Protein structure prediction

TQVAKKILVTCALPYANGSIHLGHMLEHIQADVWVRYQRMRG
HEVNFICADDAHGTPIMLKAQQLGITPEQMIGEMSQEHQTDF
AGFNISYDNYHSTHSEENRQLSELIYSRLKENGFIKNRTISQLY
DPEKGMFLPDRFVKGTCPKCKSPDQYGDNCEVCGATYSPTEL
IEPKSVVSGATPVMRDSEHFFFDLPSFSEMLQAWTRSGALQEQ
VANKMQEWFESGLQQWDISRDAPYFGFEIPNAPGKYFYVWLD
APIGYMGSFKNLCDKRGDSVSFDEYWKKDSTAELYHFIGKDI
VYFHSLFWPAMLEGSNFRKPSNLFVHGYVTVNGAKMSKSRGT
FIKASTWLNHFDADSLRYYYTAKLSSRIDDIDLNLEDFVQRVN
ADIVNKVVNLASRNAGFINKRFDGVLASELADPQLYKRFTDA
AEVIGEAWESREFGKAVREIMALADLANRYVDEQAPWVVAK
QEGRDADLQAIAQWGINLFRVLMTYLKPVLPKLTERAEAFLN
TELTWDGIQQPLLGHKVNPFKALYNRIDMRQVEALVEASKEE
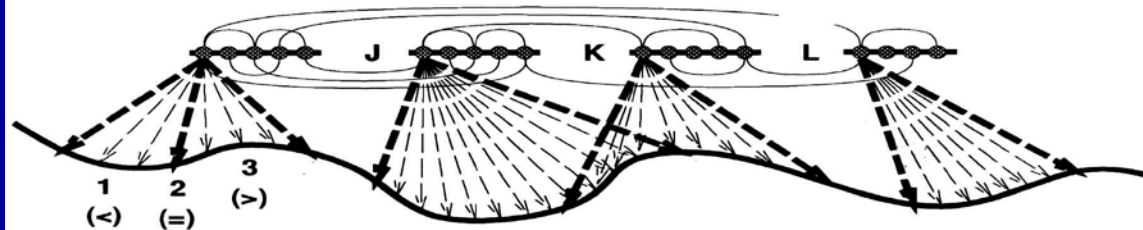VKAAAAPVTGPLADDPQDGCGRHDRVVDSGSK
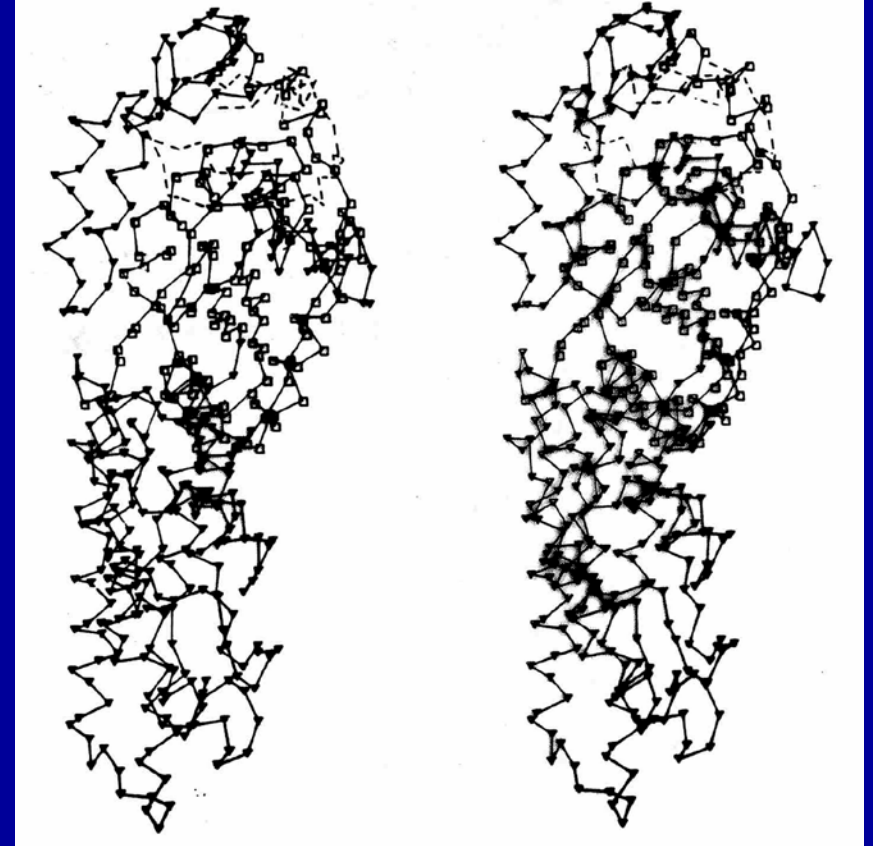
"Protein Threading"
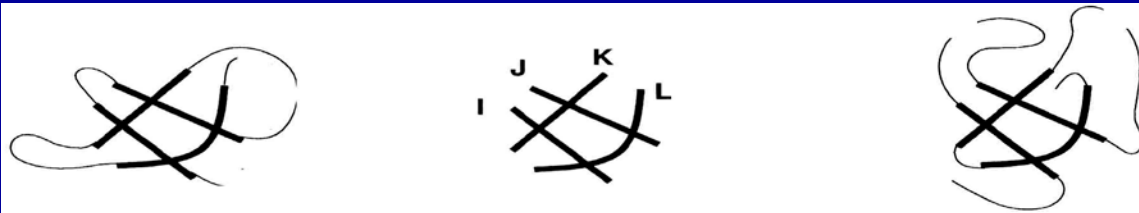
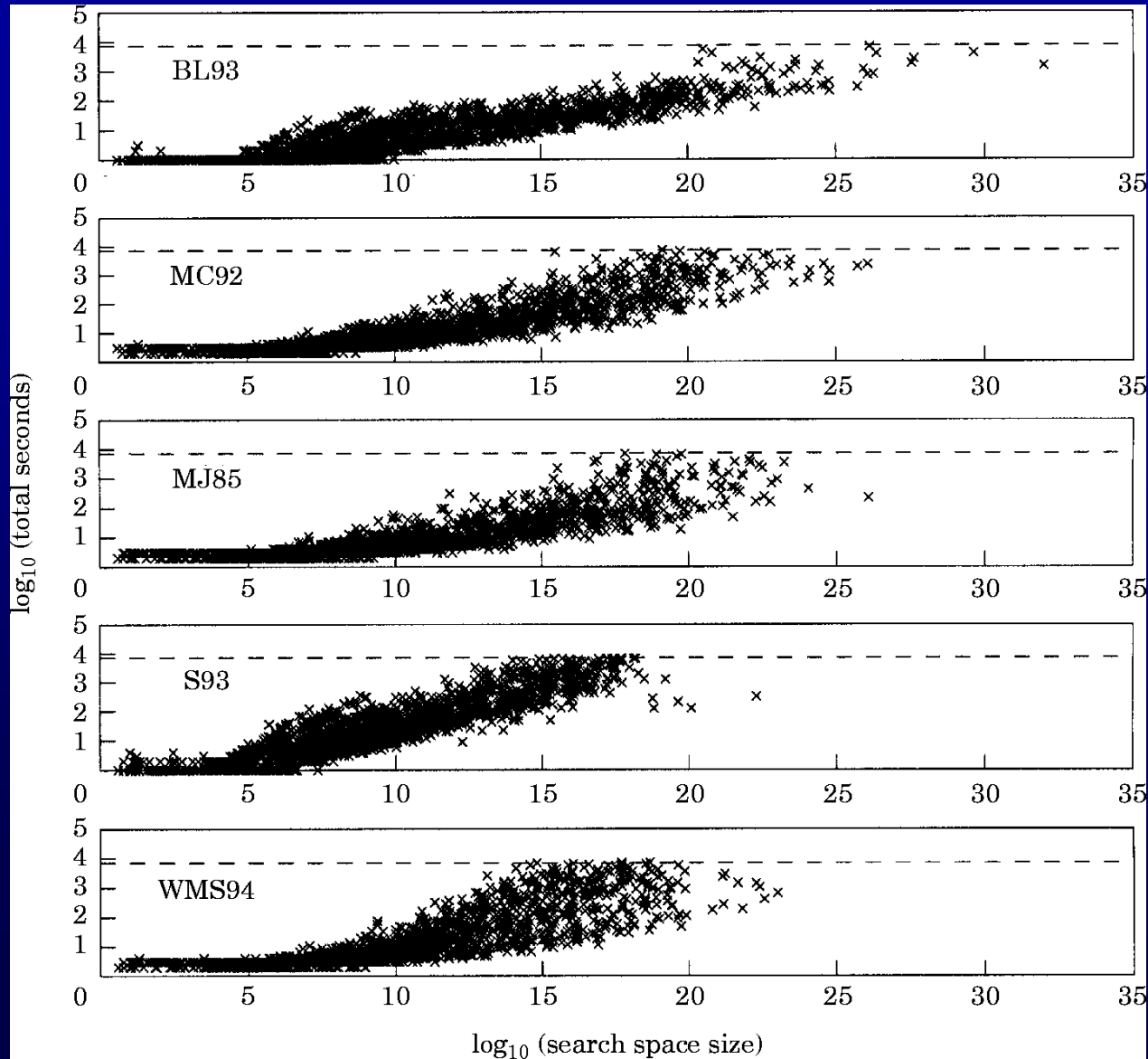(A) Two Structurally Similar Proteins and a Core of Four Segments
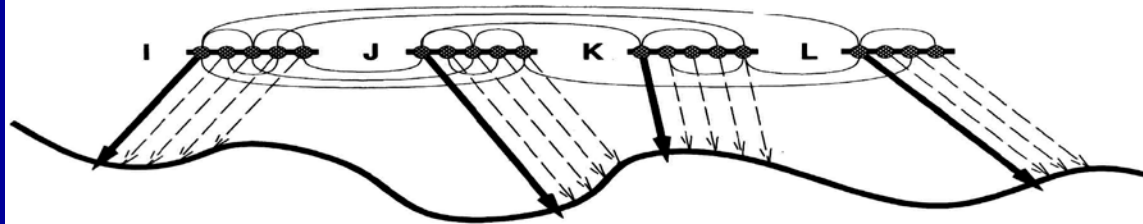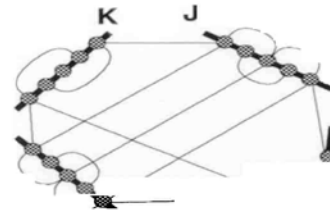
(C) One Possible Threading with a Novel Sequence

(D) This Set of Possible Threadings Will Be Split Into Three Subsets at Segment I

The time required to find the global minimum is shown on log-log axes as a function of search space size.
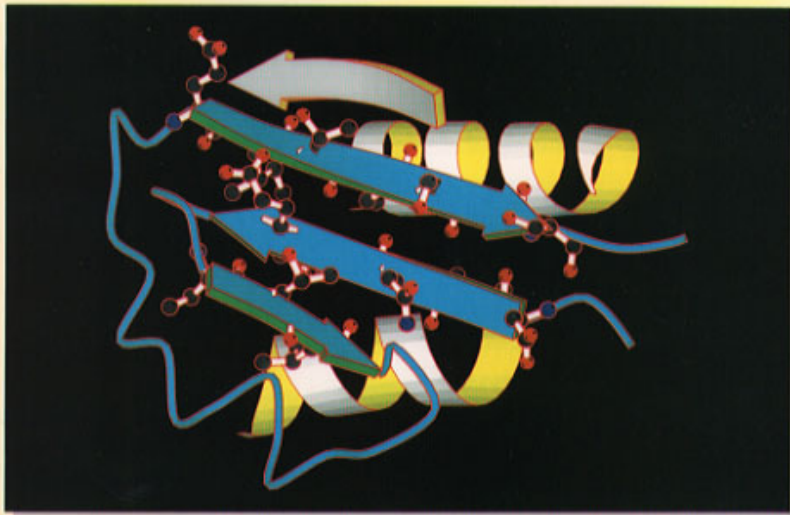
Heuristic (Objective) Function

We have explored several alternative forms of the lower bound (Lathrop & Smith, 1994). Our current version, denoted $\text{lb}(\mathcal{T})$, is:

$$\min_{t \in \mathcal{T}} f(\mathbf{t}) \geq \text{lb}(\mathcal{T})$$

$$= \min_{t \in \mathcal{T}} \sum_i \left[g_1(i, t_i) + g_2(i-1, i, t_{i-1}, t_i) \right.$$

$$\left. + \min_{\substack{\mathbf{u} \in \mathcal{T} \\ l_j^{\max} = +\infty}} \sum_{|j-i|>1} \tfrac{1}{2} g_2(i, j, t_i, u_j) \right] \quad (5)$$

"**Global Optimum Protein Threading with Gapped Alignment and Empirical Pair Score Functions**"

Lathrop and Smith

J. Mol. Biol. 255(1996)641-665