**02.4-8   PROTEIN-PROTEIN RECOGNITION: THE INTERFACES OF PROTEIN INHIBITORS AND THE SERINE PROTEINASES.** By Michael N.G. James, Anita R. Sielecki, Randy J. Read, Masao Fujinaga, Catherine A. McPhalen and Harry Greenblatt, Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2H7.

There are 10 known families of protein inhibitors of the serine proteinases. Many of the inhibitors are relatively small proteins (50-60 amino acids). Others are small domains of larger multidomain proteins. We have determined the crystal structures of four of these inhibitors complexed to their cognate enzymes: the third domain of the turkey ovomucoid (OMTKY3) complexed to SGPB (1) and to α-chymotrypsin (2), the chymotrypsin inhibitor CI-2 from barley seeds complexed and inhibitor CI-2 from barley seeds complexed to subtilisin NOVO (3), eglin-c from leeches complexed to subtilisin Carlsberg (4) and the potato chymotrypsin inhibitor 1 (PCI-1) complexed to SGPB (see Figure). Eglin-c and CI-2 are from the same inhibitor family and have similar three dimensional structures. They are distinguished by the absence of disulfide bridges. The inhibitors of the serine proteinases form one of the best examples of convergent evolutionary pressures giving rise to a common function but with distinctly different 3-dimensional tertiary structures. All of these inhibitors are characterized by extremely tight binding ($K_d < 10^{-9}$ M) to their cognate enzyme. The common structural feature among the inhibitors is a segment of polypeptide chain spanning 6-7 residues which binds to the active site region of the enzyme. This segment has a very similar tertiary structure among all the inhibitors and resembles the conformation of an enzyme-substrate Michaelis complex. There is one peptide bond, between residues P_1 and P_1', which can be cleaved but only at a very slow rate, $k_{cat} \sim 10^{-6}$ sec^{-1}. There are several common hydrogen-bonding and electrostatic features that the different inhibitor families display in this region which contribute to these molecules acting as inhibitors and not good substrates. Water molecules play important roles at the interface between the inhibitors and the enzymes and most likely contribute to the recognition.

PCI-1
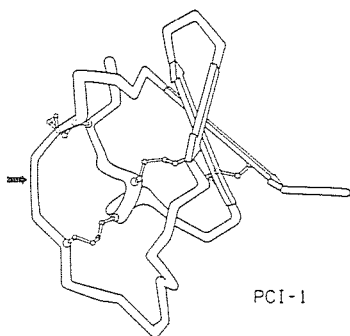
Potato chymotrypsin inhibitor. The 4 disulfide bridges in the molecule are indicated. The scissile bond between Leu-Asn is indicated by a solid arrow.

References

(1) Read et al., Biochemistry 22, 4420, 1983.
(2) Fujinaga et al., J. Mol. Biol., 1987, in press.
(3) McPhalen et al., Proc. Natl. Acad. Sci. 82, 7242 1985.
(4) McPhalen et al., FEBS Lett. 188, 55, 1985.

02.5-1   STRUCTURAL STUDIES OF DNA FRAGMENTS WITH BASE PAIR MISMATCHES. By O. Kennard and W. N. Hunter, University Chemical Laboratory, Lensfield Road, Cambridge, U.K.

The concept of base pair mismatches as a mechanism for introducing variations into genetic coding is embedded in the concept of DNA as the carrier of genetic information. The fidelity of transmitting the genetic code rests on specific pairing between components of the double helix: the Watson-Crick hydrogen pairs of normal DNA. Different combinations of the bases, however, if not detected and excised can result in modification of the genetic code and may eventually lead to mutational changes. The paper presents the results of X-ray diffraction studies of DNA fragments containing different types of base pair mismatches. It discusses the molecular structure of the ismatched base pairs, the way they influence the conformation of the double helix and possible mechanisms whereby such errors are recognised and corrected.
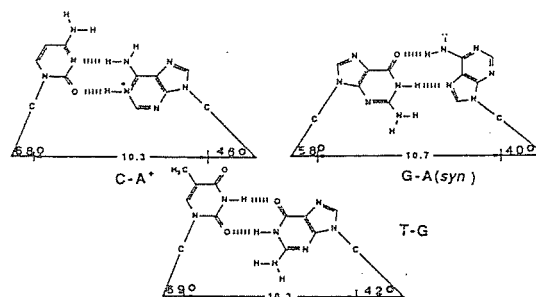
02.5-2   POSSIBLE ROLE OF PROTONATION - DEPROTONATION IN NUCLEIC ACID BREATHING: A HALF-SANDWICH OF WATER MOLECULE IN A TRINUCLEOTIDE COMPLEX OF $(A^+p\ Ap\ A^+)\cdot(Ap\ A^+p\ A^+)$. R. Parthasarathy, T. Srikrishnan and S. M. Fridey, Center for Crystallographic Research and Department of Biophysics, Roswell Park Memorial Institute, 666 Elm Street, Buffalo, NY, 14263.

Crystallization of A3'p5'A3'p5'A (ApApA) at acidic pH yields two crystalline forms I and II in space groups $P4_12_12$ and $P4_1$ (Z=8 in both forms), respectively. An earlier study (D. Suck , P.C. Manor & W. Saenger, Acta Cryst. B32,1727,1976) of I and a preliminary study of II indicated that the molecule exists in the crystal as $Ap\ A^+p\ A^+$. Our detailed examination of II shows that this crystal form actually contains the dimeric complex (A1$^+$p A2p A3$^+$)·(A4p A5$^+$pA6$^+$), one of the six possible combinations of protonated dimers. The assignments of these charges are based on a combination of criteria involving difference electron density maps, ring angles at N(1) and the stereochemistry of hydrogen bonding. The two molecules of ApApA have significant conformational differences which are directly coupled to the differences in protonation of the two molecules in the complex. The microenvironment around the adenine bases decides which two of the three bases in each molecule will be protonated. The dimer is stabilized by two self pairs between adenines A2:A6 and A3:A5 and stacking of A1,A2,A5 and A4 on top of one another. The A2:A6 self pair is between a charged and uncharged adenine and the A3:A5 self pair is between two charged

adenines. The bases in the base pairs are not in a plane and are bent about the center of the duplex forming a pleat. A3 and A6 which swing out are within stacking distances but have only limited partial stacking. A water molecule OW2 hydrogen bonded to N(7) of A1 is stacked on top of A6 forming a half-sandwich (R. Parthasarathy, T. Srikrishnan & S. L. Ginell in Biomolecular Stereodynamics, R. H. Sarma, Ed. Adenine, New York, 1983). On the other hand, the base A3 of the dimer is not involved in such a stacking interaction with a water molecule since N(7) of A4 is hydrogen bonded to a neighboring A1 and not to any water molecule. Crystals of II are tetragonal, $P4_1$ with cell constants a = b = 14.070(1), c = 43.906(5) Å, Dobsd. = 1.56 g.cm^{-3} for $(C_{30} H_{37} N_{15} O_{16} P_2)_2$. 11 H_2O. Using complete three-dimensional intensity data to the limit of Cu sphere (10077 reflections, 7044>3σ), the structure was solved using MULTAN 80 and weighted Fourier syntheses to an R-value of 0.064. All the six adenosines show _anti_ conformation across the glycosidic bond and a range of puckers of the C3'-_endo_ family. The two helical segments A1pA2 and A4pA5+ show the preferred (g$^+$,g$^+$) conformation whereas the non-helical loop segments A2pA3$^-$ and A5$^+$pA6$^+$ show (g$^+$,g$^+$) conformation. The agreement index R_{symm} between hkl and khl for form II is 0.38 indicating the lack of crystallographic two fold axis. We also have collected data from other crystals for which R_{symm} had values of 0.08, 0.12 and 0.32. Work is in progress on these crystals. Work supported in part by NIH GM 24864.

02.5-3   THE   Z-DNA   COMPATIBLE   SEQUENCE d(GTGTACAC) CRYSTALLIZES AS A-DNA.

S. Jain[1], G. Zon[2], M. Sundaralingam[1]
1. Department of Biochemistry, University of Wisconsin, Madison, WI 53706, 2. Applied Biosystems, Foster City CA 94404.

The molecular structure of a self complementary DNA octamer, GTGTACAC has been solved to 2.5 A resolution using single crystal x-ray diffraction methods. The cell constants of the crystal are, a=42.43 A, c=24.75 A, space-group, $p4_32_12$, Z=4 (duplexes). The molecule forms a right-handed A-DNA-like double-helical structure with characteristic shallow and deep grooves, but does not have the typical base-pair tilt of A-DNA. Currently the residual R-factor for the reflections out to 2.5 A is 24.5% and the waters of hydration have still not been included. Further refinement of the structure is in progress.

The results are of much interest since (GT)n/(CA)n sequences are widely distributed in natural DNA (E.N. Trifonov, et.al., 1985, FEBS 185, 197). Moreover, the GTG/CAC triplet occurs very frequently in genomic regulatory regions, and also has some unusual physical properties (P. Lu, et.al., 1983, J. Biomolec. Struc. Dyn. 1, 509). Some investigators have suggested that these sequences may be involved in formation of the left-handed Z-DNA structure (A. Nordheim, A. Rich., 1983, Proc. Nat. Acad. Sci. USA 80, 1821), but in-vivo experiments have yielded no evidence in favor of this hypothesis (A. Rodriguez-Campos, et.al., 1986, EMBO 5, 1727; D.S. Gross, et.al., 1985, J. Mol. Biol. 183, 251). Our results also demonstrate that such a sequence adopts a right-handed rather than a left-handed conformation.

02.5-4   NUCLEIC ACID JUNCTIONS AND MACROMOLECULAR DESIGN.

By N.C. Seeman, C.J. Newton, M.L. Petrillo, J.-H. Chen, J.E. Mueller, J.A. Maiorella, and R.D. Sheardy, Department of Biology, SUNY/Albany, Albany, NY 12222; R.-I. Ma and N.R. Kallenbach, Dept. of Biology, University of Pennsylvania, Philadelphia, PA 19104, USA.

The Watson-Crick base-pairing interactions of nucleic acids constitute a particularly powerful system for controlling the structure and connectivity of these polymers. For many years, molecular biologists have formed specific linear double helical molecules by mixing pairs of complementary single strands. Recently we have shown that stable branched nucleic acid complexes, called junctions, can be formed from mixtures of 3, 4 or 5 oligonucleotide single strands with carefully selected sequences (N.R. Kallenbach, R.-I. Ma and N.C. Seeman, 1983 Nature 305 829-831); these complexes contain multiple double helical arms all coming from a central point. The fundamental rule in implementing the sequence selection algorithm is the minimization of sequence symmetry.

Nucleic acid junctions are analogs of ephemeral intermediates seen in the processes of replication and recombination, and we are exploring the structural, dynamic and thermodynamic properties of these structures from that perspective. From circular dichroism spectroscopy, it is clear that the junction does not perturb the structure of the arms. Nuclear magnetic resonance spectroscopy has indicated that the bases which flank the junction are paired. Combined gel electrophoretic and oligomerization-ligation studies have indicated that a large range of 3-dimensional structures are available to junctions. Crystallization is in progress.

Besides being important objects of study, junctions may be regarded as macromolecular valence clusters with specifically addressable ends, particularly if asymmetric sticky-ended associations are employed. The idea is to construct geometrical figures and N-connected networks, in which the edges are double helical nucleic acids, while the vertices are nucleic acid junctions. Preliminary experiments indicate that hydrogen-bonded base-pairing of cohesive ends can be used to direct the formation of linked clusters as well. It is worth noting that this system is more complicated than simple valence clusters composed of atoms connected by bonds: In this larger system, the twist of the double helices plays an important role in determining the shape of the products. Indeed, changing the number of nucleotide pairs between junctions can alter both the geometry and the topology of an array of junctions. In addition, it appears that the system can respond to torsional stress by altering the structure of the junction itself.

This junction-association system offers a useful paradigm for understanding crystal formation, since the intermolecular contacts are pre-determined: It has prompted the recent suggestion of an entropic driving force to explain the small number of molecules (typically one) seen in the asymmetric units of molecular crystals. In particular, we have put forward the idea that the maximization of entropy in reciprocal space is similar to the expansion of a gas in direct space, $\Delta S = R \ln (\bar{V}_2/\bar{V}_1)$ (N.C. Seeman, 1985 J. Biomol. Str. Dyns. 3, 11-34). The goals of this part of the work also include the rational design of crystals and the fabrication of mechanical and electronic devices on the nanometer scale.