Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 8880–8884, October 1991
Biochemistry

Distribution of glutamine and asparagine residues and their near neighbors in peptides and proteins

ARTHUR B. ROBINSON AND LAURELEE R. ROBINSON*

Oregon Institute of Science and Medicine, 2251 Dick George Road, Cave Junction, OR 97523

Communicated by Martin D. Kamen, July 18, 1991

ABSTRACT **In a statistical study of neighboring residues in 1465 peptides and proteins comprising 450,431 residues, it was found that the preferences for residues neighboring to glutamine and asparagine residues are consistent with the hypothesis that the rates of deamidation of these residues are of biological significance. Some dipeptide and tripeptide structures have special usefulness and some are especially undesirable. More such structures exist for amide residues than for other residues, and their specific types are those most relevant to the deamidation of amide residues under biological conditions.**

Glutamine (Gln) and asparagine (Asn) residues in peptides and proteins have the unique property that they are inherently unstable under physiological solvent conditions. Deamidation of these residues to become glutamic and aspartic residues occurs in a wide variety of peptides and proteins. These deamidation rates are dependent upon pH, ionic strength, temperature, and other solvent conditions (1).

Deamidation is also dependent upon the structure of the peptide or protein near each amide residue. Deamidation half-times for Gln and Asn residues under physiological solvent conditions vary over a range of at least 1 day to 9 years depending upon the primary sequence of residues near the amides and upon secondary and tertiary structure effects (1–3). The first experimental demonstration that deamidation depends upon three-dimensional protein configuration *in vitro* and *in vivo* was in cytochrome *c* (4, 5).

It has been proposed that structure-dependent deamidations of peptides and proteins serve as *in vivo* molecular clocks that control the rates of protein turnover, development, and aging (1, 6–8). In addition to other evidence, distribution functions of deamidation rates, amide contents of proteins, and residues neighboring to amide residues were found to be consistent with the amide molecular clock hypothesis (refs. 1 and 6; A.B.R., F. C. Westall, and L.R.R., unpublished work done in 1973). The reliable protein sequences available at the time of these calculations in the early 1970s were, however, too few for reliable detailed analysis.

We have now carried out an analysis of the distribution functions of dipeptides, tripeptides, and polypeptides with 1–9 intervening residues between residue pairs for all 20 amino acid residues. This analysis was performed on the sequences of 1465 molecules with 450,431 amino acid residues. These were selected from a set of 7396 sequenced peptides and proteins by eliminating duplicate protein types. The results of this analysis of sequence distribution functions are consistent with the molecular clock hypothesis and agree with predictions made on the basis of that hypothesis 20 years ago (6, 7).

MATERIALS AND METHODS

Selection of Peptides and Proteins. Peptide and protein sequences were drawn from the Protein Sequence Data Bank of the Protein Identification Resource (PIR) listing of March 31, 1988. PIR is a service of the National Biomedical Research Foundation, Georgetown University Medical Center, 3900 Reservoir Road, N.W., Washington, DC 20007.

From the 7396 sequences in the listing, 1465 were selected to eliminate statistical bias from the inclusion of closely related sequences. The 1465 selected sequences comprised 450,431 amino acid residues. The selected residues included 35,155 Ala, 8669 Cys, 24,161 Asp, 28,354 Glu, 17,367 Phe, 33,229 Gly, 9906 His, 23,161 Ile, 25,872 Lys, 40,625 Leu, 10,101 Met, 20,212 Asn, 23,435 Pro, 19,208 Gln, 23,105 Arg, 32,070 Ser, 26,311 Thr, 29,012 Val, 5990 Trp, and 14,488 Tyr.

The selection was made before any statistical calculations. Once selected, the list of included proteins was not modified in the course of this work. A complete listing of the included sequences is available from the Oregon Institute of Science and Medicine, 2251 Dick George Road, Cave Junction, OR 97523.

Separate sets of proteins were also evaluated where only human, only animal, and only plant and bacterial proteins were included. These studies are the subject of a separate report.

Calculation of Results. All calculations were carried out in a 386-16 computer (Zeos) with use of an ORACLE database program and utility programs written in the C programming language.

The selected sequences also included 90 Asx and 79 Glx residues where the amide or acid was not specified and 98 Xaa where the residue was not identified. There were also 1465 pairs of peptide and protein ends in the sequence set. These entities were treated identically with the 20 naturally occurring residues during the calculations of expected and found sequences. All pairs and triplets involving these entities were then omitted before cumulative distribution functions were drawn and the other reported data analysis was carried out.

The expected numbers of pairs and triplets were calculated as the appropriate products of the numbers of each residue with normalization to the total numbers of residues and ends. In this way the calculation assumed one pool of residues and ends from which all possible sequences could be drawn. It had been intended also to perform the calculations based on the summation of independent calculations of each of the 1465 peptides and proteins separately. This work was, however, unexpectedly terminated* before those calculations were completed.

The numbers of pairs and triplets found were determined by counting of occurrences in the complete set. Probabilities were then assigned to these found values by assuming gaussian distributions centered on the expected values and having standard deviations equal to the square root of the expected values. The gaussian distributions were then com-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

*Laurelee Robinson carried out the computer work described in this paper. All data analysis was completed before July 1988. Before her death in November 1988, Laurelee Robinson saw and approved these calculations and results. She did not read the final manuscript of this publication.

puter-integrated in the regions up to the found values in order to determine the value of P. Values of P near 0 correspond to sequences found more often than expected, and values of P near 1 correspond to sequences found less often than expected.

These calculations were performed for all 8000 possible triplet sequences. The triplet results were displayed and analyzed in 20 sets of 400 in which the central amino acid residue defined each set. For each of the 20 sets, a cumulative distribution function of probabilities was drawn for the 400 constituent triplets. The automatic computer output consisted of a point-by-point graph of each cumulative distribution function and a listing of each triplet with its corresponding expected number of occurrences, found number of occurrences, and P value with the listings ordered by increasing P.

These calculations were also performed for the 4000 sequence pairings that included 10 sets of 400 pairs separated according to the number of intervening residues. In this way all pairings from side-by-side out to and including pairings with 9 intervening residues were evaluated. These results consisted of 10 cumulative distribution functions and 10 listings of expected values, found values, and probabilities, P, like those described above for the triplet evaluations.

RESULTS

Amino Acid Residue Pairs. Fig. 1 shows the cumulative distribution function of P for side-by-side pairs of residues with no intervening residue. If the observed numbers of pairings were random, the distribution would be a straight diagonal line extending from the origin to the point at $P = 1$ and $n = 400$. The accumulation of preferred pairings at low P and of rejected pairings at high P is evident. Pairs that are random in occurrence are evident in the central straight portion of this curve. If the number of residues were great enough, this random portion might diminish to an inflection point that represents zero residues, since all structures may have at least a slight bias in frequency. Extrapolation of this straight portion to the $P = 0$ and $P = 1$ axes gives estimates of 130 preferred pairings, 150 rejected pairings, and 120 random pairings. Reliable separate listings of all of these three types are not possible, however, because the limited data do not allow rigorous separation of the weaker preferences and rejections from the random occurrences. We chose, therefore, to continue analysis only with pairs having $P < 0.001$ or $P > 0.999$. This ensured that less than one listing in each of our sets of preferred or rejected pairings is expected to occur by chance.

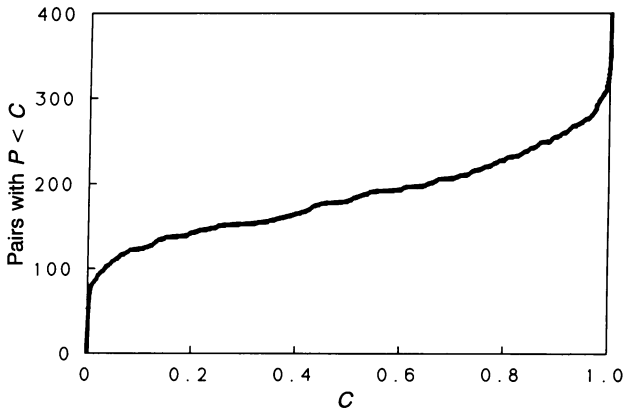

FIG. 1. Cumulative distribution function of probabilities, P, for the observed frequencies of side-by-side residue pairings in a set of 450,431 residues in 1465 unique peptides and proteins. The P values were calculated by reference to normal error functions centered on the frequencies expected from amino acid residue abundances in the set. C is the maximum P for each plotted group of pairs.

Fig. 2 shows the number of preferred pairings at $P < 0.001$ for each of the 10 sets of 400 pairs with 0–9 residues intervening between the pairs. As expected, the correlations diminish as the residues are farther separated in the peptide chain. The figure also shows these values after omission of all pairs with two identical residues. Except for Met, pairs with identical residues have $P < 0.001$ in most of the 10 cases. After these correlations, which are probably compositional in nature, are omitted, the curve in Fig. 2 shows an expected tendency to decrease toward zero correlations as the distance between the residues increases. The pairings Lys-Glu, Glu-Lys, Pro-Gly, Gly-Pro, Ile-Asn, Asn-Ile are also unusual in that they have $P < 0.001$ in 60% or more of their 20 possibilities. These 6 pairs are included in all the values plotted in Fig. 2.

Fig. 2 also shows interesting maxima in the regions of 3 and 6 intervening residues. The maximum at 3 intervening residues may be caused by the structure of α-helical peptide regions. One turn of α-helix brings the side chains of residues with 3 intervening residues into close proximity with one another. This proximity apparently leads to preferences similar to those found in closer primary structure locations. The maximum at 6 intervening residues is not as simply explained, since two turns of helix require 7 intervening residues for good spatial correlation. The maximum at 6 may be the result of some other commonly occurring three-dimensional configuration.

At the $P < 0.001$ level in side-by-side pairings, 8 of the 43 pairs involve Asn and none involve Gln. Less than 0.04 of each type is expected by chance. Asn pairs are, therefore, much more often specifically selected than expected by chance, whereas Gln pairs do not show this characteristic. Of the 20 naturally occurring residues, Asn is the most preferred residue in the set. Conversely, Gln is the only one of the 20 with no preferred pairings at $P < 0.001$ (other than with itself). The preferred, nonidentical Asn pairs in order from

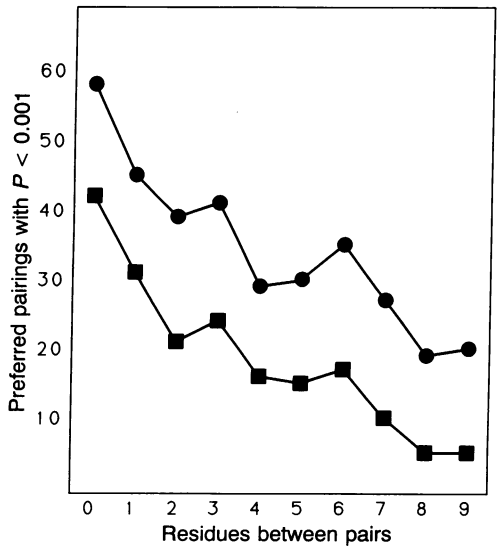

FIG. 2. Number of residue pairings with $P < 0.001$ as compared with random expectation in a set of 1465 unique peptides and proteins comprising 450,431 residues vs. the number of residues between the pairs under consideration. The 400 possible pairings for the 20 common amino acid residues were calculated for each point, comprising 4000 calculations for this graph. Only 0.4 pairing is expected by chance at each plotted point. Solid circles and line show the complete result of the calculations. Solid squares and line show this result with pairings of identical residues subtracted.

least probable to most probable are Asn-Pro, Ile-Asn, Asn-Ile, Tyr-Asn, Lys-Asn, Asn-Tyr, Glu-Asn, and Phe-Asn.

Rejected side-by-side pairings are also interesting. There are 8 Gln pairings and 4 Asn pairings rejected at $P < 0.999$, while less than 0.04 of either type of pairing is expected by chance. It was predicted 20 years ago (6, 7) that a rejection of the more rapidly deamidating peptide sequences in biologically longer-lived peptides would be found. It was suggested that this would introduce a bias against short-lived amide sequences into the distribution of sequences in proteins. At that time, the only experimental evidence consisted of a significantly lower total amide content in a few longer-lived proteins (6).

The 12 rejected sequences of side-by-side residue pairs involving amides that have now been demonstrated at the $P < 0.999$ confidence level are, in order of least probable to most probable, Gln-Ser, Asn-Thr, Thr-Gln, Gln-Asp, Gln-Phe, Asp-Gln, Val-Gln, Asn-Arg, Ser-Asn, Lys-Gln, Pro-Asn, and Gln-Thr. As predicted (6, 7), these are primarily sequences that have been shown in peptide experiments to deamidate rapidly (1, 2).

The sequence Asn-Gly is of special interest, since, as a result of its formation of a succinimide intermediate, it is the most rapidly deamidating sequence known (3). In some peptides Asn-Gly has an inherent primary sequence deamidation half-time under physiological conditions of about 1 day (3). The value $P = 0.0036$ was found for the pair Asn-Gly in the 1465 molecule set. This suggests either that the deamidation of this sequence is suppressed *in vivo* by some mechanism or that this sequence is actively used as a molecular clock. Otherwise, as a result of its marked inherent instability, it would probably be a rejected rather than a preferred sequence.

Amino Acid Residue Triplets. The distributions of residue triplets that include Asn or Gln as the center residue in the set of 450,431 residues are even more striking than the pairings described above. With the probability level chosen at $P > 0.997$ so that about 1 Gln triplet and about 1 Asn triplet are expected by chance, 21 Gln triplets and 8 Asn triplets were found in the sequences.

In order of decreasing P, with the least probable numbers of triplets listed first, the rejected Gln triplets are Asp-Gln-Ala, Asp-Gln-Pro, Pro-Gln-Lys, Glu-Gln-Ser, Asp-Gln-Asp, Arg-Gln-Asp, Gly-Gln-Ala, Lys-Gln-Asp, Phe-Gln-Ser, Leu-Gln-Val, Val-Gln-Ile, Asp-Gln-Gly, Leu-Gln-Ile, Thr-Gln-Glu, Val-Gln-Phe, Ile-Gln-Cys, Thr-Gln-Val, Lys-Gln-Ser, Gly-Gln-Asp, Asp-Gln-Thr, and Ala-Gln-Asp, and the rejected Asn triplets are Ser-Asn-Asp, Ala-Asn-Glu, Glu-Asn-Thr, Gly-Asn-Asp, Ser-Asn-Gln, Pro-Asn-Lys, Gln-Asn-Thr, and Ser-Asn-Thr.

With equal weight given to each triplet residue, the five most rejected neighbors to Asn and Gln in these triplets in order of most to least rejected are Asp, Thr, Ser, Lys, and Glu. This is the result expected from the peptide deamidation-rate experiments carried out between 1970 and 1974 (1, 2, 6) and from the predictions made in 1970 and 1974 (2, 7). The rejected neighbors are those which would decrease the deamidation half-time of peptides and proteins to times that are shorter than the optimum *in vivo* lifetimes of long-lived molecules. Listed in order of decreasing median deamidation rate, the neighboring residues to Asn in the 20 most rapidly deamidating peptides known are (Asn-Gly), Ser, Thr, His, Arg, Asp, Cys, Glu, and Lys (1–3).

This observation of rejection of unstable sequences is not dependent upon the value of P chosen as a division point. With $P < 0.98$ and, therefore, 8 triplets expected by chance, 24 additional Asn triplets are found. This gives a total of 32. In order of decreasing P, the six most common neighbor residues in these triplets are Glu, Thr, Ser, Gln, Arg, and His. An identical treatment for Gln triplets gives neighbor residues of Asp, Phe, Gly, Thr, Lys, and Ala. As expected, the Gln triplet compositions are not as clearly correlated with instability, since Asn peptides, as a group, deamidate more rapidly than Gln peptides (2).

When a parallel calculation for preferred triplets was performed, Lys and Glu shared the five-most-preferred list with Phe, Pro, and Ile. Lys and Glu appeared in both the rejected and preferred lists, depending upon which side of the amide residues they occupied. For example, the triplet Glu-Asn-Gly is preferred at $P = 0.0035$, while Gly-Asn-Glu is rejected at $P = 0.9948$. A Wilcoxon nonparametric pairwise test of the hypothesis that Glu is preferred on the amino (left) side of Asn and rejected on the carboxyl (right) side of Asn in the 38 single amide triplets, including that in which the opposite entity is actually the end of the peptide chain, gave a one-tailed $P = 0.0098$. The four triplets with two Asn or two Glu residues were excluded from this Wilcoxon test.

The Gln and Asn triplet calculations are illustrated in Figs. 3 and 4. These figures show the cumulative distribution function of probabilities for the observed frequencies of the 400 Gln and 400 Asn triplet combinations in the set of 450,431 residues in 1465 unique peptides and proteins. By drawing a straight line through the random portion of this distribution function and observing the vertical intercepts, it is possible to estimate the number of rejected and preferred combinations, although some of these are too weakly correlated to list reliably without inclusion of random sequences as well. These intercepts are at 120 triplets rejected and 70 triplets preferred for the Gln-centered triplets. For Asn triplets, the intercepts are at 50 rejected and 50 preferred. There are, therefore, about twice as many undesirable Gln triplets as there are undesirable Asn triplets. These triplet distribution functions are not as sharply defined as is the pair function shown in Fig. 1, because there are about 20 times more pairs than triplets in each calculation.

While the tendency of protein structures to incorporate fewer unstable amide sequences is evident in these observations, some notable exceptions are observed. The Asn-Gly pair, which deamidates very rapidly, is not strongly rejected. Moreover, Glu-Asn and Lys-Asn sequences are among Asn preferred triplets. We suggest that these sequences may be found to have special usefulness as molecular clocks for short time intervals. This use may tend to overcome the tendency to reject more rapidly deamidating sequences in longer-lived molecules.

The Asn triplets that include Gly show a strong preference for Gly on the right, carboxyl side in expected agreement with

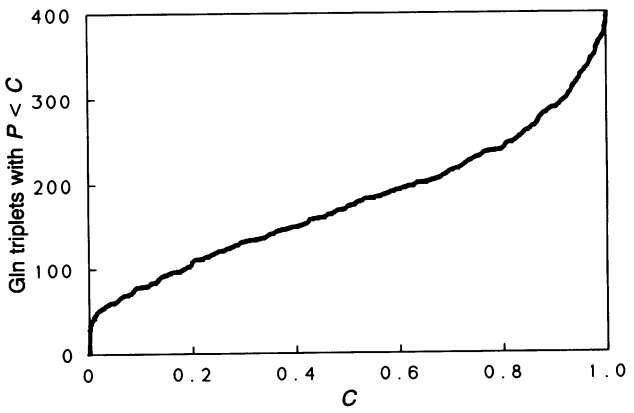

FIG. 3. Cumulative distribution function of probabilities, P, for the observed frequencies of 400 Gln triplet combinations in a set of 450,431 residues in 1465 unique peptides and proteins. The P values were calculated by reference to normal error functions centered on the frequencies expected from amino acid residue abundances in the set. C is the maximum P for each plotted group of triplets.

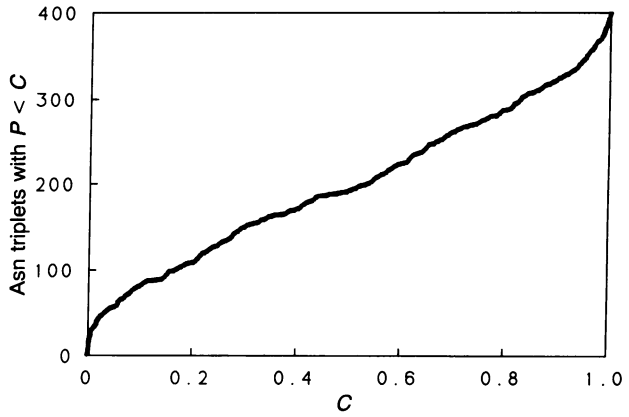

FIG. 4. Cumulative distribution function of probabilities, P, for the observed frequencies of 400 Asn triplet combinations in a set of 450,431 residues in 1465 unique peptides and proteins. The P values were calculated by reference to normal error functions centered on the frequencies expected from amino acid residue abundances in the set. C is the maximum P for each plotted group of triplets.

the calculations of pair frequencies. Moreover, the six most preferred triplets with Gly beside Asn are, in order of decreasing preference, Gly-Asn-Pro, Asn-Asn-Gly, Glu-Asn-Gly, Lys-Asn-Gly, Arg-Asn-Gly, and Ser-Asn-Gly. This preference for acidic, basic, and hydroxyl neighbors reinforces the possibility that Asn-Gly sequences may have special significance as molecular clocks.

Triplet frequencies are more informative than pair frequencies, since the behavior of the central residue is usually determined by its neighbors on both sides. For example, the first use of residue sequence distributions in an experimental study was in the determination of the essential residues in the peptide that induces experimental allergic encephalomyelitis (ref. 9; F. C. Westall, A.B.R., and L.R.R., unpublished work done in 1973). It was predicted and found that the Gly following Trp in the disease-inducing sequence Ser-Arg-Phe-Ser-Trp-Gly-Ala-Glu-Gly-Gln-Lys is an essential residue. The Trp, Gln, and Lys residues are also essential residues in this peptide. While our pair frequencies show a rejection of Trp-Gly and a preference for Gly-Trp, the three triplets Ser-Trp-Gly, Arg-Trp-Gly, and His-Trp-Gly are opposite in correlation and strongly preferred in this 1465-molecule set. Accordingly, the sequence Ser-Trp-Gly proved to be biologically essential. Change of Gly to Ala eliminated biological activity, as was later understood through molecular modeling.

DISCUSSION

These calculations show that amide residues are distributed in peptides and proteins with a higher statistical preference for and rejection of particular sequences than are the other 18 naturally occurring amino acids. Moreover, the observed distributions are understandable in terms of the inherent instability of Gln and Asn residues under *in vivo* solvent conditions.

This treatment does not imply that deamidation rates are set solely by primary structure. It has been theoretically expected and experimentally known for 20 years that secondary and tertiary structure are important in governing deamidation rates (4, 5, 7). Regardless of three-dimensional effects, however, it is possible to discern statistical correlations in the primary-structure-determined part of deamidation rates.

It was suggested (7) that peptides and proteins in general should contain fewer amide residues with short deamidation half-times than are expected by chance. This tendency would result from the requirement that amides with lifetimes shorter than those of specific macromolecules be rejected from use in those molecules. The rejection process, itself, may be one way that *in vivo* molecular lifetimes have been refined (7). The inclusion of specific amides as molecular clocks, however, may increase the overall use of these residues. In this set of 450,431 residues, 4.5% were Asn and 4.3% were Gln. This is a little higher than the 3.1% of each that is expected from random messenger RNA.

In pentapeptide measurements under physiological solvent conditions, deamidation half-times ranging from 1 day to more than 9 years have been measured. Steric hindrance by the side chain of the residue immediately preceding the amide residue was shown in the series of deamidation half-times Gly < Ala < Leu < Phe < Ile < Val with a range of 418–3278 days for Gln peptides and Gly < Ala < Val < Leu < Ile with a range of 87–507 days for Asn peptides. By contrast, a charged side chain in the residue immediately preceding the amide residue increases the rate of deamidation as was shown by the series of half-times His < Asp < Glu < Lys < Arg with a range of 96–389 days for Gln peptides and Arg < Asp < His < Glu < Lys with a range of 18–61 days for Asn peptides. Also, there are catalytic effects of neighboring hydroxyl groups from Thr and Ser (1).

A nearest-neighbor effect on deamidation can be diminished or enhanced by preferences in bond angles impressed upon the amide and its neighbors, by the availability of solvent molecules to the amide residue, by the presence of specific peptide or solvent ions, by side-chain groups that sterically hinder or enhance reaction intermediates or reaction products, and by association of the neighbor groups with other molecules or organelles in the system. These effects depend upon three-dimensional as well as primary structure. Therefore, even if we focus our attention on the residues immediately adjacent to Gln and Asn in the peptide chains, numerous higher-order factors can affect deamidation rate.

As living things are dynamic systems whose molecular and cellular structure undergoes continuous anabolism and catabolism, so also the proteins of which they are constructed are often postsynthetic mixtures of different molecules in different stages of degradation or change.

The static properties of Asn and Gln are not particularly unusual. Since deamidation involves a change of charge, it is especially disruptive of molecular structure. Why should these unstable residues be included in most protein molecules? Our hypothesis is that their instability is in fact their principal biological function—that deamidation serves as a molecular clock for biological processes (6, 7). Deamidation provides a means for changing the structures of peptides and proteins during predetermined, timed intervals after synthesis.

In general, two ideas prevail about each time-dependent biological process. Either specific molecular clocks have been developed to time specific processes or successive time-requiring processes and eventual abandonment of the molecule or organism to deterioration and death occur without specific clocks. With respect to either of these ideas, deamidation may play an important part. If either sort of molecular timing is operative: it is likely that the lifetime of each protein molecule is optimized for the good of its cell or organelle, the lifetime of each cell is optimized for the good of its organ or organism, and the lifetime of each organism is optimized for the good of its species within each competitive environment.

Consider the case of protein turnover. After the apparatus of a cell has manufactured and released a protein molecule, it is unable to evaluate the quality of each molecule or to monitor its individual activity. Overall substrate and reaction product monitoring enables the cell to oversee the protein molecules as a group, but not as individuals. The wide

diversity, large number, and subtle structure of protein molecules make it impractical for a cell to build machinery to check the molecular health and function of each protein molecule and to respond accordingly.

If a protein molecule remains in existence too long, chemical degradation of its structure may render it inefficient. An accumulation of inefficient molecules would lower the specific activity of the protein pool, to the disadvantage of the cell. Moreover, for some enzyme molecules, the amount of enzyme must be frequently adjusted to optimize cell metabolism. How is the cell to remove only certain protein molecules and preferably the older molecules? Postsynthetic deamidation, where an intrinsic lifetime is built into each protein molecule when it is synthesized, can set that lifetime at an optimum compromise between the metabolic effort required to resynthesize the molecule and the disadvantages associated with its remaining in existence too long. Where long protein life is an advantage, a converse process may operate.

Turnover of a protein species can, for example, be timed by a deamidation-mediated increase in susceptibility of the protein to catabolism. This susceptibility could be determined by a change in protein structure, cellular location, or state of aggregation. The amides would provide a planned obsolescence of individual protein molecules. In this way, the specific activity of pools of continuously needed proteins could be maintained and pools of intermittently needed proteins could be controlled.

Development of organisms also requires molecular clocks, so the various developmental processes will occur at the correct times and for the correct durations. These processes may sometimes be timed by deamidation. Aging of organisms is, in some ways, analogous to protein turnover. The intrinsic life-spans of organisms represent a planned or unplanned obsolescence that may be of value to each species. Especially if intrinsic life-span is actively determined, deamidation is a likely candidate for the timing mechanism.

Amide molecular clocks are under simple genetic control, are easily adjustable to biologically important time intervals, and are easily implemented by the changes that they can cause in the proteins in which they are included. If these timed changes are not being put to biological use, then they are surely damaging to the order of biological systems and should have been eliminated long ago.

1. Robinson, A. B. & Rudd, C. (1974) *Curr. Top. Cell. Reg.* **8,** 247–295.
2. Robinson, A. B., Scotchler, J. W. & McKerrow, J. H. (1973) *J. Am. Chem. Soc.* **95,** 8156–8159.
3. Geiger, T. & Clarke, S. (1987) *J. Biol. Chem.* **262,** 785.
4. Flatmark, T. & Sletten, K. (1968) *J. Biol. Chem.* **243,** 1623–1629.
5. Robinson, A. B., McKerrow, J. H. & Legaz, M. (1974) *Int. J. Peptide Protein Res.* **6,** 31–35.
6. Robinson, A. B., McKerrow, J. H. & Cary, P. (1970) *Proc. Natl. Acad. Sci. USA* **66,** 753–757.
7. Robinson, A. B. (1974) *Proc. Natl. Acad. Sci. USA* **71,** 885–888.
8. Robinson, A. B. (1978) *Mech. Ageing Dev.* **9,** 225–236.
9. Westall, F. C., Robinson, A. B., Caccam, J., Jackson, J. & Eylar, E. H. (1971) *Nature (London)* **229,** 22–24.