wavelength peak of *Aequorea* GFP[5,6]. We now report that simple point mutations in *Aequorea* GFP ameliorate its main problems and bring its spectra much closer to that of *Renilla*.

Serine 65 of the amino-acid sequence of *Aequorea* GFP becomes part of the *p*-hydroxybenzylideneimidazolinone chromophore. To test the hypothesis[7] that Ser 65 undergoes additional dehydration to form a vinyl side chain, we mutated that residue to Ala, Leu, Cys or Thr. If a vinyl group were formed by elimination of H_2O or H_2S, Ser and Cys should give identical spectra very different from Ala and Leu in which elimination is impossible. Serendipitously, all four mutants showed single excitation peaks, located at 470–490 nm, whose amplitudes were four- to sixfold greater than that of wild-type for equal numbers of molecules (*a* in the figure). These results exclude vinyl formation. The Ser 65→Thr mutant (S65T) was selected for further characterization because it had the longest wavelengths of excitation and emission (490 and 510 nm), which closely resembled those reported for *Renilla* GFP (498 and 508 nm). The crucial post-translational oxidation[8] to produce the fluorophore from the nascent polypeptide chain proceeded about fourfold more rapidly in S65T than in the wild-type protein (*b* in the figure). This acceleration ameliorates a poten-tially significant limitation in using GFP as a reporter protein for rapid gene inductions[8].

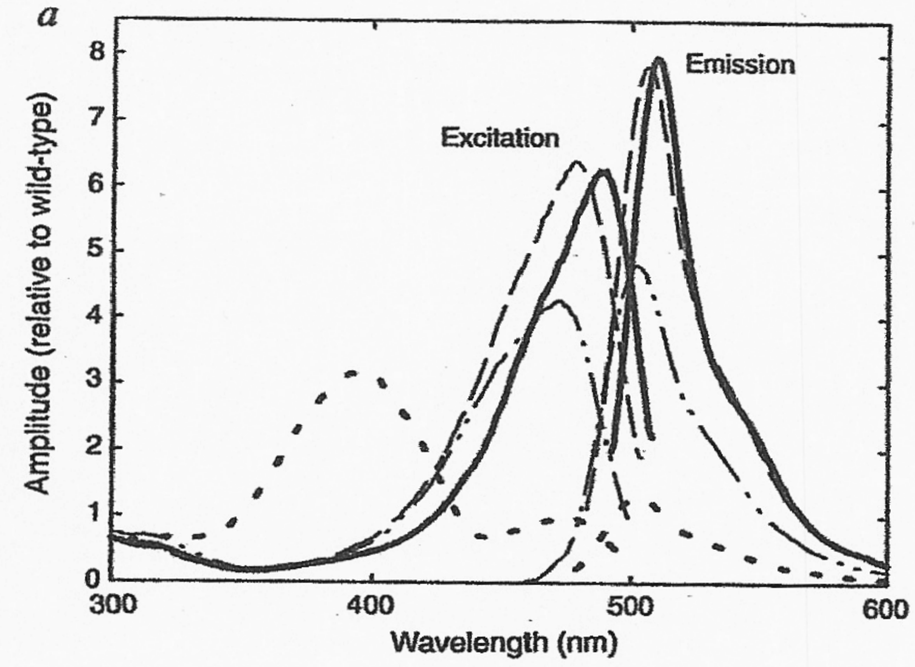
Mutations of Ser 65 to Arg, Asn, Asp, Phe, and Trp gave fluorescence intensities well below that of wild type. It remains unclear exactly how position 65 controls spectral properties or why *Aequorea* chose serine. Nevertheless, the greatly increased brightness and rate of fluorophore generation in mutants such as S65T should make them superior to wild-type *Aequorea* GFP for most experimental uses.

Note added in proof: GFP variants generated by combinatorial mutagenesis of positions 64–69 have excitation peaks near 490 nm, but their amplitudes and the kinetics of fluorophore formation have not been quantified[12].

Roger Heim
Andrew B. Cubitt
Roger Y. Tsien*
Howard Hughes Medical Institute and Department of Pharmacology, University of California, San Diego, La Jolla, California 92093-0647, USA

*To whom correspondence should be addressed.

1. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. & Prasher, D. C. *Science* **263**, 802–805 (1994).
2. Prasher, D. C. *et al. Gene* **111**, 229–233 (1992).
3. Wang, S. & Hazelrigg, T. *Nature* **369**, 400–403 (1994).
4. Ward, W. W., Cody, C. W., Hart, R. C. & Cormier, M. J. *Photochem. Photobiol.* **31**, 611–615 (1980).
5. Ward, W. W. *et al. Photochem. Photobiol.* **35**, 803–808 (1982).
6. Ward, W. W. in *Bioluminescence and Chemiluminescence* (eds DeLuca, M.A. & McElroy, W.D.) 235–242 (Academic, New York, 1981).
7. McCapra, F., Razavi, Z. & Neary, A. P. *JCS chem. Commun.* 790–791 (1988).
8. Heim, R., Prasher, D. C. & Tsien, R. Y. *Proc. natn. Acad. Sci. U.S.A.* **91**, 12501–12504 (1994).
9. Kunkel, T. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 488–492 (1985).
10. Inouye, S. & Tsuji, F. I. *FEBS Lett.* **341**, 277–280 (1994).
11. Inouye, S. & Tsuji, F. I. *FEBS Lett.* **351**, 211–214 (1994).
12. Delagrave, S., Hawtin, R. E., Silva, C. M., Yang, M. M. & Youvan, D. C. *Bio/Technology* **13**, 151–154 (1995).



Comparison of recombinant wild-type and mutant green fluorescent proteins. *a*, Fluorescence excitation and emission spectra of wild-type protein (– – –), Ser 65→Ala (– · · –), Ser 65→Cys (– –) and Ser 65→Thr (—) mutants. The coding region of *Aequorea gfp* cDNA[2] (gift of D. Prasher) was cloned into the T7 expression vector $pRSET_B$ (Invitrogen), giving a polyhistidine-tagged fusion protein expressed in *Escherichia coli*, BL21(DE3)LysS (Novagen). Oligonucleotide-directed mutagenesis at the codon for Ser 65 was performed[9] on the same construct using the Muta-Gene Phagemid *in vitro* kit version 2 (Bio-Rad). All fusion proteins were purified on nickel-chelate columns (Qiagen). Spectra were obtained from equal concentrations of GFPs controlled by densitometry of denaturing gels stained with Coomassie blue. Excitation spectra were obtained collecting emission at the respective peak wavelengths (508, 503, 507, and 511 nm for wild-type, S65A, S65C and S65T) and were corrected by a quantum counter; emission spectra were likewise measured at the respective excitation peaks (475, 471, 479, and 489 nm) and were corrected using factors from the fluorometer manufacturer. The amplitude of the 475 nm excitation peak of wild-type GFP has been defined as 1.0. The sixfold greater peak amplitude of S65T arises from a 5.5-fold higher extinction coefficient (39,200 $M^{-1}\,cm^{-1}$ at 490 nm for S65T compared with 21,000 and 7,150 $M^{-1}cm^{-1}$ at 395 and 475-nm for wild-type), similar fluorescence quantum yield (0.68 versus 0.77), and slightly narrower emission spectrum. The extinction coefficients and quantum yields reported here for recombinant wild-type protein fused to a polyhistidine tag are in good agreement with literature[6] values for GFP extracted from *Aequorea* when corrected for the revised molecular mass[2]. Recombinant expression and polyhistidine tagging are known[1,10,11] not to affect the fluorescence spectra of wild-type GFP. *b*, Rates of autoxidative fluorophore generation in wild-type (■) and S65T (▲) GFP, measured by development of fluorescence after admission of air to *E. coli* cultures anaerobically grown in GasPak pouches (Becton-Dickinson) for 3 days. Air was readmitted while transferring the cells to phosphate-buffered saline containing 8 mM NaN_3 as a metabolic inhibitor. The time course of subsequent fluorescence development measured the final oxidation step in the protein's self-modification to generate its internal fluorophore[8]. Data from two independent runs normalized by their respective asymptotic fluorescence values were pooled for each protein. The smooth curves are exponential curve fits consistent with pseudo-first-order kinetics, with time constants of 2.0 and 0.45 h for wild-type and S65T, respectively. Previously reported time constants for wild-type GFP autoxidation[8] were rather longer, probably because the protein was held in bacteria for longer periods of anaerobic growth, which seems to slow subsequent oxidation.

Kinetics of protein folding

SIR — Šali *et al.*[1] have attempted to resolve the "Levinthal paradox" of how proteins find their unique native conformations so fast. Although we agree with some of their points, we question others.

First, a model can bear on the Levinthal paradox only if the folding kinetics are run at a temperature low enough for the native state to be more stable than the denatured states. But Šali *et al.* are not studying native conditions: their molecules are mostly denatured. The temperatures they use are so high that equilibrium populations of the native states of many of their "folding sequences" are only 1–5% (ref. 2), and none exceeds 40% (ref. 1). Other model studies[3] show that native states can be accessed quickly in certain ranges of denaturing temperatures, but most of the chains will not stay there. If Šali *et al.* could not find native states under folding conditions, they have not completely addressed the Levinthal paradox.

Second, Šali *et al.* state that the "neces-

sary and sufficient condition for a sequence to fold rapidly in the present model is that the native state is a pronounced energy minimum"[2] and "the features that depend only on the lower discrete part of the spectrum can be characterized by use of the compact self-avoiding chains alone, neglecting the non-compact conformations"[2]. But this conclusion is contradicted by 11 of the sequences they studied for which lowest-energy conformations are not maximally compact[2]. The true denatured ensemble is generally much larger than the maximally compact ensemble, so neither the first excited state "energy gap" ΔE_{10}, defined by them, nor the temperature, T_x, which they define using only the maximally compact ensemble[2], is precisely related to true thermodynamic stability. Furthermore, the correlation they observed between their energy gap and folding kinetics is only a weak trend, and the energy-gap condition they use is not sufficient to discriminate between folding and non-folding sequences (see Fig. 7 of ref. 2). In general, folding rate depends on the entire energy landscape[3–6], not just the energy gap in a highly restricted ensemble.

Ultimately, the Levinthal problem is not that a protein has too many degrees of freedom. It is the shape, not the size alone, of the conformational energy landscape that matters[4–6,9]. Many large landscapes have shapes that can be quickly traversed to reach the bottom[6–9]. Šali *et al.* show that Metropolis Monte Carlo sampling can find the lowest energy of a particular parameterized potential function, but this was already clear from many earlier efforts[10,11]. The issue, therefore, is whether their potential function is better than earlier models. Baldwin in News and Views[12] has said that Šali *et al.* were using a potential function of the Miyazawa–Jernigan type, picked from the pairwise interactions in the protein database. But, as Šali *et al.* have noted[1], the terms are picked from a random gaussian distribution, not from the databank. Their potential function is not particularly physical, as correlations among contact energies of different pairs of amino-acid residues are neglected. It is unclear whether the potential is any more or less protein-like than any of the potentials used in previous works.

Baldwin[12] describes the work of Šali *et al.* as an important "new view" of protein folding. Naturally, lattice models are useful for addressing general physical principles of protein folding, even though they involve considerable simplification. However, it is clear from many earlier efforts, including some that used comparable lattice simulations, that many of the ideas Baldwin cites as "new" are already in the literature (refs 4, 10, 13–18, and refs therein, and reviewed more recently in refs 3, 5, 6).

Hue Sun Chan
Department of Pharmaceutical Chemistry, University of California at San Francisco, California 94143-1204, USA

This note was written with the assistance and concurrence of: Joseph D. Bryngelson, NIH, Bethesda, Maryland 20892, USA; Carlos J. Camacho, Facultad de Fisica, PUC, Casilla 306, Santiago 22, Chile; Ken A. Dill, Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143-1204, USA; José N. Onuchic & Nicholas D. Socci, Department of Physics, University of California at San Diego, La Jolla, California 92093-0319, USA; D. Thirumalai, Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742, USA; and Peter G. Wolynes, School of Chemical Sciences, University of Illinois, Urbana, Illinois 61801, USA.

KARPLUS *ET AL.* REPLY — Chan *et al.* raise several questions, all of which have simple answers. The object of our study[1,2] was to examine a large number of sequences and to separate those that fold from those that do not. Consequently, a temperature slightly above T_m the midpoint of the folding transition, was used to speed up the reaction. If folding to the native state were always possible under the simulation conditions, as Chan *et al.* imply there would not have been any non-folding sequences and our computer experiment would have failed. But this is not the case. Further, the same folding kinetics is observed throughout the temperature range where the true native state stability varies from 1 to 40%[19].

A pronounced energy gap between the native and first excited state (equations (3) and (10) in ref. 1) for the fully compact ensemble, is a necessary and sufficient condition for rapid folding in the model study. It is necessary because no sequence without such a minimum folds to the native state, and is sufficient because all sequences with such a minimum do fold (Fig. 7 of ref. 2). As to the 11 out of 200 sequences that have their minimum outside the fully compact set, none satisfied the energy condition nor did they fold repeatedly either to the lowest fully compact state or to the lowest energy state found by a Monte Carlo simulation. Thus, these sequences confirm and generalize the folding criterion[2]. Further, the use of the energy condition for quantitatively determining the folding rate has been demonstrated[20]. There is a strong correlation between the results from the fully compact states and the complete set of states (ref. 2 and Fig. 17 therein).

The nature of the configuration space, as well as the number of conformers, is important for the Levinthal paradox[1,2]. Surfaces can be constructed for which resolution of the paradox is trivial, but this is not true for the 27-mer since only a fraction of sequences fold rapidly. The large size of the configuration space is necessary for the existence of a paradox. The 27-mer model has 10^{16} configurations and requires fewer than 5×10^7 Monte Carlo steps to find the native state. Short oligomers that have been extensively studied on a two-dimensional square lattice[16,21] may be too small; for example, more Monte Carlo steps (10^5 or more) than there are configurations (4×10^4) were required for folding a 13-mer[16].

The aim of the lattice simulations was to use random interactions so as to determine what differentiates folding from non-folding sequences. The exact choice of parameters was not important, as long as a reasonable set was used. The 27-mer parameters[2] correspond to the Miyazawa and Jemigan set[22] in terms of the magnitude of the interaction energies and their standard deviation.

As to Baldwin's statement in News and Views that we presented a "new view" of protein folding, we agree that some of the concepts in refs 1 and 2 were presaged in earlier work of Go and Abe[23] and of others cited in refs 1 and 2. The 27-mer model studies[1,2,18] provided the first demonstration that the energy-gap condition and a detailed mechanism for resolving the Levinthal paradox could be found *a posteriori* in computer experiments without having to be introduced explicitly *a priori* to achieve folding[24].

M. Karplus
A. Šali
E. Shakhnovich
Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138, USA

1. Šali, A., Shakhnovich, E. I. & Karplus, M. *Nature* **369**, 248–251 (1994).
2. Šali, A., Shakhnovich, E. I. & Karplus, M. *J. molec. Biol.* **235**, 1614–1636 (1994).
3. Socci, N. D. & Onuchic, J. N. *J. chem. Phys.* **101**, 1519–1528 (1994).
4. Camacho, C. J. & Thirumalai, D. 19. *Proc. natn. Acad. Sci. U.S.A.* **90**, 6369–6372 (1993).
5. Chan, H. S. & Dill, K. A. *J. chem. Phys.* **100**, -9238–9257 (1994).
6. Bryngelson, J. D., Onuchic, J. N., Socci, N. D. & Wolynes, P. G. *Proteins: Struct. Funct. Genet.* (in the press).
7. Taketomi, H., Ueda, Y. & Gō, N. *Int. J. Peptide Protein Res.* **7**, 445–459 (1975).
8. Zwanzig, R., Szabo, A. & Bagchi, B. *Proc. natn. Acad. Sci. U.S.A.* **89**, 20—22 (1992).
9. Dill, K. A. *Curr. Opin. struct. Biol.* **3**, 99–103 (1993).
10. Honeycutt, J. D. & Thirumalai, D. *Proc. natn. Acad. Sci. U.S.A.* **87**, 3526–3529 (1990).
11. Skolnick, J., Kolinski, A. *A. Rev. phys. Chem.* **40**, 207–235 (1989).
12. Baldwin, R. L. *Nature* **369**, 183–184 (1994).
13. Bryngelson, J. D. & Wolynes, P. G. *Proc. natn. Acad. Sci. U.S.A.* **84**, 7524–7528 (1987).
14. Guo, Z., Thirumalai, D. & Honeycutt, J. D. *J. chem. Phys.* **97**, 525–535 (1992).
15. Leopold, P. E., Montal, M. & Onuchic, J. N. *Proc. natn. Acad. Sci. U.S.A.* **89**, 8721–8725 (1992).
16. Miller, R. *J. et al. chem. Phys.* **96**, 768–780 (1992).
17. Goldstein, R. A., Luthey-Schulten, Z. A. & Wolynes, P. G. *Proc. natn. Acad. Sci. U.S.A.* **89**, 9029–9033 (1992).
18. Shakhnovich, E., Farztdinov, G., Gutin, A. M. & Karplus, M. *Phys. Rev. Lett.* **67**, 1665–1668 (1991).
19. Karplus, M., Caflish, A., Sali, A. & Shakhnovich, E. in *Modelling of Biomolecular Structures and Mechanisms* (eds Pullman, A. *et al.*) 69–84 (Kluwer Academic, Dordrecht, 1994).
20. Abkevich, V. I., Gutin, A. M. & Shakhnovich, E. *J. chem. Phys.* **101**, 6052–6062 (1994).
21. Camacho, C. J. & Thirumalai, D. *Proc. natn. Acad. Sci. U.S.A.* **90**, 6369–6372 (1993).
22. Miyazawa, S. & Jernigan, R. L. *Macromolecules* **18**, 534–552 (1985).
23. Go, N. & Abe, H. *Avd. Biophysics* **18**, 149–164 (1984).
24. Karplus, M. & Šali, A. *Curr. Opin. struct. Biol.* **5**, 58–73 (1995).