47. C. F. Bohren and D. R. Huffman, *Absorption and Scattering by Small Particles* (Wiley, New York, 1983).
48. W. Stockton *et al.*, *J. Appl. Phys.* **70**, 4679 (1991).
49. F. Behroozi *et al.*, *ibid.* **68**, 3688 (1990).
50. A. Lackner, K. C. Kim, J. D. Margerum, E. Sherman, *Liq. Cryst.*, in press; paper presented at the 14th International Liquid Crystal Conference, Pisa Italy, June 1992.
51. J. J. Krebs *et al.*, *J. Appl. Phys.* **70**, 6404 (1991).
52. G. A. Prinz, *Science* **250**, 1092 (1990).
53. W. Roberson, *Proc. Soc. Photo-Opt. Instrum. Eng.* **453**, 320 (1983).
54. D. A. Kirkpatrick, G. Bekefi, A. C. diRenzo, H. P. Freund, A. L. Ganguly, *Phys. Fluids B* **1**, 1511 (1989).
55. R. L. Sheffield, E. R. Gray, J. S. Fraser, *Nucl. Instrum. Methods* **A272**, 222 (1988).
56. D. A. Kirkpatrick *et al.*, *Appl. Phys. Lett.* **60**, 1556 (1992).
57. C. M. Chow *et al.*, *Mater. Sci. Eng.* **A158** (1992).
58. J. M. Calvert *et al.*, *Thin Solid Films* **210/211**, 359 (1992); J. Calvert *et al.*, *Solid State Technol.* **34**, 77 (1991).
59. S. E. Shafik, in *Biopolymeric Controlled Release Systems*, D. E. Wise, Ed. (CRC Press, Boca Raton, FL, 1984), vol. 2, pp. 139–156.
60. J. M. Schnur and A. S. Rudolph, paper presented at the Sixth International Symposium on Recent Advances in Drug Delivery Systems, Salt Lake City, UT, 22 to 25 February 1993; *J. Controlled Release*, in press.
61. R. Price and M. Patchan, *J. Microencapsulation* **8**, 301 (1991).
62. R. Price, M. Patchan, A. Clare, D. Rittschof, J. Bonaventura, *Biofouling* **6**, 207 (1992).
63. R. Price, *J. Microencapsulation*, in press.
64. J. Selinger, personal communication.
65. P. L. Ahl, R. Price, J. Smuda, B. P. Gaber, A. Singh, *Biochim. Biophys. Acta* **1028**, 141 (1990).
66. The work described in this paper is the result of several years of research on a coherent theme by many people in a number of different laboratories. Space limitations make it impossible to thank each of them in this section. I thank T. Coffey for his sense of vision and foresight that led to the establishment of a multidisciplinary research group in which this type of work could be pursued. I. Skurnick was a source of excellent advice and provided a sounding board for many of the concepts. The former DARPA, the former ONT, and ONR all provided financial support. R. Shashidhar, B. Ratna, A. Singh, J. Selinger, and S. L. Schnur were especially helpful in their careful reading and critical comments.

Biological Diversity, Soils, and Economics

Michael Huston

Terrestrial biological diversity is supported by solar energy captured by plants growing in soil. This soil-based plant productivity also provides the foundation for human societies through production of food and renewable forms of energy. Variations in plant productivity, resulting from differences in inherent soil fertility, variations in climate and weather, and differences in chemical inputs and agricultural practices, produce patterns of biological diversity that are associated with the agricultural component of economic productivity. Ecological processes lead to a generally negative relation between the diversity of plant species and potential agricultural productivity at both local and global scales. One implication of this negative relation is that preservation of areas of high plant biodiversity does not require the sacrifice of productive agricultural lands.

The relation of biological diversity to economics is generally seen as an impediment imposed on economic development or as a difficult-to-value externality determined by the opportunity-cost of extinction. Both of these views miss the true economic significance of biological diversity, and their perpetuation has negative implications both for the preservation of biodiversity and for rational economic development. Biological diversity can be understood as the consequence of organisms struggling for survival in the same environment in which humans, as individuals and societies, struggle for survival.

The fundamental unit of biodiversity is the species. Extinction of species, as well as of subspecies and populations, is of international concern for many reasons, including the loss of genetic material of potential value for medicine and agriculture (*1*). Biodiversity is considered to be highest where there are the most species living in the same area, as exemplified by tropical rain forests. Thus, biodiversity is an issue of coexistence, and extinction can be seen as the extreme manifestation of a failure to coexist. Extinction can result from (i) the failure of organisms to survive natural disasters such as floods, droughts, ice ages, and asteroid impacts; (ii) the failure of organisms to coexist with one another as a result of competition, predation, or disease; and (iii) the failure of organisms to coexist with humans as the result of overharvesting, habitat destruction, or direct competition.

Local Patterns of Biodiversity

The processes leading to "survival of the fittest" form the basis for understanding the economic implications of biological diversity. The "fittest" species, which in many cases are the best competitors, are able to dominate more rapidly and completely under some environmental conditions than under others. Dominance by a single species often occurs more rapidly under favorable conditions, where all organisms can potentially grow rapidly. It is under these conditions that the best competitors are able to eliminate most other species by monopolizing an essential resource. Abiotic disturbances, such as fires or hurricanes, as well as predators or herbivores can prevent competitive dominance by killing some proportion of the dominant competitors, particularly under favorable conditions when dominance would otherwise occur rapidly (*2*). Thus, under the most favorable conditions, diversity among competing species can be reduced by competition. However, under somewhat poorer conditions, the superior competitors are not able to dominate as rapidly, if at all, and many species are able to coexist, leading to higher diversity among individuals that are smaller in size, fewer in number, or both. Of course, under extremely poor conditions, few or no species are able to survive, and biodiversity drops to very low levels.

The pattern described above appears most strongly among plants, primarily because of the importance of a single limiting resource, sunlight. The conclusion from the above argument is that plant biodiversity should be higher on unproductive, poor soils than on fertile, productive soils where plants are taller and total plant mass is higher. This pattern of highest plant diversity on poor soils and low plant diversity on the best soils is found throughout the world under a wide variety of conditions (*3–5*). The diversity of plant species in temperate prairies and grasslands (Fig. 1, A to C) (*3–6*) shows the same pattern in relation to soil fertility that is found among trees in tropical forests (Fig. 1, D to F) (*4*, *7–9*). Addition of nutrients to plant communities nearly always leads to a reduction in plant diversity (*3–6*, *10*), as demonstrated by fertilization experiments in temperate pastures and grasslands (*10*) and tropical forests (*11*). Anthropogenic nutrient pollution also reduces plant diversity, both in aquatic systems, where the effects of eutrophication are well known, and in terrestrial systems, such as the European heathlands that are being converted to grasslands by atmospheric deposition of nitrogen (*12*). A variety of theoretical models predict this pattern and explain why there are exceptions to the general rule (*2*, *3*, *5*, *6*, *12*).

Agriculture, Biodiversity, and Conservation

Soil is also the key to explaining much of the variation in agricultural productivity. Situations where agricultural productivity is in-

---

The author is in the Environmental Sciences Division, Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, TN 37831.

herently high are situations where the natural biodiversity of plants would be expected to be relatively low. Thus, preserving a large proportion of the plant biodiversity in the world should not have much economic or social cost because it does not require a sacrifice of agriculturally productive land. On the other hand, areas with high plant biodiversity are most likely to be environments where agricultural productivity is naturally low, as a result of low nutrient availability, shallow rocky soils, insufficient rainfall, or unfavorably high temperatures. Consequently, regions with high plant biodiversity are unlikely to be able to support productive agriculture without intensive management and chemical inputs (*13*).

This phenomenon has a number of important consequences. One is that societies tend to preserve areas of high plant biodiversity purely as a result of market economics. Successful farmers do not waste resources trying to grow crops on land that is not capable of sustained high productivity agriculture with cost-effective management and chemical inputs (*14*). Nearly all of the most productive land around the world has been under some form of cultivation for many centuries, and sometimes thousands of years. Whatever ecosystems existed on these soils have long since disappeared, except for a few accidental remnants. While many plant and animal species undoubtedly became extinct as a result, we can expect that the loss of plant diversity was relatively low, because most lands taken for agriculture were likely to have had low plant diversity. In contrast, much of the recent destruction of biodiversity in the tropics and elsewhere has been the result of government subsidies and tax incentives that allowed the deforestation or plowing of low productivity (and high diversity) lands that would not support agriculture in a true market economy and usually are abandoned or sold after the initial subsidies have been obtained (*15*).

The market-driven pattern of biodiversity preservation can be clearly seen in the distribution of public land in the United States (Fig. 2A). The relative amount of land owned by the federal government is inversely related to the average value of agricultural products produced by the land on a per area basis. These lands are the national parks, national forests, wildlife refuges, and the other public lands that receive varying degrees of protection from exploitation and are the primary refuges for the remaining biodiversity of the United States.

This same pattern can also be seen on a global scale (Fig. 2B), where a relatively large proportion of land in some countries is set aside for conservation in spite of potentially high agricultural value. Among countries with agricultural production values over $200 per hectare, between 0.5 and 18% of the land area is protected, whereas in countries with production values less than $200 per hectare from 0 to 38% of the land is officially protected for conservation purposes. The potentially higher conservation set-asides that are economically justifiable in countries with low agricultural productivity are often not achieved.

Global Biodiversity and Soil Fertility

The inverse or unimodal association between plant diversity and soil fertility that is found along local or regional gradients (Fig. 1) and in fertilization experiments also appears on a global scale. One of the major patterns in the distribution of life on Earth is the latitudinal gradient of species diversity. Among plants, as well as among most groups of animals, the number of species found in an area of a given size (for example, 1 hectare) increases from either pole toward the equator, reaching a maximum in wet lowland tropical rain forests (*16*).

Several independent types of data suggest that a latitudinal gradient of soil fertility parallels the global pattern of species diversity. Global soil maps show that poor soils dominate the tropical latitudes (although soils with similar properties are also found in parts of the temperate zone), whereas the most fertile soils are limited to certain areas of the temperate zone (Fig. 3). Most areas of tropical rain forest are underlain by soils classified as infertile (*17*).

Many tropical soils are low in inherent fertility because they formed from geological parent materials that were low in essential elements or because they have lost most of their nutrients and become

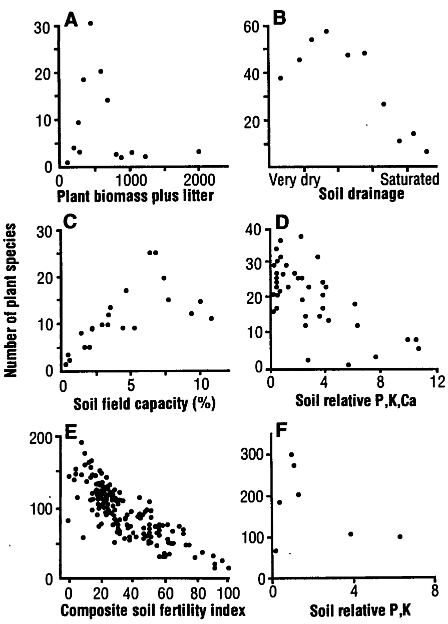

Fig. 1. Association of plant species number (the simplest measure of plant diversity) with parameters of soil fertility, showing the unimodal pattern of plant diversity along gradients of soil fertility and plant productivity. Position of the mode (maximum diversity) varies among the graphs, at least partially as a result of different scalings imposed by use of different parameters as correlates of soil fertility. The association between species number and productivity is negative along part or all of each productivity gradient. (**A**) British herbaceous plants, plant productivity indexed by sum of living and dead plant biomass [reprinted from (*23*)]; (**B**) North American prairies, plant productivity indexed by soil drainage conditions [based on (*24*)]; (**C**) annual plants in an Israel desert, plant productivity indexed by soil water-holding (field) capacity [reprinted from (*25*)]; (**D**) number of Costa Rican rain forest trees species per 0.1 ha, soil fertility indexed as the sum of relative soil P, K, and Ca among 46 sites [reprinted from (*4*)]; (**E**) number of vascular plant species per 0.1 ha in Ghanaian closed forest, soil fertility indexed by a principal components axis negatively correlated with total exchangeable bases and other measures of soil fertility [reprinted from (*9*)]; (**F**) number of tree species per hectare in West Malesian rain forest, soil fertility indexed as sum of relative soil P and K [based on (*26*)].

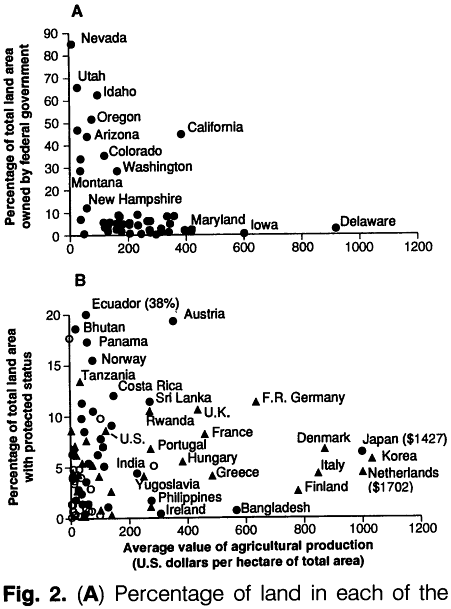

Fig. 2. (**A**) Percentage of land in each of the lower 48 states in the United States that is owned by the federal government as public lands (such as national forests, national parks, and national rangelands) in relation to the average value of agricultural production per hectare of total land area in each state [reprinted from (*5*)]. (**B**) Percentage of land with officially protected status (such as national parks or refuges) in various countries in relation to the average value of agricultural production per hectare of total land area in each country. Values are given in parentheses for countries that are off the scale. Symbols indicate approximate annual rainfall; open circles indicate less than 55 cm of rainfall per year, triangles indicate 55 to 130 cm per year, and closed circles indicate over 130 cm per year [based on data from (*27*)].

acidic as a result of warm temperatures that accelerate chemical and biological processes, and high rainfall that leaches nutrients out of the soil (*17, 18*). Geological stability and the lack of glaciation in much of the lowland tropics have reduced the input of fresh, mineral-rich substrate for soil formation. High fertility soils in the tropics are generally limited to areas of active volcanism or alluvial sediments from young mountain ranges (*19, 20*).

An additional measure of soil fertility is the amount of crop that can be harvested from a specific area of land. Data on both the standardized monetary value of agricultural production (Fig. 4A) and the actual weight of root crops (such as potatoes, yams, and manioc) harvested from a specific area of land (Fig. 4B) show a clear decrease in maximum agricultural productivity from countries in the temperate zone to countries near the equator. The data in Fig. 4 are divided into three classes on the basis of each country's average annual rainfall. While high root crop productivity can be achieved regardless of annual rainfall, the highest productivity at any latitude is found in countries that are able to afford irrigation and fertilizer.

Fertilization and good agricultural practices can increase productivity in tropical as well as temperate soils (*20*). Nonetheless, the inherent physical and chemical properties of a soil impose limits on the increases in yield that can be achieved (*21*). Other factors in addition to poor soils, such as low light owing to cloudiness and high plant respiration rates caused by elevated temperatures, may also contribute to low productivity of agriculture and forestry in parts of the tropics (*22*).

Global Economic Implications

The latitudinal gradient of soil fertility has global economic significance because most developing countries are located in the tropical latitudes. One tragic consequence of this pattern is that rural poverty is likely to be much more severe in tropical than temperate countries. The general pattern of tropical poverty is reflected by the latitudinal gradient of per capita gross national production (GNP) (Fig. 5A), which is the opposite of the latitudinal diversity gradient. The above arguments suggest that both of these patterns may have the same root cause.

This global pattern may suggest a sort of environmental determinism that dooms some tropical countries to poverty, but such a pattern is not inevitable. The problem for much of the Third World is not their climate or poor soils per se, but rather that many Third World economies are based disproportionately on agriculture (Fig. 5B). Weak economies and inherently low soil fertility are unable to support the investment needed to improve agricultural productivity. Many developed nations of the temperate zone and arid regions achieve

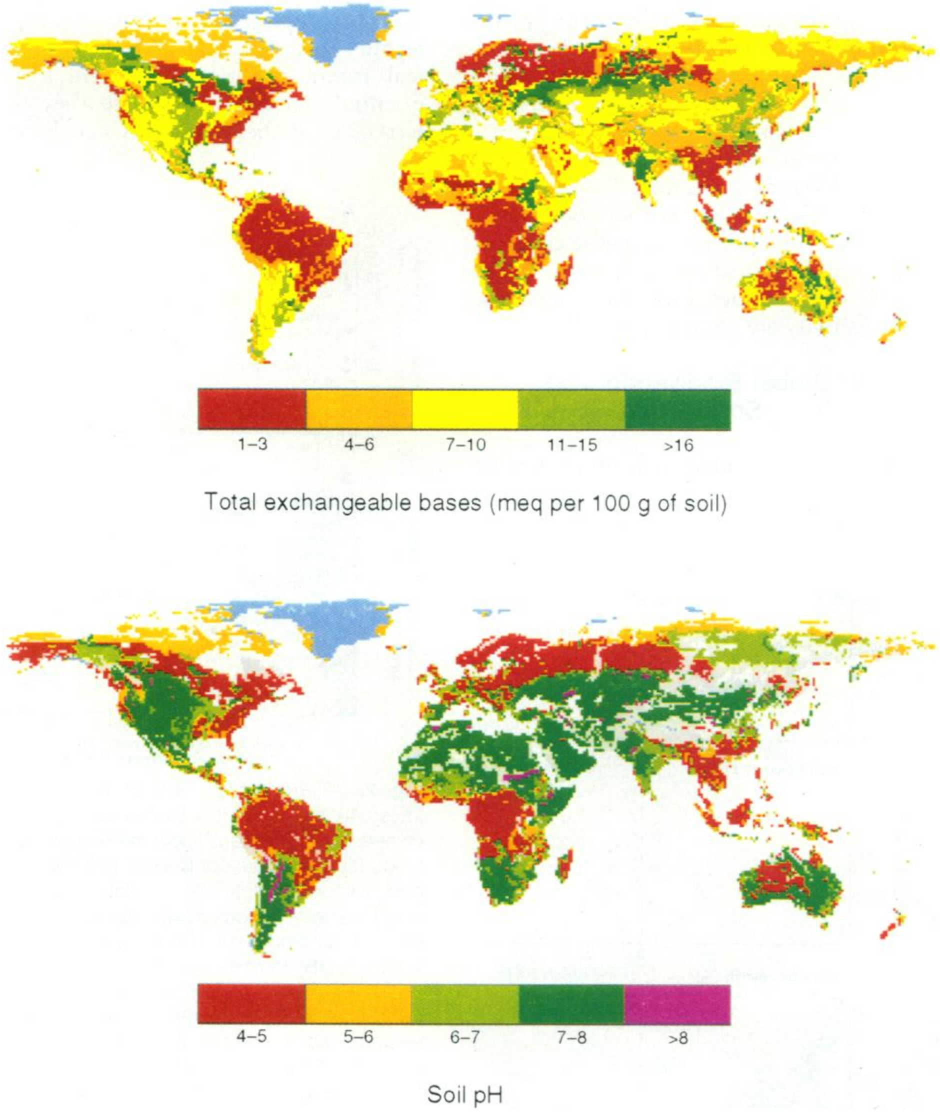

Fig. 3. Global patterns of soil properties related to fertility. (**Top**) Estimated total exchangeable bases (K, Mg, Ca, Na) of soils in 1° by 1° cells; (**bottom**) average pH in upper 100 cm of soil. Information on each of these soil properties, which are expected to be highly correlated, is derived from independent data (*28*). Gray indicates soil types for which pH data are inadequate. Light blue indicates ice.

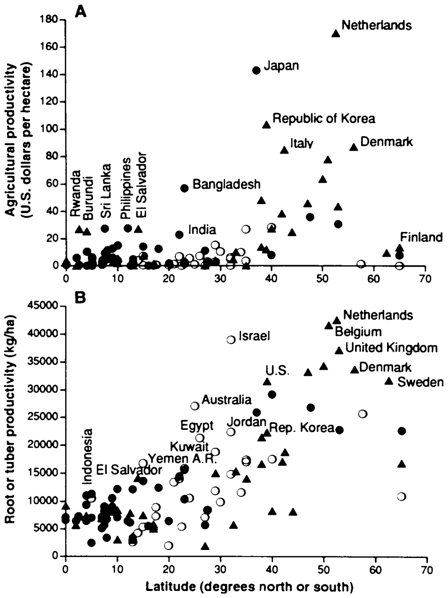

Fig. 4. (**A**) Global latitudinal gradient of agricultural productivity by country in terms of the value of agricultural products in U.S. dollars per hectare of total land area. (**B**) Global latitudinal gradient of the yield of root and tuber crops (such as yams, potatoes, and manioc) in kilograms per hectare of land used for these crops. Symbols indicate approximate annual rainfall as in Fig. 2 [reprinted from (*5*) based on data from (*27*)].

their high agricultural productivity as a result of heavy investment in fertilization and irrigation (see Fig. 4), which is an option not available to most underdeveloped countries under current global economic conditions.

Agricultural productivity based on inherent soil fertility is probably inadequate to finance the transition from an agrarian to an industrial society in many tropical countries, as it did for the United States and much of Europe. Fortunately, such an agriculturally driven transition should not be necessary in this age of global trade and demand for a wide variety of resources and services. Unlike the global pattern of soil fertility, there is no such pattern for mineral resources such as coal, uranium, oil, and gas (Fig. 5C). These, and presumably all other mineral resources, are distributed without regard to latitude. A stronger emphasis on mineral resources and industrialization on the basis of these natural resources seems more likely to lead to successful economic development than dependence on agriculture.

The predictable patterns of plant biodiversity in relation to potential agricultural productivity on both local and global scales suggest that much of the Earth's remaining plant biodiversity can be preserved with relatively little cost in terms of agricultural productivity. The relation of other components of biodiversity to soil fertility is not as well understood. The diversity of organisms that feed on specific plants or plant parts, such as many insect and bird species, is likely to be closely related to plant diversity. However, there are some important components of biodiversity, such as large marine vertebrates, and predatory birds and mammals, that seem to reach their highest diversity in productive areas. Conservation of these components of biodiversity may require a decision to set aside land that could otherwise be used for productive agriculture. These global patterns also imply that the optimal solution to preserving the biological resources contained in the Earth's biodiversity, while simultaneously maximizing the sustainable production of food and energy provided by plants, cannot be achieved within any single country. Movement toward an integrated global economy based on efficient production of agricultural and industrial products and unrestricted trade might actually improve conservation of biodiversity.

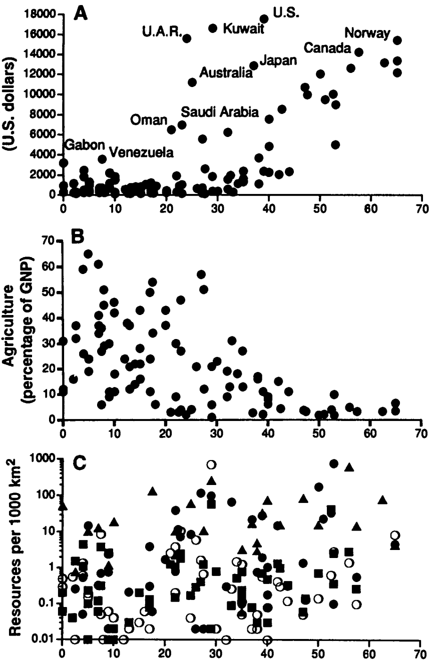

Fig. 5. (**A**) Global latitudinal gradient of per capita GNP. (**B**) Global latitudinal gradient of the proportion of GNP contributed by the agricultural sector. (**C**) Global pattern of mineral energy resources. Closed circles indicate bituminous coal in units of million metric tons. Open circles indicate crude oil in units of million tons. Squares indicate natural gas in units of billion cubic meters. Triangles indicate uranium in units of metric tons [reprinted from (*5*) based on data from (*27*)].

REFERENCES AND NOTES

1. N. Myers, *The Sinking Ark* (Pergamon, Oxford, 1979); *A Wealth of Wild Species* (Westview, Boulder, CO, 1983); E. O. Wilson, *The Diversity of Life* (Harvard Univ. Press, Cambridge, MA, 1992); P. Ehrlich and A. Ehrlich, *Ambio* **21**, 219 (1992).
2. R. T. Paine, *Am. Nat.* **100**, 65 (1966); J. H. Connell, *Science* **199**, 1302 (1978); M. A. Huston, *Am. Nat.* **113**, 81 (1979).
3. J. P. Grime, *J. Environ. Manage.* **1**, 151 (1973); *Plant Strategies and Vegetation Processes* (Wiley, Chichester, United Kingdom, 1979).
4. M. A. Huston, *J. Biogeogr.* **7**, 147 (1980).
5. ———, *Biological Diversity: The Coexistence of Species on Changing Landscapes* (Cambridge Univ. Press, Cambridge, in press).
6. D. Tilman, *Resource Competition and Community Structure* (Princeton Univ. Press, Princeton, NJ, 1982); *Plant Strategies and the Dynamics and Structure of Plant Communities* (Princeton Univ. Press, Princeton, NJ, 1988).
7. D. G. Campbell, D. C. Daly, G. T. Prance, U. N. Maciel, *Brittonia* **38**, 369 (1986).
8. C. F. Jordan, *Nutrient Cycling in Tropical Forest Ecosystems* (Wiley, Chichester, United Kingdom, 1985).
9. J. B. Hall and M. D. Swaine, *J. Ecol.* **64**, 913 (1976).
10. J. B. Lawes and J. H. Gilbert, *Philos. Trans. R. Soc. London* **171**, 189 (1880); ———, M. T. Masters, *ibid.* **173**, 1181 (1882); R. A. Kempton, *Biometrics* **35**, 307 (1979); M. A. Huston, *Bull. Ecol. Soc. Am.* **61**, 110 (1980); D. Tilman, *Ecol. Monogr.* **57**, 189 (1987).
11. P. A. Harcombe, *Ecology* **58**, 1375 (1977); M. A. Huston, thesis, University of Michigan (1982).
12. F. Berendse and W. T. Elberse, in *Perspectives on Plant Competition*, J. Grace and D. Tilman, Eds. (Academic Press, New York, 1990), pp. 93–116; F. Berendse, M. J. M. Oomes, H. J. Altena, W. T. Elberse, *Biol. Conserv.* **62**, 59 (1992).
13. The success of large-scale plantation forestry efforts is strongly influenced by soil conditions [J. Evans, *Plantation Forestry in the Tropics* (Clarendon, Oxford, 1992)]. Some plantations, such as Jari, may have failed for soil-related reasons [P. M. Fearnside and J. M. Rankin, *Interciencia* **7**, 329 (1982); *ibid.* **10**, 121 (1985)]. Soil conditions may also underlie the failures of numerous large-scale resettlement schemes in the Amazon basin, where rain forest has been cleared for agriculture then abandoned within a few years because the lands were unable to sustain continuous subsistence farming [P. M. Fearnside, in *The Geophysiology of Amazonia*, R. E. Dickinson, Ed. (Wiley, New York, 1987), pp. 37–61; in *The Earth in Transition: Patterns and Processes of Biotic Impoverishment*, G. M. Woodwell, Ed. (Cambridge Univ. Press, Cambridge, 1990), pp. 211–238].
14. E. Matthews, *J. Climatol. Appl. Meteorol.* **22**, 474 (1983); E. Veldkamp, A. M. Weitz, I. G. Staritsky, E. J. Huising, *Land Degrad. Rehabil.* **3**, 71 (1992).
15. D. J. Mahar, *Government Policies and Deforestation in Brazil's Amazon Region* (World Bank Wildlife Fund, Conservation Foundation, Washington, DC, 1989); R. Repetto, in *Public Policies and the Misuse of Forest Resources*, R. Repetto and M. Gillis, Eds. (Cambridge Univ. Press, New York, 1988).
16. T. Dobzhansky, *Am. Sci.* **38**, 209 (1950); G. G. Simpson, *The Geography of Evolution* (Chilton, Philadelphia, 1965); E. R. Pianka, *Am. Nat.* **100**, 33 (1966).
17. D. D. Richter and L. I. Babbar, *Adv. Ecol. Res.* **21**, 316 (1991).
18. H. Jenny, *The Soil Resource* (Springer-Verlag, New York, 1980); N. Craig and P. Halais, *Emp. J. Exp. Agric.* **2**, 349 (1934).
19. E. J. Fittkau, J. Irmler, W. J. Junk, F. Feiss, G. W. Schmidt, in *Tropical Ecological Systems*, F. B. Golley and E. Medina, Eds. (Springer-Verlag, New York, 1975); W. J. Junk and K. Furch, in *Amazonia*, G. T. Prance and T. Lovejoy, Eds. (Pergamon, Oxford, 1985).
20. P. A. Sanchez, *Properties and Management of Tropical Soils* (Wiley, New York, 1976); P. A. Sanchez and S. W. Buol, *Science* **188**, 598 (1975).
21. R. A. Olson, in *Handbook of Agricultural Productivity*, M. Rechigal, Ed. (CRC Press, Boca Raton, FL, 1982); O. P. Engelsted *et al.*, Eds., *Fertilizer Technology and Use* (Soil Science Society of America, Madison, WI, 1985); S. L. Tisdale and W. L. Nelson, *Soil Fertility and Fertilizers* (Macmillan, New York, 1985).
22. H. C. Dawkins, *Emp. For. Rev.* **38**, 175 (1959); *Oxford Univ. for. Soc. J. Fifth Series* **12**, 1 (1964); P. W. Richards, in *Biometeorology*, S. W. Tromp, Ed. (Macmillan, New York, 1962); J. Chang, *Geogr. Rev.* **58**, 333 (1968); J. L. Monteith, *J. Appl. Ecol.* **9**, 747 (1972).
23. M. M. Al Mufti, C. L. Sydes, S. B. Furness, J. P. Grime, S. R. Band, *J. Ecol.* **65**, 759 (1977).
24. R. Dix and F. Smeins, *Can. J. Bot.* **45**, 21 (1967).
25. P. Kutiel and A. Danin, *Vegetatio* **70**, 45 (1987).
26. P. S. Ashton, *Ann. Mo. Bot. Gard.* **64**, 694 (1977).
27. The World Bank, *World Tables, 1988–89 Edition* (Johns Hopkins Univ. Press, Baltimore, 1989); World Resources Institute, *World Resources 1990–1991* (Oxford Univ. Press, New York, 1990); U.S. Department of Commerce, *Statistical Abstract of the United States 1990*. The monetary value of agricultural productivity is based on the agricultural portion of GNP expressed in U.S. dollars and divided by the total land area of the country. While this statistic averages the productivity of farmlands with nonagricultural lands (Figs. 2 and 4A), the weight of root crops per hectare is based only on the amount of land that is actually used to produce these crops (Fig. 4B).
28. Maps are based on the FAO-UNESCO Soil Map of the World with soil types assigned to 1° by 1° cells according to Zobler's World Soil Data File [FAO-UNESCO Soil Map of the World, vols. 2 to 10 (UNESCO, Paris, 1971 to 1981); L. Zobler, *A World Soil File for Global Climate Modeling*, NASA Technical Memorandum No. 87802 (1986)]. Total exchangeable bases (TEB, the sum of soil cations Ca, Mg, Na, and K) are estimated on the basis of the properties that define the units in the Soil Groups as estimated by the AGLS Soil Resources Group of FAO. TEB is roughly approximated by solving the equation for percent base saturation, BS% = [(TEB/

CECsoil) × 100] with an equation approximating soil cation exchange capacity (soil CEC), which gives TEB = (BS%/100) × {3.5 × OC% + [(Clay% × CECclay)/100)]}; where OC% is percent organic carbon in the soil, Clay% is the percent clay content, and CECclay is the approximate cation exchange capacity for the dominant clay mineral (for example, kaolinite, allophane) or an average value of 40 meq per 100 g for mixed clay mineralogy. Soil pH values are based on the average weighted value of measured pH over the top 100 cm of 400 soil profiles from the FAO soil units. [AGLS Soil Resources Group, *Soil Properties and Qualities Estimation Based on Soil Groups by the Soil Map of the World* (FAO, Rome, Italy, 1992)].

29. I thank S. Buol, D. DeAngelis, G. Marland, J. McBrayer, W. Post, and D. Richter for critical review of the manuscript. F. Nachtergaele of the FAO Soil Resources Working Group in Rome provided the information on soil properties used to generate Fig. 3. Supported by the Program for Ecosystem Research of the U.S. Department of Energy's Office of Health and Environmental Research under contract DE-AC05-84OR21400 with Martin Marietta Energy Systems, Inc. This is Oak Ridge National Laboratory Environmental Sciences Division publication 4172.

RESEARCH ARTICLE

Protein Design by Binary Patterning of Polar and Nonpolar Amino Acids

Satwik Kamtekar, Jarad M. Schiffer,* Huayu Xiong,
Jennifer M. Babik, Michael H. Hecht†

A general strategy is described for the de novo design of proteins. In this strategy the sequence locations of hydrophobic and hydrophilic residues were specified explicitly, but the precise identities of the side chains were not constrained and varied extensively. This strategy was tested by constructing a large collection of synthetic genes whose protein products were designed to fold into four-helix bundle proteins. Each gene encoded a different amino acid sequence, but all sequences shared the same pattern of polar and nonpolar residues. Characterization of the expressed proteins indicated that most of the designed sequences folded into compact α-helical structures. Thus, a simple binary code of polar and nonpolar residues arranged in the appropriate order can drive polypeptide chains to collapse into globular α-helical folds.

The essential features for the design of proteins are suggested by two familiar properties of natural proteins. First, globular, water-soluble proteins invariably fold into structures that bury extensive hydrophobic surface area while simultaneously exposing polar side chains to solvent (*1*). Second, these structures typically contain an abundance of hydrogen bonded secondary structure (α helices and β strands).

The dual constraints of forming regular secondary structure while burying hydrophobic side chains (and exposing hydrophilic ones) are met most directly by designing novel amino acid sequences capable of forming amphiphilic α helices, β strands, or both. For a segment of polypeptide chain to form an α helix or a β strand with one hydrophilic face and one hydrophobic face, the sequence must be designed with a periodicity of polar and nonpolar residues that matches the repeat for that type of secondary structure. For example, for the design of a stable β sheet protein, the sequence must be composed predominantly of alternating polar and nonpolar residues. Conversely, for an α-helical protein, the periodicity of polar and nonpolar residues must approximate the 3.6-residue repeat that is characteristic of α helices.

We describe a general strategy for protein design that is based on the assumption that the ability of a sequence to form amphiphilic secondary structures may actually suffice to drive a designed polypeptide chain to fold into a compact native-like structure. Our strategy is based on the premise that formation of stably folded structures does not require the explicit design of specific interresidue contacts; the precise packing of the three-dimensional jigsaw puzzle need not be specified a priori. Only the sequence location, but not the identity, of the polar and nonpolar residues must be specified explicitly.

We produced a collection of protein sequences that satisfy these criteria. This collection was generated by constructing a degenerate family of synthetic genes. Each gene encoded a different amino acid sequence, but all sequences shared the same periodicity of polar and nonpolar residues. The sequence degeneracy in this family of genes was made possible by the organization of the genetic code: Wherever a nonpolar amino acid was required, the degenerate codon NTN was used (where N represents a mixture of A, G, T, and C). Wherever a polar amino acid was required, the degenerate codon NAN was used. With these degenerate codons, positions requiring a nonpolar amino acid were filled by Phe, Leu, Ile, Met, or Val, whereas positions requiring a polar amino acid were filled by Glu, Asp, Lys, Asn, Gln, or His.

Design criteria. To test our design strategy, we focused on a structural motif that was small and simple, yet large enough to form a globular structure with a well-defined hydrophobic core and an abundance of secondary structure. To simplify the design and characterization of the novel proteins, we chose a fold that was composed of only one type of secondary structure, α helices. Previous work on the design of peptides and proteins has indicated that helices are easier to design than β structures (*2–5*). This probably reflects both the greater modularity of helical structures and the greater tendency of β structures to form insoluble aggregates. For these reasons we designed four-helix bundles. The four-helix bundle is a common fold among natural proteins (*6*) and has also been the target structure in previous efforts directed toward the design of novel proteins (*2–4*). Two representations of an idealized four-helix bundle are shown in Fig. 1.

To sample a large section of sequence space, we used degenerate nonpolar codons at each of the buried positions and degenerate polar codons at each of the surface positions in the α helices of our designed proteins. In the four helices shown in Fig. 1B, there are 24 buried (hydrophobic) positions and 32 surface (hydrophilic) positions. Because each buried position can be occupied by one of five different nonpolar side chains (Phe, Leu, Ile, Met, or Val) and each surface position can be occupied by any one of six different polar side chains (Glu, Asp, Lys, Asn, Gln, or His), a total of $5^{24} \times 6^{32} = 4.7 \times 10^{41}$ different amino acid sequences are theoretically possible. Our design strategy is based on the premise that a substantial fraction of the sequences fitting this pattern will actually fold into proteins that are compact, stable, and α-helical.

In our collection of sequences (Fig. 2) the interhelical turn sequences were not degenerate. It was necessary to define explicitly some regions of our sequence in order to facilitate gene construction, and

---

S. Kamtekar, H. Xiong, and M. H. Hecht are in the Department of Chemistry and J. M. Schiffer and J. M. Babik are in the Department of Molecular Biology at Princeton University, Princeton, NJ 08544.

*Present address: Department of Biology, University of California, San Diego, La Jolla, CA 92093.

†To whom correspondence should be addressed.