

Early Selection of the Amino Acid Alphabet Was Adaptively Shaped by Biophysical Constraints of Foldability

Mikhail Makarov, Alma C. Sanchez Rocha, Robin Krystufek, Ivan Cherepashuk, Volha Dzmitruk, Tatsiana Charnavets, Anneliese M. Faustino, Michal Lebl, Kosuke Fujishima, Stephen D. Fried,* and Klara Hlouchova*



Cite This: *J. Am. Chem. Soc.* 2023, 145, 5320–5329



Read Online

ACCESS |



Metrics & More

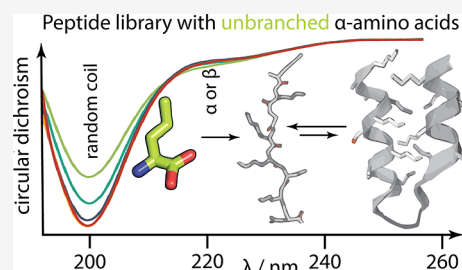


Article Recommendations



Supporting Information

ABSTRACT: Whereas modern proteins rely on a quasi-universal repertoire of 20 canonical amino acids (AAs), numerous lines of evidence suggest that ancient proteins relied on a limited alphabet of 10 “early” AAs and that the 10 “late” AAs were products of biosynthetic pathways. However, many nonproteinogenic AAs were also prebiotically available, which begs two fundamental questions: Why do we have the current modern amino acid alphabet and would proteins be able to fold into globular structures as well if different amino acids comprised the genetic code? Here, we experimentally evaluate the solubility and secondary structure propensities of several prebiotically relevant amino acids in the context of synthetic combinatorial 25-mer peptide libraries. The most prebiotically abundant linear aliphatic and basic residues were incorporated along with or in place of other early amino acids to explore these alternative sequence spaces. The results show that foldability was likely a critical factor in the selection of the canonical alphabet. Unbranched aliphatic amino acids were purged from the proteinogenic alphabet despite their high prebiotic abundance because they generate polypeptides that are oversolubilized and have low packing efficiency. Surprisingly, we find that the inclusion of a short-chain basic amino acid also decreases polypeptides’ secondary structure potential, for which we suggest a biophysical model. Our results support the view that, despite lacking basic residues, the early canonical alphabet was remarkably adaptive at supporting protein folding and explain why basic residues were only incorporated at a later stage of protein evolution.



INTRODUCTION

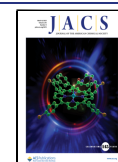
The role of peptides and proteins in life’s emergence has historically been sidelined as they have been, at times, perceived as only relevant after the evolution of nucleotide-based polymers, the genetic code, and a translation apparatus. However, the prebiotic abundance of amino acids and their ease of condensation resulting in polymers capable of creating functional hubs along with various cofactors (such as metal ions and organic compounds) has prompted some chemists to reconsider peptides’ role during life’s early evolution.^{1–4} While extant proteins are built from a sequence space spanned by the 20 canonical amino acids (cAAs) and rely on the specificity of the Central Dogma, early peptides (and peptide-like polymers) likely arose from a larger pool of prebiotically plausible monomers.

The modern protein alphabet was most likely selected during the first 10–15% of Earth history (~4.5–3.9 Ga), but the factors guiding its chemical and biological evolution remain unclear.⁵ Some of the cAAs are thought to be “late” additions to the genetic code, enabled only (or enriched significantly) by the evolution of their biosynthetic pathways.⁶ On the other hand, the “early” amino acids (Ala, Asp, Glu, Gly, Ile, Leu, Pro, Ser, Thr, and Val) have been found abundant in several different prebiotic sources, including atmospheric spark

discharge (Miller–Urey type) experiments, hydrothermal vent syntheses, and even meteorites.^{7–10} Although some studies have detected possible prebiotic synthesis pathways for other cAAs under very specific conditions,^{11,12} a consensus “early” canonical alphabet has been defined by two independent meta-analyses based on reports of the prebiotic abundance of amino acids and the order with which they entered the genetic coding.^{6,13} The ranking derived by these two meta-analyses has since provided a base for multiple independent studies recapitulating the protein consequences of the evolving alphabet.^{14–16} Nevertheless, besides the “early” cAAs, many other noncanonical amino acids (ncAAs) and their alternatives (such as β- and γ-amino acids, hydroxy acids or dicarboxylic acids) were apparently available through prebiotic synthesis.^{2,3,7,8,17} Why certain amino acids were selected over others (and when) has been repeatedly discussed in the literature over the last few decades.^{5,8,17–19} This question can

Received: December 6, 2022

Published: February 24, 2023



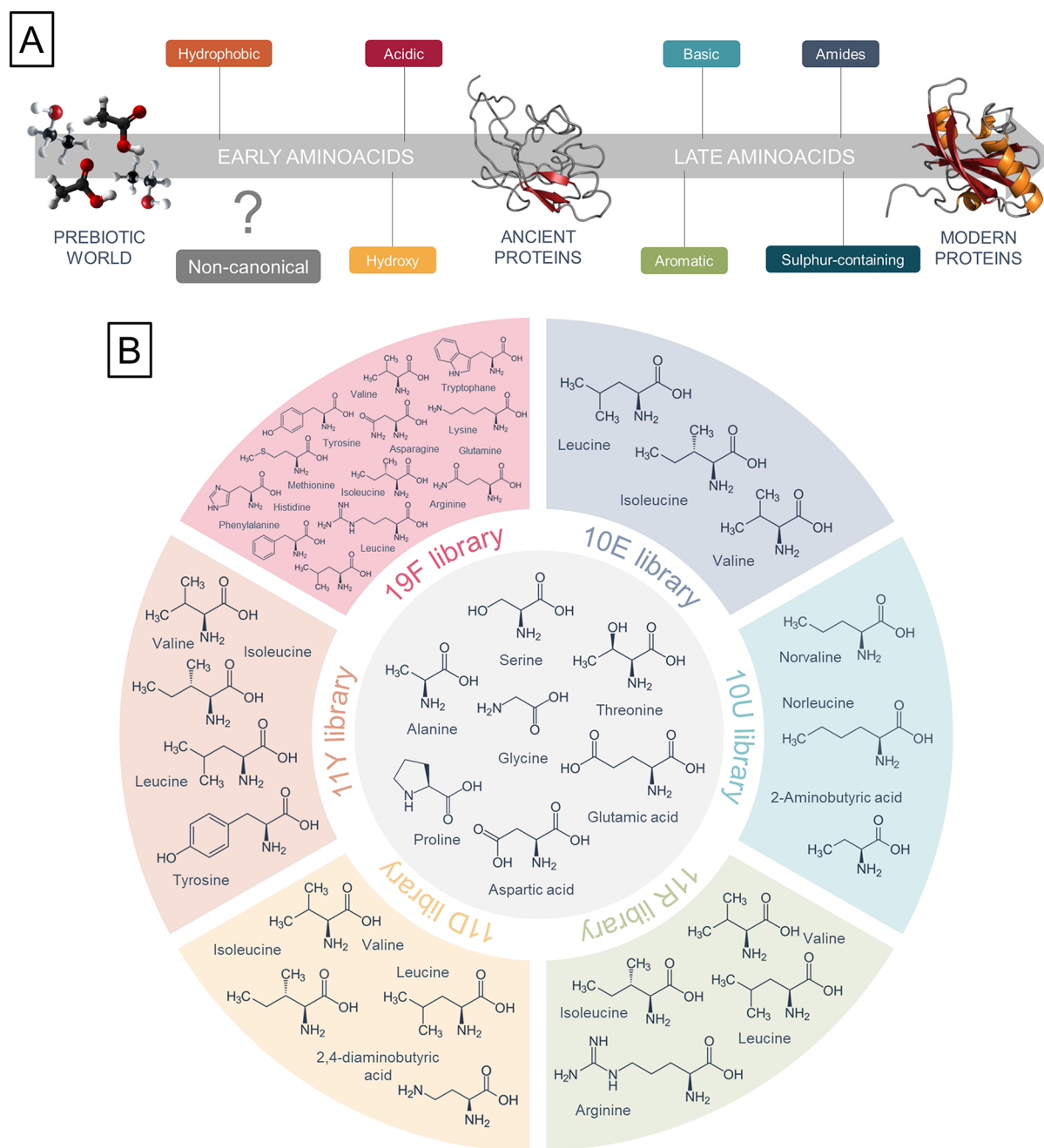


Figure 1. (A) Assumed early and late stages of amino acid alphabet incorporation during protein evolution and (B) design of peptide libraries based on this order.

be further subdivided into two separate but related questions, namely: (i) Why were the 10 early cAAs selected from the prebiotic environment and (ii) what factors drove the selection of the additional residues in the following era? Has protein evolution been successful as a consequence of the selected cAAs, or could similar structural and functional spaces be formed with alternative alphabets?⁸

While some of the choices for the early amino acids are quite expected (especially for the C2 and C3 amino acids), others have been repeatedly considered thought-provoking. Most strikingly, the early amino acid alphabet lacked positively charged residues, even though diaminopropionic and diamino-butyric acids (C3 and C4 amino acids, DAP and DAB, respectively) have been shown to be prebiotically plausi-

ble.^{20,21} Neither lysine nor arginine are considered prebiotically plausible, while ornithine is known to cyclize, promoting peptide chain scission.⁸ Moreover, the canonical alphabet does not include some of the most abundant linear aliphatic amino acids such as α -amino-*n*-butyric acid (ABA), norvaline (Nva), and norleucine (Nle) (C4, C5, and C6, respectively) while their branched alternatives (such as Val, Leu, Ile) were selected⁸ despite being available at much reduced abundances. It has been hypothesized that perhaps the earlier version of the amino acid alphabet included some nCAAs that were later replaced by late cAAs.^{2,22} This is an intriguing hypothesis especially in light of the fact that some of these nCAAs have been reported as promiscuous substrates of some aminoacyl-tRNA synthetases.²³

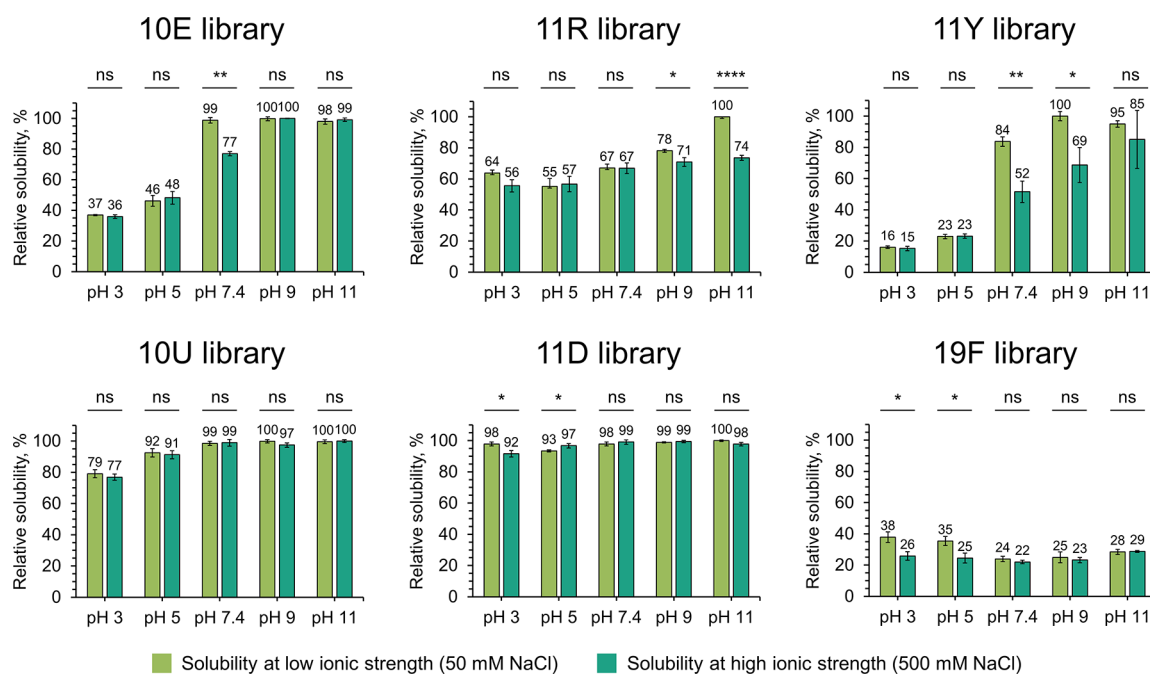


Figure 2. Solubility of 25-mer combinatorial peptide libraries at different pHs and ionic strengths. Solubility was measured for 0.5 mg/mL nominal peptide library solutions in 20 mM ABP buffers (pH 3–11) at 50 and 500 mM NaCl. Solubility of 19F peptide library was measured by fluorescence spectroscopy, and solubility of the other peptide libraries was measured by UV–vis absorption spectroscopy. Solubility measurements were conducted in technical triplicate ($n = 3$) on separate preparations of each library, and error bars correspond to standard deviations. Student's t tests were performed to assess significance of salt on solubility (ns, $P > 0.05$; *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$). For detailed statistical analysis, see Table S2.

This study experimentally determines the properties that would have accompanied some of the most feasible nAA candidates from the prebiotic pool. We incorporated the selected nAAs into combinatorial peptide libraries along with (or replacing) other early amino acids to evaluate their effect on fundamental physicochemical properties such as solubility and secondary structure formation. The outcomes of our study show that an inclusion of some of the most feasible alternatives would *not* have supported the twin requirements of rich secondary structure potential and solubility. Hence, our study demonstrates that the biophysical requirement of foldability (by which we mean “ability to fold into a soluble defined conformation”) provided a selective pressure shaping the building block selection in the early alphabet and therefore explains why certain early AAs (and not others) were selected to construct proteins.

RESULTS

Design and Synthesis of Peptide Libraries from Alternative Alphabets. Twenty-five-mer combinatorial peptide libraries were synthesized using solid-support chemical synthesis with isokinetic mixtures of amino acids (see Table S1).

To model the sequence space available to different subsets of the amino acid alphabets, the libraries included the entire canonical alphabet without Cys (19F; F = full), its prebiotically available subset of 10 (10E; E = early), an alternative of 10E where the branched aliphatic amino acids were substituted with their unbranched prebiotically abundant alternatives (10U; U = unbranched), the 10E library supplemented with Arg as a representative of a modern cationic cAA (11R; R = Arg); or DAB as a representative of a potentially early cationic

AA (11D; D = 2,4-diaminobutyric acid); or Tyr as a representative of an aromatic AA (11Y; Y = Tyr) (Figure 1).

MALDI spectra confirmed the expected molecular weight range and distribution of the combinatorial libraries, reflecting their respective compositions (Figure S1). Apart from the DAB in 11D library, the amino acid analyses of all libraries confirmed the expected composition (Figure S2). The analysis protocol could not detect DAB whose presence in the 11D library (contrary to the 10E library) was instead determined using fluorescamine assays (Figure S3). All the libraries passed quality control and were lyophilized to establish Cl^- as the counterion for all downstream analyses.

Solubility and Aggregation Profiles. The ability to remain soluble under different conditions could represent an important selection factor during the formation of the early protein alphabet, as prebiotically relevant environments spanned from alkaline hydrothermal vents to acidic lakes.^{24,25} We measured the solubility profiles of these peptide libraries in the pH range 3–11 and also at low vs high ionic strength, spectrophotometrically by absorption of peptide bonds at 215 nm and fluorometrically in case of library 19F due to its very poor solubility (Figure 2). In general, ionic strength did not significantly modulate solubility, with the exception of the 11Y library (see below).

At a fixed nominal peptide concentration (0.5 mg/mL), the 19F library is significantly less soluble than the libraries with narrower amino acid repertoires. This result is in line with previous studies reporting that subsets of the canonical library that are enriched with the early amino acids are significantly more soluble than the full alphabet version.^{14,16,26} This phenomenon could be partly explained by the physicochemical nature of the full alphabet (such as the presence of aromatic amino acids) leading to increased aggregation propensity²⁷ or

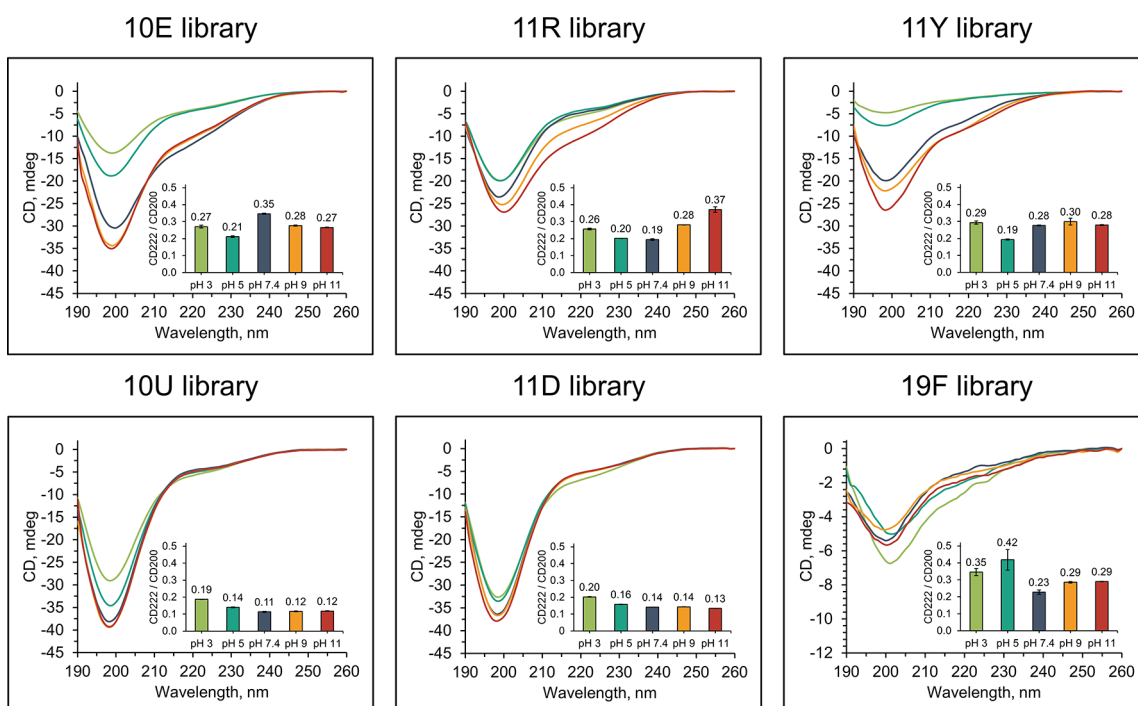


Figure 3. Effect of pH on the secondary structure of 25-mer combinatorial peptide libraries. CD spectra were recorded at a nominal concentration of 0.2 mg/mL peptide library (however is less for aggregation prone libraries, e.g., 19F) in a series of 10 mM ABP buffers (pH 3–11) at 25 mM NaCl. The inset graphs show the ratios of CD signals at 222 to 200 nm. CD spectra were acquired in duplicate and averaged and the 222/200 ellipticity ratio calculated. Measurements were then conducted in technical duplicate ($n = 2$) on separate preparations of each library, and error bars correspond to standard deviations. For detailed statistical analysis, refer to Table S3.

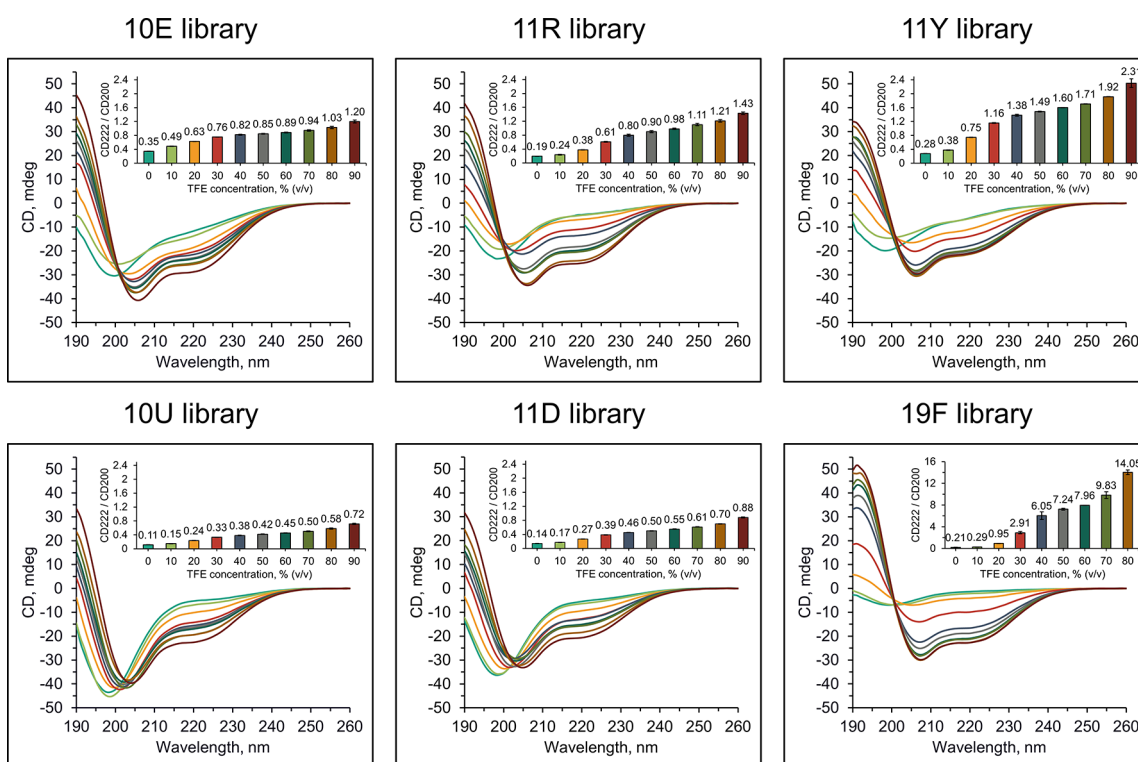


Figure 4. Effect of 2,2,2-trifluoroethanol to induce secondary structure of 25-mer combinatorial peptide libraries. CD spectra were recorded at 0.2 mg/mL nominal concentration in 10 mM ABP buffer (pH 7.4) with 0–90% (v/v) 2,2,2-trifluoroethanol. The inset graphs show the ratio of CD signal at 222–200 nm. The y-axis has the same scale for all except for 19F library. CD spectra were acquired in duplicate, averaged, and the 222/200 ellipticity ratio calculated. Measurements were then conducted in technical duplicate ($n = 2$) on separate preparations of each library, and error bars correspond to standard deviations. For detailed statistical analysis, refer to Table S4.

the greater “search complexity” of a larger alphabet (leading to a more rugged landscape). Regardless of the mechanism, the results support the view that peptides composed of the full alphabet are more reliant on chaperones and/or translation to fold.^{2,27}

The solubilities of the canonical library subsets (10E, 11R, and 11Y) trend upward with increasing pH (going from pH 3 to pH 11) and 11Y (which includes Tyr as a representative aromatic amino acid) is the least soluble of these three as expected. This result can be potentially explained by pointing out that all of these alphabets (as do modern proteomes)²⁸ tilt toward being acidic; hence, at alkaline pH, the peptides will accrue larger net negative charge and will repel one another electrostatically (inhibiting aggregation). This is broadly consistent with the practice of using mildly alkaline buffers in protein refolding experiments, which tend to facilitate refolding by preventing aggregation.²⁹ In general, high salt concentrations (500 mM NaCl) did not perturb solubility levels to the same extent as pH change, with the exception of the 11Y library for which high salt significantly lowered solubility at the alkaline pH's where it was soluble. We speculate this may be due to cation- π interactions with the nonphysiologically high tyrosine content (ca. 9.4% in 11Y vs 2–3% in natural proteomes).

The two libraries that include the noncanonical alternatives of aliphatic (10U) and basic (11D) amino acids are strikingly more soluble than 10E and 11R throughout the examined pH range (Figure 2). Both of these results are unexpected for distinct reasons. Based on the physical chemistry of elementary hydrocarbons, one might expect that the 10U library would be less soluble because linear hydrocarbons are more hydrophobic than branched hydrocarbons of equal carbon-count because they create cavities with greater surface area.³⁰ Moreover, the inclusion of a basic amino acid would raise the average isoelectric point of the library (relative to 10E) to be closer to neutral, which would decrease intermolecular repulsion and thereby increase aggregation.

The soluble fractions of the libraries were additionally screened for the occurrence of soluble aggregates using size-exclusion chromatography (Figures S4 and S5, in low and high ionic strength, respectively). No significant amount of such phenomenon was observed with the exception of 11Y and 10E libraries where a minor (up to ~10%) fraction of soluble aggregates was detected (Figure S4). Overall, the elution profiles correlated with the spectrophotometric solubility measurements in which the 11R library was the most soluble of the canonical subset libraries, and the 10U and 11D libraries were fully soluble across the entire pH range.

Secondary Structure Propensity. Natural nonfolding sequences as well as random sequences with low secondary structure content have been previously reported to be highly soluble.^{16,31} The solubility profiles of the 11D and 10U libraries therefore suggest that the incorporation of the respective ncAAs may decrease the potential for secondary structure formation. CD spectroscopy was used to estimate average secondary structure content of the libraries over the studied pH range (Figure 3) and upon its induction with 2,2,2-trifluoroethanol (TFE) in pH 7.4 (Figure 4). Because no dominant secondary structure motif is expected in these highly diverse libraries and we are rather interested in average structural propensities, the secondary structure content was evaluated as the ratio of CD signal at 222 nm to the CD signal at 200 nm (shown by inset graphs in Figures 3 and 4), because

both α helices and β sheets exhibit higher ellipticities than random coils at 222 nm.³² We did not observe any significant induction of secondary structure upon changing the metal ion concentration (Figure S6).

The CD spectra of the 10U and 11D libraries showed significantly lower relative signals at 222 nm, indicating that these peptides possessed profoundly decreased secondary structural propensity when compared with the respective canonical libraries 10E and 11R (Figure 3). At pH 7.4, the CD₂₂₂/CD₂₀₀ ratio was 3-fold lower for the 10U library relative to 10E (P -value = 0.0004 by Student's t test ($n = 2$)) and 2.4-fold lower for the 11D library relative to 10E (P -value = 0.0003 by Student's t test ($n = 2$)). These differences are also significant relative to 11R (P -values of 0.0056 and 0.0101, respectively). Such significant and large differences in secondary structure propensity between 10U and 11D relative to 10E carried across the entire pH range (Figure 3 and Table S3). Hence, the 10E library with branched aliphatic amino acids has a higher structural propensity than the 10U library with unbranched aliphatic alternatives in general. The addition of DAB in the 11D library also significantly decreases secondary structure propensity while the addition of canonical amino acids, Arg (in 11R) or Tyr (in 11Y), increases the secondary structure propensity mildly under some pH values. This trend becomes more pronounced upon TFE titration, which is often used to induce helical structure and increase protein stability.^{32–34} Because the overall folding stability of our 25-mer peptide libraries is small, we sought to test whether the trends we observed would extrapolate to a scenario in which folding stabilities are uniformly increased. TFE addition potentiates secondary structure content robustly in the 10E library, an effect that is further amplified in alphabets that include Arg and Tyr (Figure 4). Remarkably, the inclusion of DAB (in the 11D library) significantly decreases the potentiating capacity of TFE, implying that the inclusion of DAB still reduces the secondary structure potential even when they have been artificially stabilized. While the poor solubility of the 19F library is reflected in the overall low intensity of its CD spectra, its elevated structural propensity (when compared with all the other library subsets) becomes evident in the TFE titration (Figure 4).

DISCUSSION

The absence of ABA, Nva, and Nle in the proteinogenic alphabet has been considered striking given their high prebiotic abundance.¹⁰ Some authors have commented that this universal feature of biochemistry is a potential example of a “frozen accident” due to the absence of an obvious advantage of the canonical aliphatic residues over these ncAAs. At the same time, Weber and Miller hypothesized that inclusion of linear aliphatic amino acids would increase the side chain mobility and hence would not promote formation of ordered tertiary structure.⁸ Our experimental observations support this hypothesis. Moreover, structural prediction with PEPstrMOD (a molecular dynamics based predictor that generalizes to noncanonical amino acids) on 1,000 randomly chosen peptides from each library show that the 10U library has considerably less secondary structure content (11%) than the 10E library (26%), confirming the same trend (Figure 5).³⁵

To explain these findings, we start with the accepted model that the driving force for folding (and therefore secondary structure acquisition) is hydrophobic packing interactions.^{36,37} Isolated secondary structures are generally not stable in neutral

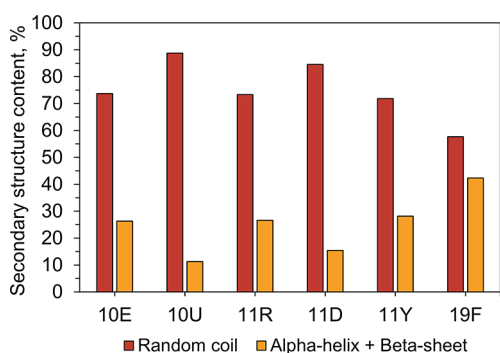


Figure 5. PEPstrMOD prediction of secondary structure content of 1,000 sequences selected randomly from the 25-mer combinatorial peptide libraries.

water³⁸—this is because the interamide hydrogen bonds associated with secondary structure are equally stabilizing as amide-solvent hydrogen bonds in the unfolded state.^{39–41} As a consequence, observation of secondary structure in our CD spectra can be treated as a proxy for peptides possessing hydrophobic packing interactions to stabilize those secondary structures.

Figure 6A explains why the 10U library has lower secondary structure-forming potential. Because unbranched hydrocarbons have more conformational freedom than their branched isomers, the ordering that is implicit in packing is expected to be more entropically costly. The lower packing efficiency of these side chains simultaneously explains the higher solubility of the 10U library. Hence, the requirement of foldability most likely acted to purge ABA, Nva, and Nle from the pool of amino acids that became proteinogenic, and ultimately canonical.

The more prebiotically abundant diamino acids, DAP and DAB, are also strikingly absent in the modern amino acid alphabet. One possible scenario is that DAP and DAB were utilized in an ancient alphabet and then were ultimately displaced with Arg, Lys, and His.¹⁸ Ornithine has also been proposed as a prebiotically plausible diamino acid that could have assisted in nucleic acid binding.^{42–44} However, it should be noted that ornithine is at least 1 order of magnitude less abundant in prebiotic environments and significantly more unstable compared to DAP and DAB.^{20,43,44} Another school of thought holds that proteins lacked positively charged residues altogether until biosynthetic pathways for the prebiotically unavailable Arg, Lys, and His evolved.^{6,13} While the second theory might seem unlikely at first glance, proteins lacking positively charged amino acids have recently been shown capable of folding and even binding with nucleic acids, assisted by metal ion cofactors.^{9,27,45,46} Our results concerning the 11D library further support the second theory. Hence, this body of work is building a compelling case that primordial proteins were highly acidic.

We find that, when DAB is incorporated into the 10E alphabet, it reduces secondary structure potential. Hence, its inclusion would have been a step *backward* rather than a step *forward* toward the goal of assembling foldable polypeptides. The same trend is observed in the PEPstrMOD prediction on 1,000 randomly chosen peptides from the 11D library (Figure 5). The 11D library is also exceptionally soluble, so the combination of high solubility and low foldability implies that 11D polypeptides are mostly unfolded in aqueous solution. By way of contrast, the 19F library that has the most folding capacity (as seen in PEPstrMOD and in CD) is also the least soluble (cf. Figure 2), because its bulky aromatic amino acids can alternatively lead to folding or aggregation depending on

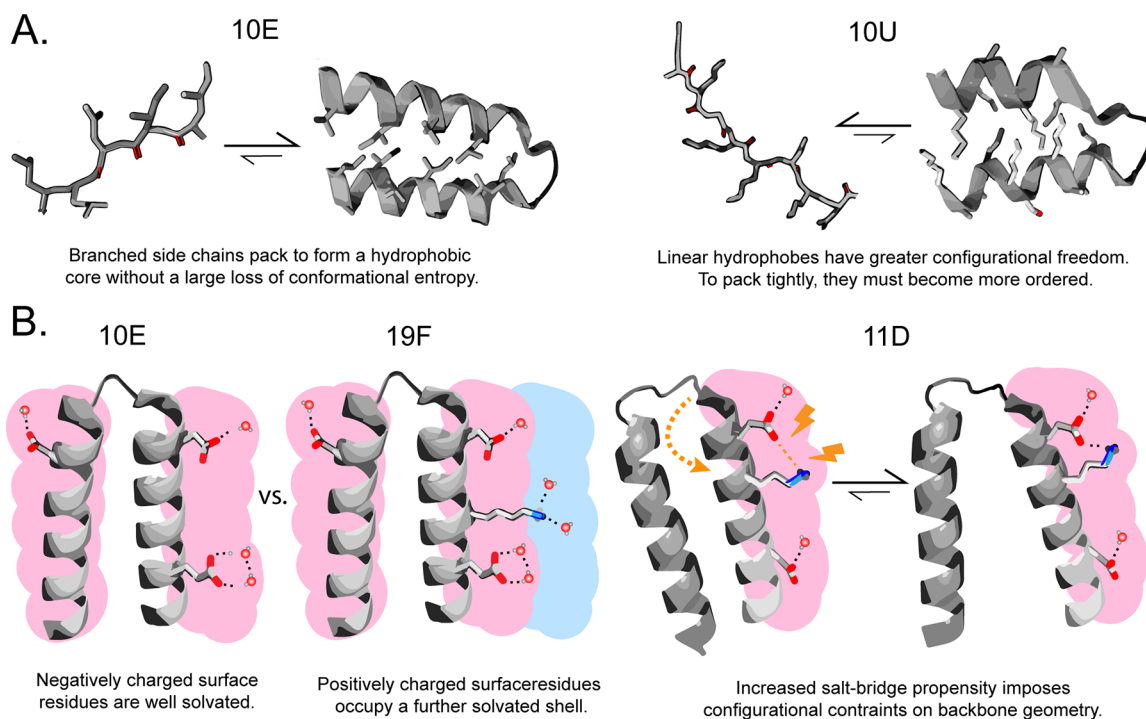


Figure 6. Biophysical models for the detrimental effect of unbranched hydrophobic side chains (A) and short-chained basic residues (B) on folding stability of small proteins. (A) Unbranched hydrophobic residues render core packing more entropically expensive. (B) Short-chained basic residues occupy the same solvation sphere as canonical acidic residues, increasing the frequency of salt bridges over solvation with water.

whether they are placed in appropriate positions in the peptide sequences.

Why does DAB harm protein structure propensity so significantly? It has been previously pointed out that these short-chain amines suffer from lactamization and acyl migration in peptides.^{8,47} This point is important in relation to the long-term stability of peptide chains, but it cannot feasibly explain the results of our study given that lactamization is a slow reaction in relation to the time scale of our experimental work. Moreover, the MALDI spectra of the library did not show any detectable scission products. It has also been proposed that charged groups close to the backbone interfere with backbone H-bonding,⁴⁸ though this does not explain the high foldability of short-chained acidic residues in 10E. A third possibility that we propose (Figure 6B) is that DAB disrupts structural potential because of a conflict associated with its simultaneous presence with Asp and Glu, i.e., negatively charged AAs with short side chains. Especially in these short 25-mer polypeptides, charged residues are likely to be on the surface. However, by having surface-anions and surface-cations in such close proximity, a polypeptide would be inclined to fold in such a way to favor formation of ion-pair salt bridges (Figure 6B, right). The backbone angles required to form such salt bridges would create additional constraints that might be hard to simultaneously satisfy along with hydrophobic packing, hence the ion-pairs would stabilize the protein in an unpacked (and therefore unfolded) conformation. Naturally, there are two ways to avoid this folding problem: Either (1) place positive charges further away from the backbone (on a longer side chain) so that they will primarily interact by hydrogen bonding with water in a distinct solvation shell or (2) do not include positively charged residues at all (Figure 6). We hypothesize that early proteins avoided ion pair-induced unfolding using the second strategy, whereas modern proteins adopted the first, though this strategy only became available once Arg and Lys could be biosynthesized. One experiment that could further test this hypothesis is the synthesis of a counterfactual peptide library in which basic residues use short side chains (diaminobutyric acid) while acidic residues use long side chains (e.g., 2-amino adipic acid). Our model would predict that such a combination would also support secondary structure formation. Such a finding addresses the central challenge for rational design of proteinogenic xeno alphabets: even a small library of candidate amino acid structures implies functionally infinite possible alphabets.^{49,50} Limiting alphabets that contain both positive and negatively charged side chains to those with unequal side chain lengths could be a crucial principle guiding protein design in synthetic sequence space.

CONCLUSION

In summary, our study on the solubility and secondary structure propensity of several prebiotically relevant amino acid alphabets supports the assertion that foldability played a critical early role in governing which amino acids ultimately became part of the canonical alphabet.^{51,52} Unbranched amino acids and short-chain basic amino acids were excluded, despite their prebiotic abundance because they “oversolubilized” polypeptides by stabilizing their unfolded conformations. Polypeptides with propensity to cooperatively fold *via* hydrophobic packing would have then been better adapted for binding cofactors, recognizing RNA, and performing enzymatic catalysis.² More broadly, this study supports the

view that the early canonical alphabet (10E), despite its deficiencies in relation to the modern canonical alphabet, was remarkably adaptive at supporting folding for the earliest proteins.

EXPERIMENTAL PROCEDURES

All chemicals if not stated otherwise were purchased from Sigma-Aldrich.

Design and Preparation of Combinatorial Peptide Libraries.

Split and mix synthesis of peptide libraries has its limitations.⁵³ A complete library of 25-mers composed of 10 amino acids would in theory require 10^{25} beads. One gram of solid phase support contains about 10^6 beads—therefore, it would cover only 10^{-17} % of possible structures. To increase the probability of completeness of the possible structures, it is necessary to conduct the synthesis considering individual molecules.

Average molecular weight of a 25-mer peptide composed of 10 amino acids is about 2500 Da. Therefore, 2.5 kg of these peptides would potentially contain 6×10^{23} individual peptides. In order to have a chance to synthesize a complete library, we would need ~250 kg of peptides, where each peptide would be present only once. But, the chance of finding any given peptide in the mixture is about 70%. Therefore, studying a complete library is practically impossible.

We decided to synthesize the peptide mixture on 200 mg of the solid support and therefore create theoretically only $\sim 10^{-4}$ % of possible structures. The approach using individual molecules is still capable of production of 10^{17} × more possible structures than split and mix strategy. However, considering complete randomness of the synthetic process, we concluded that the sample is a representative collection of possible structures. One milligram of synthetic peptide mixture contains about 2.4×10^{17} individual peptides.

Synthesis of the peptide mixture was accomplished by coupling mixtures of amino acids in which the ratio of individual amino acids was adjusted according to their reactivities. This approach was used by several authors, and various amino acid ratios were reported.⁵⁴ We have used as the basis the ratios reported by Santi et al.^{55,56} and adjusted the ratios for unnatural amino acids based on their structure and our previous experience (Table S1).

Synthesis was performed on the automatic peptide synthesizer Spyder Mark IV, using standard Fmoc peptide synthesis protocol.⁵⁷ Fmoc amino acids were individually weighted (see Table S1), mixed together, and dissolved to create 0.3 M solution in 0.3 M *N*-hydroxybenzotriazole (HOBt) in dimethylformamide (DMF). Rink resin (200 mg, 0.42 mmol/g) was distributed into 10 mL syringes of the synthesizer and swelled in DMF for 10 min. The mixtures from Table S1 were defined as extra amino acids 1–6 and placed into amino acid containers 21–26. Synthetic protocol was as follows: 2 × 1 min washing with DMF, 1 × 1 and 1 × 20 min treatment with 20% piperidine in DMF, 4 × 1 min wash with DMF, 1 × 1 min wash with 0.3 M HOBt, 2 × 60 min coupling with amino acid mixture, and 1 M diisopropylcarbodiimide in DMF. After 25 cycles of the synthesis, the resin was washed with DMF and dichloromethane and dried *in vacuo*. Dried resin was treated with 5 mL of Mixture K (trifluoroacetic acid–thioanisole–water–phenol–ethanedithiol, 82.5:5:5:5:2.5, v/v) for 2 h.⁵⁸ Resin was filtered off and peptide mixture was precipitated with diethyl ether three times and dried *in vacuo*. The pellet was then dissolved in 20% acetic acid and lyophilized. Prior to further experiments, all the peptide libraries were lyophilized three times with 1 mM HCl overnight.

Quality Control of Peptide Libraries. The molecular weight distributions of combinatorial peptide libraries were confirmed by mass spectrometry using UltrafleXtreme MALDI-TOF/TOF mass spectrometer (Bruker Daltonics, Bremen, Germany) according to the standard procedure.

Prior to the amino acid analysis, the library samples were hydrolyzed in 6 M hydrochloric acid at 110 °C for 20 h and the hydrolysate was evaporated and reconstituted with 0.1 M hydrochloric acid containing the internal standard. Amino acid analysis was performed on Agilent 1260 HPLC (Agilent Technologies, Wald-

bronn, Germany) equipped with a fluorescence detector using automated *o*-phtalaldehyde/2-mercaptopropionic acid (OPA/MPA) or ninhydrin derivatization in the case of 10E, 10U, 11R and 11D, 11Y, and 19F libraries, respectively.

The presence of 2,4-diaminobutyric acid in the 11D library was confirmed by fluorescamine assay.⁵⁹ For the assay, peptide library solution (75 μ L) in PBS buffer was incubated with 25 μ L of fresh stock of 3 mg/mL fluorescamine in DMSO at room temperature for 1 h in a 96-well plate (Greiner 650209). Fluorescence intensity ($\lambda_{\text{ex}} = 365$ nm, $\lambda_{\text{em}} = 470$ nm) was recorded using Tecan Spark plate reader. Primary amine concentrations were determined by linear interpolation of measured intensities for peptide H-GTIQPYPFWSWGY-NH₂ in concentration range 87.5–2.7 μ M. Obtained amine concentrations were then divided by the concentration of the analyzed peptide determined by amino acid analysis to calculate the amount of amine equivalents per molecule. Experiments were conducted in duplicate.

Solubility and Aggregation Propensity Measurements. A series of 10 Britton–Robinson buffers at pH 3.0, 5.0, 7.4, 9.0, and 11.0 were prepared by mixing 20 mM acetic acid, 20 mM boric acid, and 20 mM phosphoric acid and adjusting pH to the desired value with 5 M NaOH. The ionic strength was adjusted to either 50 mM (low ionic strength) or 500 mM (high ionic strength) with NaCl, and all Britton–Robinson buffers were filter-sterilized using 0.22 μ m PVDF membrane before use. Lyophilized peptide libraries were thoroughly resuspended in autoclaved Milli-Q water to 5 mg/mL and subsequently diluted 10 times into Britton–Robinson buffer to the final concentration of 0.5 mg/mL. The samples were gently shaken at room temperature for 30 min and then centrifuged at maximum speed (21,300g) for 15 min at 4 °C in order to remove insoluble fraction. The supernatant (soluble fraction) was taken and used for solubility and aggregation propensity analyses.

The relative amount of peptides in soluble fractions was estimated spectrophotometrically by absorption of peptide bonds at 215 nm for all peptide libraries except the 19F peptide library. Due to the extremely low solubility, solubility of the 19F peptide library was estimated by fluorescence spectroscopy using Pierce quantitative fluorometric peptide assay (Thermo Fisher Scientific), and 0.5 mg/mL solution in DMSO was used as a standard. Each peptide library was prepared at least three times independently. For each preparation, spectrophotometric measurements were performed five times, whereas fluorometric measurements were performed in triplicates.

The relative amount of soluble aggregates was estimated by size-exclusion chromatography. The 100 μ L aliquot of supernatant was loaded onto a Superdex 75 Increase 10/300 GL column (Cytiva) that was pre-equilibrated with two bed volumes of the corresponding 20 mM Britton–Robinson buffer. The peptides were eluted from the column by one bed volume of the buffer at 0.5 mL/min flow rate at room temperature, and the eluted peptides were detected by absorption at 215 nm.

Circular Dichroism Spectroscopy. The CD spectra were recorded using a Chirascan-plus spectrophotometer (Applied Photo-physics, Leatherhead, UK) over the wavelength range 190–260 nm in steps of 1 nm with an averaging time of 1 s per step. Cleared samples at 0.2 mg/mL nominal concentration in 1 mm path-length quartz cells were placed into a cell holder, and spectra were recorded at room temperature. The CD signal was obtained as ellipticity in units of millidegrees, and the resulting spectra were averaged from two scans and buffer-spectrum subtracted. All CD measurements were performed twice, and the resulting ratio of ellipticity at 222 nm to ellipticity at 200 nm was averaged and used to estimate the secondary structure content in peptide libraries.

To estimate the effect of pH on secondary structure, 5 mg/mL suspensions of peptide libraries in Milli-Q water were diluted into 10 mM Britton–Robinson buffers at pH 3.0, 5.0, 7.4, 9.0, and 11.0 to the final nominal concentration of 0.2 mg/mL, gently mixed at room temperature for 30 min, and then centrifuged at maximum speed (21,300g) for 15 min at 4 °C in order to remove insoluble fraction.

To estimate the effect of 2,2,2-trifluoroethanol on secondary structure, 5 mg/mL suspensions of peptide libraries in Milli-Q water

were diluted in a series of 10 mM Britton–Robinson buffer (pH 7.4) containing 0–90% (v/v) 2,2,2-trifluoroethanol.

To estimate the effect of metal ions on secondary structure, 5 mg/mL suspensions of peptide libraries in Milli-Q water were diluted in a series of 10 mM Tris buffer at pH 7.4 supplemented with 0, 10, 100, and 1,000 μ M mixtures of NaCl, KCl, MgCl₂, MnCl₂, and ZnCl₂.

Bioinformatic Analysis of in Silico Peptide Libraries. One thousand 25-mer sequences were randomly generated *in silico* for the libraries, and their structures were predicted using PEPstrMOD.³⁵ PEPstrMOD uses force field libraries Force field_NCAA and Force field_PTMM and integrates them in Molecular Dynamics (MD) simulations using AMBER. Additionally, it implements the force field library SwissSideChain with the MD package GROMACS. The parameters employed in the MD simulations were 50 ps of simulation time, and the selected peptide environment was set to “vacuum”.

DSSP⁶⁰ was employed to determine the secondary structure elements from the peptide tertiary structures generated by PEPstrMOD. The obtained annotations of secondary structure were grouped into three larger classes: helix (H, α helix; G, 310 helix; and I, π -helix), sheet (B, isolated β -bridge and E, extended strand), and loop (T, hydrogen bonded turn; S, bend; and “_” blank spaces).

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/jacs.2c12987>.

Figures of molecular weight distribution, amino acid composition, quantification of free amino acid groups, aggregation propensity, and effect of metal ions on the secondary structure propensities and tables of weights and amounts of protected amino acids, statistical analysis of relative solubilities, and statistical analysis of CD222/CD200 values (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Stephen D. Fried – Department of Chemistry and T. C. Jenkins Department of Biophysics, Johns Hopkins University, Baltimore, Maryland 21218, United States; orcid.org/0000-0003-2494-2193; Email: sdfried@jhu.edu

Klara Hloučková – Department of Cell Biology, Faculty of Science, Charles University, BIOCEV, Prague 12843, Czech Republic; Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, Prague 16610, Czech Republic; orcid.org/0000-0002-5651-4874; Email: klara.hlouckova@natur.cuni.cz

Authors

Mikhail Makarov – Department of Cell Biology, Faculty of Science, Charles University, BIOCEV, Prague 12843, Czech Republic

Alma C. Sanchez Rocha – Department of Cell Biology, Faculty of Science, Charles University, BIOCEV, Prague 12843, Czech Republic

Robin Krystufek – Department of Physical Chemistry, Faculty of Science, Charles University, Prague 12843, Czech Republic; Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, Prague 16610, Czech Republic; orcid.org/0000-0003-4096-1159

Ivan Cherepashuk – Department of Cell Biology, Faculty of Science, Charles University, BIOCEV, Prague 12843, Czech Republic

Volha Dzmitruk – Institute of Biotechnology of the Czech Academy of Sciences, BIOCEV, Vestec 25250, Czech Republic

Tatsiana Charnavets – Institute of Biotechnology of the Czech Academy of Sciences, BIOCEV, Vestec 25250, Czech Republic

Anneliese M. Faustino – Department of Chemistry, Johns Hopkins University, Baltimore, Maryland 21218, United States

Michal Lebl – Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, Prague 16610, Czech Republic

Kosuke Fujishima – Earth-Life Science Institute, Tokyo Institute of Technology, Tokyo 1528550, Japan; Graduate School of Media and Governance, Keio University, Fujisawa 2520882, Japan; orcid.org/0000-0002-8844-812X

Complete contact information is available at:

<https://pubs.acs.org/10.1021/jacs.2c12987>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Vyacheslav Tretyachenko, Dr. Valerio Guido Giacobelli, Dr. Lucie Bednarova, and Prof. Stephen Freeland for helpful discussions about this project. The authors would also like to acknowledge Dr. Radko Souček and Dr. Martin Hubálek for their assistance with amino acid analysis and MALDI analyses, respectively. This work was supported by the Human Frontier Science Program grant HFSP-RGY0074/2019 to K.F., S.D.F., and K.H. M.M., A.C.S.R., and I.C. acknowledge support by the project the “Grant Schemes at CU” (reg. no. CZ.02.2.69/0.0/0.0/19_073/0016935), project no. START/SCI/148. I.C. is supported by the Visegrad Fund Scholarship (no. 52110039). V.D., T.Ch., K.H., M.M., A.C.S.R., and I.C. acknowledge CMS CF Biophysical techniques of CIISB, Instruct-CZ Centre, supported by MEYS CR Infrastructure project LM2018127 and European Regional Development Fund-Project, “UP CIISB” (No. CZ.02.1.01/0.0/0.0/18_046/0015974). S.D.F. acknowledges support from the NIH Director’s New Innovator Award (DP2-GM140926).

REFERENCES

- (1) Milner-White, E. J.; Russell, M. J. Functional Capabilities of the Earliest Peptides and the Emergence of Life. *Genes (Basel)* **2011**, *2* (4), 671–688.
- (2) Fried, S. D.; Fujishima, K.; Makarov, M.; Cherepashuk, I.; Hlouchova, K. Peptides Before and During the Nucleotide World: An Origins Story Emphasizing Cooperation Between Proteins and Nucleic Acids. *J. R. Soc. Interface* **2022**, *19* (187), 20210641.
- (3) Frenkel-Pinter, M.; Samanta, M.; Ashkenasy, G.; Leman, L. J. Prebiotic Peptides: Molecular Hubs in the Origin of Life. *Chem. Rev.* **2020**, *120* (11), 4707–4765.
- (4) Freeland, S. Undefined Life’s Biochemistry: Implications for Abiogenesis. *J. R. Soc. Interface* **2022**, *19* (187), 20210814.
- (5) Cleaves, H. J. The Origin of the Biologically Coded Amino Acids. *J. Theor. Biol.* **2010**, *263* (4), 490–498.
- (6) Higgs, P. G.; Pudritz, R. E. A Thermodynamic Basis for Prebiotic Amino Acid Synthesis and the Nature of the First Genetic Code. *Astrobiology* **2009**, *9* (5), 483–490.
- (7) Kitadai, N.; Maruyama, S. Origins of building blocks of life: A review. *Geosci. Front.* **2018**, *9* (4), 1117–1153.
- (8) Weber, A. L.; Miller, S. L. Reasons for the Occurrence of the Twenty Coded Protein Amino Acids. *J. Mol. Evol.* **1981**, *17* (5), 273–284.

(9) Longo, L. M.; Blaber, M. Protein Design at The Interface of The Pre-Biotic and Biotic Worlds. *Arch. Biochem. Biophys.* **2012**, *526* (1), 16–21.

(10) Zaia, D. A. M.; Zaia, C. T. B. V.; De Santana, H. Which Amino Acids Should Be Used in Prebiotic Chemistry Studies? *Orig. Life Evol. Biosph.* **2008**, *38* (6), 469–488.

(11) Ménez, B.; Pisapia, C.; Andreani, M.; Jamme, F.; Vanbellingen, Q. P.; Brunelle, A.; Richard, L.; Dumas, P.; Réfrégiers, M. Abiotic Synthesis of Amino Acids in the Recesses of the Oceanic Lithosphere. *Nature* **2018**, *564* (7734), 59–63.

(12) Patel, B. H.; Percivalle, C.; Ritson, D. J.; Duffy, C. D.; Sutherland, J. D. Common Origins of RNA, Protein and Lipid Precursors in a Cyanosulfidic Protometabolism. *Nat. Chem.* **2015**, *7* (4), 301–307.

(13) Trifonov, E. N. Consensus Temporal Order of Amino Acids and Evolution of The Triplet Code. *Gene* **2000**, *261* (1), 139–151.

(14) Newton, M. S.; Morrone, D. J.; Lee, K. H.; Seelig, B. Genetic Code Evolution Investigated through the Synthesis and Characterisation of Proteins from Reduced-Alphabet Libraries. *ChemBioChem* **2019**, *20* (6), 846–856.

(15) Solis, A. D. Reduced Alphabet of Prebiotic Amino Acids Optimally Encodes the Conformational Space of Diverse Extant Protein Folds. *BMC Evol. Biol.* **2019**, *19* (1), 158.

(16) Tretyachenko, V.; Vymětal, J.; Bednářová, L.; Kopecký, V.; Hofbauerová, K.; Jindrová, H.; Hubálek, M.; Souček, R.; Konvalinka, J.; Vondrášek, J.; Hlouchová, K. Random Protein Sequences Can Form Defined Secondary Structures and Are Well-Tolerated In Vivo. *Sci. Rep.* **2017**, *7* (1), 15449.

(17) Freeland, S. “Terrestrial” Amino Acids and Their Evolution. *Amino Acids, Peptides and Proteins in Organic Chemistry: Origins and Synthesis of Amino Acids*, 1st ed.; Wiley-VCH Verlag GmbH & Co. KGaA, 2010; Vol. 1, pp 43–75.

(18) Vázquez-Salazar, A.; Lazcano, A. Early Life: Embracing the RNA World. *Curr. Biol.* **2018**, *28* (5), R220–R222.

(19) Raggi, L.; Bada, J. L.; Lazcano, A. On the Lack of Evolutionary Continuity between Prebiotic Peptides and Extant Enzymes. *Phys. Chem. Chem. Phys.* **2016**, *18* (30), 20028–20032.

(20) Meierhenrich, U. J.; Munoz Caro, G. M.; Bredehoff, J. H.; Jessberger, E. K.; Thiemann, W. H.-P. Identification of Diamino Acids in the Murchison Meteorite. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (25), 9182–9186.

(21) Wolman, Y.; Haverland, W. J.; Miller, S. L. Nonprotein Amino Acids from Spark Discharges and Their Comparison with the Murchison Meteorite Amino Acids. *Proc. Natl. Acad. Sci. U. S. A.* **1972**, *69* (4), 809–811.

(22) Alvarez-Carreño, C.; Becerra, A.; Lazcano, A. Norvaline and Norleucine May Have Been More Abundant Protein Components during Early Stages of Cell Evolution. *Orig. Life Evol. Biosph.* **2013**, *43* (4–5), 363–375.

(23) Apostol, I.; Levine, J.; Lippincott, J.; Leach, J.; Hess, E.; Glascock, C. B.; Weickert, M. J.; Blackmore, R. Incorporation of Norvaline at Leucine Positions in Recombinant Human Hemoglobin Expressed in *Escherichia Coli*. *J. Biol. Chem.* **1997**, *272* (46), 28980–28988.

(24) Martin, W.; Russell, M. J. On the Origin of Biochemistry at an Alkaline Hydrothermal Vent. *Philos. Trans. R. Soc. London B. Biol. Sci.* **2007**, *362* (1486), 1887–1925.

(25) Grotzinger, J. P.; Kasting, J. F. New Constraints on Precambrian Ocean Composition. *J. Geol.* **1993**, *101* (2), 235–243.

(26) Tanaka, J.; Doi, N.; Takashima, H.; Yanagawa, H. Comparative Characterization of Random-Sequence Proteins Consisting of 5, 12, and 20 Kinds of Amino Acids. *Protein Sci.* **2010**, *19* (4), 786–795.

(27) Tretyachenko, V.; Vymětal, J.; Neuwirthová, T.; Vondrášek, J.; Fujishima, K.; Hlouchová, K. Modern and Prebiotic Amino Acids Support Distinct Structural Profiles in Proteins. *Open Biol.* **2022**, *12* (6), 220040.

(28) Kozłowski, L. P. Proteome-pI: Proteome Isoelectric Point Database. *Nucleic Acids Res.* **2017**, *45* (D1), D1112–D1116.

- (29) Wang, Y.; Van Oosterwijk, N.; Ali, A. M.; Adawy, A.; Anindya, A. L.; Dömling, A. S. S.; Groves, M. R. A Systematic Protein Refolding Screen Method using the DGR Approach Reveals that Time and Secondary TSA are Essential Variables. *Sci. Rep.* **2017**, *7* (1), 9355.
- (30) Reynolds, J. A.; Gilbert, D. B.; Tanford, C. Empirical Correlation Between Hydrophobic Free Energy and Aqueous Cavity Surface Area. *Proc. Natl. Acad. Sci. U. S. A.* **1974**, *71* (8), 2925–2927.
- (31) Vymětal, J.; Vondrášek, J.; Hlouchová, K. Sequence Versus Composition: What Prescribes IDP Biophysical Properties? *Entropy (Basel)* **2019**, *21* (7), 654.
- (32) Kelly, S. M.; Jess, T. J.; Price, N. C. How to Study Proteins by Circular Dichroism. *Biochim. Biophys. Acta* **2005**, *1751* (2), 119–139.
- (33) Buck, M. Trifluoroethanol and Colleagues: Cosolvents Come of Age. Recent Studies with Peptides and Proteins. *Q. Rev. Biophys.* **1998**, *31* (3), 297–355.
- (34) Starzyk, A.; Barber-Armstrong, W.; Sridharan, M.; Decatur, S. M. Spectroscopic Evidence for Backbone Desolvation of Helical Peptides by 2,2,2-Trifluoroethanol: An Isotope-Edited FTIR Study. *Biochemistry* **2005**, *44* (1), 369–376.
- (35) Singh, S.; Singh, H.; Tuknait, A.; Chaudhary, K.; Singh, B.; Kumaran, S.; Raghava, G. P. PEPstrMOD: Structure Prediction of Peptides Containing Natural, Non-Natural and Modified Residues. *Biol. Direct.* **2015**, *10* (1), 73.
- (36) Tanford, C. The Hydrophobic Effect and The Organization of Living Matter. *Science* **1978**, *200* (4345), 1012–1018.
- (37) Dill, K. A. Dominant Forces in Protein Folding. *Biochemistry* **1990**, *29* (31), 7133–7155.
- (38) Soman, K. V.; Karimi, A.; Case, D. A. Unfolding of an Alpha-helix in Water. *Biopolymers* **1991**, *31*, 1351–1361.
- (39) Klotz, I. M.; Franzen, J. S. Hydrogen Bonds between Model Peptide Groups in Solution. *J. Am. Chem. Soc.* **1962**, *84* (18), 3461–3466.
- (40) Koh, J. T.; Cornish, V. W.; Schultz, P. G. An Experimental Approach to Evaluating The Role of Backbone Interactions in Proteins using Unnatural Amino Acid Mutagenesis. *Biochemistry* **1997**, *36* (38), 11314–11322.
- (41) Baldwin, R. L. In Search of The Energetic Role of Peptide Hydrogen Bonds. *J. Biol. Chem.* **2003**, *278* (20), 17581–17588.
- (42) Longo, L. M.; Despotović, D.; Weil-Ktorza, O.; Walker, M. J.; Jabłońska, J.; Fridmann-Sirkis, Y.; Varani, G.; Metanis, N.; Tawfik, D. S. Primordial emergence of a nucleic acid-binding protein via phase separation and statistical ornithine-to-arginine conversion. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117* (27), 15731–15739.
- (43) Hay, R. W.; Morris, P. J. Proton Ionisation Constants and Kinetics of Base Hydrolysis of Some α -Amino-acid Esters in Aqueous Solution. Part III. Hydrolysis and Intramolecular Aminolysis of $\alpha\omega$ -Diamino-acid Methyl Esters. *J. Chem. Soc., Perkin Trans. 2* **1972**, 1021–1029.
- (44) McGee, W. M.; McLuckey, S. A. The Ornithine Effect in Peptide Cation Dissociation. *J. Mass. Spectrom.* **2013**, *48* (7), 856–861.
- (45) Despotović, D.; Longo, L. M.; Aharon, E.; Kahana, A.; Scherf, T.; Gruic-Sovulj, I.; Tawfik, D. S. Polyamines Mediate Folding of Primordial Hyperacidic Helical Proteins. *Biochemistry* **2020**, *59* (46), 4456–4462.
- (46) Giacobelli, V. G.; Fujishima, K.; Lepšík, M.; Tretyachenko, V.; Kadavá, T.; Makarov, M.; Bednárová, L.; Novák, P.; Hlouchová, K. In Vitro Evolution Reveals Noncationic Protein-RNA Interaction Mediated by Metal Ions. *Mol. Biol. Evol.* **2022**, *39* (3), msac032.
- (47) Poduska, K.; Katrukha, G. S.; Silaev, A. B.; Rudinger, J. Amino Acids and Peptides. LII. Intramolecular Aminolysis of Amide Bonds in Derivatives of Alpha,Gamma-Diaminobutyric Acid, Alpha,Beta-Diaminopropionic Acid, and Ornithine. *Collect. Czechoslov. Chem. Commun.* **1965**, *30*, 2410–2433.
- (48) Padmanabhan, S.; York, E. J.; Stewart, J. M.; Baldwin, R. L. Helix Propensities of Basic Amino Acids Increase with the Length of the Side-chain. *J. Mol. Biol.* **1996**, *257*, 726–734.
- (49) Mayer-Bacon, C.; Agboha, N.; Muscalli, M.; Freeland, S. Evolution as a Guide to Designing xeno Amino Acid Alphabets. *Int. J. Mol. Sci.* **2021**, *22* (6), 2787.
- (50) Brown, S.; Voráček, V.; Freeland, S. What Would an Alien Amino Acid Alphabet Look Like and Why?. *UMBC GABS Symposium 2022*, 2023.
- (51) Ilardo, M.; Bose, R.; Meringer, M.; Rasulev, B.; Grefenstette, N.; Stephenson, J.; Freeland, S.; Gillams, R. J.; Butch, C. J.; Cleaves, H. J., 2nd Adaptive Properties of the Genetically Encoded Amino Acid Alphabet Are Inherited from Its Subsets. *Sci. Rep.* **2019**, *9* (1), 12468.
- (52) Ilardo, M.; Meringer, M.; Freeland, S.; Rasulev, B.; Cleaves, H. J., 2nd Extraordinarily Adaptive Properties of the Genetically Encoded Amino Acids. *Sci. Rep.* **2015**, *5*, 9414.
- (53) Lam, K. S.; Lebl, M.; Krchňák, V. The “One-Bead-One-Compound” Combinatorial Library Method. *Chem. Rev.* **1997**, *97* (2), 411–448.
- (54) Ostresh, J. M.; Winkle, J. H.; Hamashin, V. T.; Houghten, R. A. Peptide Libraries: Determination of Relative Reaction Rates of Protected Amino Acids in Competitive Couplings. *Biopolymers* **1994**, *34* (12), 1681–1689.
- (55) Rutter, W.; Santi, D. Peptide Mixtures. US5420246A, 1995.
- (56) Ivanetich, K. M.; Santi, D. V. Preparation of Equimolar Mixtures of Peptides by Adjustment of Activated Amino Acid Concentrations. *Methods Enzymol.* **1996**, *267*, 247–260.
- (57) Lebl, M.; Flegelova, Z.; Poncar, P.; Mudra, P.; Paces, O.; Knor, M.; Buzek, M.; Smrz, J.; Pokorný, V.; Pesek, V. Device for Parallel Oligomer Synthesis, Method of Parallel Oligomer Synthesis and Use Thereof. WO-2019115514-A1, 2017.
- (58) King, D. S.; Fields, C. G.; Fields, G. B. A Cleavage Method Which Minimizes Side Reactions Following Fmoc Solid Phase Peptide Synthesis. *Int. J. Pept. Protein Res.* **1990**, *36* (3), 255–266.
- (59) Udenfriend, S.; Stein, S.; Böhlen, P.; Dairman, W.; Leimgruber, W.; Weigle, M. Fluorescamine: a Reagent for Assay of Amino Acids, Peptides, Proteins, and Primary Amines in The Picomole Range. *Science* **1972**, *178* (4063), 871–872.
- (60) Kabsch, W.; Sander, C. Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolymers* **1983**, *22* (12), 2577–2637.