Why We Stop Synthesizing Essential Amino Acids: The Extracellular Protein Hypothesis

Genshiro Esumi[1]

1 University of Occupational and Environmental Health, Japan

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.

Abstract

Humans cannot synthesize nine of the twenty amino acids that constitute proteins, known as essential amino acids. It has been traditionally considered that this inability arose because humans could obtain these amino acids in sufficient quantities through their diet. However, recent advances in life sciences have shown that all eukaryotic organisms with the ability to ingest external protein resources have uniformly lost the ability to synthesize almost identical amino acids, including those belonging to branches of the evolutionary tree entirely different from humans, such as *Dictyostelium* and *Tetrahymena*. Yet, the reasons behind their essentiality and the commonality of these essential amino acids remain elusive and unexplained. In this paper, I propose a novel and simple explanation that **organisms can maintain their amino acid balance by solely synthesizing amino acids that are more abundant in extracellular proteins compared to intracellular proteins.** This explanation is based on two previously unrecognized assumptions. The first assumption is that intracellular proteins act as amino acid buffers for subsequent protein synthesis, facilitated by the continuous recycling of their amino acids during the degradation and synthesis cycle. The second assumption is that there are consistent differences in amino acid composition between extracellular and intracellular proteins, economically driven by the lower synthesis costs for extracellular structures. Despite the limited data available for examining these assumptions, the evidence lends support to their validity. Therefore, this "Extracellular Protein

Hypothesis" provides a novel and convincing explanation to the nearly century-old mystery: the origin of essential amino acids.

Keywords: essential amino acids, non-essential amino acids, extracellular protein synthesis, evolutionary biology.

Introduction

Humans are unable to synthesize nine out of the twenty amino acids that constitute proteins; these are known as essential amino acids [1][2][3][4][5][6]. The historical and simplest explanation for this phenomenon is that humans did not need to synthesize these amino acids due to their abundance in their diet [3][4][5]. However, subsequent observations and research have revealed that the loss of synthesis capabilities for these amino acids is not unique to humans; it is also present in other animals. This suggests an origin at the level of a common ancestor shared by humans and these animals, with this trait being inherited by their descendants [2][3][4][5][6]. More recent advancements in life sciences, however, have shown that similar losses of amino acid synthesis capabilities have independently occurred across multiple branches of the eukaryotic evolutionary tree [3][4][5][6]. This indicates that the loss of amino acid synthesis ability occurred multiple times throughout eukaryotic evolution, consistently involving similar amino acids each time [3][4][5][6]. To date, the underlying reasons for this enigmatic pattern of individual and independent losses of similar amino acid synthesis capabilities in different eukaryotic lineages remain elusive and still unexplained. This paper aims to explore these questions by proposing a novel hypothesis that seeks to unravel the complexities behind these observations.

Current Understanding

History and Definition of Essential Amino Acids in Nutrition

The exploration of amino acid nutrition began around the mid-19th century, reaching a significant milestone in 1935 when Rose and his team identified threonine as the last of the 20 amino acids that make up proteins [7]. In their pioneering experiments with rats, they were the first to demonstrate that weight gain could be achieved with a diet consisting exclusively of amino acids, rather than proteins, as the nitrogen source. This groundbreaking observation established the foundation for the field of practical amino acid nutrition [7]. Further research by Rose on humans showed that deficiencies in specific amino acids led to the breakdown of body proteins and disrupted nitrogen balance. Conversely, the absence of other amino acids did not produce such effects, thus maintaining nitrogen balance [1]. Subsequently, amino acids were divided into essential, necessary for body protein maintenance, and non-essential, which do not impact this critical balance.

Hidden Complexities of Essential Amino Acid Evolution

Later genomic analyses have revealed that in humans and some animals, mutations have inactivated several enzymes responsible for synthesizing essential amino acids [3][4][5][6]. These genetic discoveries confirm humans' inherent inability to produce these amino acids internally, and it is speculated that the loss of these amino acid synthesis capabilities occurred at the stage of a common ancestor, with descendants inheriting this trait. However, challenging previous assumptions, subsequent genomic analyses have unveiled unexpected revelations. Research indicates that not only humans but also a wide array of metazoans and diverse eukaryotic organisms from various phylogenetic branches, including cellular slime molds (*Dictyostelium*; Amebozoa) and *Tetrahymena* (a protozoan), have similarly lost the ability to synthesize almost identical sets of amino acids (Table 1) [3][4][5][6]. Given the evolutionary divergence of these organisms and humans from common ancestors before the divergence from plants, which can synthesize all required amino acids, these observations suggest independent losses of common amino acid synthesis capabilities across various evolutionary lineages. Moreover, as far as we can observe, all current organisms, without exception, seem to have concurrently lost the ability to synthesize common essential amino acids upon acquiring each feeding capability [3]. This concept has not been proven but is considered empirically correct. Given the complexity of this phenomenon, it is no surprise that this widespread and striking commonality, observed independently across lineages, remains a significant mystery even in recent literature [6].

Common name	Scientific name	Essential Amino Acids
Human	*Homo sapiens*	Met, Phe, Lys, His, Trp, Ile, Leu, Val, Thr
Green spotted puffer	*Tetraodon nigroviridis*	Met, Phe, Lys, His, Trp, Ile, Leu, Val, Thr
Vase tunicate	*Ciona intestinalis*	Met, Phe, Lys, His, Trp, Ile, Leu, Val, Thr, Arg
Fruit fly	*Drosophila melanogaster*	Met, Phe, Lys, His, Trp, Ile, Leu, Val, Thr, Arg
African malaria mosquito	*Anopheles gambiae*	Met, Phe, Lys, His, Trp, Ile, Leu, Val, Thr, Arg
Nematode worm	*Caenorhabditis elegans*	Met, Phe, Lys, His, Trp, Ile, Leu, Val, Thr, Arg
Cellular Slime Mold	*Dictyostelium discoideum*	Met, Phe, Lys, His, Trp, Ile, Leu, Val, Thr, Arg, Ser
Tetrahymena	*Tetrahymena thermophila*	Met, Phe, Lys, His, Trp, Ile, Leu, Val, Thr, Arg

Table 1. Essential Amino Acids in Representative Organisms. Table 1 compiles the essential amino acids for selected organisms from various evolutionary backgrounds, as documented in the referenced studies [3][4][5][6][8]. Despite their phylogenetic differences, the species listed exhibit remarkably similar profiles of essential amino acids. Notably, the organisms at the bottom of the table, Cellular Slime Mold and Tetrahymena, belong to completely distinct evolutionary lineages compared to the others listed.

Do Dietary Sources Determine the Essentiality of Amino Acids?

The question of whether an organism's diet dictates the essential amino acids is one of the initial and simplest inquiries when considering the factors that define essential amino acids. This primarily stems from the fact that autotrophic organisms, such as plants and fungi, which lack the capability to ingest, do not require amino acids [2][3][4][5][6][9], whereas eukaryotic organisms that have gained the ability to ingest food uniformly demonstrate a common set of essential amino acids [3]. This suggests a potential simple correlation between the acquisition of feeding capabilities and the consequential loss of amino acid synthesis abilities. However, the dietary sources (food resources) that organisms consume vary significantly by species, habits, and environmental contexts, inherently introducing a diversity (variability) in amino acid composition. Taking human clinical nutrition as an example, it underscores the importance of dietary choices in daily life and as a fundamental concept in nutrition science, emphasizing the complexity in dietary amino acid sources. Given these factors, it is highly unlikely that a universally stable and consistent amino acid composition exists across the vast diversity of organisms, sufficient to cause a uniform and universal loss of the ability to synthesize nearly half of the 20 amino acids. Therefore, it is considered impractical to define the boundary between essential and non-essential amino acids solely based on the diet source of organisms.

Do the Characteristics of Each Amino Acid Determine Its Essentiality?

The question of whether the characteristics of each amino acid determine its essentiality is a natural inquiry to follow the consideration of dietary sources. The 20 amino acids that constitute proteins each have unique characteristics, and these underpin the diversity of biological proteins. On the other hand, these amino acids are composed of elements that are relatively common in the body. The synthesis of amino acids takes place using metabolic products within the body as basic materials, but this synthesis requires energy. Akashi and Gojobori's paper, which estimates this synthesis cost in units of high-energy phosphate bonds [10], demonstrates a disparity of more than sixfold between the simplest amino acids, glycine and alanine, and the most complex, tryptophan. Generally, amino acids that are higher in cost tend to be larger in size, have greater hydrophobicity, and involve more steps and enzymes in their synthesis. It has long been observed that essential amino acids are generally high-cost, whereas non-essential amino acids are low-cost (Figure 1) [10][11][12]. Thus, while the boundary does not align perfectly with the disparity in synthesis costs of each amino acid, the hypothesis that the synthesis cost defines the boundary between essential and non-essential amino acids seems to maintain a certain level of validity. The validity will be examined in the following subsection.

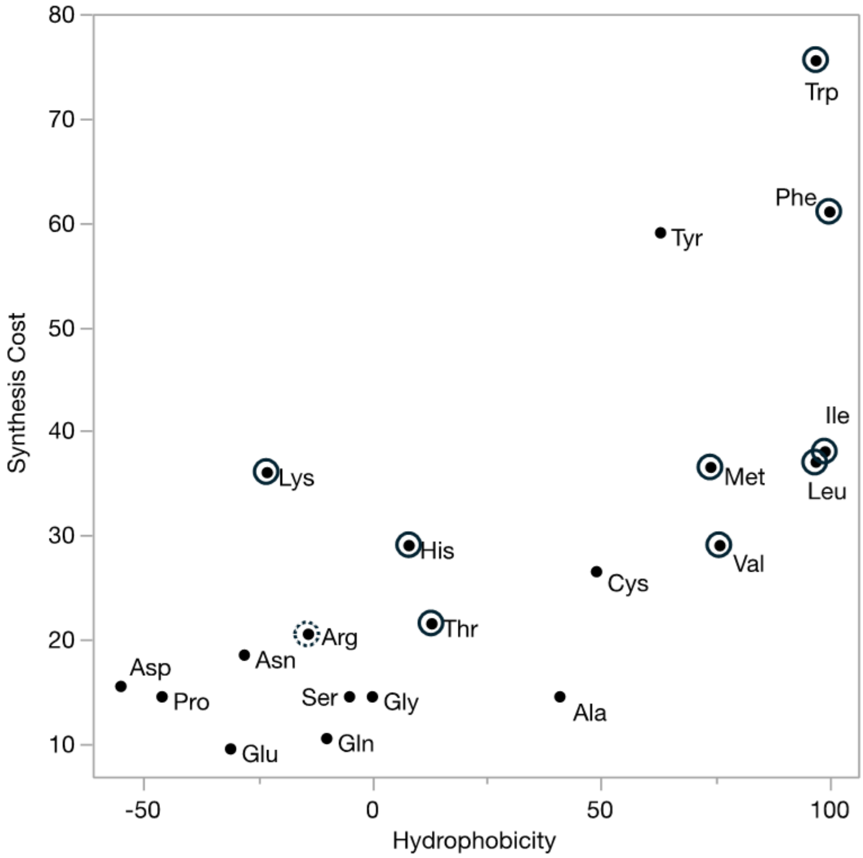

Figure 1. Correlation between Amino Acid Synthesis Cost, Hydrophobicity, and Essentiality. Figure 1 depicts the relationship between the synthesis cost and hydrophobicity of amino acids, along with their classification as essential or non-essential. The vertical scale represents the amino acid synthesis cost, measured in units of high-energy phosphate bonds [10], whereas the degree of hydrophobicity for each amino acid is quantified on the horizontal axis [11][12]. Essential amino acids are denoted with ringed plots. Arginine, however, is marked with a dashed ring to reflect its status as essential in most organisms but not in humans. The plot reveals a moderate correlation between synthesis cost and hydrophobicity. There is also a related trend concerning amino acid essentiality. However, this boundary between essential and non-essential amino acids does not perfectly correlate with these factors.

Note: The hydrophobicity value for proline, not available in the primary literature [11], was sourced from an alternate study [12].

What Makes Essential Amino Acids Essential?: Initial Insights

What fundamentally renders amino acids essential? Throughout life's history, evolutionary changes are widely recognized to be driven by random genetic mutations that lead to phenotypic diversity. Within this diversity, phenotypes that fail to adapt are subjected to natural selection, facilitating evolution through the survival and reproduction of adaptable phenotypes. Reflecting on animal evolution, it is notable that the inability of animals to synthesize essential amino acids has not led to selective disadvantages, despite being a deficiency phenotype. This suggests that organisms can survive and reproduce without synthesizing these essential amino acids. However, these organisms cannot tolerate the loss of the ability to synthesize amino acids termed 'non-essential.' In humans, the loss of amino acid degradation capabilities leads to recognized congenital metabolic disorders, but the failure to synthesize non-essential amino acids is not categorized as a disease, indicating that such a loss prevents viable development. The primary driver of natural selection against individuals who have lost amino acid synthesis capabilities would be the deficiency symptoms resulting from the absence of those amino acids. Therefore, the distinction between essential and non-essential amino acids should be based more on their use within the organism rather than on factors such as synthesis costs or the number of enzymes involved in their synthesis. We can conclude that the demarcation between essential and non-essential amino acids is not determined solely by the individual properties of each amino acid but significantly by the characteristic ways in which organisms utilize these amino acids. The next section will examine how organisms employ essential and non-essential amino acids distinctly, reflecting their unique roles and the implications for amino acid synthesis capabilities.

Hypothesis Development

This section explores the question: Why do all eukaryotic organisms that have acquired the ability to ingest consistently lose similar amino acid synthesis capabilities?

Principal Component Analysis of Food Composition Table

Prior to the current study, I conducted a statistical analysis using Principal Component Analysis (PCA) on the "STANDARD TABLES OF FOOD COMPOSITION IN JAPAN" published by the Ministry of Education, Culture, Sports, Science, and Technology in Japan [13]. Unexpectedly, I found that the eigenvector of the first principal component aligned with the boundary between essential and non-essential amino acids (Figure 2b) [14]. Foods are broadly classified into animal and plant groups, known to have significant differences in amino acid composition. However, the first principal component did not distinguish between animal and plant foods, and similar eigenvectors were also observed in the subgroup analyses of both animal and plant foods (Figures 2a, 2c, and 2d) [14]. PCA, a statistical method for extracting trends from high-dimensional data in order of their statistical significance, suggested that the alignment of the first principal component with the boundary between essential and non-essential amino acids was not coincidental but indicative of an underlying, yet unknown, correlation.

Unknown Correlation Between Food Compositions and Essential Amino Acids

Why then did they align? Foods and their ingredients are essentially parts of the body of eukaryotic organisms. In the analysis of the first principal component within animal foods, meats and gelatins were positioned at each extreme (Figure 2a). Meats are largely composed of intracellular proteins, while gelatins are identical to collagen and represent extracellular matrix proteins. The fact that the first principal component divides the amino acid composition of animal foods into meats and gelatins suggests a disparity in amino acid composition between intracellular and extracellular compartments. This observation led to the concept that biological body parts are composed of two types of proteins: intracellular proteins, which are relatively rich in essential amino acids, and extracellular proteins, which are comparatively rich in non-essential amino acids.

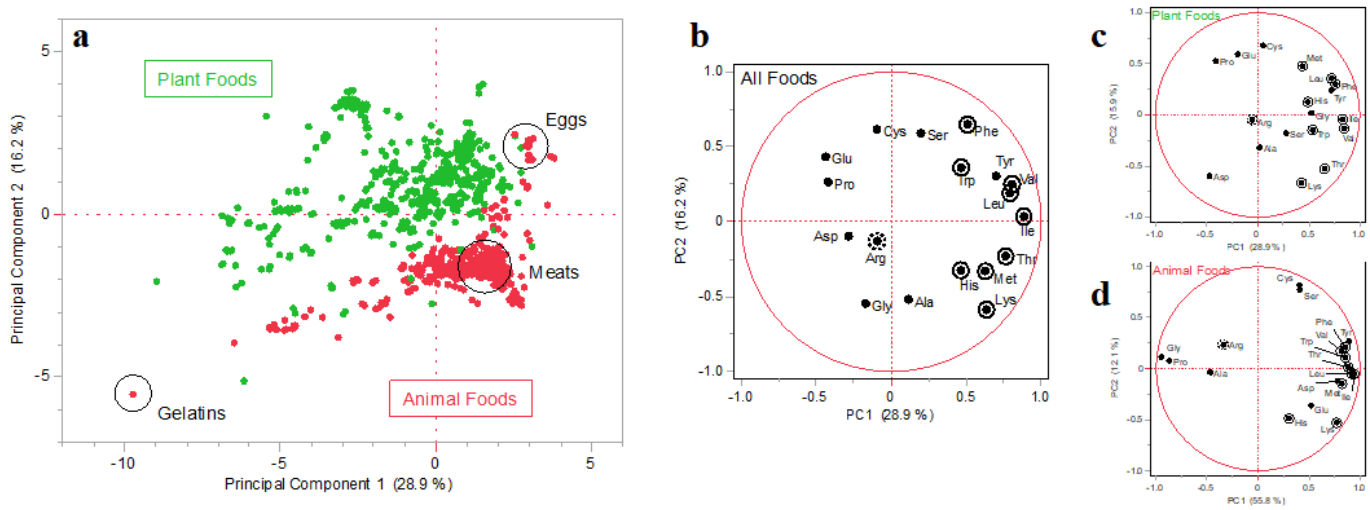

Figure 2. PCA Plots of Food Amino Acid Compositions and Their Eigenvectors

Figure 2a: Principal component analysis (PCA) plot of the amino acid composition of food items (n=1558) from a food composition table [13][14]. The horizontal axis represents the first principal component, and the vertical axis represents the second principal component, with animal foods plotted in red and plant foods in green. The general areas for Meats, Eggs, and Gelatins are demarcated.
Figure 2b: Eigenvectors from the PCA of the amino acid composition of food items (n=1558) are displayed [13][14]. The horizontal axis corresponds to the first principal component (PC1), and the vertical axis to the second principal component (PC2). The direction and length of each amino acid's eigenvector are plotted, and essential amino acids are marked with rings for distinction. Arginine is marked with a dashed ring to denote its conditional essentiality in most organisms. Essential amino acids tend to cluster towards the positive end of PC1. Notably, tyrosine, while not an essential amino acid, is also located in proximity to this cluster of essential amino acids, yet it is not marked differently to reflect its non-essential status.
Figure 2c and 2d: Eigenvectors for plant and animal foods (n=657 and n=569, respectively) from the food composition table are presented [13][14], following the format of Figure 2b. In both plots, essential amino acids are oriented towards the positive direction of the first principal component, indicating their commonality in the dataset. Tyrosine is included within this essential amino acid group without special marking, reflecting its position in the dataset.

Note: Other food items, including processed foods, are not displayed in these figures.

Can the Amino Acid Composition Disparity Between Intracellular and Extracellular Explain Essential Amino Acids?: A Hypothesis

If the difference in amino acid composition between intracellular and extracellular compartments corresponds to the boundary between essential and non-essential amino acids, theoretically, it could either represent a cause, a consequence, or simply a coincidental alignment with the distinction between essential and non-essential amino acids. In this context, my speculation leads me to conclude that this difference acts as a cause and serves as a background factor in defining the boundary, as detailed below. Considering the continuous cycle of protein synthesis and degradation that occurs within cells—the fundamental units of life—it is reasonable to assume that the primary source of amino acids for subsequent protein synthesis is derived from the degradation of intracellular proteins [15]. Therefore, intracellular proteins would essentially serve as reservoirs, acting as buffers for the amino acid supply during subsequent protein synthesis. Under such conditions, if extracellular proteins consistently exhibit distinct amino acid compositions, reliance primarily on the degradation of intracellular proteins for amino acid resources could lead to a deficiency in certain amino acids during their synthesis. Consequently, a consistent disparity in amino acid composition between intracellular and extracellular compartments could be instrumental in delineating essential from non-essential amino acids and might be the cause of their separation.

Two Essential Assumptions for the Hypothesis

Based on these observations and extrapolations, I postulate two conditions for the hypothesis: first, that intracellular proteins act as an amino acid buffer for protein synthesis; and second, that a consistent set of amino acids is used more frequently outside the cell than within cellular proteins. Under these assumptions, I observe a dichotomy in the need for amino acid synthesis based on the difference in amino acid composition between the intracellular and extracellular compartments. This is what I have termed the "Extracellular Protein Hypothesis."

In this section, I have explained the development and rationale behind the Extracellular Protein Hypothesis, which proposes an explanation for the origin of essential amino acids by focusing on the potential disparities in amino acid composition between intracellular and extracellular compartments. In the next section, we will examine the two underlying assumptions and assess the validity of the hypothesis itself.

Hypothesis Validation

This section examines the Extracellular Protein Hypothesis, which posits that the disparity in amino acid composition between intracellular and extracellular proteins correlates with the necessity to maintain amino acid synthesis capabilities. To evaluate the validity of this hypothesis, it is necessary to investigate the typical amino acid compositions of both intracellular and extracellular proteins. This section discusses four main aspects: the amino acid composition of intracellular proteins, the amino acid composition of extracellular proteins, the recycling of intracellular amino acids, and the amino acids required for the synthesis of extracellular proteins.

Intracellular Protein Amino Acid Composition

The amino acid compositions of intracellular and extracellular proteins are determined by the nucleotide sequences of genes within the organism's genome. While some genes are responsible for synthesizing extracellular proteins, the majority encode intracellular proteins. Analysis of amino acid residues in the proteomes of various organisms, representing the complete list of proteins encoded by an organism's genome, shows that amino acid distributions typically follow bell-shaped, single-peaked normal distributions, also similar to binomial distributions [15][16]. I speculated that this pattern suggests that the distributions may be constrained by the composition of the organism's intracellular protein degradation products [15]. Conversely, the actual amino acid (residue) composition of intracellular contents would be inevitably constrained by the amino acid composition of protein genes within the proteome. Given the supposed mutual constraints between proteome genes and cellular amino acid compositions, it naturally follows that the proteome's composition induces convergence and leads them within a narrow range, which, I hypothesized, might account for the bell-shaped distributions observed [15]. Moreover, the universal genetic code shared by all organisms, along with their genome's adherence to Chargaff's second parity rule [17][18], is hypothesized to impose additional constraints on the proteome’s composition. As a result, these constraints likely ensure that the composition of intracellular proteins' amino acids remains within a certain range across different organisms, and consequently, cells universally maintain a specific level of essential amino acids within themselves.

Extracellular Protein Amino Acid Composition

Extracellular proteins are synthesized inside the cell and then localized outside the cell membrane or secreted from the cell. While all proteins can serve as valuable amino acid resources when broken down, these proteins will not easily be recycled back into the cell or repurposed as resources for new proteins as intracellular proteins. Therefore, it is quite plausible that extracellular proteins are composed of amino acids that are less costly to synthesize. In fact, analyses of protein genes in various bacterial species have shown that extracellular proteins uniformly utilize amino acids with lower synthesis costs [19]. Similarly, in humans, major components of the extracellular matrix, such as collagen, elastin, and keratin-related proteins that constitute body hair, exhibit a pronounced preference for non-essential amino acids, which have lower synthetic costs, in their amino acid composition [20]. For several years, I have been searching for studies that specifically examine the amino acid composition in both intracellular and extracellular compartments. Ultimately, I found only one such publication. This study presented data on the amino acid composition of chicken muscle in both compartments [21]. A comparison of these data, despite being limited, revealed that the disparity in amino acid composition between intracellular and extracellular compartments aligns almost completely with the boundary between essential and non-essential amino acids (Table 2). Considering these observations, it would be reasonable to infer that the amino acid composition of the extracellular compartment, in comparison to that of intracellular compositions, consistently contains a higher proportion of non-essential, lower-cost amino acids.

Amino Acids	Intracellular compositions				Extracellular compositions				LN(Extra/Intra)					Essentiality
	6mo Leg	1.2yr Leg	6mo Breast	1.2yr Breast	6mo Leg	1.2yr Leg	6mo Breast	1.2yr Breast	6mo Leg	1.2yr Leg	6mo Breast	1.2yr Breast	Averages(⇓)	
Pro	0.0429	0.0410	0.0425	0.0413	0.2064	0.2117	0.2077	0.2006	1.572	1.642	1.587	1.581	1.595	
Gly	0.0559	0.0561	0.0572	0.0547	0.2638	0.2851	0.2702	0.2689	1.552	1.626	1.553	1.592	1.580	
Ala	0.0763	0.0785	0.0771	0.0776	0.0967	0.1011	0.0979	0.0940	0.237	0.253	0.240	0.192	0.231	
Arg	0.0537	0.0544	0.0545	0.0535	0.0556	0.0554	0.0522	0.0505	0.034	0.018	-0.043	-0.058	-0.012	(+)
Ser	0.0491	0.0500	0.0492	0.0499	0.0354	0.0331	0.0375	0.0403	-0.327	-0.413	-0.271	-0.214	-0.306	
Glu	0.1422	0.1384	0.1401	0.1404	0.0912	0.0891	0.0904	0.0961	-0.444	-0.440	-0.438	-0.379	-0.425	
Asp	0.0964	0.0982	0.0985	0.0970	0.0582	0.0553	0.0562	0.0603	-0.503	-0.573	-0.560	-0.475	-0.528	
Cys	0.0112	0.0067	0.0069	0.0111	0.0061	0.0033	0.0069	0.0043	-0.612	-0.728	0.011	-0.936	-0.566	
Thr	0.0521	0.0517	0.0515	0.0522	0.0242	0.0211	0.0244	0.0271	-0.767	-0.894	-0.746	-0.657	-0.766	+
Phe	0.0344	0.0330	0.0339	0.0336	0.0163	0.0154	0.0165	0.0145	-0.745	-0.763	-0.720	-0.837	-0.767	+
Lys	0.0872	0.0874	0.0872	0.0873	0.0382	0.0390	0.0353	0.0391	-0.825	-0.807	-0.903	-0.803	-0.835	+
Val	0.0629	0.0675	0.0659	0.0644	0.0285	0.0232	0.0299	0.0265	-0.792	-1.067	-0.789	-0.889	-0.884	+
Leu	0.0899	0.0898	0.0896	0.0900	0.0349	0.0308	0.0349	0.0346	-0.946	-1.068	-0.943	-0.956	-0.978	+
Ile	0.0555	0.0563	0.0557	0.0560	0.0190	0.0153	0.0185	0.0183	-1.070	-1.301	-1.101	-1.116	-1.147	+
Met	0.0266	0.0270	0.0263	0.0273	0.0083	0.0073	0.0065	0.0086	-1.163	-1.313	-1.405	-1.156	-1.259	+
Tyr	0.0300	0.0291	0.0302	0.0288	0.0093	0.0071	0.0077	0.0085	-1.172	-1.406	-1.365	-1.220	-1.291	
His	0.0256	0.0277	0.0267	0.0266	0.0078	0.0065	0.0071	0.0077	-1.187	-1.446	-1.324	-1.235	-1.298	+
Trp	0.0081	0.0074	0.0071	0.0084	-	-	-	-	-	-	-	-	-	+

Table 2. Difference in Amino Acid Composition between Intracellular and Extracellular Compartments of Avian Skeletal Muscles. This table presents the molar composition of amino acids within the intracellular and extracellular compartments of leg and breast skeletal muscle tissues in chickens at 6 months and 1.2 years of age. The data, derived from referenced literature [21], have been converted from mass to molar quantities, with the total molar composition of amino acids in each compartment normalized to equal one. The analysis compares intracellular and extracellular profiles using the natural logarithm of their ratios, and the amino acids are ordered such that the average logarithmic ratios for all four tissues are presented in descending order. Furthermore, the rightmost column displays the amino acids' essentiality with a plus sign. Arginine, marked with a plus sign in parentheses, indicates its essentiality for chickens but not for humans. Except for arginine and tyrosine, there is complete agreement between the average differences in amino acid composition within cellular compartments and the delineation between essential and non-essential amino acids.

Note1: The blue/red bars for each logarithmic value in this table are specifically biased to demarcate their essentiality threshold.

Note2: Glutamine and asparagine are reported as glutamic acid and aspartic acid, respectively, due to the processing methods used at the time of measurement. As a result, the analysis is based on 18 amino acids, reflecting these substitutions.

Recycling of Intracellular Amino Acids

Studies using radioactive isotopes have estimated and reported that humans synthesize about 200g of protein per day while consuming about 40g of protein [22]. Thus, even if the ingested proteins are entirely utilized for the synthesis of new proteins, the source for the synthesis of the remaining 160g difference must rely either on the synthesis of new amino acids or the degradation of self-proteins. In the process of recycling these self-proteins, it is believed that cells continuously degrade and resynthesize their own proteins, maintaining a state known as proteostasis. During this process of continuous amino acid recycling, intracellular proteins are likely used as a buffer to enhance the efficiency of protein synthesis. Simultaneously, it is probable that cells have evolved under selective pressure to minimize amino acid wastage in protein synthesis. Therefore, it is hypothesized that intracellular proteins are utilized as an amino acid resource buffer during protein synthesis, and the efficiency of amino acid recycling has been maximally optimized through evolution.

Amino Acids Required for Synthesis of Extracellular Proteins

On the other hand, if such a highly optimized system for the recycling of intracellular proteins were to synthesize proteins with an extremely biased amino acid composition in large quantities, it would encounter a discrepancy in amino acid resource supply. This is particularly true for extracellular proteins, which are often required in significant amounts for the structural composition outside the cell and generally have a composition that is biased compared to somatic proteins. This can be inferred from data in studies comparing the amino acid composition of eggs and pre-hatching chicks [23]. Compared to eggs, chicks consistently have more glycine and proline, with glycine increasing more than twofold and proline over 1.5 times (Table 3). These amino acid changes are speculated to be associated with the massive synthesis of extracellular proteins such as collagens. Therefore, particularly during the transition from egg to chick, these amino acids are thought to be newly synthesized from their precursors, and their synthetic capabilities appear essential for successful hatching. This phenomenon reflects and supports the notion that the ability to synthesize amino acids, which correspond to non-essential amino acids in extracellular proteins, needs to be maintained and is evidence thereof.

Amino Acids	High Weight Line			Low Weight Line			Averages (⇓)	Essentiality
	Eggs	Chicks	LN(Chick/Egg)	Eggs	Chicks	LN(Chick/Egg)		
Gly	0.186	0.402	0.770	0.145	0.355	0.894	0.832	
Pro	0.233	0.366	0.449	0.184	0.321	0.557	0.503	
Arg	0.353	0.448	0.238	0.322	0.382	0.172	0.205	(+)
Glu	0.751	0.776	0.033	0.585	0.676	0.145	0.089	
His	0.144	0.166	0.138	0.130	0.133	0.023	0.081	+
Ala	0.328	0.333	0.015	0.265	0.302	0.132	0.073	
Tyr	0.174	0.180	0.035	0.142	0.154	0.077	0.056	
Lys	0.416	0.432	0.038	0.376	0.373	-0.008	0.015	+
Thr	0.250	0.243	-0.028	0.205	0.214	0.044	0.008	+
Leu	0.498	0.491	-0.015	0.417	0.429	0.028	0.006	+
Asp	0.575	0.532	-0.077	0.483	0.477	-0.013	-0.045	
Phe	0.321	0.285	-0.118	0.263	0.253	-0.039	-0.079	+
Val	0.390	0.340	-0.136	0.328	0.320	-0.023	-0.080	+
Ile	0.324	0.276	-0.159	0.282	0.261	-0.077	-0.118	+
Cys	0.076	0.062	-0.207	0.065	0.062	-0.060	-0.133	
Ser	0.339	0.271	-0.227	0.277	0.244	-0.126	-0.177	
Met	0.156	0.092	-0.529	0.148	0.091	-0.479	-0.504	+

Table 3. Difference in Amino Acid Quantities between Eggs and Chicks. This table provides a comparison of amino acid quantities in chicken egg contents and chicks, based on molar amounts. The original data, reported as mass amounts in the referenced literature [23], have been converted to molar quantities for this analysis. Measurements were made for both high-weight and low-weight strain lines within the chicken species, and the data are presented as average molar quantities of each amino acid per egg and per chick. The comparison between the amino acid quantities of eggs and chicks was performed by calculating the logarithmic ratios to identify significant differences. Amino acids in the table are sorted in descending order based on the average logarithmic values for both high-weight and low-weight strains. Similar to Table 2, the rightmost column displays the amino acids' essentiality with a plus sign. Additionally, arginine, which is essential in chickens but not in humans, is indicated with a parenthesized plus sign. The ordering did not show a strong correlation with the essential amino acids in chickens; however, it is notable that glycine and proline levels, both non-essential amino acids, significantly increased in both strains, which is believed to result from extensive synthesis of collagen in the extracellular matrix. Additionally, the transition from egg to chick demonstrated an increase in the quantities of arginine and histidine, which are considered essential amino acids in chickens, suggesting that while these amino acids are classified as essential, there may be a retained capacity for their synthesis within the organism.

Note: Consistent with the measurement methods used, glutamine and asparagine were reported as glutamic and aspartic acids, respectively, due to the processing methods used at the time of measurement. As a result of these conversions and the absence of threonine measurements in this study, the analysis is based on 17 amino acids instead of the standard set of 20.

This section demonstrates that a disparity in amino acid composition exists between intracellular and extracellular compartments, and this disparity is likely a determining factor for the necessity of maintaining amino acid synthesis capabilities, thereby forming the basis for the division between essential and non-essential amino acids. These considerations demonstrate that the Extracellular Protein Hypothesis has substantial validity.

Discussion

Introduction of the Extracellular Protein Hypothesis

In this paper, I propose the Extracellular Protein Hypothesis, which suggests that the disparity in amino acid composition between intracellular and extracellular compartments across multiple organisms could explain the basis for essential amino acids. Previous theories did not adequately explain how the boundary between essential and non-essential amino acids originated. My hypothesis introduces a novel perspective by considering this amino acid composition disparity between the inside and outside of cells.

Re-evaluation of Amino Acid Essentiality

Since the introduction of Rose's concept of essential amino acids, their importance has been widely recognized in nutrition, while non-essential amino acids have received less attention. Recent reports, however, have begun acknowledging the nutritional importance of non-essential amino acids [9]. Nevertheless, the focus on essential amino acids remains predominant in clinical nutrition. In contrast to this traditional view, the Extracellular Protein Hypothesis posits that non-essential amino acids are crucial for the synthesis of extracellular proteins, suggesting a new paradigm in understanding amino acid essentiality.

The Ideals and Realities of the Extracellular Protein Hypothesis

This paper introduced the Extracellular Hypothesis. However, upon examination, several discrepancies between the ideals proposed by this hypothesis and the realities revealed by analysis have been identified. Initially, the hypothesis assumed that low-cost amino acids are more frequently used extracellularly. Although the analysis showed a significant correlation between the boundary of intracellular and extracellular compartments and the gradient of their cost disparities, it was not a complete match (Figure 1, Table 2). This suggests the presence of factors other than cost that dictate the amino acid compositions inside and outside the cell. Furthermore, analysis has consistently shown that the distinctions between essential and non-essential amino acids, as inferred from the disparities in amino acid compositions inside and outside the cell, differ notably for two specific amino acids. Arginine, typically classified as essential in many organisms except humans, was found to fall within the non-essential group in this study (Figure 1, Figures 2b, 2c, 2d, and Table 2). In contrast, tyrosine, which is generally considered a non-essential amino acid, consistently appeared within the essential amino acid group (Figure 1, Figures 2b, 2c, 2d, and Table 2).

Considering the increase in arginine levels during the transition from egg to chick (Table 3), it is possible that the synthetic capability for arginine is not completely lost. If so, the classification of arginine as essential may not stem from a lack of synthetic capability but rather from the increased demand within the urea cycle for processing ammonia, a byproduct of extensive protein degradation during events such as starvation or development. This excess demand might underscore the functional essentiality of arginine under such physiological conditions.

Regarding tyrosine, its classification as non-essential might be misleading due to its synthesis pathway being contingent upon phenylalanine, an essential amino acid. This dependency on phenylalanine suggests that tyrosine already lacks its independent synthetic capability. Although tyrosine was not deemed essential in Rose's 'minus one' experiments—a methodology used to determine essential amino acids—theoretically, if tyrosine, along with all nine essential amino acids, were removed from the diet, then tyrosine, despite being considered a non-essential amino acid, would also become deficient. On the other hand, from my own incidental observations, which are neither frequent nor extended over long periods, a reduction in pigmentation, such as in hair, presumably due to a deficiency of tyrosine leading to decreased melanin production, has been noted in children receiving total parenteral nutrition. These explorations and personal observations lend support to the argument. Therefore, taking into account the results of testing the hypothesis, it might be considered plausible to reclassify tyrosine as an essential amino acid.

Domain-Specific Amino Acid Requirement Profiles: A Comparative Discussion

The Extracellular Hypothesis finds key evidence in the nearly uniform composition of essential amino acids in feeding eukaryotes. However, this uniformity is absent in prokaryotes, such as bacteria, which exhibit varied amino acid requirements [8]. This variability likely stems from differences in amino acid utilization among biological domains. While eukaryotic organisms lose consistent and similar amino acid synthesis capabilities, prokaryotic organisms can adaptively lose the ability to synthesize amino acids that are abundant in their environment, serving as an environmental adaptation. This adaptability in prokaryotes might be attributed to their smaller cell size and the absence of autophagy capabilities, leading to an insufficient function as amino acid reservoirs for intracellular proteins. Furthermore, prokaryotic organisms are less inclined to produce extracellular proteins, including the extracellular matrix found in animals. In contrast, eukaryotic organisms, with their larger size and acquisition of autophagy capabilities alongside normal protein degradation pathways, are believed to possess enhanced amino acid storage abilities [24]. This enhanced capacity could contribute to improved starvation resistance and stability in amino acid supply for protein synthesis. It is this optimized buffering function that is thought to have spurred the development of the Extracellular Protein Hypothesis, identified exclusively in eukaryotic organisms.

The Evolutionary Basis of Amino Acid Essentiality: A Theoretical Exploration

In this paper, I hypothesize that the disparity in amino acid composition between intracellular and extracellular protein compartments is the origin of amino acid essentiality, a trait consistently observed in all eukaryotic organisms capable of ingestion. To understand the background of this phenomenon, I propose the following speculation: Initially, by acquiring the ability to ingest external nutrients, organisms gained the potential to utilize external amino acid resources, potentially reducing their need to synthesize all amino acids. However, ingestion required coordination with locomotion abilities, necessitating the development of extracellular protein structures. Primarily due to economic constraints, these extracellular proteins differed in composition from intracellular proteins, favoring amino acids with lower synthesis costs. According to the extracellular protein hypothesis, the synthesis capabilities for amino acids predominantly used in extracellular proteins could not be lost. Paradoxically, this speculation explains the origin of essential amino acids: The evolutionary process likely first optimized intracellular protein synthesis, followed by the acquisition of heterotrophy and an increase in extracellular protein synthesis. This sequence suggests that the loss of synthesis capabilities for common essential amino acids was not an accident but a natural consequence of evolutionary adaptations. This might also explain why organisms across various branches of the evolutionary tree have convergently lost the ability to synthesize nearly uniform sets of amino acids.

Limitations of the Study

This study is primarily limited by two significant deficiencies. The first deficiency is the lack of empirical data on the specific amino acid compositions within the intracellular and extracellular compartments. The absence of detailed data makes it challenging to accurately assess the variations in amino acid compositions between these two compartments. The second deficiency involves our limited understanding of the overall flow of amino acids within and across cellular boundaries. These deficiencies complicate our efforts to evaluate whether the disparities in amino acid compositions between the intracellular and extracellular spaces indeed act as the decisive factor in distinguishing between essential and non-essential amino acids.

The Extracellular Protein Hypothesis: Future Research Considerations

The scarcity of research on the disparity in amino acid composition between intracellular and extracellular compartments highlights a gap in our current scientific understanding. To validate and confirm the Extracellular Protein Hypothesis, future research will likely need to accurately measure the specific amino acid compositions of these compartments or simulate entire biological cell systems computationally, both of which pose significant challenges due to the current limited focus in this area. Moving forward, such research endeavors could deepen our understanding of amino acid essentiality and provide comprehensive verification of the Extracellular Protein Hypothesis.

Conclusion

In this paper, I have presented the Extracellular Protein Hypothesis, which explains that the need to synthesize non-essential amino acids for extracellular proteins has paradoxically led to the common "essential amino acids" found in eukaryotic organisms that have acquired the ability to ingest. This hypothesis challenges the traditional concept of amino acid nutrition, which has been heavily biased toward "essential" amino acids. It has the potential to initiate a paradigm shift, influencing not just our current understanding of amino acid nutrition, but also redefining how we perceive the roles of amino acid composition within the broader context of biology.

Statements and Declarations

Acknowledgments

I am deeply grateful to Associate Professor Hiroaki Oda from Nagoya University. Without his Japanese paper on the enigma of essential amino acids [5], this research would not have been possible. Further appreciation is extended to all experts who, through online and in-person discussions, provided pertinent comments and advice that shaped my initially unrefined hypothesis.

Competing Interest Statement

No competing interests are declared.

Data and Materials Availability

The raw data for the food composition table analyzed in this study can be downloaded from the link in reference[13]. Details on the amino acid compositions of chicken meats and eggs/chicks are found in references [21] and [23], respectively.

References

1. a, b *ROSE W. C. (1957). The amino acid requirements of adult man. Nutrition abstracts and reviews, 27(3), 631–647. http://www.ncbi.nlm.nih.gov/pubmed/18146542*
2. a, b, c *Gutiérrez-Preciado, A., Romero, H. & Peimbert, M. (2010). An Evolutionary Perspective on Amino Acids. Nature Education 3(9):29. https://www.nature.com/scitable/topicpage/an-evolutionary-perspective-on-amino-acids-14568445/ (Accessed on Jan 28, 2024.)*
3. a, b, c, d, e, f, g, h, i, j, k *Payne, S. H., & Loomis, W. F. (2006). Retention and Loss of Amino Acid Biosynthetic Pathways Based on Analysis of Whole-Genome Sequences. Eukaryotic Cell, 5(2), 272–276. https://doi.org/10.1128/EC.5.2.272-276.2006*
4. a, b, c, d, e, f, g, h, i *Guedes, R., Prosdocimi, F., Fernandes, G., Moura, L., Ribeiro, H., & Ortega, J. (2011). Amino acids biosynthesis and nitrogen assimilation pathways: a great genomic deletion during eukaryotes evolution. In BMC Genomics (Vol. 12, Issue S4). Springer Science and Business Media LLC. https://doi.org/10.1186/1471-2164-12-s4-s2*
5. a, b, c, d, e, f, g, h, i, j *Oda, H. (2007). Essential Amino Acids and Nonessential Amino Acids in Evolution. Nippon Eiyo*

Shokuryo Gakkaishi, 60(3), 137–149. https://doi.org/10.4327/jsnfs.60.137 (In Japanese).

6. [a, b, c, d, e, f, g, h, i] *Trolle, J., McBee, R. M., Kaufman, A., Pinglay, S., Berger, H., German, S., Liu, L., Shen, M. J., Guo, X., Martin, J. A., Pacold, M. E., Jones, D. R., Boeke, J. D., & Wang, H. H. (2022). Resurrecting essential amino acid biosynthesis in mammalian cells. ELife, 11. https://doi.org/10.7554/eLife.72847*
7. [a, b] *McCoy, R. H., Meyer, C. E., & Rose, W. C. (1935). FEEDING EXPERIMENTS WITH MIXTURES OF HIGHLY PURIFIED AMINO ACIDS. In Journal of Biological Chemistry (Vol. 112, Issue 1, pp. 283–302). Elsevier BV. https://doi.org/10.1016/s0021-9258(18)74986-7*
8. [a, b] *Gietzen, D. W., & Rogers, Q. R. (2006). Nutritional homeostasis and indispensable amino acid sensing: a new solution to an old puzzle. Trends in Neurosciences, 29(2), 91–99. https://doi.org/10.1016/j.tins.2005.12.007*
9. [a, b] *Reeds, P. J. (2000). Dispensable and Indispensable Amino Acids for Humans. The Journal of Nutrition, 130(7), 1835S-1840S. https://doi.org/10.1093/jn/130.7.1835S*
10. [a, b, c] *Akashi, H., & Gojobori, T. (2002). Metabolic efficiency and amino acid composition in the proteomes of Escherichia coli and Bacillus subtilis. Proceedings of the National Academy of Sciences, 99(6), 3695–3700. https://doi.org/10.1073/pnas.062526999*
11. [a, b, c] *Monera, O. D., Sereda, T. J., Zhou, N. E., Kay, C. M., & Hodges, R. S. (1995). Relationship of sidechain hydrophobicity and α-helical propensity on the stability of the single-stranded amphipathic α-helix. In Journal of Peptide Science (Vol. 1, Issue 5, pp. 319–329). Wiley. https://doi.org/10.1002/psc.310010507*
12. [a, b, c] *Sereda, T. J., Mant, C. T., Sönnichsen, F. D., & Hodges, R. S. (1994). Reversed-phase chromatography of synthetic amphipathic α-helical peptides as a model for ligand/receptor interactions Effect of changing hydrophobic environment on the relative hydrophilicity/hydrophobicity of amino acid side-chains. In Journal of Chromatography A (Vol. 676, Issue 1, pp. 139–153). Elsevier BV. https://doi.org/10.1016/0021-9673(94)00371-8*
13. [a, b, c, d, e] *Office for Resources, Council for Science and Technology, Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan. (2015). STANDARD TABLES OF FOOD COMPOSITION IN JAPAN - 2015 - (Seventh Revised Edition) - Amino Acids -. https://www.mext.go.jp/en/policy/science_technology/policy/title01/detail01/1374030.htm (Accessed on Sep 1, 2020.)*
14. [a, b, c, d, e] *Esumi, G. (2020). Autophagy: possible origin of essential amino acids. Cambridge Open Engage. https://doi.org/10.33774/coe-2020-lll03 This content is a preprint and has not been peer-reviewed.*
15. [a, b, c, d] *Esumi, G. (2023). The Distributions of Amino Acid Compositions of Proteins in an Organism's Proteome Uniformly Approximate Binomial Distributions. Jxiv. https://doi.org/10.51094/jxiv.408 This content is a preprint and has not been peer-reviewed.*
16. [^] *Esumi, G. (2023). The Synonymous Codon Usage of a Protein Gene Is Primarily Determined by the Guanine + Cytosine Content of the Individual Gene Rather Than the Species to Which It Belongs To Synthesize Proteins With a Balanced Amino Acid Composition. Jxiv. https://doi.org/10.51094/jxiv.561 This content is a preprint and has not been peer-reviewed.*
17. [^] *Crick, F. H. C. (1968). The origin of the genetic code. In Journal of Molecular Biology (Vol. 38, Issue 3, pp. 367–379). Elsevier BV. https://doi.org/10.1016/0022-2836(68)90392-6*
18. [^] *Fariselli, P., Taccioli, C., Pagani, L., & Maritan, A. (2020). DNA sequence symmetries from randomness: the origin of*

the Chargaff's second parity rule. In Briefings in Bioinformatics (Vol. 22, Issue 2, pp. 2172–2181). Oxford University Press (OUP). https://doi.org/10.1093/bib/bbaa041

19. [^]*Smith, D. R., & Chapman, M. R. (2010). Economical Evolution: Microbes Reduce the Synthetic Cost of Extracellular Proteins. In T. J. Silhavy & J. Tiedje (Eds.), mBio (Vol. 1, Issue 3). American Society for Microbiology. https://doi.org/10.1128/mbio.00131-10*
20. [^]*Esumi, G. (2022). Origin of Essential Amino Acids: Extracellular Matrix Hypothesis. Jxiv. https://doi.org/10.51094/jxiv.121 This content is a preprint and has not been peer-reviewed. (In Japanese).*
21. [a, b, c]*KARATZAS, C. N., & ZARKADAS, C. G. (1989). Comparison of the Amino Acid Composition of the Intracellular and Extracellular Matrix Protein Fractions Isolated from Avian Skeletal Muscles. In Poultry Science (Vol. 68, Issue 6, pp. 811–824). Elsevier BV. https://doi.org/10.3382/ps.0680811*
22. [^]*YOUNG, V. R., STEFFEE, W. P., PENCHARZ, P. B., WINTERER, J. C., & SCRIMSHAW, N. S. (1975). Total human body protein synthesis in relation to protein requirements at various ages. In Nature (Vol. 253, Issue 5488, pp. 192–194). Springer Science and Business Media LLC. https://doi.org/10.1038/253192a0*
23. [a, b, c]*Lepore, P. D., Siegel, P. B., & King, K. W. (1963). Proximate and amino acid composition of eggs and chicks from growth-selected lines of white rocks. Life Sciences, 2(8), 584–593. https://doi.org/10.1016/0024-3205(63)90111-5*
24. [^]*Onodera, J., & Ohsumi, Y. (2005). Autophagy Is Required for Maintenance of Amino Acid Levels and Protein Synthesis under Nitrogen Starvation. In Journal of Biological Chemistry (Vol. 280, Issue 36, pp. 31582–31586). Elsevier BV. https://doi.org/10.1074/jbc.m506736200*