Protein Energy

Assignment: Given a sequence for a certain protein and two protein models (given in separate files), determine which of the two protein model is closest to the protein sequence by computing the energy of each.

Sequence:

DDKTWTTILLFAHEKSQDDTKKQIKVLQDAGAGGDKVPPIGARPEHGKIVVQKCPSVLRAAALVPWKNEN
AGIDGTKAVKVFQTAGMNWAVVGLHYLESSESWDLSLLGKIFSVAQKKGEGALVCISGKRERVENGHIYE

EHARKVKLFIAFGCGFGAALLAMVATIALDQGGDAPDVFLRSCAKIIHRKVQNYLSESVAKTIAIPAGHEVE
NGTGQAVGRADEETGVVIWGEGFDESYVKIAKIGH

Formula used to calculate model1 and model2's energy:

$$U = \sum_{i=1}^{N} \sum_{j=i+1}^{N} flag(i,j) \left(\epsilon_{ij} \left(\left(\frac{s_{ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{s_{ij}}{r_{ij}} \right)^{6} \right) + \frac{1}{4\pi\epsilon_0\epsilon_r} \frac{q_i q_j}{r_{ij}} \right) + \sum_{i=1}^{N} ASP(i) ASA(i)$$

- N = number of atoms a molecule has
- U = total energy (sum of its vdW energy, electrostatics energy, and solvation energy.
- flag(i,j) = 0, if i and j are chemically bonded, flag(i,j) = 1 otherwise.
- $\varepsilon_{ij} = \varepsilon_i \varepsilon_j$ and $\sum_i$ are given for each atom
- $s_{ij} = (s_i + s_j)/2$ and s_i are given for each atom
- r_{ij} is the distance between the two atoms i and j
- q_i and q_j are the charges of atoms *i* and *j*, respectively (given)
- $(1/4\pi\varepsilon_0) = 332$
- $\varepsilon_0 = 4$. (dielectric constant for water)
- *ASP(i)* : atomic solvation parameter of atom *i* (given)
- *ASA(i):* accessible surface area of *i*)

$$ASA_i = 0.2 * 4 * \pi * (r_i + R_{H2o})^2$$

- r_i = vdW radius of atom I
- R_{H2o} = 1.4 Angstrom (radius of water molecule)