Chapter 12 What makes knots strong

Benzinger on the basis of his discovery of the large compensation of standard enthalpy by standard entropy in thermal denaturation of ribonuclease A proposed in 1971 that the folded stability of ribonuclease A depends on the strength of only small fraction of all secondary interactions. The molecular justification for his unpopular proposal became obvious only with the discovery of substructures and the dominant role of knots in stability. In turn that rationalization raised the question of the high stability of knots as opposed to matrices. How is it that knots of mesophiles with 12% of the total number of residues establish the native state against the destabilizing stress from matrices. That it does explains why most mesophiles fall on Pohl's compensation plot for the activation enthalpy and entropy in the melting rate. Because these reflect not only the disruption of the knots but also the release of tension from the matrices, the ratio of knot to matrix residues must be fixed in this very large class of proteins. The remaining sources of the overall free-energy change are very little altered in the transition step because geometry changes are very small. Gregory and Lumry located the strength in special hydrogen bonds enhanced electrostatically by being set in a region of low permanent polarization. This seems to be correct as far as it goes but there are fine details still not well described that must be understood before the quantitative significance of the electrostatic synergism can be accurately estimated. Why, for example, are knot residues so often aromatic? Why do the helix sections of knots benefit in strength from alanyl residues? What is the synergism between knots and disulfide bonds; do the latter explain melting temperatures above 354 K? The spider silks especially the heavy duty dragline silks will probably provide the explanations since they are the strongest substances in tension known as pure polypeptides even stronger than Kevlar and much better than any metal. The dragline spider

silks are very elastic, piezoelectric and very resistant to high temperatures, acid and salt solutions. They are cleverly managed by spiders who can weave and unweave them from soft polymers to the fibers of still unknown composition except that the major component is antiparallel β sheet. Evolution has apparently developed protein knots along those lines although with less extreme properties. The major virtue of this discovery it the separation of folded stability problems from physiological functions but there are several more of similar importance to life..

Knot formation is highly cooperative and very sensitive to the sidechains of the residues contributing the peptide O and N atoms for the critical hydrogen bonds. In BPTI six of those residues have aromatic side chains of which two are free to undergo ring flips even at subzero temperatures still attached to their knots at those temperatures The one of the latter characterized in detail by Skalicky, Mills, Sharma and Szyperski (J., J.,S .,and T). J. Am. Chem. Soc. 123 (2000), 388) has an activation enthalpy of 14 kcal/M and activation entropy of –4 cal/MK . It falls on the Pohl compensation plot at a very low position and zero activation heat capacity demonstrating that total knot disruption is not necessary for ring flips even when held tightly in the knot. The single BPTI knot is a very small section of β sheet. Nuclease T-1 and T-4 lysozyme have two knots as usual for enzymes but only one is β sheet and although the sheet parts are much bigger than 6%, the knot is a small part near the center. of the sheet structure.. Levitt found that the knot H bonds of BPTI are very short possibly shorter than those in the knots of the G protein of streptococcal virus. Accurate H bond lengths are available only from B factors as already discussed. Cornilescu et al found the knot H-bonds in the G protein knots to have relatively large nmr j coupling. as is consistent with proton exchange rates and relatively high thermal stability of the protein. The through-bond j coupling suggests some electron and perhaps proton delocalization with a covalent contribution partially responsible

for the knot strength. It is important to test the generality of this finding as can now be done with B factors.

Another important example of knot variability was provided by the thermal-denaturation rates of kanamycin nucleotidyltransferase (253 residues) by M. Matsumura and coworkers. Three of the residues positions of knot residues were modified by exchanges with the activation enthalpy and entropy consequences shown in Fig.@ (from ref. @). As detailed in the caption very large changes in the activation thermodynamic quantities were produced by exchanges of one, two or three positions. Even the single substitution of tyrosine for aspartate at position 80 was sufficient to increase the activation enthalpy by 70 kcal/M and the activation free energy at 298K by 6.5 kcal/mole, much larger than the wild type. The changes tend to be cumulative and in these substitutions most increase knot stability. That they all fall on Pohl's compensation plot is very important since it shows that the indexing of proteins by the number of residues discussed in chapter 12 breaks down in residue exchange producing deviations from the sequence of their wild type reference. Apparently wild types are very highly selected in evolution so that residue exchange in SDM experiments away from the wild type reflect evolutionary consequences far more subtle than has been apparent.

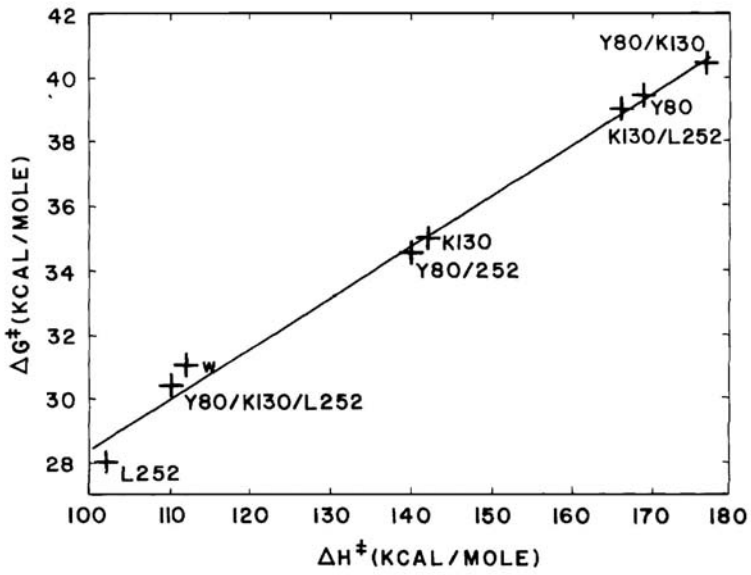


Figure 1. The compensation plot for thermal denaturation rates of kanamycin nucleotidyltranferase computed from the rate data supplied by M. Matsumura. The slope is the Pohl slope 354K. W wild type, Y80 asp to try at 80, K130 thre to lys, L255 leu for pro. The data illustrate the extreme sensitivity of the mechanical properties of knots to single residue substitutions. Single knot substitutions can destroy a knot or greatly strengthen it. Such substitions may explain the conversion of the non-infective state of prion to the infective state

Collapsed polypeptide structures gain unusual thermodynamic stability in at least two ways. Dehydration is very effective for proteins as first shown by Fujita and Noda but thermal denaturation as a phase transition still occurs at very low hydration even with little if any plasticizing water has been removed so this alternative is not a likely one. A second is suggested by the great strength and thermal stability of the spider silks. Although single BPTI knot has moostly aromatic residues rather than alanine and glycine in spider silks, it does contains three eight-membered rings each with two short peptide-peptide hydrogen bonds with very low B factors in a short antiparallel β sheet. This with the probable help from the three disulfide bonds prevents melting up to 373K, well above the common 354K limit for mesophiles. The larger knots of the streptococcal G protein has the same sheet configuration and short hydrogen bonds although maximum melting temperature appears to be no greater than

354K found by Alexander et al. The residues forming short peptide-peptide hydrogen bonds as reported for this protein are large and mostly aromatic so the permanent polarization over the two knots is low. This arrangement strengthens the nearby H bonds and may be the major basis for the short H bonds. The dramatic conversion of the protein by Regan and coworkers to the ROP protein by addition and rearrangement of aromatic and aliphatic side chains follows the same logic as discussed by Gregory and Lumry in connection with the curvature of helices embedded in protein surfaces.

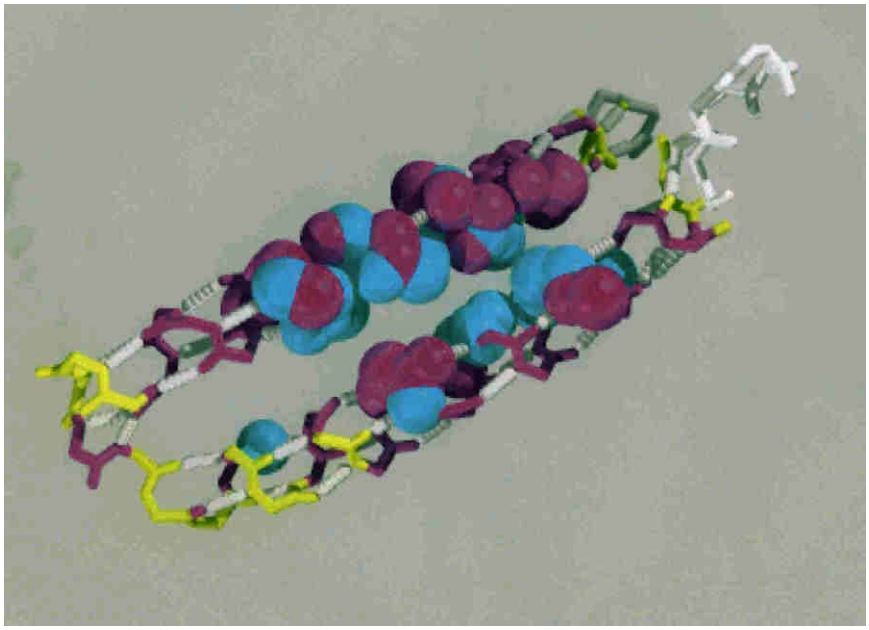


Figure 2 The ROP protein as constructed by Regan et al by residue exchange in the B protein of streptococcus. The original knots are replaced by the single long knots as shown. Helices and sheet secondary structures have variable strength depending on their residue composition and the dielectric properties of their contiguous regions. Knots that are pure secondary structures are rare and sections of secondary structures can be matrix or knot. Excessive attention has been given to secondary structures since they are easily detected and the alternative (knots and matrices in general) have not been obvious.

The cohesive strength of these "hydrophobic" clusters is not much enhanced by contraction in the clusters nor does it contribute very much to the free energy stabilizing the native species. As possible exception would be unusual contraction to increase dispersion and oribital overlap reductions in potential energy. Considering the small strength of the clusters and the relatively small

compression that seems unlikely. Instead the enhancement in strength appears to be due in largest part from the favorable dipole-dipole interaction of the H-bonding groups in an environment of low dielectric constant but rhe effect is not a simple additive one as discussed later in this chapter.

In Chap. @ the estimate of the reduction in potential energy in each cluster is made using the activation enthalpy per mole of residue compute for all resides in a mesophilic protein as 1530 cal but since only one in eight peptide-peptide H bonds form such cluster the value per knot H bond is 8 times larger, 12.7 kcal. Although this is the value for a cluster and not a single hydrogen bond, it is very large relative to the usual enthalpy of formation of hydrogen bonds but probably quite accurate and we have previously tried to rationalize it on the bases of proton delocalization reported by Kearly and coworkers. However since then Jeurgen and coworkers report that the normal-mode analysis used by the French workers is incorrect. Nevertheless there is a significant possibility that those silks have some special construction of peptide-peptide H bonds with exceptional properties as appear to be required for the very high thermal and chemical stability of the dragline silks and the strengths and insolubilities of the amyloid filaments and placks all dominated by anti-parallel β.filaments.

The low solubility of polyalanine and polyglycine as antiparallel β structures illustrated by those filaments does not appear to be explicable on the basis of current knowledge of simple polypeptides. Obviously missing is a full explanation of the cooperativity and the electron rearrangements likely to be responsible. The filaments have eight-membered and ten-membered rings each closed by interchain H bonds and arrays of these will resemble the aromatic rings of graphite if there is sufficient delocalization of electrons across the hydrogen bonds to support some of the in-plane resonance strength of graphite. These are peptide-peptide H bonds not be confused with the interdomain H

bonds involved that participate through chemical change in catalysis of trypsin and similar cofactor-independent enzymes. The latter derive their proton mobility from domain closure in the chapters on enzymes.. Short interpeptide hydrogen bonds in knots are produced by general knot contraction but the degree of covalency so produced is not known with any accuracy. Covalency is required to produce the graphite strength we suggest as a possible basis of spider silk properties and also perhaps of Kevlar, the competitor in strength with the dragline spider silks also with possibilities for graphite-like conjugation. The short hydrogen bonds in the peptide structures suggest some proton delocalization and some covalency but there is much uncertainty. In some extremeophiles from archaea there are examples of exceptionally short knot H bonds consistent perhaps with special cooperative electronic properties. Many of the potentially important consequences of the proton delocalization for proteins have been discussed in Chapter 29 or Methods in Enzymology 259 stimulated by the large proton delocalization reported by Kearley and coworkers for polyglycine. Jeurgen et al criticized the normal-coordinate analysis of the French workers for polyglycine correctly but the sheet substructures whether in matrices or knots eight- and ten-membered rings closed by pairs of peptide-peptide H bonds as shown in the following figure and it seems possible that under sufficient cohesive contraction the protons migrate across these H bonds to convert some of the H-bonded rings to primary-bond links as also shown below. The smallest B values in the sheet structures of the streptococcus G protein, nuclease T-1 and T-4 lysozyme are found near the center of these structures just as shown in the figure of Kearley et al. producing new primary bond clusters and short chains .that are the actual knots. This is illustrated for the G protein below.

The B factors of helix secondary structures in proteins do not make this point or at least we have not found examples. However it is now quite likely that this proton rearrangement s common in strong sheet knots, spider silks and

amyloid filaments explaining at least in part their exceptional strength.Zundel and coworkers find widespread proton polarization in side-chain H bonds of proteins some with proton hopping from one side to another in intermediate states of function This is the probable explanation of the mysterious behavior of low-barrier H bonds in proteins consistent with the formation of strong knots without special explanations not testable with the large coordinate errors in x-ray-diffraction data.

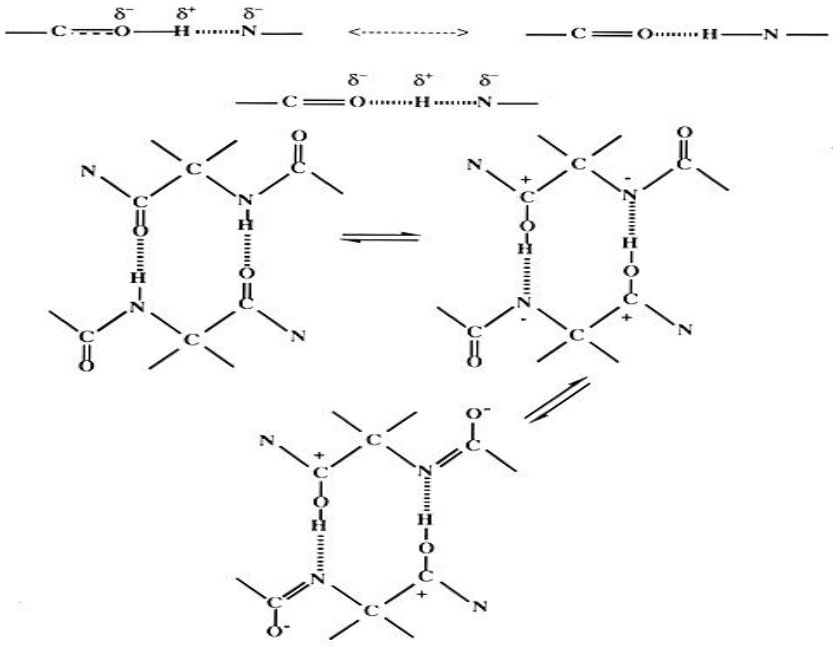
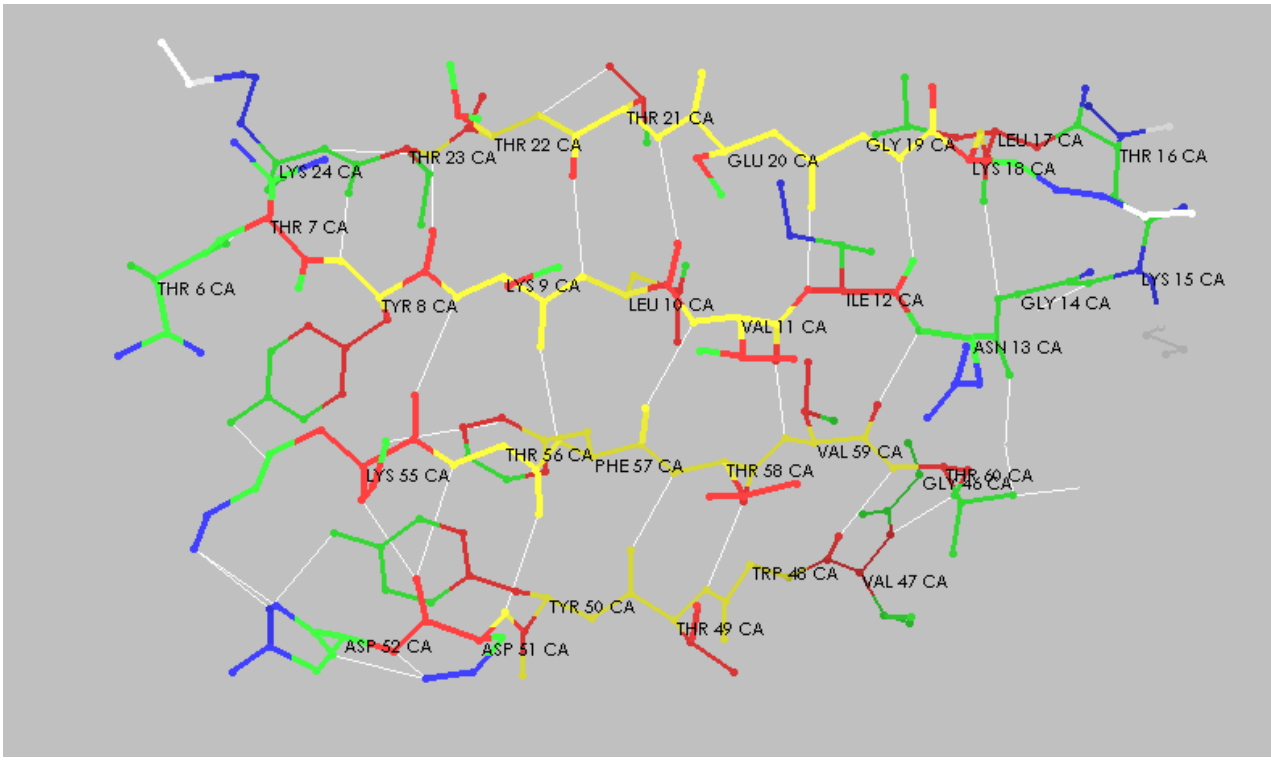


Figure 3. The electron and proton delocalization structures based on the proton delocalization deductions by Kearley et al from the H-bond proton vibrational spectrum of crystalline polyglycine. The eight-membered ringsappear favorable to contraction of their two H bonds toward something resembling the phenyl ring, the construction element of graphjtes . From Kearley and coworkers who used infra-red spectra.

Cornilescu et al found that the large "through-bond" (scalar) coupling coefficients in the peptide-peptide H bonds of the antiparallel sheet substructures of the streptococcal G protein are anticorrelated with the H-bond lengths established with unusual accuracy by x-ray diffraction. Furthermore the B factors from the PDB entries @ also anticorrelate with the j coupling. coefficients for the knot H bonds. These in turn are correlated with the proton-

exchange rates. The aggregate of these results establishes close relationships among bond length, coupling coefficient, free volume from the knot B factors and proton-exchange rates. All the short H bonds are found in the two knots, shown in yellow in $Fig. B @ as forming a single large sections of large antiparallel β sheets. Note however that the maximum melting temperature for this protein determined by Anderson and coworkers is 354K the general limit for small mezophilic proteins and several hundred degrees below the melting or decomposition temperature of the dragline silks..



The G protein is not an enzyme and has more perfect C-2 symmetry in its matrices than do most enzymes although its higher precision in this connection may be due the higher x-ray-diffraction resolution of the G-protein study. The knots are fused and appear to be too stiff for much bending of the H-bond hinge, also part of the β knot assembly. That knot structure is an important discovery insofar as it establishes the presence of some extensive multi-H bond cooperativity in peptide-peptide hydrogen bonds responsible for the unusually

strength of knots. However, the single small knot of BPTI with its short knot H bonds probably profits from the same advantage without cooperativity so the strength and possible covalent character may be due to clustering rather than inter-H bond cooperation due then to puzzling involvement of side chains. The compression is unlikely to be due to contraction of the matrices since knots appear always to have some short peptide-peptide hydrogen bonds.

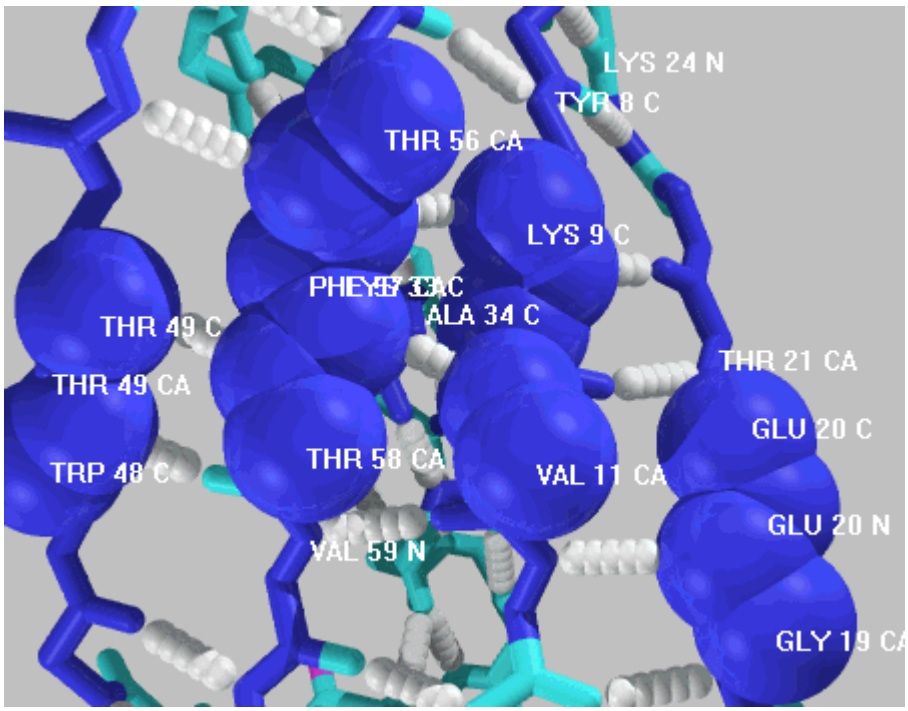


Fig. @. The two knots of the G protein of streptococcus. Twelve residues as shown may not be the complete knots since there are somoe attached atoms with only slightly larger B values. Those shown are 5. Nuclease T-1 and T-4 lysozyme have single knots constructed in this way and the pattern for sheet knots may be common.

The streptococcus G protein is a mesophile probably a very stable one and it may illustrate a quantitative difference in knots rather than a qualitative one. The work required to destroy knot cooperativity is the work done again the contraction forces in the knots and can be estimated from activation enthalpies in thermal denaturation or from the standard enthalpy of thermal denaturation

of dry proteins since both are dominated by potential-energy changes. Thus for ribonuclease A Pohl found 65 ±5 kcal/M for the activation energy in normal solution and Battistel and Bianchi found the standard enthalpy change to be 55±15 kcal/mole at 0.12 g water/g protein. The number of special H-bonds in the knot suggest activation enthalpies equivalent to about 10kcal/mole for each of the seven knot H bonds thus about the same as the enthalpy required to break one mole of water-water H bonds?

Fig. @ shows that six of the very short peptide-peptide H bonds are in three of the eight-membered rings consistent with the ring hypothesis and the considerable strength of this single-knot protein.

Large contracted antiparallel β sheet substructures are found in some archaea proteins and by analogy with spider silks give partial explanation for their extreme properties. The DNA binding protein (Fig. 13 @) is an example. It is a bacterial @ providing protection against thermal destruction. The two sections of its large, single knot fit into the two double-helix grooves. Neither B factors nor coordinates from x-ray-diffraction measurements have been reported. for the sheet knots. Helices might be expected to have similar properties but in the G-protein study of Cornilescu et al the j-coupling coefficients within the single long α helix were found to be small with H-bond lengths not shorter than those in model compounds. The α-helix H bonds show neither electron delocalization nor compression in the staphylococcus G protein.

Regan et al replaced sheet knots by helix knots by residue exchanges that reduced the dielectric constant between two helices. The large β-sheet knots were converted to helix knots very similar to those in the ROP protein by localizing leucine, norleucine, and phenylalanine groups in a continuous cluster between the two helices (Fig. 23). The resulting reduction in the local dielectric

constant caused contraction of the helix hydrogen bonds in contact with the cluster. The two knots so produced form a slightly curved sandwich of low-B atoms inserted cluster of sidechains. Adjustment of helix B values and the fine details of structure by such residue selection is quite common and most easily detected by the curvature of helices that have a polar environment on one side and a non-polar one on the other. This produces curvature toward the non-polar side and is common in helices that lie in protein surfaces. Cooperative packing of groups of residues to control local dielectric constants is obviously the basis of protein construction. Nevertheless almost all research on cohesive interactions in proteins is confined to attempts to assign helix-stability parameters to individual residues not likely to be very profitable unless focused on whole matrices.