Temperature dependence of amino acid hydrophobicities

Richard Wolfenden, Charles A. Lewis Jr., Yang Yuan, and Charles W. Carter Jr.

Department of Biochemistry and Biophysics, University of North Carolina, Chapel Hill, NC 27599

Contributed by Richard Wolfenden, April 20, 2015 (sent for review February 16, 2015; reviewed by Robert L. Baldwin, George D. Rose, and D. Peter Tieleman)

The hydrophobicities of the 20 common amino acids are reflected in their tendencies to appear in interior positions in globular proteins and in deeply buried positions of membrane proteins. To determine whether these relationships might also have been valid in the warm surroundings where life may have originated, we examined the effect of temperature on the hydrophobicities of the amino acids as measured by the equilibrium constants for transfer of their side-chains from neutral solution to cyclohexane (K_{w,c}). The hydrophobicities of most amino acids were found to increase with increasing temperature. Because that effect is more pronounced for the more polar amino acids, the numerical range of K_{w,c} values decreases with increasing temperature. There are also modest changes in the ordering of the more polar amino acids. However, those changes are such that they would have tended to minimize the otherwise disruptive effects of a changing thermal environment on the evolution of protein structure. Earlier, the genetic code was found to be organized in such a way that—with a single exception (threonine)—the side-chain dichotomy polar/nonpolar matches the nucleic acid base dichotomy purine/pyrimidine at the second position of each coding triplet at 25 °C. That dichotomy is preserved at 100 °C. The accessible surface areas of amino acid side-chains in folded proteins are moderately correlated with hydrophobicity, but when free energies of vapor-to-cyclohexane transfer (corresponding to size) are taken into consideration, a closer relationship becomes apparent.

The equilibrium conformations of proteins in neutral solution are strongly influenced by interactions between their constituent amino acids and solvent water. Early work on the crystal structure of hemoglobin and related proteins showed that the side-chains of the more polar amino acid residues tend to be exposed to solvent, whereas less polar side-chains tend to be buried within the interior of globular proteins (1). Later, those tendencies were put to a quantitative test by measuring equilibria of transfer of amino acid side-chains from neutral aqueous solution into less polar environments, such as the vapor phase (2, 3) or a nonpolar solvent such as cyclohexane (4), which dissolves only ~2 × 10^{-3} M water at saturation (5) and appears to be devoid of specific interactions with solutes. The water-to-cyclohexane distribution coefficients (K_{w,c}) of the 20 common side-chains [here termed “hydrophobicities” (6, 7) and expressed in concentration units of mol/L in each phase; SI Appendix] were found to span a range of 15 orders of magnitude at pH 7 and 25 °C. Values of K_{w,c} have been shown to be related to their outside-to-inside distributions in globular proteins (4, 8) and to their tendencies to appear within the buried sequences of transmembrane proteins (9–11). Those solvent distribution experiments were conducted at what we would consider ordinary temperatures. However, there is widespread (12, 13)—if not universal (14)—agreement that life originated when the earth was warmer—perhaps much warmer—than it is today, and some modern organisms thrive at temperatures approaching the boiling point of water. The literature discloses little concrete information about how the hydrophobicities of the individual amino acids respond to changing temperature. Would the rules of protein folding be expected to differ significantly for an organism such as Pyrococcus horikoshii, which grows optimally at 98 °C (15), and might such differences have affected the early stages of evolution of life on Earth?

Earlier work also uncovered the existence of a pronounced bias in the relationship between amino acid hydrophobicity and the genetic code. Thus, a pyrimidine at the second codon position signals amino acids whose average hydrophobicity is much greater than those coded by a purine at the same position (2, 3). The values reported here allow a more detailed analysis, described in a companion paper (16), which reveals that two separate codes for amino acid size and hydrophobicity appear to be embedded in different parts of their tRNA sequences, with size (represented by vapor-to-cyclohexane equilibria) encoded in the acceptor stem, and hydrophobicity (represented by water-to-cyclohexane equilibria) embedded in the anticodon. Would one expect these relationships to be maintained at elevated temperatures?

In the present work, we sought to address these questions by determining the effect of temperature on equilibria of transfer of molecules representing the side-chains of each amino acid from neutral aqueous solution to cyclohexane (K_{w,c}).

Materials and Methods

For molecules representing the side-chains of Gly, Ala, Val, Leu, and Ile (hydrogen, methane, propane, butane, and isobutane), values of K_{w,c} were obtained by combining the best values for the molar solubilities of the gases in water and in cyclohexane, compiled from the literature in the Solubility Data Series (17). These values are considered accurate to within ±0.2 kcal/mol in ΔG at 25 °C ±0.4 kcal/mol in ΔH, and ±0.6 kcal/mol in ΔTS at 25 °C.

For molecules representing most of the other amino acid side-chains, we used 1H-NMR as described earlier (4) to determine values of K_{w,c} for the uncharged forms. In most cases, the solute was dissolved at a known concentration, established gravimetrically (0.01 M), in the more-favored solvent (water or cyclohexane). An aliquot (1.5 mL) was introduced into a scintillation vial and mixed vigorously with the less-favored solvent (15 mL) for 30 min using a magnetic stirring bar, immersed in a Haake K20 water bath at the desired temperature, with the water bath adjusted precisely at temperatures in the range 25–100 °C. For some measurements, the temperature was increased in steps of 5 °C.

Systematic relationships have long been recognized between the hydrophobicities of amino acids and (i) their tendencies to be located at the exposed surfaces of globular and membrane proteins and (ii) the composition of their triplets in the genetic code. Here, we show that the same coding relationships are compatible with the high temperatures at which life is widely believed to have originated. An accompanying paper reports that these two properties appear to be encoded separately by bases in the acceptor stem and the anticodon of tRNA.

Author contributions: R.W. and C.W.C. designed research; R.W., C.A.L., Y.Y., and C.W.C. performed research; R.W. analyzed data; and R.W., C.A.L., and C.W.C. wrote the paper.

Reviews: R.L.B., Stanford University; G.D.R., Johns Hopkins University; and D.P.T., University of Calgary.

The authors declare no conflict of interest.

1To whom correspondence should be addressed. Email: water@med.unc.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1507565112/-/DCSupplemental.

Significance

Systematic relationships have long been recognized between the hydrophobicities of amino acids and (i) their tendencies to be located at the exposed surfaces of globular and membrane proteins and (ii) the composition of their triplets in the genetic code. Here, we show that the same coding relationships are compatible with the high temperatures at which life is widely believed to have originated. An accompanying paper reports that these two properties appear to be encoded separately by bases in the acceptor stem and the anticodon of tRNA.
bath in which the temperature was maintained within 0.1 °C as measured with an ASTM International thermometer. The mixture was allowed to settle for 15 min, and a sample (0.1 mL) of the aqueous layer was removed using a pipette with a gel-loading tip and mixed with 99% D$_2$O (0.8 mL) to which a known concentration of pyrazine (0.01 M, δ 8.60 ppm) had been added as a proton integration standard for determining the concentration of solute.  

$^1$H-NMR spectra were acquired using a Bruker 500-MHz spectrometer equipped with a cryoprobe, using a water suppression pulse sequence with two transients. When the favored solvent was cyclohexane, $K_{wext}$ was calculated from the concentrations of solute in the aqueous phase before and after extraction. When the favored solvent was water, the concentration of solute in cyclohexane was determined by back-extraction from the cyclohexane phase (10 mL) with D$_2$O (1 mL).

Values of $K_{wext}$ for the uncharged side-chains of Asp and Glu (acetic and propionic acids) were determined in the presence of 0.03 M HCl to suppress ionization. Values for the uncharged side-chain of His (1-methylimidazole) and Lys (1-butylamine) were determined in the presence of 0.03 M KOH to suppress ionization, and Lys (1-butylamine) was determined in the presence of 0.03 M KOH to suppress ionization. Values for the uncharged side-chain of Arg (N-propylguanidine) were obtained by conducting measurements in the presence of 0.3, 1, and 3 M KOH to suppress ionization, and no significant variation was observed. To circumvent problems associated with the volatility of methanethiol (the side-chain of Cys) at elevated temperatures, distribution measurements were conducted in the presence of potassium phosphate buffer (0.1 M, pH 12.35), to maintain this solute in a constant state of ionization. $K_{wext}$ values of molecules representing each side-chain (19) (e.g., acetic acid representing the side-chain models at elevated temperatures from the corresponding experiments. To obtain the distribution coefficient of uncharged methanethiol, these results were corrected to reflect the fraction of uncharged thiol that was present in the aqueous phase at various temperatures. In the case of methanol, distribution measurements were conducted using 1-14C-methanol by determining the concentrations of radioactivity in both the aqueous and the cyclohexane phase. In each case, experiments were conducted over at least a fivefold range of solute concentrations. The results showed no significant variation in $K_{wext}$ values, such as would have been expected if self-association of the solute had been present in either phase. Table 1 shows the temperature range over which the positions of transfer equilibria were measured. In all cases, van't Hoff plots were linear within experimental error over the ~50 °C temperature range examined, yielding values with an estimated error of ±0.3 kcal/mol in ΔG at 25 °C and ±0.5 kcal/mol in ΔH, and ±0.8 kcal/mol in TΔS at 25 °C. Differences in heat capacity appeared not to exert an important influence on the response of these transfer equilibria to temperature, in view of the absence of significant curvature in the van’t Hoff plots over the 40–50 °C temperature range examined.

The side-chains of Asp, Glu, His, Lys, and Arg enter cyclohexane in detectable amounts only in their uncharged forms, as expected from the extremely large heats of solution of ions by water (20). Thus, the overall equilibrium of transfer $K_{wext(total)}$ of these ionizable solutes (involving both the charged and uncharged forms of the side-chains of Asp, Glu, His, Lys, and Arg that are present at pH 7) include a term for the water-to-cyclohexane transfer of the uncharged species, divided by the fraction ($\alpha$) of the solute molecules that are present in the uncharged form in aqueous solution at pH 7. For an amine side-chain (A), that fraction was calculated from the $pK_a$ value of the amine’s conjugate acid (AH$^+$) using Eq. 1. For a carboxylic acid side-chain (CH), $\alpha$ was calculated using Eq. 2.

\[
\text{fraction uncharged} = \frac{\alpha = ([A]_{\text{uncharged}} + [A^+])}{[A]_{\text{total}}} = \frac{1}{1 + ([A^+] / [A]_{\text{total}})}.
\]

To estimate values of $\alpha$ at elevated temperatures, enthalpies of ionization of molecules representing each side-chain (19) (e.g., acetic acid representing Asp and n-butylamine representing Lys) were used to estimate $pK_a$ values of the side-chain models at elevated temperatures from the corresponding values (21) at 25 °C (Table 2). The resulting values of the equilibrium constant $K_{wext(total)}$, which describes the distribution of the solute in both its charged and uncharged forms from neutral aqueous solution to cyclohexane, are shown in Table 3.

Table 1. Thermodynamics of transfer (kcal/mol at 25 °C) of molecules representing uncharged amino acid side-chains from water to cyclohexane

| Amino acid | Solute        | $\log K_{wext}$ 25 °C | $\Delta G$ 25 °C | $\Delta H$ 25 °C | $\Delta S$ 25 °C | $\log K_{wext}$ 100 °C | Range, °C |
|------------|---------------|-----------------------|-----------------|-----------------|-----------------|------------------------|-----------|
| Ile        | Butane        | 4.24                  | -5.77           | 1.03            | 6.80            | 4.39                   | 10–50     |
| Leu        | Isobutane     | 4.24                  | -5.77           | 0.60            | 6.37            | 4.33                   | 10–50     |
| Val        | Propane       | 4.09                  | -5.56           | 1.48            | 7.04            | 4.31                   | 10–50     |
| Pro        | Triangulation | 3.80                  | -6.10           | 2.10            | 8.24            | 4.13                   | 10–50     |
| Phe        | Toluene       | 2.64                  | -3.59           | -0.25           | 5.34            | 2.30                   | 10–50     |
| Ala        | Methane       | 2.11                  | -2.87           | 2.55            | 5.42            | 2.49                   | 10–50     |
| Met        | Ethylmethylsulfide | 1.91             | -2.60           | 0.80            | 3.40            | 2.03                   | 10–50     |
| Trp        | 3-Methylindole | 1.83                 | -2.49           | -0.13           | 2.36            | 1.81                   | 10–50     |
| Cys        | Methanethiol  | 1.53                  | -2.08           | 3.12            | 5.20            | 1.99                   | 10–49     |
| Gly        | Hydrogen      | 0.20                  | -0.27           | 2.24            | 5.22            | 0.53                   | 10–50     |
| Lys        | n-Butylamine  | -0.27                 | 0.37            | 5.31            | 4.94            | 0.51                   | 15–45     |
| Tyr        | p-Cresol      | -0.31                 | 0.42            | 4.35            | 3.93            | 0.33                   | 15–44     |
| Thr        | Cysteine      | -0.42                 | 0.92            | 5.72            | 4.23            | 0.23                   | 15–52     |
| Glu        | Proionic acid | -2.26                 | 3.07            | 10.35           | 7.28            | -0.74                  | 10–50     |
| Ser        | Methanol      | -2.82                 | 3.84            | 6.21            | 2.37            | -1.88                  | 15–50     |
| Asp        | Acetic acid   | -3.29                 | 4.47            | 8.34            | 3.86            | -2.06                  | 10–40     |
| His        | 4-Methylimidazole | -3.49               | 4.75            | 11.59           | 6.84            | -1.78                  | 15–45     |
| Gln        | Propionamide  | -4.07                 | 5.54            | 8.79            | 3.25            | -2.78                  | 10–40     |
| Arg        | n-Proarguidine| -4.32                 | 5.88            | 13.74           | 7.86            | -2.30                  | 10–50     |
| Asn        | Acetamide     | -4.88                 | 6.64            | 7.14            | 0.51            | -3.83                  | 10–55     |

*Data from Goldberg et al. (19).

Values for proline, an amino acid that does not have a side-chain in the ordinary sense, were estimated indirectly by first determining equilibrium constants ($K_{wext}$ and $K_{wext}$) for transfer of N-acetylpyrrolidine and N-butylacetamide. To obtain $K_{wext}$, the ratio between those equilibrium constants was multiplied by the values observed for propane, the side-chain for norvaline (38).

This work.
(expressed as mol/L in each phase) (6, 7) and the corresponding values of their free energies, enthalpies, and entropies of transfer at 25 °C. Most of the values at 25 °C resemble values reported earlier (4), with minor discrepancies arising from new information about the critically evaluated solubilities of the aliphatic hydrocarbon side-chains in water and in cyclohexane in the International Union of Pure and Applied Chemistry Solubility Data Series (15). Values of $K_{w=<(\text{uncharged})}$ for the uncharged side-chains at 100 °C were estimated by combining values of $K_{w=<(\text{uncharged})}$ at 25 °C with the present values of $\Delta H$ for water-to-cyclohexane transfer.

Table 2 shows the fractions ($\alpha$) of molecules representing the ionizing side-chains of Asp, Glu, His, Lys, and Arg (acetic acid, propionic acid, 4-methylimidazole, butylamine, and $n$-propylguanidine) that remain uncharged at pH 7 and 25 °C, based on $K_w$ values from the literature (21). Values for the heats of ionization of molecules representing the ionizing side-chains of Asp, Glu, His, Lys, and Arg were used to calculate their $\alpha$ values at 100 °C as described above at pH 6.14, the pH value of a neutral solution at 100 °C (22). Table 3 shows values of log $K_w$ at 25 °C for all species (charged plus uncharged), obtained by dividing $K_{w=<(\text{uncharged})}$ by the value of $\alpha$ at 25 °C. A similar procedure was used to obtain values of $K_w$ at 100 °C for all species (charged plus uncharged).

**Table 3.** Logarithmic equilibrium constants [log$_{10}$ $K_{w=<(\text{total})}$] for transfer of molecules in both their charged and uncharged forms from neutral aqueous solution to cyclohexane at 25 °C and 100 °C

| Amino acid | Solute | $pK_a$ (25 °C) | $\log_{10} \alpha$ (25 °C) | $\Delta H$ ionization (kcal/mol) | $pK_w$ (100 °C) | $\log_{10} \alpha$ (100 °C) |
|------------|--------|---------------|----------------|-----------------|----------------|----------------|
| Asp        | Acetic acid (39) | 4.76 (39) | −2.24 | −0.33 (19) | 4.78 | −1.38 |
| Glu        | Propionic acid (39) | 4.82 (39) | −2.18 | −0.36 (19) | 4.84 | −1.42 |
| His        | 4-Methylimidazole (39) | 6.95 (39) | −0.30 | 8.77 (19) | 5.66 | −0.60 |
| Lys        | $N$-Butylamine (39) | 10.95 (39) | −3.95 | 13.84 (19) | 8.91 | −2.77 |
| Arg        | $n$-Propylguanidine (37) | 13.6 (37) | −6.60 | 18 (37) | 10.95 | −4.81 |

**Discussion**

**Enthalpy of Transfer of Uncharged Side-Chains.** Uncharged side-chains tend to become more hydrophobic with increasing temperature, as indicated by their positive $\Delta H$ values for water-to-cyclohexane transfer (Table 1). Because H-bonds become weaker with increasing temperature (23), an increase in enthalpy would be expected when water is released from close contact with the polar groups of a solute when it passes from water into cyclohexane; positive $\Delta H_w$ values (5–14 kcal/mol) are observed for the uncharged, but polar, side-chains of Arg, Asp, Asn, Glu, Lys, Gln, His, Ser, and Thr. Much smaller $\Delta H$ values (−0.25 to 3.1 kcal/mol) are observed for the nonpolar side-chains of Ile, Leu, Val, Pro, Phe, Ala, Met, Trp, Cys, and Gly, consistent with their inability to participate in strong H-bonding or other polar interactions with water. The enthalpy of transfer of the Tyr side-chain occupies an intermediate position ($\Delta H_w$ = 4.3 kcal/mol), consistent with its intermediate position in the overall scale of hydrophobicities (Fig. 1).

**Entropies of Transfer of Uncharged Side-Chains.** An increase in entropy would be expected when structured water is released from close contact with nonpolar groups around which it was organized (Table 1). One would therefore anticipate an increase in entropy when a nonpolar molecule leaves water and enters a nonpolar solvent like cyclohexane (23). However, because water is also organized in the vicinity of polar substituents, a reduction in the local organization of solvent water would also be expected when a polar molecule leaves water and enters cyclohexane. It is therefore not surprising that substantial positive values of $\Delta T\Delta S$ are observed for water-to-cyclohexane transfer of side-chains of Ile, Leu, and Val, which are nonpolar, and for the side-chains of Arg and Glu, which are strongly polar.

**Overall Equilibria of Transfer of Amino Acid Side-Chains (Including Both Charged and Uncharged Species) from Water to Cyclohexane at pH 7.** Because amino acid side-chains enter cyclohexane only in their uncharged forms, the effect of ionization of the side-chains of His, Lys, Arg, Asp, and Glu is to draw these molecules preferentially into the aqueous phase [as indicated by $K_w$], to an extent that varies with the extent to which the $pK_w$ value of the ionizing side-chain differs from 7 (Table 3). Values of the overall equilibrium constant [$K_w$] for transfer of each side-chain, taking into account the sum of the concentrations of both the charged and uncharged forms (Eqs. 1 and 2), span a range of $10^{15.2}$-fold at 25 °C, with Arg occupying an extreme position (Table 3 and Fig. 1).

At elevated temperatures, that range of values becomes smaller, for at least two reasons. First, the effect of temperature in enhancing $K_w$ is more pronounced for the more polar than for the less polar amino acids, so that the overall range of $K_w$ values for the uncharged side-chains shrinks with increasing temperature, from $10^{6.8}$-fold at 25 °C to $10^{4.5}$-fold at 100 °C (Table 1 and Fig. 1). Second, the fraction ($\alpha$) of the side-chains of His, Lys, Arg, Asp, and Glu that remains uncharged decreases with increasing temperature (Table 2).

**Relationship Between Side-Chain Transfer Equilibria and Protein Folding.** Considerable effort has been devoted to the search for a general scale that might relate the physical properties of the side-chain of each of the 20 amino acids to its solvent-accessible surface area in folded proteins (ASA$_\text{fold}$), as determined using a probe that is often a water-sized sphere (24–32). Values of ASA$_\text{fold}$ depend on the size and shape of the probe, the presence or absence of steric constraints on the approach of a probe that
Hydrophobicities [K_{w>c}] at 25 °C and 100 °C, colored according to the middle base of the corresponding codon. Changes in the ordering of side-chain hydrophobicities at 25 °C and 100 °C, colored according to the second base of the corresponding codon as indicated. Hydrophobic residues, near the top, tend to be associated with purines (red and purple), whereas hydrophilic residues, near the bottom, tend to be associated with pyrimidines (green and blue).

Fig. 1. Hydrophobicities [K_{w>c}] of amino acid side-chains at 25 °C and 100 °C, colored according to the middle base of the corresponding codon. Changes in the ordering of side-chain hydrophobicities at 25 °C and 100 °C, colored according to the second base of the corresponding codon as indicated. Hydrophobic residues, near the top, tend to be associated with purines (red and purple), whereas hydrophilic residues, near the bottom, tend to be associated with pyrimidines (green and blue).

The one-dimensional correlations described in Fig. 2A indicate the more prominent role for K_{w>c} values in both sets of correlations (SI Appendix, Fig. S3). The K_{w>c} branch of the phase-transfer cycle (Fig. 2B) is related to amino acid size as represented by its surface area in a tripeptide (R^2 = 0.86), its volume (R^2 = 0.83), or its mass (R^2 = 0.89). It is uncorrelated with K_{c,v} (R^2 = 0.01; SI Appendix) and constitutes a linearly independent source of information that was omitted from previous attempts to understand how the properties of amino acids are related to protein folding. Its apparent helpfulness in contributing to prediction of the locations of the common amino acids in folded proteins agrees with the well-established close-packing of amino acid side-chains in proteins, in which large and small amino acids differ in their structural requirements (35). The finding that the core/surface distributions of the 20 amino acids can be better approximated using two complementary types of solvent transfer free energies seems likely to be useful in protein design and phylogenetics (36), by identifying suitable amino acid substitutions.

Effects of Temperature on Protein Folding and the Genetic Code. As noted earlier, most side-chains become more hydrophobic with rising temperature, and those increases are greater for the more polar amino acids. For the uncharged forms of the side-chains, the numerical range of ΔG_{w>c} values (11.6 kcal/mol at 25 °C) shrinks to 6.7 kcal/mol at 100 °C. When both the charged and uncharged side-chains are evaluated, the decrease is 8.0 kcal/mol at 100 °C, compared to 7.9 kcal/mol at 25 °C. The core/surface distributions of the 20 amino acids in folded proteins agrees with the well-established close-packing of amino acid side-chains in proteins, in which large and small amino acids differ in their structural requirements (35). The finding that the core/surface distributions of the 20 amino acids can be better approximated using two complementary types of solvent transfer free energies seems likely to be useful in protein design and phylogenetics (36), by identifying suitable amino acid substitutions.

Fig. 2. Correlations between amino acid properties and their behavior in folded proteins. (A) Predictions of the logarithm of the distribution constant (K_{surf}), derived from the exposed surface area of amino acids in folded proteins based on distribution coefficients of their side-chains. Rose-colored ellipses denote 95% of the area necessary to include all points in the scatterplots. Squared correlation coefficients are given above the scattergrams. (B) Vapor phase analysis of amino acid transfer free energies (4). Colored lines indicate pairs of values used to calculate log K_{surf} according to the background colors in C. (C) Scatterplots between observed values of log K_{surf} and values calculated using the three independent pairs of experimental phase transfer free energies in B, two at a time, along with their synergistic interaction (Eq. 3). These correlations average R^2 = 0.9, whereas those in A average R^2 = 0.4. Correlations between estimates based on each pair of vapor phase cycle values (panels with white background), average 0.95.
uncharged forms of all of the side-chains are taken into consideration, the details change. The large heat of deprotonation of the guanidinium side-chain of Arg (18 kcal/mol) (37) causes its pK\textsubscript{a} value to decrease by 2.7 units when the temperature rises from 25 °C to 100 °C, whereas the pK\textsubscript{a} values of the side-chains of Lys (10.95) and His (6.95) decrease by 2.0 and 1.3 units, respectively. In contrast, the heats of ionization of acetic and propionic acids are ~0.1 kcal/mol (19), so that the pK\textsubscript{a} values of the side-chains of Asp and Glu do not change significantly when the temperature increases from 25 °C to 100 °C. As a result, the numerical range of ΔG\textsubscript{w-c} values for water-to-cyclohexane transfer from neutral aqueous solution to cyclohexane shrinks from 15.2 kcal/mol at 25 °C to 10.74 kcal/mol at 100 °C. There is also a modest reshuffling of relative hydrophobicities, so that the order His > Glu > Lys > Gln > Asn > Asp at 25 °C is reordered to the series Glu > Lys > His > Gln > Asp > Asn at 100 °C. However, their K\textsubscript{w-c}(total) values at 100 °C fall within a factor of 60, an extremely narrow range compared with a range of 1.5 × 10\textsuperscript{24}-fold for all of the amino acid side-chains (SI Appendix, Table S1, last column). The limited magnitude of these changes, for the more polar amino acids, would presumably have tended to minimize the potentially disruptive effects of changes in thermal environment on the structures of globular proteins and also on the relative tendencies of amino acids to be found in deeply buried sequences of transmembrane proteins. It seems reasonable to infer that the rules of protein folding, insofar as they involve solvation by water, are not very different in a thermophilic organism such as Pyrococcus horikoshii, which grows optimally at 98 °C, from those in a mesophilic organism growing at ordinary temperatures.

Earlier work revealed a pronounced bias in the genetic code (2, 3). Our recent reanalysis indicates that two separate codes, for amino acid size and hydrophobicity, appear to be embedded in tRNA sequences, with size encoded in the acceptor stem and hydrophobicity embedded in the anticodon (16). One aim of the present work was to determine whether these relationships are likely to have been maintained at elevated temperatures. In Fig. 1, which shows the relative hydrophobicities of the 20 common amino acids at 25 °C and 100 °C, the second base of each coding triplet is colored to indicate whether it is a purine (A or G, warm colors) or pyrimidine (C or U, cool colors). Fig. 1 shows that there is a single departure (Thr) from perfect correspondence between the dichotomy polar/nonpolar and the dichotomy purine/pyrimidine. That relationship is maintained at both high and low temperatures, implying that the meaning of a genetic code established in a hot environment would not have changed very much as the surroundings cooled.

**ACKNOWLEDGMENTS.** We thank Jan Hermans and Daniel Iosm for helpful discussions during the preparation of this manuscript. This work was supported by National Institutes of Health Grants GM-18325 (to R.W.) and GM-78227 (to C.W.C.).