorter, wider), and left-handed Z-DNA.

iii) RNA (Ribonucleic acid)

Ribonucleic acids consist of four ribonucleotides, except that the thymine nucleobase is replaced by uracil. Generally, RNA adopts a single-stranded configuration and folds upon itself to generate unique three-dimensional conformations using inter-molecular base pairing, producing:

•Structural elements like hairpins, loops, bulges, and pseudoknots that form the active site.

c. Proteins:

Proteins are the most versatile biopolymers, being α-L-amino acid chains forming about half of the dry mass of the cell. Proteins play different roles, including structural (collagen), enzymatic, regulatory, and protective (antibodies).

i)Amino Acids

An amino acid contains an alpha-carbon attached to an amino group ($-NH_3^+$), carboxyl group ($-COO^-$), and R-group, which differs among various amino acids.

In aqueous media, they function as zwitterions..

ii)The Peptide Bond

A peptide bond forms between an amino group and a carboxyl group to form a peptide bond and produce a molecule of water. Peptide bonds have a partial double bond characteristic making them stiff and planar.

d. Carbohydrates

Carbohydrates constitute nature's largest population of organic compounds, serving as both energy sources and structural components.

i) Classification

Table2: Classification of Carbohydrates:

Class	Composition	Examples	Primary Role
Monosaccharides	Single sugar unit	Glucose, fructose	Energy source, building block
Oligosaccharides	2–10 sugars	Maltose, sucrose, lactose	Energy, cell recognition
Polysaccharides	Many sugars	Starch, cellulose, glycogen	Storage, structure

1.6 PHYSICS OF PROTEIN

i)Charge on protein molecules

i)Charge on Protein Molecules

Protein acts as a polyampholyte; it bears acidic and basic sites that depend on pH. Below the protein isoelectric point, it has positive net charge; above the isoelectric point, it carries negative charge; at the isoelectric point, the net charge is zero.

ii)Protein folding and helix–coil transition

Protein folding involves transitions between ordered helical states (low entropy, stabilized by hydrogen bonds) and disordered coil states (high entropy). The Zimm–Bragg model describes this using:

- a propagation parameter s for adding a residue to an existing helix, and
- a nucleation parameter σ for initiating a helix segment.

At the transition temperature T_m, the population of helix and coil is equal (50% each).

iii)Kinetics of protein folding

Folding pathways are often classified by how many kinetic states are involved:

- Two-state model: U ↔ N, where unfolded (U) converts directly to native (N) without detectable intermediates.
- Two-step model: U ↔ I → N, with a single intermediate I on the pathway.
- Three-step or multi-step models: U ↔ I_1 ↔ I_2 → N, with multiple intermediates.

These schemes capture whether the energy landscape has a simple single barrier or multiple wells and barriers, and they are probed experimentally by stopped-flow, temperature-jump, or single-molecule methods.

iv)Energetics of ligand binding

When a ligand L binds to a protein P to form a complex PL (P + L ↔ PL), the equilibrium is described by an association constant K_a. The thermodynamics obey:

$$\Delta G = -RT\ln\ K_a = \Delta H - T\Delta S \qquad \rightarrow (1.13)$$

where ΔG is the free energy change, ΔH the enthalpy change (e.g., hydrogen bonding, van der Waals, electrostatics), and ΔS the entropy change (e.g., ordering of water, conformational restriction).

CHAPTER 2

OLD AND NEW KINETIC MODEL

2.1 OLD AND NEW KINETIC MODEL

- **Model 1: The Simple Test Tube (Closed System)**

Model 1: Clumping test-tube experiment (closed system)

At the beginning of studying protein clumping, scientists considered the simplest glass bottle (in vitro). Researchers created complex kinetic models explaining three basic mechanisms:

1.Primary nucleation – spontaneous formation of aggregates by protein monomers

2.Elongation - addition of monomers to ends of fibrils

3.Secondary nucleation – creation of new aggregates induced by existing fibrils (fragmentation or surface-catalytic nucleation)

Such models described the mechanism of aggregation well enough for proteins in a pure buffer. These findings allowed to understand molecular bases of mutations causing disease and mechanisms of action for future drugs.

Mathematically,

- **Monomer** (μ): The single, free protein brick.
- **Aggregate Mass** (ν): The total mass of all the big clumps.

Equation: $\mu+\nu=$ Constant (Monomer + Clumps = Fixed amount). $\rightarrow (2.1)$

- Works in test tubes
- Fails in living cells

The Critical Gap

But in this case, the assumption of mass conservation applied. That is, the amount of protein (free monomers + fibrils) should remain unchanged. In biology, this approach fails due to:

i) continuous synthesis of proteins by the cell
ii) decomposition and clearance of proteins
iii) transitions between aggregation-prone and aggregation-resistant forms
iv) existence of protein complexes, which should dissociate before aggregation.

Model2: Biological Cell (open system)

To make this model more accurate, the researchers introduced source-term S(T) into the equation. Mathematically, it looks like**:**

$$\mu(T)+\nu(T)=S(T) \qquad \rightarrow \textbf{(2.2)}$$

The amount of protein depends on time (T)

2.2 THE SOURCE-TERM CONCEPT

The researchers extended the classic models by adding source term function s(t), describing protein synthesis in time:

From equation 2.2,

$$\mu(T)+\nu(T)=S(T)$$

Dimensionless Analysis

In order to explore all kinds of possible behaviour, the research team introduced the following variables in their analysis::

- **K** = k_{SOURCE}/k (production rate in dimensionless units)
- **ε** = $\lambda^2/(2k^2)$ (relative intensity of primary nucleation)
- **T** = $k \cdot t$ (time in dimensionless units)

Such a mathematical trick helps study the whole parameter space and find universal patterns regardless of particular protein properties.

2.3 EXPERIMENTAL CASE STUDY

Aggregation mechanism of p53 protein

Biological background

P53 protein plays an important role in our bodies as a "guardian of the genome". It prevents cancer development by causing death to damaged cells. When protein aggregates, its activity stops and it can acquire harmful properties, making it an important prognostic marker in oncology.

The Aggregation Mechanism:

Native p53 doesn't aggregate; it must first unfold. This unfolding process creates a natural "source-term" where:

- Non-aggregating species = folded p53
- Aggregation-prone monomer = unfolded p53
- Source process = protein unfolding

Experimental Data Analysis:

Using published data from Wilcken et al. (2012), the team analyzed p53 aggregation at multiple concentrations (3, 6, 9, 12 μM, plus seeded reactions).

Critical Observation:

When normalized by total concentration, all curves collapsed onto a single master curve—a hallmark of first-order kinetics where $M(t) \approx s(t)$.

CHAPTER 3

PROPOSED RESEARCH WORKS

3.1 THESIS PROPOSAL

The basic idea of traditional structural function paradigm in molecular biology is that a protein must fold into a stable, three-dimensional native structure to perform its biological role, which has been significantly expanded by the recognition of intrinsically disordered proteins (IDPs) [1]. IDPs possess intrinsically disordered regions don't have a single stable structure under physiological conditions, instead they are dynamic ensembles of interconverting arrangements. This causes IDPs to structural changes caused by misfolding, aggregation, and clumping. This protein aggregation is a sign of various pathological conditions, including neurodegenerative diseases like Alzheimer's, Parkinson's, and Huntington's disease.

Healthy cellular function maintenance depends on proteostasis which is defined as critical balance between protein synthesis, correct folding, and the efficient clearance of damaged or misfolded proteins [2]. This balance is critically examined by the biological complexity of aging for IDPs. As we age several protective systems of the body begin to deteriorate. Thymic involution reduces immune regulation and clearance efficiency starting in early adulthood, proteasomal activity declines due to oxidative stress, and melatonin production decreases sharply after middle age. Moreover, genetic factor, like the fragility factor (F) in carrier states, can set a higher baseline risk for protein clumping, which makes individuals more volatile to these age-related declines.

The self-organization of proteins and polypeptides into biologically active secondary structures, post their biosynthesis, is mediated by several factors. This is a thermodynamic process replete with several metastable states until the equilibrium state with minimum free energy is achieved, which is the native state of the protein. For the minimization of the free energy hydrophobic forces play a major role which entails localization of nonpolar residues in the core of the molecular structure, and simultaneous placement of hydrophilic residues on the surface thereby balancing the hydrophobic-hydrophilic interactions with water. Furthermore, hydrogen bonding, van der Waals forces and screened Coulombic interactions play a key role in stabilizing the secondary structure. In addition, the environmental factors like pH, ionic strength, temperature and presence of other moieties contribute to this stabilization. However, any imbalance in the aforesaid molecular interactions may lead to misfolding, aggregation or clumping of proteins that causes neurodegenerative diseases like amyloidosis, Alzheimer's, Parkinson's and prion diseases. Cell stress induced by the misfolded insoluble fibrillar protein molecules are often responsible for neuronal death. A general correlation between protein aggregation dynamics and neurodegeneration needs to be established which is the objective of this work.

Current comprehension of neurodegeneration is predicated on fragmented, age-centric viewpoints that do not incorporate multifactorial nonlinear biological mechanisms [3],[4]. This study provides a simplistic systems level modelling that highlights the dynamic relationship between aging variables, hormone control, genetic predisposition, and the gut-brain axis. With the application of a constitutive differential equation for protein clumping kinetics, this research introduces the "Gut Factor" (G(t)) as a significant kinetic multiplier. This factor defines the condition of the intestinal barrier and microbial health, where gut health balance and the release of pro-inflammatory endotoxins like LPS (lipopolysaccharides) serve as major drivers of protein misfolding, and thus causing diseases. In addition to pointing out the key factors, this conceptual framework provides a solution to the issues associated with inherent abnormalities in proteins and polypeptides.

3.2 METHODOLOGY

3.2.1 The IDP Vulnerability Factor

The old protein models' assumptions revolve around the stability of 'lock and key'. However, this model focuses on the Disorder-Function perspective [5]. Considering IDPs state is largely dependent on the molecular physics of their surrounding environment, they appear as dynamic

structural arrays (or protein clouds). This fundamental shift in research is crucial because IDPs cannot achieve a single stable structure under physiological circumstances. They adopt many exchangeable states, instead. Organized proteins have amino acid composition with higher hydrophobic content in polar and charged residues than IDPs which further leads to their environmental sensitivity [6]. IDPs are resistant to near boiling temperatures and acidic pH conditions because of their residue composition whereas regular proteins denature under this condition [7].

3.2.1.1 Modelling Approach

i. Kinetic Framework

The major insight of this study is the formulation of a non-linear differential equation that goes beyond simple linear aging models [8]. We model the change in the concentration of clumped/aggregated protein (P_C) over time (t) using a systems level differential framework which serves as the constitutive equation given by

$$\frac{dP_C}{dt} = \left(\boldsymbol{k_{prod}} \cdot \boldsymbol{F} \cdot \boldsymbol{f(A,H)} \cdot \boldsymbol{G(t)} \cdot \boldsymbol{P_{total}}\right) - \left(\boldsymbol{k_{clear}} \cdot \boldsymbol{M(t)} \cdot \boldsymbol{G(t)^{-1}} \cdot \boldsymbol{P_C}\right) \quad (3.1)$$

where each parameter represents a biologically relevant factor (or lever) namely,

- P_C : Concentration of clumped protein which accumulates when formation exceeds clearance
- k_{prod} and k_{clear} : Rate constants for aggregated protein production and clearance. For aggregation to dominate in the system, $k_{prod} \gg k_{clear}$.
- F: Fragility Factor; represents genetic carrier state where $F > 1$ (e.g., $F \approx 1.5$), which indicates increased likelihood of misfolding and aggregation due to reduced functional protein reserve [9]. (Figure 1)

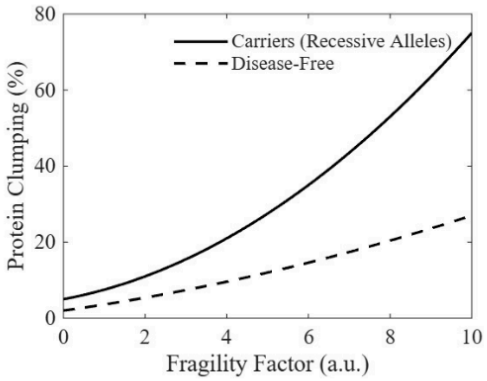


Figure 1: Study of Protein Clumping based on Fragility Factor.

- A: Age/Inflammation Factor; increases with age due to thymic involution and inflammaging processes.
- H**:** Hormonal/Gender Factor; it is modulated by estrogen, testosterone, and other hormones.
- M(t): Melatonin/Clearance Factor; decreases with age, representing declining efficiency of antioxidant defence and waste removal efficacy.
- G(t): Gut Factor; acts as a bidirectional catalyst that amplifies aggregation by 180–200% in the age range of 45–65 years.
- $P_{total} = P_c + P_{free}$: Total protein concentration, where $\boldsymbol{P_{free}}$ represents the sum of monomers and oligomers.

ii. Time-Dependent Parameter Specification

The model recognizes that the age-related factors are inherently time-dependent levers.

- A(t): Grows exponentially during early life (0–40 years), accelerating most significantly during young adulthood and middle age, then plateaus after age of 60 years as maximum values approach saturation. Hence the risk of IDPs being clumped due to thymic involution increases exponentially after 20 years as the thymus gland's

protective immunological and regulatory functions that peak during early development and begin to decline and ultimately vanish in early 20s [10]. (Figure 2)

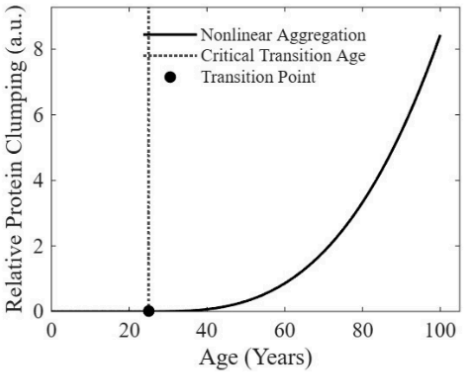


Figure 2: Protein Clumping Risk dependency on age.

- M(t): Demonstrates strong negative correlation with age; starting at peak at birth, melatonin levels consistently decline throughout the lifespan, reaching near zero at the age of 100 years due to lifestyle and stress. All of which in total are influenced by hormones as well [11]. (Figure 3)

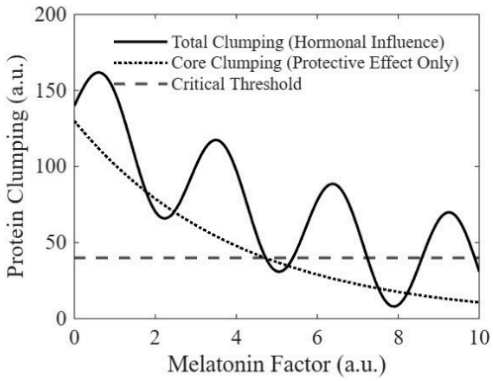


Figure 3: Protein Clumping Risk based on Melatonin Factor.

- $f(A,H)$ (The Proteostatic Buffer): Since this function depends on both age and hormones and both being time dependent; it is itself time dependent. This function

accounts for the simultaneous protection of estrogen and the thymus. In youth, this function acts as a kinetic suppressant, holding the aggregation flux at a negligible baseline and neutralizing the pro-aggregatory potential of IDPs, and with age it fluctuates, but equalizes for both gender around the age of 47 years [12],[13] (Figure4).

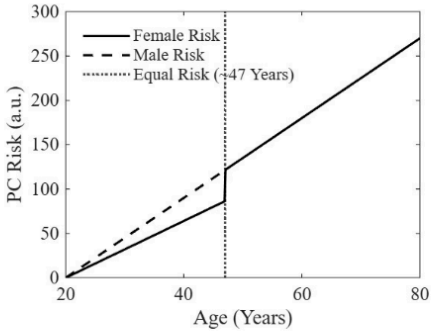


Figure 4: Protein Clumping Risk based on age, gender and hormone dependence.

- G(t) (The Gut-Multiplier): A high G(t) (representing dysbiosis and LPS leakage) lowers the entropic barrier for IDP aggregation while simultaneously poisoning the proteasomal clearance pathways.

iii. Equilibrium Analysis*:*

At equilibrium, when formation rate equals the clearance rate

$$(P_{total}\text{-}P_{free}) = \frac{k_{prod}}{k_{clear}} \times \frac{F}{M} \times f(A,H) \times G^2 \times \left(\mathbf{1} + \left(\frac{\boldsymbol{P_{free}}}{\boldsymbol{P_c}}\right)\right) \qquad (3.2)$$

This equilibrium equation reveals that the total clumping profile follows a sigmoid curve, showing sharp transitions around the age of 35 years that represent global state transitions between proteostatic phases (Figure 5).

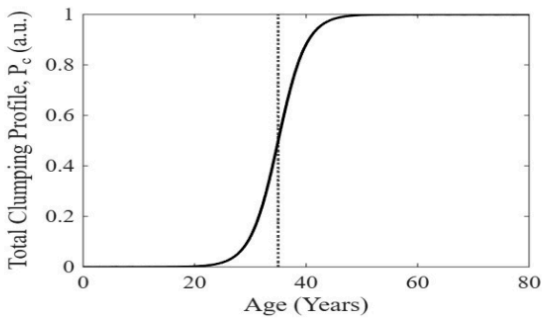


Figure 5: Equilibrium dynamics of the total clumping profile across the human lifespan derived from Eq. (3.2).

From equation (3.2) it is easy to be understood that that proteostatic clumping does not move in a simple straight line, instead it moves in a S-shaped path, showing a phase-dependent condition. During the first thirty years of life, the profile remains near-zero, that means a strong proteostatic balance which inhibits aggregation formation. However, the curve shows an important universal state transition at age of 35 years, at which point the system splits. This sudden alteration shows a proteostatic instability due to stress, which results in the fast nonlinear growth of protein aggregates. Following this the growth profile approaches a saturation level, which defines equilibrium at the specific age of 35 years that demarcates the boundary between accurate protein stabilization, and its instability zones.

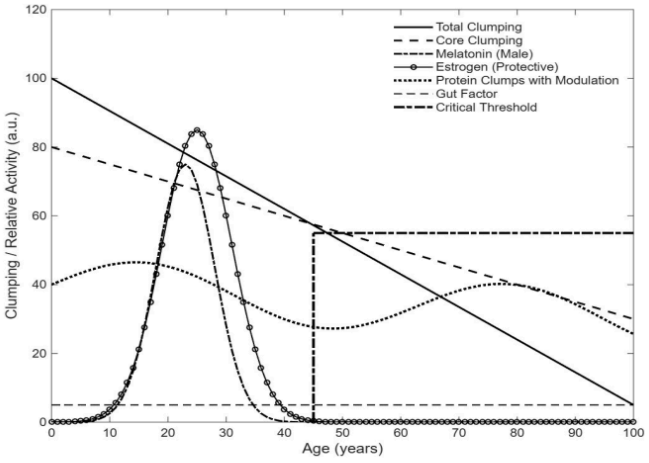


Figure 6: This figure shows kinetics of protein aggregation and self-assembly considering all the factors discussed in this work.

Figure 6 gives a consolidated summary of relative contributions of all biological factors to the total aggregation of proteins with respect to age. The overall clumping corresponds to a S-shaped sigmoid curve, commencing at almost 100% and declining weirdly around the age of 35 years, signifying a global state change from one phase to another. The low clumping perspective is governed by thymic level reaches its highest point at birth and its significance begins to deplete at around 25 years of age before progressively diminishing further. The estrogen hormone factor acts as a shield to inhibit protein clumping, and it exhibits a sinusoidal oscillatory variation with exponential decay, displaying hormonal cycles that stabilize and diminish post-menopause. The melatonin factor exhibits a pronounced Gaussian peak at the age of 22 years followed by a rapid decline at age of 40 years, signifying factors active solely during peak physical and reproductive maturity in both males and females. Interestingly, the estrogen protective nature shows a dual-phase behaviour with a Gaussian peak at 26 years and a piecewise shift at the age of 35 years, which infers the change towards the onset of menopause. An important step function escalating to 55% at the age of 45 years clearly establishes a pathological limit. The model uses a bifurcation analysis where age 45 acts as a phase transition. By setting a 55% activity threshold, the point where linear metabolic decline (1%/year) meets the exhaustion of hormonal buffers, we demonstrate that the age of 47 years represents the mathematical 'Point of No Return' where proteostatic failure becomes statistically inevitable. The gut factor G(t) is an important health factor rising between the age of 55-60 years to indicate gut imbalance and increased biological noise inside the human body. At last, protein clumping major dependency on G(t) gives an important conclusion, showcasing how gut health effects protein aggregation by forming a hump that decreases at the age of 70 years.

iv. Gut-Brain Axis as a Kinetic Multiplier

In the constitutive equation, the Gut Factor (G) represents the state of the intestinal barrier and microbial health through the Gut-Brain correlation [14]. It is a time-dependent variable G(t) that ranges from 1 (Healthy) to a maximum G_{max} (Severe Imbalance). It is given by

$$G(t) = \frac{\text{LPS}(t) * \text{Permeability}(t)}{\text{Beneficial Metabolites (SCFA)}} \qquad (3.3)$$

LPS is a pro-inflammatory toxin found in the outer membrane of "Gram-negative" bacteria. It is often referred to as an endotoxin. In a healthy gut, LPS stays trapped in the intestines and is excreted. It acts as the primary driver of formation (k_{prod}) of the following:

- Systemic Alarm: Once in the blood, LPS triggers a massive immune response.
- Neuroinflammation: LPS can cross a weakened and directly activate microglia, which creates an environment where proteins (like Alpha-synuclein) misfold and clump much faster.
- The Vicious Cycle: High LPS levels further damage the gut lining, leading to more LPS leakage.

On the other hand, SCFAs are the primary beneficial metabolites produced when good gut bacteria ferment dietary fibres [15].

The essence of the present work lies in the observation of aggregation amplification observed between age of 45-65 years. In the context of IDPs, we model the gut factor as a source of macromolecular crowding. Other salient features of the model are: (i) mechanism protocol that entails when the intestinal barrier fails, the systemic influx of inflammatory markers increases the density of the intracellular environment, and (ii) for the IDPs this increased density forces the disordered monomers to collapse into seeds of nucleation. This explains why GI symptoms precede clinical neurodegeneration by approximately 15 to 20 years with the gut provides the kinetic energy for the initiation of IDP miss-steps. Accumulation outpaces the clearance capacity, resulting in a rapid shift from a soluble monomeric state to an insoluble aggregated fibrillar phase. An intact intestinal barrier and a balanced gut microbiota, characterized by robust production of short-chain fatty acids, mitigate neuroinflammation and preserve brain functioning, thereby counteracting the aggregation cascade [16],[17].

v. Modelling Universal Tipping Points

The model reveals a critical phenomenon of the ‘risk equalization’ clearly defined as the following:

- During Phase 1 (Age 0-25 years), high thymic defence and melatonin levels helps to maintain a Low Clumping conditions inside the human body.

- During Phase 2 (25-45 years), gender-specific division occurs. Females are benefitted from estrogen whereas males have a sharper 85% drop in melatonin, leading to faster threshold crossings.
- Around the age of 47 years, the model predicts a universal proteostatic breakdown when protective layers start to degrade, specifically estrogen levels and gut stability decreases (Phase 3). This change is caused by the excluded volume effects of macromolecular overcrowding, which limits the space available for normal protein folding, and as a result, abnormal intermolecular interactions ensue leading to uncontrolled aggregation [18]. As aggregation kinetics speed up, the system reaches at an intersection point. At this point, the rate of aggregate formation exceeds the rate of clearance, and the system transits from a stable state into a self-harming condition of instability. This results in a positive feedback loop where protein clumps trigger microglial cells and neuroinflammation, which in turn promotes protein aggregation, that in turn replicates the conditions seen in neurodegenerative disorders [19].

3.2.1.2 IDP-Specific Characterization Methodology:

i. Experimental Characterization Framework:

The kinetic modelling was validated through experimental characterization of IDPs which used standard biophysical techniques that takes advantages on the distinct conformational behaviour of IDPs compared to globular proteins. These techniques include circular dichroism spectroscopy to detect the characteristic lack of secondary structure, nuclear magnetic resonance for residue-specific conformational dynamics, and fluorescence-based methods to monitor hydrodynamic radius and compaction states.

3.3 Environmental Sensitivity Assessment

Important aspect of this model is understanding IDP response to environmental disturbances:

- pH Effects: Neutralizing acidic groups at lower pH which helps to reduces the net charge on IDPs (or regions), leading to the increased solubility and a more compact structural state. This contrasts with the globular proteins, in which it is found that acidic conditions cause protonation of negatively charged side chains, that leads to charge imbalances and aggregate formation [20].

- Temperature Effects: While high-temperature conditions reveal the hydrophobic core of ordered proteins that leads to aggregation, IDPs shows resistance to increased

temperatures due to their lower hydrophobic residue content. Instead, increased temperature often enhances the conformational dynamics of IDPs, potentially aiding in the sampling of various structural states, including those that might facilitate functional interactions [21].

3.4 FUNCTIONAL CHARACTERIZATION INTEGRATION

The kinetic model incorporates IDP functional categories relevant to aggregation and disease:

- Molecular Assemblers: IDPs with high disorder content interact with multiple binding partners to encourage formation of higher-order complexes. The disorder content increases with the size of the protein complexes, and upon binding to partners, disordered assemblers maintain their open structure, allowing multiple proteins to bind to a single IDR. This multivalent binding capacity is structurally facilitated by the extended conformational behaviour of IDPs, which possess hydrodynamic volumes significantly larger than globular proteins of equivalent mass, thereby maximizing the accessible surface area for intermolecular interactions [22]. This inherent flexibility and lack of fixed three-dimensional structure enable IDPs to participate in diverse binding mechanisms, including coupled folding and binding, as well as highly dynamic and multivalent interactions [23].
- Scaffolding Function: IDPs can regulate spatiotemporal assembly of signalling partners by acting as scaffolds, which includes formation of biomolecular condensates. Scaffolding regions contain the highest degree of disorder among functional categories [24].

This functional characterization aligns with the D2 (Disorders in Disorders) concept, as IDPs are associated with neurodegenerative diseases including Alzheimer's disease (deposition of α-synuclein, tau, and amyloid-β proteins) and Parkinson's disease (α-synuclein accumulation) where the accumulation of aggregates containing intrinsically disordered proteins serves as a defining pathological feature [25].

3.5 RESULTS AND DISCUSSION

This study of the kinetics of self-aggregation proposed in this work reveals a complex, non-linear environment that goes beyond traditional models of simply aging framing

neurodegeneration as a systemic collapse of proteostatic balance. It becomes clear that the tilt to a pathological state is given by opposed pleiotropy, by analysing the concentration of aggregated protein (P_c) through an algorithm that explains for the tension between production and clearance. Here, specific alleles that offer reproductive benefits early in life through stress-response pathway mechanisms contribute to the buildup of detrimental products in post-reproductive stages

The kinetic trajectory displays a sigmoid profile with a critical turning point around age 35, where the statistical bottleneck for IDP clustering is lowered while at the same time compromising proteasomal clearance pathway. This transition is accelerated by the age-dependent decrease of the Melatonin/Clearance Factor (M), which shows a sharp decrease from birth, reaching near-zero by age 100. This decline significantly reduces the usefulness of antioxidant defences as well as the lymphatic elimination systems required for the clearance of disorganized oligomers.

An important component of this research is the finding of the Gut-Brain Axis as an essential kinetic multiplier (G(t)), especially throughout the midlife window of 45–65 years. During this phase, poor intestinal barrier integrity permit for the systemic import of inflammatory signals, which results in macromolecular saturation. This biophysical activity limits the excluded volume becomes available for natural protein folding, causing disordered monomers to collapse and start nucleation, thus increasing the aggregation [26].

This increase in levels produces the proteostatic tipping point at about 47 years old, which is the equilibrium stage at which gender-specific risks fade away. Although the Hormonal Factor (H) is regulated by estrogen and testosterone, which acts as a good kinetic suppressant in youth, the midlife drop in these endocrine buffers, together with a 85% decline in melatonin levels by age 50, which drives the system toward a 55% clumping threshold. At this splitting point, the pace of aggregate accumulation surpasses the rate of cellular clearance, resulting in a positive feedback loop that stimulates microglial cells and causes inflammation in the brain. The subsequent uncontrolled aggregation weakens IDPs' critical fundamental roles, which are required for the creation of higher-order complexes and spatiotemporal regulation of biomolecular condensates. By analysing these physiological shifts, the model creates a 15-20 year preclinical window where gastrointestinal dysbiosis and endocrine volatility act as early indicators of proteostatic failure, allowing a significant lead time for therapeutic intervention before the onset of irreversible motor or memory. The salient features of various relevant

models are presented in Table-1. Data for previous models are taken from the literature cited in this work.

Table-3: Comparison of theoretical frameworks

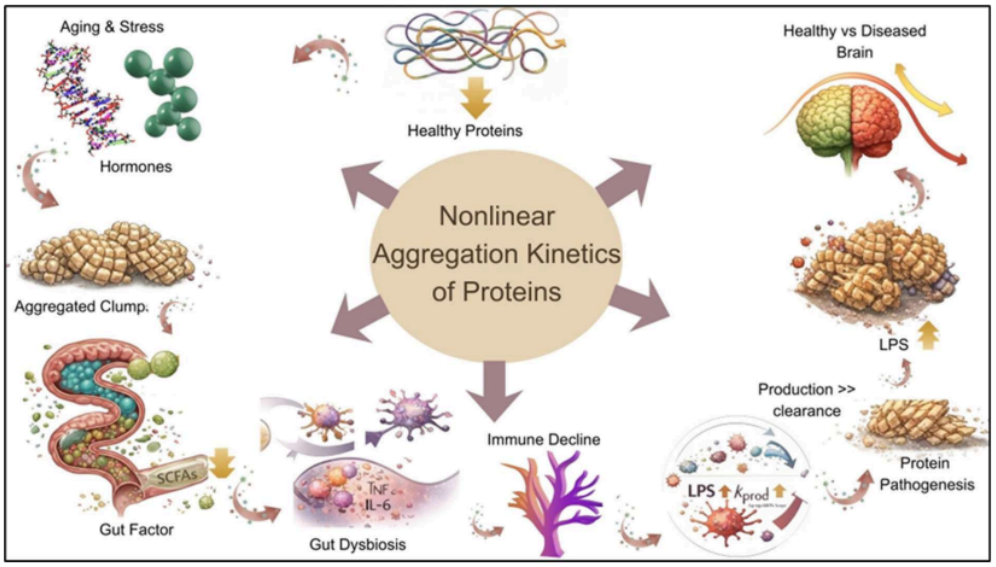
Aspect	Previous Models	This Model
Primary Driver	Genetic mutations, linear aging	Midlife gut imbalance G(t): 180-200% aggregation amplification (ages 45-65 years)
Aggregation Kinetics	Simple: $\frac{dP}{dt} = (k_{prod} - k_{clear})$	$\frac{dP_C}{dt} = \left(k_{prod} \cdot F \cdot f(A,H) \cdot G(t) \cdot P_{total}\right) - (k_{clear} \cdot M(t) \cdot G(t)^{-1} \cdot P_C)$ **:** Gut as kinetic multiplier
Gender Difference	Males > Females (genetic)	Convergence at the age of 47 years: Estrogen loss + universal dysbiosis equalizes risk
Critical Threshold	None defined	55% clumping at age of 45 years: Universal tipping point regardless of early peaks
Thymus Role	Early decline only	Low clumping window (0-25years): Thymic protection delays aggregation onset
Melatonin Effect	General antioxidant	Rapid male decline: Melatonin factor crashes 85% by age of 50 years, first threshold breach
Therapeutic Target	Clearance enhancement	Gut restoration primary: SCFAs reduce G(t) from 2.8→1.0 breaks vicious cycle
Preclinical Window	5-10 years	15-20 years: GI symptoms precede motor PD, matching G(t) peak at age of 55 years
Feedback Loops	None modelled	LPS vicious cycle: Gut→brain→gut amplification via microglial activation

Intervention Timing	Symptomatic	Three clear windows: Early (20-35 years, thymus), Peak (45-65 years, gut), Late (>70 years, clearance)

Evidence of important nonlinear changes in the molecular structure of proteins is rapidly challenging the view of human aging as a continuous linear decrease. While recent research[3] has identified a generic molecular change around age of 44 years, in this work a more detailed and mechanistically integrated view of the tippling point near the age of 47 years is presented. Unlike other omics-based observation that focus on various metabolic markers, our model identifies the convergence of gender-specific protections and the initiation of widespread proteostatic decline. It also recognises how the gut imbalance acts as a prominent driver of protein clumping specifically between the ages of 45 and 65 years, that has not been adequately investigated in the general multi-omics literature. Earlier studies[8] have identified generic turning points for neurodegeneration at the age of 40 years, while our research gives a predicted mathematical framework that explains risk equalization at the age of 47 years governed by the vanishing of the estrogen protective and thymic regulation factors simultaneously.

Finally, identifying these tipping points moves the treatment focus away from late-stage symptom treatment and towards proactive, stage-specific therapies. Recognition of age 47 years as a universal convergence point enables targeted midlife gut regeneration and early immunological assistance, which has the potential to prevent neurodegeneration decades before symptoms appear. This systems-level approach bridges the gap between solitary metabolic measurements and the complex relationship of factors that influence human health.

A schematic depiction of various lever parameters influencing protein aggregation kinetics, and their biophysical consequences are given in Scheme 1.



Scheme 1: Progressive transition of proteins into pathological aggregates driven by the nonlinear interplay of biological stress that act as levers.

4. CONCLUSION

With the help of transfer of the attention from late-stage protein accumulation to midlife intestinal imbalance as the primary cause of disease pathway, the outcomes of this study essentially reshape the development of neurodegeneration. In this simplistic, and yet predictive

model, it is shown that the path to neurodegeneration is not a linear progression of aging, but it lies between the ages of 45 and 65 years, there is a systemic kinetic surge of protein aggregation. The age of 35 years marks an important transition point where proteostatic stability collapses, initiating the transition into a permanent state of protein instability and aggregation. Next to this shift, the estimation of the Gut-Brain Kinetic multiplier is critical to this perspective shift, with a surge from 1.0 to 2.8 serving as the primary catalyst for breaking the 55% clumping threshold which identifies age-adjusted imbalance as the major pathogenic factor, despite early-life genetic or gender-specific characteristics. While males have an earlier melatonin driven vulnerability and females have a post-menopausal, these diverse trajectories converge due to universal midlife microbiome alterations, this model clarifies the merging of pathological risk at age of 47 years.

Herein lies an important roadmap in preventative medicine by the creation of a preclinical window of 15-20 years, during time gastrointestinal symptoms occur, followed by the peak in G(t). Strategies aiming at disrupting the LPS-mediated vicious cycle by restoring the gut and consuming Short-Chain Fatty Acids (SCFAs), such as butyrate therapy, present a theoretical means of resetting the G(t) multiplier back to 1.0. Finally, these findings suggest that proactive midlife gut stabilization, rather than downstream aggregate clearance, is the most effective strategy for preventing neurodegenerative onset, allowing for a shift from reactive symptomatic management to true disease-modifying interventions.

Based on this model's three unique biological windows, future research and clinical trials should move away from a one-size-fits-all approach to neuroprotection. Instead, we propose a tiered method that focuses on the dominating kinetic driver at each life stage. Table-2 summarizes the key results of our work.

Table-4: Key ages when major biological changes occur in human body.

Phase	Focus years	Biological Driver	Strategy
Primordial	20–35	Thymus & Sex Hormones	Immunological/Endocrine buffering

Preventative	45–65	Gut Imbalance (G(t))	LPS reduction & SCFA restoration
Stabilizing	70+	Impaired Clearance	Clearance enhancement & G(t) recovery

We can shift the clinical targets from managing the symptoms to maintaining the integrity of a healthy gut-brain axis, by adopting this chronobiological framework. Developing a general understanding of the neurodegenerative diseases by moving beyond the isolated age-based paradigms is one of the remaining challenges of molecular biophysics. The current work was a step in that direction.

AMISHI_2.docx

ORIGINALITY REPORT

5%	4%	3%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.ncbi.nlm.nih.gov Internet Source	2%
2	www.coursehero.com Internet Source	<1%
3	patentimages.storage.googleapis.com Internet Source	<1%
4	J. M. Goldwasser, A. Rudin. "Analysis of Block and Statistical Copolymers by Gel Permeation Chromatography: Estimation of Mark-Houwink Constants", Journal of Liquid Chromatography, 2006 Publication	<1%
5	alevelchemistry.co.uk Internet Source	<1%
6	learningspot.co Internet Source	<1%
7	www.worldpharmanews.com Internet Source	<1%
8	synapse.koreamed.org	

	Internet Source	<1%
9	Yinlin Zhu, Yuheng Wang, Liang Zhao. "Research on the ensemble forecasting method for green tide paths in the Yellow Sea based on parameter perturbation", Marine Pollution Bulletin, 2025 Publication	<1%
10	J.M. Russell, R.N. Cooper. "Flocculation of slaughterhouse effluents with aluminium salts", Environmental Technology Letters, 2008 Publication	<1%
11	Submitted to La Trobe University Student Paper	<1%
12	www.biorxiv.org Internet Source	<1%
13	Underhill, P.T.. "On the coarse-graining of polymers into bead-spring chains", Journal of Non-Newtonian Fluid Mechanics, 20040920 Publication	<1%
14	api-depositonce.tu-berlin.de Internet Source	<1%
15	Submitted to Dudley College Student Paper	<1%
	Submitted to HRUC	

16	Student Paper	<1%
17	Submitted to Macquarie University Student Paper	<1%

Exclude quotes	On	Exclude matches	< 8 words
Exclude bibliography	On		

AMISHI_2.docx

GRADEMARK REPORT

FINAL GRADE

/0

GENERAL COMMENTS

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9

PAGE 10

PAGE 11

PAGE 12

PAGE 13

PAGE 14

PAGE 15

PAGE 16

PAGE 17

PAGE 18

PAGE 19

PAGE 20

PAGE 21

PAGE 22

PAGE 23

PAGE 24

PAGE 25

PAGE 26

PAGE 27

PAGE 28

PAGE 29

PAGE 30

PAGE 31

PAGE 32

PAGE 33