

Supporting information for Keating *et al.* (2001) *Proc. Natl. Acad. Sci. USA* **98** (26), 14825–14830. (10.1073/pnas.261563398)

Details of the Computational Methods

Stage 1. Dead End Elimination (DEE). The DEE procedure used was based on that of Goldstein (1) and consisted of the following sequence, repeated until no more rotamers or pairs were eliminated: (i) iterative elimination of single rotamers by using Eq. **1a** below, taking into account dead ending pairs where these had been previously identified (2); (ii) a single round of pairs elimination using Eq. **1b**; (iii) use of dead ending pairs to identify dead ending single rotamers according to criteria outlined by Lasters *et al.* (3). The problem of repacking a core of 12 residues is not challenging for DEE, so little effort was made to optimize performance.

Goldstein's original criteria are given by Eqs. **1a** and **1b** with $E_{\text{cutoff}} = 0$, where $E(i_u)$ is the energy of rotamer u of residue i , $E(i_u, j_v)$ is the energy of interaction between rotamer u of residue i and rotamer v of residue j , $E(i_u j_v, k_x) = E(i_u, k_x) + E(j_v, k_x)$, and $E(i_u j_v) = E(i_u) + E(j_v) + E(i_u, j_v)$. If the inequality of Eq. **1a** is satisfied, rotamer i_u can be eliminated; if Eq. **1b** is satisfied, the pair $(i_u j_v)$ is identified as dead-ending. We used an E_{cutoff} of 30 kcal/mol (1). Leach and Lemon have pointed out that this ensures that no solution within E_{cutoff} of the global minimum solution will be eliminated (4).

$$E(i_u) - E(i_v) + \sum_j \min_w [E(i_u, j_w) - E(i_v, j_w)] > E_{\text{cutoff}} \quad [1a]$$

$$E(i_u j_v) - E(i_x j_y) + \sum_k \min_w [E(i_u j_v, k_w) - E(i_x j_y, k_w)] > E_{\text{cutoff}} \quad [1b]$$

A* Search. A* is a branch and bound algorithm that uses a lower bound on the total energy of the protein to guide a search for the global minimum. In our implementation, A* is used to search the conformational space remaining after DEE and return solutions, in order of increasing energy, until (i) all solutions within the DEE energy cutoff have been reported, or (ii) 10,000 solutions have been reported. We implemented A* with energy bounds as described by Leach and Lemon; we also used their criterion for choosing the order of residue expansion in the A* search tree (4).

The DEE/A* search was carried out independently on each library backbone template, yielding up to 10,000 structures per backbone. Prior to minimization (Stage 2), a list of *unique a-* and *d-*position side-chain configurations was collated. Each unique conformation that survived DEE was minimized only once.

Side-Chain Rotamer Library. The following library was used for DEE/A* search. It contains rotamers from backbones with $\phi = -60$, $\psi = -50$ that occur with a frequency $>0.1\%$ in the rotamer library `bbdep99.Aug.sortr12lib` (for more recently updated libraries see www.fccc.edu/research/labs/dunbrack/bbdep.html) (5, 6).

Subrotamer Method. The combination of rigid rotamers and a Lennard--Jones potential to describe van der Waals repulsion can severely penalize conformations that are only slightly incorrect. To compensate for this, we reduced van der Waals radii for the DEE and A* calculation to 90% of their original values. We found that performance was further improved by replacing rigid rotamers with an ensemble of subrotamers according to the method of Mendes *et al.* (7). Simple rotamer energies and pair-wise energies were replaced with the following expressions, where u and v are rotamers of i and j , respectively, and s and t are subrotamers of u and v .

$$\bar{E}(i_u) = -RT \ln \sum_s \exp[-E(i_{us})/RT] \quad [2a]$$

$$\bar{E}(i_u, j_v) = \{-RT \ln \sum_s \sum_t \exp[-\frac{E(i_{us}) + E(j_{vt}) + E(i_{us}, j_{vt})}{RT}]\} - \bar{E}(i_u) - \bar{E}(j_v) \quad [2b]$$

Stage 2. Solutions returned from the A* search were minimized using the PARAM19 potential with 100% van der Waals radii, electrostatics turned off, and an explicit hydrogen bonding term added, as described previously (8–10). The minimization procedure was adapted from ref. 10. A coiled-coil geometry was imposed on the model by using a framework force constant to restrain the C α s to a curve described by Crick's parameterization (11). Side-chain dihedral angles were initially constrained to library values with a force constant of 100 kcal/mol \cdot rad² while the helices were subjected to 700 steps of adapted basis Newton--Raphson (ABNR) minimization with a force of 5 kcal/mol \cdot Å compressing the two helices towards one another. After removing the side-chain constraints and the compression force, ABNR minimization was continued until the

difference in energy between steps was less than 0.01 kcal/mol. Different values of ω_0 were sampled by fixing the ω_0 parameter to a new value and carrying out steepest descent and ABNR minimization until the energy decreased by less than 0.005 kcal/mol between steps. The best minimized ω_0 structure was taken as the global minimum.

Timings. Calculations were performed on a single R10,000 processor (SGI ORIGIN 2000) (Silicon Graphics, Mountain View, CA). Pair-wise energy calculations took ≈ 95 seconds, and DEE/A* calculations took ≈ 6 seconds per backbone. Full Stage 2 minimizations took ≈ 15 seconds per side-chain configuration; this time was not optimized.

Energy Corrections. Energies from the molecular mechanics calculations were converted to relative unfolding free energies for different sequences using a thermodynamic cycle derived from figure 1 of ref. 11. Figure 7 (ref. 11) illustrates this approach by comparing the unfolding free energies of two sequences: CC_x with x in a core d-position and CC_y with y substituted for x . The top row represents folded coiled coils, in axial projection. The bottom row represents unfolded peptides. CC_x and CC_y have unfolding free energies ΔG_x^u and ΔG_y^u . Two additional coiled coils are included in the cycle: CCf_x and CCf_y , differ only in the substitution of x or y at an exterior **f** position of the coiled coil. CCf_x and CCf_y have unfolding free energies $\Delta G_{f_x}^u$ and $\Delta G_{f_y}^u$. With these definitions, because the cycle is closed, $\Delta\Delta G^u$ can be expressed as $\Delta\Delta G^u = \Delta G_x^u - \Delta G_y^u = k + \Delta G_{f_x}^u - \Delta G_{f_y}^u$. k is calculated from predicted structures of CC_x , CC_y , CCf_x , and CCf_y by using CHARMM and a solvation correction, as described in the text. $\Delta G_{f_x}^u$ and $\Delta G_{f_y}^u$ are taken from ref. 13. By using this cycle, we avoid modeling the unfolded state explicitly.

1. Goldstein, R. F. (1994) *Biophys. J.* **66**, 1335–1340.
2. Lasters, I. & Desmet, J. (1993) *Protein Eng.* **6**, 717–722.
3. Lasters, I., De Maeyer, M. & Desmet, J. (1995) *Protein Eng.* **8**, 815–822.
4. Leach, A. R. & Lemon, A. P. (1998) *Proteins* **33**, 227–239.

5. Dunbrack, R. L., Jr. & Karplus, M. (1994) *Nat. Struct. Biol.* **1**, 334–340.
6. Dunbrack, R. L., Jr. & Karplus, M. (1993) *J. Mol. Biol.* **230**, 543–574.
7. Mendes, J., Baptista, A. M., Carrondo, M. A. & Soares, C. M. (1999) *Proteins* **37**, 530–543.
8. Brooks, B. R., Brucoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983) *J. Comp. Chem.* **4**, 187–217.
9. Nilsson, L. & Karplus, M. (1986) *J. Comput. Chem.* **7**, 591–616.
10. Harbury, P. B., Tidor, B. & Kim, P. S. (1995) *Proc. Natl. Acad. Sci. U S A* **92**, 8408–8412.
11. Crick, F. H. (1953) *Acta Crystallogr.* **6**, 685–689.
12. Harbury, P. B., Plecs, J. J., Tidor, B., Alber, T. & Kim, P. S. (1998) *Science* **282**, 1462–1467.
13. O'Neil, K. T. & DeGrado, W. F. (1990) *Science* **250**, 646–651.