

quences of hantaviruses Hantaan strains 76118 (GenBank numbers M14627 and Y00386), Lee (D00377), and Hojo (D00376); Seoul strains B-1 (X53861), SR-11 (M34882), and R22 (S68035); Puumala strains CG18-20 (M29979) and Sotkamo (X61034); and Prospect Hill strain PHV-1 (X55129) were aligned with the GAP, PILEUP, and LINEUP programs of the GCG software package (Genetics Computer Group, Madison, WI) run on a VAX computer. Predicted conserved positions for the synthesis of nested RT-PCR oligonucleotide primers for HTN-SEO viruses or PUU-PH viruses were as follows. HTN-SEO first-round primers: +2548 GATATGAATGATTG(T/C)TTTGT and -2859 CCATCAGGGTCT(T/C)TCCA; second-round: +2590 TGTATAATTGGGAC(T/A)GTATCTAA and -2751 GCAAAGTTACATTT(T/C)TTCCT (position numbering of the oligonucleotide 3' terminus was relative to the total aligned and gapped sequence length of 3722 nucleotides); PUU-PH first-round primers: +2671 TTTAAGCAATGGTG(C/T)ACTAC(T/A)AC and -3108 CCATAACACAT(AT)GCAGC; second-round: +2770 AGAAAGAAATGTGCATTTGC and -3012 CCTGAACCCATGC(AT/C)CCATC. Because of the hazardous nature of the agent,

all steps of the homogenization of autopsy tissue samples and the total RNA extraction and purification were performed under biosafety level 3 conditions. RNA extraction, first-round RT-PCR reactions, and subsequent product DNA gel electrophoresis analysis were performed essentially as described [L. L. Rodriguez, G. J. Letchworth, C. F. Spiropoulou, S. T. Nichol, *J. Clin. Microbiol.* **31**, 2016 (1993)], except the following cycle profile run on a Perkin-Elmer 9600 thermocycler was used: 41°C for 1 hour, followed by 40 cycles at 94°C for 40 s, at 38°C for 45 s, and at 72°C for 60 s. Second-round reaction conditions used 3% of the first-round reaction product, no RT step; the following profile was used: 35 cycles at 94°C for 40 s, at 43°C for 45 s, and at 72°C for 60 s.

12. PAUP: Phylogenetic Analysis Using Parsimony, version 3.0 s; D. L. Swofford (Illinois Natural History Survey, Champaign, 1991). The software was run on a SUN SPARC 10 workstation. Two equally parsimonious trees were obtained (with minor branch order differences among the SEO viruses) with the use of the BANDB option. Bootstrap confidence limits were obtained with 10,000 repetitions, and the ALLTREES option was used to

- examine the tree-length frequency distribution.
13. T. G. Ksiazek and P. E. Rollin, unpublished data.
 14. S. Zaki, unpublished data.
 15. *Morb. Mortal. Wkly. Rep.* **42**, 477 (1993).
 16. R. R. Arthur, R. S. Lofts, G. E. Glass, J. W. LeDuc, J. E. Childs, *Am. J. Trop. Med. Hyg.* **47**, 210 (1992).
 17. J. W. LeDuc, *Lab. Anim. Sci.* **37**, 413 (1987).
 18. D. G. Higgins and P. M. Sharp, *CABIOS* **5**, 151 (1989).
 19. We thank the reporting physicians, the Navajo Nation, and the staff of the Indian Health Service, the New Mexico, Arizona, and Colorado state health laboratories, and the Centers for Disease Control and Prevention epidemiologic team and environmental assessment teams for provision of samples and case information. We are also grateful to M. Monroe and S. Trappier for excellent technical assistance and to B. Mahy for support and encouragement. C.F.S. and S.M. are supported by a U.S. Department of Agriculture Animal Molecular Biology National Research Initiative grant 90-37266-5473 through the University of Nevada, Reno. H.F. is supported by a National Research Council fellowship.

9 August 1993; accepted 4 October 1993

TECHNICAL COMMENTS

Alpha Helix Propensity of Amino Acids

Michael Blaber *et al.* state that they have found a correlation between changes in stability for substitution mutants of T4 lysozyme at two positions in α helices and the amount of nonpolar surface buried by the substituted side chain when it folds to the native structure (1). If this correlation is real, their work represents a significant advance in our understanding of the structural basis for the different α helix propensities of the amino acids. However, before the validity of the correlation [figure 2 of (1)] can be properly evaluated, several technical issues should be addressed.

For a statistical correlation involving a small data set to have scientific significance, it must be generally true. A provisional test of the proposed correlation for a subset of the amino acid residue (2) substitutions (L, V, I, S, T, K, E, and N) at position 44 can be made from examination of the substitutions at a second site, position 131. Although data are presented for only five substitutions at this position (L, V, I, S, and T), there is no significant correlation between free energy differences ($\Delta\Delta G$) and buried surface area.

The correlation between the free energy of folding and buried surface area for substitutions at position 44 depends on the exclusion of three data points in addition to those for A, G, and P. The exclusion of R because of a crystal contact seems justifiable (though K should then be excluded for the same reason). However, the logic for discarding the critical F and W points [described in note 22 of (1)] is unclear. Blaber *et al.* appear to argue that, because these

two mutants crystallize in space groups that are different from that of the other mutants, the observed trans χ_1 angle (which distinguishes these side chains from all others except the wild-type S residue) may be an artifact of the different crystal environments. The fact that the conformation of tryptophan is similar in all four asymmetric environments (and is similar to that of phenylalanine in each of its two asymmetric environments) suggests that this rotamer is not an artifact of crystallization.

Finally, we would caution against the deceptive ease with which statistical correlations can be made. Most of the data presented by Blaber *et al.* falls into a fairly narrow range of values for both $\Delta\Delta G$ and for nonpolar surface buried. Nine of 17 residues (excluding G, A, and P) could have, within the experimental error of ± 0.1 kcal/mol, the same $\Delta\Delta G$ value of +0.63 kcal/mol. Even more striking is the narrow range into which most surface area values fall. Particularly noteworthy are the similarities of buried area for A, V, E, T, and S and of calculated areas for extended side chains of Q, R, K, Y, H, and F. This might be expected because, in many cases, the β carbon atom often makes the greatest contribution to buried nonpolar surface. As a consequence of this tight clustering of values within a narrow range, it only takes one or two outlying data points to establish an apparent correlation. This reduction in the number of data points that significantly contribute to a correlation increases the ever-present danger that an apparent correlation might be found by omitting data

points from an otherwise random-looking scatterplot. Only when a rigorous and compelling argument is at hand can one, in the search for a better correlation, safely exclude a subset of the data.

While burial of side chain hydrophobic surface may play a role in determining the rank order and magnitude of helix propensity, the correlation reported by Blaber *et al.* does not convincingly establish that it is the structural basis of amino acid α helix propensity.

David Shortle

Department of Biological Chemistry,
Johns Hopkins University School of Medicine,
Baltimore, MD 21205-2185

Neil Clarke

Department of Biophysics and
Biophysical Chemistry,
Johns Hopkins University School of Medicine

REFERENCES AND NOTES

1. M. Blaber, X. Zhang, B. W. Matthews, *Science* **260**, 1637 (1993).
2. Abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

9 July 1993; accepted 14 September 1993

Response: Shortle and Clarke correctly state that if a statistical correlation involving a small data set is to have scientific significance, it must be generally true. We have tested (1) the correlation in question by comparing the area buried on helix formation [as calculated by Richards and Richmond (2)] with experimentally determined scales of helix propensity that are based on substitutions in proteins (1, 3, 4), model peptides (5), host-guest experiments (6), and frequencies observed in known protein structures (7). In every case the correlation was positive, with an average value of 0.49

excluding glycine, and 0.64 if glycine is included (1). Thus, the suggested correlation between buried surface area and helix propensity extends well beyond the data presented in our report (3).

With respect to the consistency between the site 131 and site 44 substitutions, as we noted in our report (3), the respective $\Delta\Delta G$ values are in excellent agreement (correlation of 0.97, excluding Asp¹³¹ and Glu¹³¹, which appear to form a salt bridge). The site 131 variants for which crystal structures are available is limited and, as noted by Shortle and Clarke, in several cases their surface areas and $\Delta\Delta G$ values cluster together. Nevertheless, the site 131 data, insofar as they are available, are in good agreement with those at site 44.

Shortle and Clarke suggest that we discarded data for Phe⁴⁴ and Trp⁴⁴ because the *trans* conformations ($\chi_1 \sim 180^\circ$) observed for these side chains may be an artifact of different crystal environments. To the contrary, as we have confirmed from a survey of 100 well-determined structures in the Brookhaven Protein Data Bank (1), the conformations adopted by Phe⁴⁴ and Trp⁴⁴ in the mutant lysozymes correspond to the most commonly observed conformations for Phe and Trp within α helices in general. Because Trp⁴⁴ and Phe⁴⁴ adopt the *trans* conformation, they make contacts with the side chains of Glu⁴⁵ and Lys⁴⁸. These contacts are tenuous [note 22 in (3)], and we therefore showed the result of both including and deleting them from the surface area calculations (3).

We emphasized in our report that calculations of surface area are very sensitive to a number of uncertainties. We also noted that the expected changes in solvent-accessible surface area and attendant hydrophobic stabilization are small (less than 1 kcal/mol). The major uncertainty in the surface area calculations probably does not come from errors in the crystallographic coordinates, as suggested by Shortle and Clarke, but from other factors. In surface area calculations (2, 8), a static model is assumed, whereas proteins are well known to be mobile. A static model may suggest that a pair of mobile side chains on the surface of a protein are in contact and therefore are partly inaccessible to solvent. This contact may, however, be transient, in which case the calculated inaccessibility to solvent would not represent the average behavior of the side chains (compare with the discussion above regarding Trp⁴⁴ and Phe⁴⁴). Another major uncertainty relates to the appropriate estimation of solvent-exposed area in the unfolded protein (3).

The generally good agreement between "helical propensity" scales obtained by a variety of different methods suggests that there is an underlying physical basis for these values. In our report (3), the lack of an ideal site at which to perform the analysis, the inherently narrow range of thermodynamic values, and the assumptions made with regard to the nature of the folded and unfolded states are rightly expected to cause perturbations from ideality. The critical test of any conclusion, therefore, relies

upon consistent observation within a larger body of data. The available data (1) indicate that for many amino acids the hydrophobic effect is a primary contributor to observed helix propensity values. Other factors, including side chain entropy, appear to be less important for the majority of residues.

Michael Blaber
Xue-jun Zhang
Brian W. Matthews

Institute of Molecular Biology and
Howard Hughes Medical Institute,
University of Oregon,
Eugene, OR 97403

REFERENCES

1. M. Blaber *et al.*, *J. Mol. Biol.*, in press.
2. F. M. Richards and T. Richmond, *Ciba Found. Symp.* **60**, 23 (1978).
3. M. Blaber, X. Zhang, B. W. Matthews, *Science* **260**, 1637 (1993).
4. A. Horovitz, J. M. Matthews, A. R. Fersht, *J. Mol. Biol.* **227**, 560 (1992).
5. A. Chakrabarty and R. L. Baldwin, in *Protein Folding: In Vivo and In Vitro*, J. Cleland and J. King, Eds. (American Chemical Society Symposium Series 526, Washington, DC, 1993), pp. 166-177; E. Stellwagen, S.-H. Park, W. Shalongo, A. Jain, *Biopolymers* **32**, 1193 (1992); P. C. Lyu, M. I. Liff, L. A. Marky, N. R. Kallenbach, *Science* **250**, 669 (1990); K. T. O'Neil and W. F. DeGrado, *ibid.*, p. 646.
6. J. Wojcik, K.-H. Altmann, H. A. Scheraga, *Biopolymers* **30**, 121 (1990).
7. R. W. Williams, A. Chang, D. Juretic, S. Loughran, *Biochim. Biophys. Acta* **916**, 200 (1987).
8. B. Lee and F. M. Richards, *J. Mol. Biol.* **55**, 379 (1971).

26 July 1993; accepted 14 September 1993

AAAS-Newcomb Cleveland Prize

To Be Awarded for a Report, Research Article, or an Article Published in *Science*

The AAAS-Newcomb Cleveland Prize is awarded to the author of an outstanding paper published in *Science*. The value of the prize is \$5000; the winner also receives a bronze medal. The current competition period began with the 4 June 1993 issue and ends with the issue of 27 May 1994.

Reports, Research Articles, and Articles that include original research data, theories, or syntheses and are fundamental contributions to basic knowledge or technical achievements of far-reaching consequence are eligible for consideration for the prize. The paper must be a first-time publication of the author's own work. Reference to pertinent earlier work by the author may be included to give perspective.

Throughout the competition period, readers are invited to nominate papers appearing in the Reports, Research Articles, or Articles sections. Nominations must be typed, and the following information provided: the title of the paper, issue in which it was published, author's name, and a brief statement of justification for nomination. Nominations should be submitted to the AAAS-Newcomb Cleveland Prize, AAAS, Room 924, 1333 H Street, NW, Washington, DC 20005, and **must be received on or before 30 June 1994**. Final selection will rest with a panel of distinguished scientists appointed by the editor of *Science*.

The award will be presented at the 1995 AAAS annual meeting. In cases of multiple authorship, the prize will be divided equally between or among the authors.