Thermodynamic Characterization of DNA Triple-Helical Three-Way Junctions

Paul. L. Hüsler

August 1995

Department of Biochemistry
University of Cape Town RSA.
Thesis Submitted in Fulfillment of the Degree of Doctor of Philosophy

Acknowledgment

I wish to thank the Foundation for Research and Development, Pretoria, RSA, for funding this project, my supervisor Professor H.H. Klump for all his help, and Dr Jens Völker for his constructive input.

Abstract

Watson-Crick (WC) and Hoogsteen (HG) triple-helical three-way junctions were constructed from three 33-mer oligonucleotides. The same sub-set of sequences have been used in the arms of the junctions. The junctions differ primary in the arrangement of the branch point and the ends of the arms. In the case of the HG triple-helical three-way junction, the three 33-mer oligonucleotides can fold into hairpin structures, linked by a four membered cytosine loop. Each of the hairpins contain a homo-pyrimidine 10-mer single strand extension which interacts with a neighboring hairpin to form a triple-helix on lowering the pH (between 6 and 4), via Hoogsteen (HG) hydrogen bonding. Collectively this process results in the formation of the branch point and the triple-helical arms. In the case of the WC triple-helical three-way junction, the three 33-mer oligonucleotides interact with a neighboring oligonucleotide to form a duplex. Collectively this process leads to the formation of a double-helical three-way junction with each arm containing a 9-mer homo-pyrimidine extension connected by four cytosines. Each of the single strand extensions can mutually fold back onto the duplex arms converting each arm into a triplex.

The WC and HG triple-helical three-way junctions were characterized by gel electrophoresis, temperature gradient gel electrophoresis, circular dichroism (CD), uv melting, and differential scanning calorimetry (DSC). In both structures; arm A contained exclusively TAT triad bases, while arms B and C contained an increasing number of CGC^+ triads, respectively. To counteract possible crowding at the branch point the 5' sequence was shortened by one base. The assembly of the completely folded structure was found to be spontaneous if an appropriate ionic strength and pH range was chosen.

A separate set of isolated arms has been investigated to elucidate the role each arm plays in the complete structure. Comparing the summed-up properties of these arms with the data obtained for the integral three-way junctions, it is obvious that the WC three-way junction is partly distorted at the branch point, in line with observation obtained from double-helical three-way junctions, while the HG three-way junction is completely ordered.

A set of mathematical models has been developed to describe the thermal unfolding of the multi-strand DNA structure and to identify the intermediate states. Presented is a formalism, starting from the grand partition function, that describes the effects of pH on the thermal stability of triple-helices. The formalism can be used over a wide pH range. It covers nearest neighbor electrostatic effects of closely spaced cytosines in the Hoogsteen and Watson & Crick strands. A procedure is employed to predict enthalpy and entropy changes for triplex formation. The obtained values are in good agreement with the results obtained by differential scanning calorimetry. It is the first time that multistrand, branched DNA structures of this complexity were constructed, completely described, and characterized thermodynamically.

Table of Contents

Introduction

Design and Theoretical Background

Materials and Methods

Results and Discussion

Abbreviations

A	Adenine
Abs	Absorbance
C	Cytosine
C^+	Protonated Cytosine
cal	Calorie (1 calorie = 4.18J)
CD	Circular Dichroism
Cp	Molar excess heat capacity
DSC	Differential scanning calorimetry
EDTA	Ethylendiaminetetra acetic acid
G	Guanine
ΔH	Molar enthalpy change (cal/mol)
H^+	Proton cation
HG	Hoogsteen
HPLC	High performance liquid chromatography
J	Joules
LOG	Logarithm to the base 10
Mg^+	Magnesium cation
Na^+	Sodium cation
NMR	Nuclear magnetic resonance
PAGE	Polyacrylamide gel electrophoresis
R	Molar gas constant
ΔS	Molar entropy change ($cal.mol^{-1}.k^{-1}$)
T	Thymine
Tm	Temperature mid point
Tmax	Temperature maximum (where $\partial Abs/\partial T = 0$)
U	Uracil
UV	Ultra-violet
WC	Watson-Crick

Chapter 1.

Quaternary Structure of Three and Four-Way Junctions

Prologue

In 1953 (a & b), Watson and Crick postulated a model for the secondary structure of DNA in which two complemetary DNA strands interact to form a double helix. In their first paper they showed that "the novel feature of the structure is the manner in which the two chains are held together by the purine and pyrimidine bases," from which they deduced that "only specific pairs of bases can bond together." Subsequently, they were able to postulate the relationship between DNA structure and the manner in which genes are coded: "It follows that in a long molecule many different permutations are possible, and it therefore seems likely that the precise sequence of the bases is the code which carries the genetic information."

Since this historic event, it has been shown that deviations from linearity in DNA structures occur naturally during a number of cellular activities, such as genetic recombination; replication; transcription and transiently during post transcriptional modification of nucleic acids. During recombination, genetic material may undergo insertions, deletions, changes in sequence, rearrangement or exchanges of flanking sequences. Even though these anomalies occur at relatively low frequencies and are

usually short lived as intermediate states, their biological function is essential for cellular processes. The best known representative in site specific recombination and in homologous recombination is thought to be a four-way junction (Holliday junction) first proposed by Holliday (1964). Other transiently stable nonlinear structures are known, such as three-way junctions in the fork of replicating DNA and in tRNA.

Holliday junctions are formed by strand exchange between homologous duplex regions in DNA. The branch point or junction formed in this way is free to migrate in either direction along the homologues sequence. These naturally occurring junctions are intrinsically difficult to study as a consequence of branch point migration, their infrequency, and transient life span. To overcome this difficulty, synthetic, stable, immobile DNA junctions, in which the two-fold sequence symmetry have been eliminated, have been proposed (Seeman and Kallenbach., 1983; Kallenbach *et al.*, 1983; Seeman., 1985).

Branched structures occur frequently in RNA secondary and tertiary structures. These three-way junctions occur commonly in many biologically active RNA molecules, such as hammerhead rybosomes; 5S, 16S and 23S ribosomal RNAs, tRNA, and in the single stranded RNA and DNA genomes of certain viruses.

Besides their biological importance, stabie DNA helices in conjunction with stable DNA branched junctions can be used to create designed closed structures (Seeman., 1981; 1982). These structures are intended to serve as model scaffolds for synthetic chemistry. The possible field of application is wide, ranging from the formation of

macromolecular zeolite lattices to the caging of active biological molecules leading to the possible development of drug delivery systems, mechanical control on the manometer scale, and the assembly of macromolecular electronic components. The potential of such structures was suggested by Zhong and Seeman (1992): "In principal, the attachment of a series of cohesive (sticky) ends to a DNA branch junction converts it into a valence cluster whose ends are specifically available for binding only complementary ends".

Barton *et al.* and Meade *et al.* (Clery 1995) have recently demonstrated that DNA molecules are able to channel electrons along the DNA axis. It is not known if DNA junctions have the same features, but they could serve as 'soldering points' in the design of more complex and elaborate macromolecular circuits.

To date, branched DNA junctions have only been incorporated into WC double-helices, but in principle junctions can also be assembled from non-WC DNA structures like triple and quadruple-helices. A thorough thermodynamic understanding of DNA/RNA folding is therefore crucial for the design of new secondary and tertiary structures.

1.1. Design of Immobile Branched DNA Structures

Naturally occurring Holliday junctions are inherently two-fold sequence symmetrical and are consequently free to undergo branch point isomerisation (migration), resulting in a population of structures occupying different branch positions. For a complete characterization of the structural, dynamic, and thermodynamic properties of branched junctions, nucleic acid molecules have been assembled into stable branched structures, in which the branch point has been frozen and alternative competing structures have been eliminated. To this aim, stable immobile junctions have been constructed from oligonucleotides by minimizing sequence symmetry elements in the vicinity of the branch points. This holds for three-way and four-way junctions alike (Seeman and Kallenbach 1983). Symmetry minimization can be introduced by ensuring that any of the tetramer sequences incorporated into the structure is unique and that tetramer sequences spanning the branch point are not complementary. Removing the linear complement of a sequence ensures that no stable alternative pairing can take place with another sequence on any of the other arms. In four-way junctions any two fold sequence symmetry around the branch point must also be eliminated to prevent branch point migration. If the arms of the junction are extended, sequences larger than the tetramer have to be considered as building blocks in their design. Restriction sites can also be incorporated into branched DNA. This is usually done by first placing the restriction sites into the desired position of the junction and then minimizing the sequence symmetry around these sites.

1.2. Pairwise Coaxial Stacking of Helical Arms in Four-Way Junctions

As early as 1972 Sigal and Alberts had proposed a structure for the Holliday junction based on model-building studies. They suggested that pairs of helical arms stack on top of each other, resulting in two continuous helical domains containing two cross-over strands and two continuous strands. The stacked arrangement in the model also suggests that the bases abutting on the branch point remain paired and stacked. By minimizing sequence symmetry in four-way junctions Kallenbach *et al.* (1983) have demonstrated that stable, immobile DNA four-way junctions can be assembled from four complementary hexadecadeoxynucleotides. NMR studies on these synthetic immobile four-way junctions by Wemmer *el al.* (1985) and Chen *et al.* (1991) further demonstrated that base pairing around the branch point did indeed take place. These findings were later confirmed indirectly by differential scanning calorimetric studies (Marky *et al.* 1987) and chemical footprinting experiments (Churchill *et al.* 1988; and Lu *et al.* 1990).

Linear DNA molecules containing a short branched junction have been characterized by gel electrophoresis to establish whether branching confer a bending element on DNA. It was shown that the mobility of the DNA is strongly dependent on the relative position of the branch (Gough and Lilley 1985). The extent to which the electrophoretic mobility is reduced has been taken as being indicative of the degree of bending. Other groups (Cooper and Hagerman 1987; Duckett *et al.* 1988; Seeman *et al.* 1989; Duckett and Lilley 1991) have used this feature to investigate the preferred

stacking arrangement in branched DNA, by shortening or adding extensions to the ends of two of the arms.

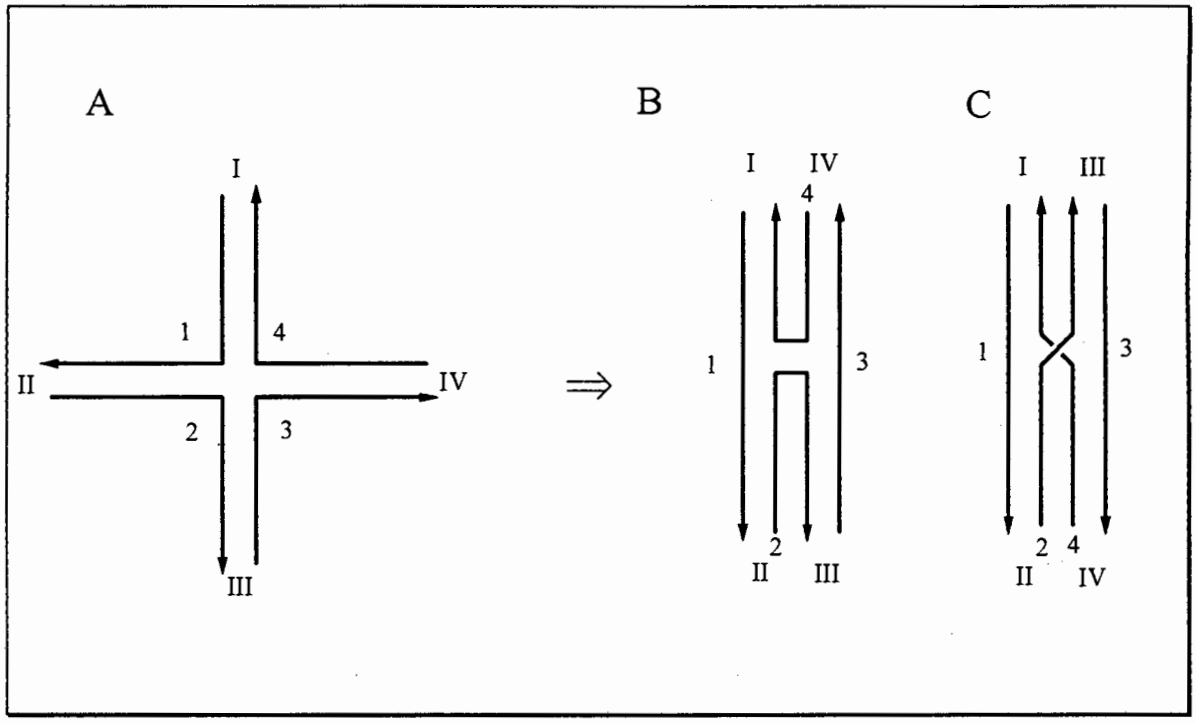

Fig. 1.1 Schematic representation of an immobile junction showing the extended structure (A) and the two stacked isomers corresponding to the antiparallel and parallel structures (B and C, respectively). The arrows indicate the 5' to 3' orientation of the oligonucleotides 1-4. Each of the oligonucleotides participate in forming the double-helical arms of the junction. In this scheme only two of the four possible stacked isomers are shown, where arm I is stacked onto arm II, and arm III onto arm IV.

Comparative studies on the relative electrophoretic mobility of four-way junctions with extended arms (in different combinations) have shown that the geometry of the junction is consistent with an antiparallel stacking arrangement, with the collinear strands lying in opposite directions to each other (cf. Fig. 1.1). Churchill *et al.* (1988) confirmed these findings in footprinting experiments using hydroxyl radicals generated by $Fe(EDTA)^{2-}$. It was shown that the cross-over strands are protected in the vicinity of the branch, while the continuous strands exhibit similar protection patterns to control duplexes. Models of the two-domain stacked arrangement predict that phosphate negative charges associated with the cross-over strands will lie in close proximity to each other at the branch point (von Kitzing *et al.* 1990). Footprinting experiments with differently charged probes are in accordance with this idea. For example, the reactivity of the cross-over strands to Fe^{2+} (Lu *et al.* 1990b) were shown to be enhanced at the branch site relative to the neutral and/or negatively charged probes Fe(EDDA) and $Fe(EDTA)^{2-}$, respectively.

Cooper and Hagerman (1989) have calculated the angle between pairs of arms using transient electric birefringence experiments. Their findings are in accordance with the proposed model of an antiparallel stacked arrangement. Murchie *et al.* (1989), using fluorescence energy transfer experiments, have suggested that the junction is more closely related to a right-handed X-like structure in which the helical stacked domains are tilted out of the plain in opposite directions (Clegg *et al.* 1992; Cooper and Hagerman 1990). Their model is also supported by DNase I protection patterns (Muchie *et al.* 1990) along the continuous strands, in particular where the phosphate

backbone turns into the major grove of the opposite helix (cf. Fig. 1.2). The X-ray diffraction pattern of dodecanucleotide crystals have also shown similar interaction patterns in the packing arrangement of the double-helices within the crystal lattice (Timsit and Moras 1991).

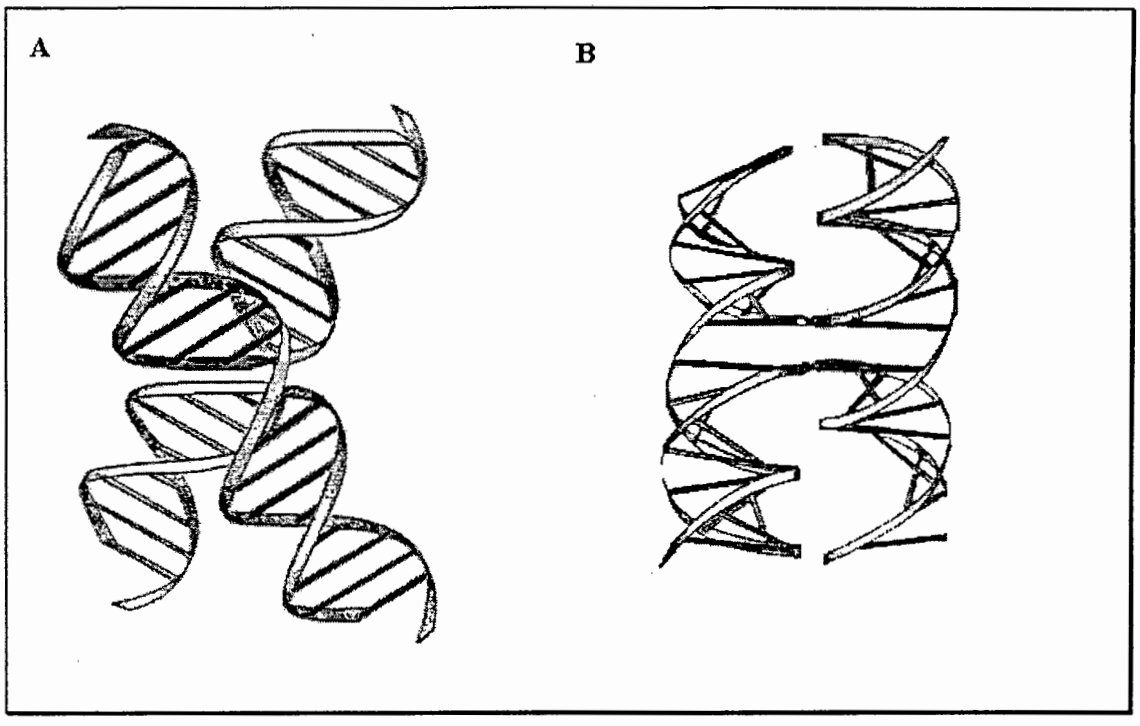

Fig. 1.2 Top and side views of the antiparallel stacked X like four-way junction, adopted from Murchie, A. I. H., Portugal, J., and Lilley, D. M. J. (1991) *EMBO.* **10**, 713-718.

The continuity of coaxial stacking of helical arms has been investigated by a type 1 restriction enzyme *Mbo*II (Murchie *et al.* 1991). The recognition site was positioned in one of the helical arms so that cleavage would take place in the adjoining coaxially stacked arm. Cleavage was indeed observed in the expected coaxially stacked arm,

inplying that the stacked arms form a continuous helical domain through the branch point. However, it is not clear whether the enzyme confers this type of structure or whether it is an inherent property of the junction. The results further suggest that an alternately stacked isomer may be in equilibrium with the former. This was concluded from an additional, but much weaker, cleavage site in one of the other arms.

1.3. Stacking of Helical Arms is Influenced by Bases Flanking the Junction.

The helical stacked arms around the junction can potentially undergo isomerisation. Only one of the isomeric arrangements is, however, usually formed. The preferred stacking arrangement (isomer) was shown by Duckett *et al.* (1988) to be influenced by the bases at branch point. The equilibrium can be shifted in favor of another isomer or stacked arrangement by simply changing the sequence of the bases abutting on the branch point.

Mismatches at the branch point can occur during recombination between heteroduplex DNA sequences. Duckett and Lilley (1991) have considered the effect of such mismatches in model four-way junctions. A range of effects was observed, from abolition of folding, through to normal branch formation. In some cases gel electrophoresis results suggest that there is a dynamic equilibrium between the folded and unfolded conformations. It was, furthermore, shown that the equilibrium could be shifted in favor of the fully folded structure by simply raising cation concentrations. In

the case of a single C·A mismatch, one of two alternate isomeric conformations was observed depending on the pH of the solution.

1.4. Counter Ion Dependence of Four-Way junction Folding

The folding properties of many biological nucleic acids require additional counterions to neutralize the excess charge accumulation where the phosphates are in close proximity to each other. This is particularly prominent for tRNA, ribozymes, and Holliday junctions, where Mg^{2+} and other divalent cations have been shown to play an important role in the tertiary folding of these structures. In particular, the tertiary structure of four-way junctions has been shown to be strongly dependent on the presence of divalent cations, especially Mg^{2+}. In the absence of Mg^{2+}, pairwise stacking of the arms is abandoned for a simpler square planar configuration. This holds for structures with arms of 25 base pairs or longer, while smaller junctions are incapable of forming stable structures at room temperature. This is not surprising in view of the large negative charge density accumulation at the branch point. It is therefore apparent that stable junctions require extensive neutralization of the negative charge associated with the cross-over strands.

Few studies have examined the detailed geometry of ion condensation in branched DNA. Most computational studies on higher order DNA structures have assumed that ion binding is similar to that of linear DNA. This will in general result in an underestimation of the extent of ion condensation in branched DNA structures. To demonstrate this, Olmsted and Hagerman (1994) have computed the number of

additional counterions associated with a model four-way junction relative to linear DNA. Using a grand canonical Monte Carlo computational procedure, they have shown that on average, typically at least 24 excess counterions can associate with four-way junctions in the long branch limit. There is not much known about the situation in three-way junctions. It can, however, be concluded from the close proximity of phosphates at the branch point that extra cations will also be required.

1.5. Antiparallel Arrangement is Preferred in Four-Way Junctions

Four-way junctions with two cross-over strands and two stacked helical domains can in theory exist as four isomeric forms. The four isomers represent junctions in which either set of strands can cross-over. The non cross-over strands can lie either with a parallel or antiparallel orientation to one another. Junctions are in principal free to undergo isomerization, but show a strong bias for the antiparallel isomers. Kimball *et al.* (1990) have demonstrated, using hydroxyl radial footprinting, that all four isomers can be constructed by connecting pairs of arms in different ways, with a short tether of thymines. The free energy difference between the antiparallel- and parallel-tethered junctions, with identical sequences in the arms, is relatively small and amounts to a mere -1.1 kcal/mol junction (Lu *et al.* 1991).

1.6. Monomobile and Bimobile Junctions

Monomobile and bimobile junctions have been constructed in which one- or two-step migration is possible (Chen *et al.* 1988; Lu *et al.* 1990a). Hydroxyl radical protection assays of these junctions show that, in the case of the monomobile junction, two antiparallel stacked isomers are present depending on the sequences of the bases at the branch point. In the case of the bimobile junction, protection is observed in each of the four strands, indicating that there is no dominant cross-over bias. This finding suggests that the cross-over bias may disappear as more steps of branch migrational freedom become accessible.

1.7. Thermodynamics of Four-Way Junction Formation

The thermodynamics of forming a branched junction relative to unbranched duplex DNA has been investigated by Lu *et al.* (1992) using competition assays and titration microcalorimetry. Two immobile junctions were constructed in these experiments, J_1 and its compliment J_1^C, in addition to the four DNA duplexes formed by combining different combinations of the same set of complementary strands. In these competition studies, labelled strands are allowed to partition between the junction and double-helices. The ratio of partitioning of the labelled strand was quantified by native gel electrophoresis. In the titration experiments, the enthalpy of branching was determined by forming either the junction or the double-helices directly from the single strands. It was shown that the difference in stability between the junction and the corresponding duplexes is relatively small at 18 °C, amounting to about 1.1 (±0.4)

kcal mol^{-1}. This value is in accordance with the results obtained from distribution studies between the duplexes and junctions by gel electrophoresis competition assays. The enthalpy and entropy changes associated with branch formation amounts to 27 (±1.3) kcal mol^{-1}, and 89 (±30) cal mol^{-1} K^{-1}, respectively. Furthermore, it was demonstrated that branch formation is accompanied by an apparent ΔCp of 0.97 (±0.05) kcal mol^{-1} k^{-1}. The positive ΔCp value is an oddity which is not usually observed for nucleic acids. Unlike nucleic acids, proteins usually have large ΔCp values as a result of buried hydrophobic groups becoming exposed to the solvent (Freire 1994). In four-way junctions, exposure of bases to the solvent is unlikely. The origin of the apparent ΔCp may be changes in the hydration properties of DNA on forming a branch (cf. review, Seeman and Kallenbach, 1994).

1.8. Double Cross-over Junctions

DNA molecules with two cross-over regions have been proposed as possible intermediates in double-strand break mechanisms. Because of their involvement in recombination it is of interest to study these molecules. They also provide an alternative means to study the less stable parallel configuration of four-way junctions, as double cross-over molecules can be designed with either parallel or antiparallel stacked helical domains. Thus, design of double cross-over molecules is considerably more complex than single cross-over molecules.

Double cross-over complexes constructed from oligonucleotides have been characterized by Fu and Seeman (1993). In general, increased tortional strain will result unless the distance between the cross-over regions is separated by half helical turns. It is possible to construct either parallel or antiparallel double cross-over molecules (cf. Fig. 1.3). For the antiparallel molecule there are only two stable structures possible, with either an even (DAE), or an odd number of helical turns (DAO). Four strands are required in the assemblage of antiparallel double cross-over structures with an even number of half helical turns, while five strands are required in molecules containing an odd number of turns. There are three types of parallel double cross-over molecules. Two types can be constructed with an odd number of half helical turns (DPOW and DPON), and only one for an even number (DPE).

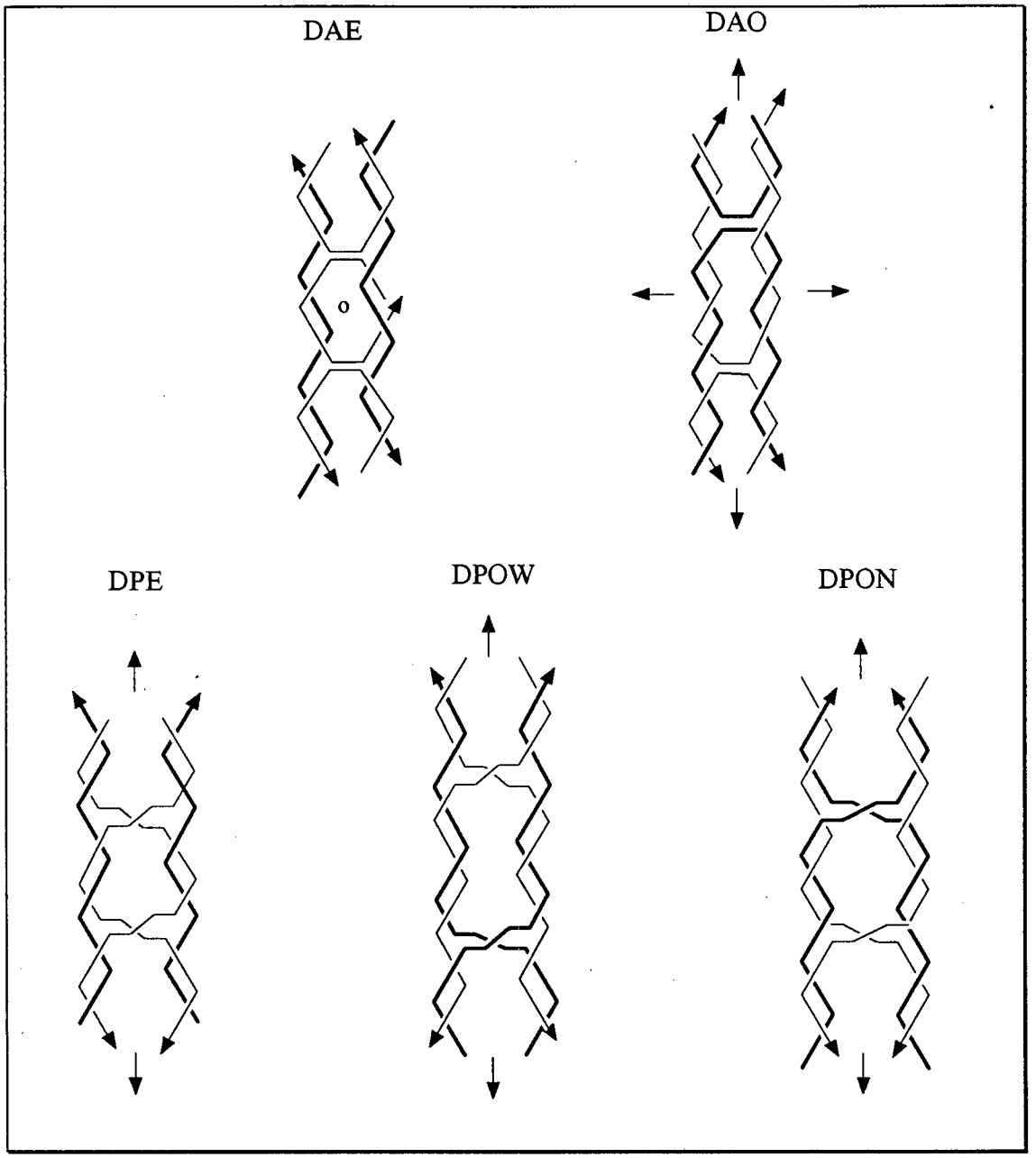

Fig. 1.3 Schematic drawing of the five double cross-over DNA structures. The structures are named according to their basic characteristics. All the abbreviations start with the letter D for double cross-over. The letter that follows indicates the orientation of the double-helices, A for antiparallel and P for parallel. The third letter refers to the number of half helical turns between the double cross-over regions, O for odd and E for even. In parallel double cross-over structures with an odd number of half helical turns two structures are possible. The last letter refers to a major half helical turn separation (W for wide) or a minor half helical separation (N for narrow). Figure adopted from Fu, T., and Seeman, N. C. (1993) *Biochemistry.* **32**, 3211-3220.

All require four strands to form the complexes. The parallel complexes are inherently unstable, they either fall apart or, at high concentrations, oligomerize into multimers. The instability of these complexes may be attributed to the crowding of phosphate residues in the central regions. Antiparallel junctions on the other hand are well behaved, each forming only one complex, presumably because its phosphates are not in direct opposition to each other. Hydroxyl radical footprinting of these molecules shows considerable protection in the cross-over regions. Central helical domains were shown to be protected as well, but to a lesser extent for parallel molecules with more than 1,5 turns between cross-overs. The weaker protection pattern seen in parallel molecules with more than one and a half turns suggested that the central helical regions may bow, so as to accommodate charge accumulation near the central helical axis. Antiparallel complexes, on the other hand, seem to remain linear without any helix distortions. The results obtained for double cross-over molecules are in accordance with single cross-over tethered junctions, where it was shown that parallel junctions are less stable than their antiparallel analogs.

1.9. Antijunctions and Mesojunctions

Studies of the general behavior of knotted DNA systems (Seeman 1992 & 1993) and their design, unveiled a more general class of structures which include antijunctions and mesojunctions. It is worth commenting on some of their properties and how they are related to junctions and double cross-over DNA molecules.

Fig. 1.4 demonstrates this principle with three- and four-stranded complexes. Antijunctions can be constructed from double-helices, arranged circumferentially around the branch point, while mesojunctions arise from a mixture of circumferential and radial helices. The lines representing the single strands are drawn flanking a central polygon with the arrows indicating the 5'→3' polarity of each strand. The point at which the lines intersect correspond to the vertices of the branched structure, while the parallel stacked thin lines represent the base pairs between complementary strands. The helical axis of each duplex is marked by an arrow perpendicular to the base pairs. It is evident from this figure that the helical axes can be arranged circumferentially or radially to the polygon or as a combination of both orientations. It is also apparent that certain topological limitations arise should all nucleotides be paired. For example, no three-armed antijunctions can be constructed with all helices circumferential. Branches with three strands can, therefore, only be represented by radial three-way junctions or mesojunctions. The convention introduced by Du *et al.* (1992) to name these structures denotes the number of strands, followed by a subscript to indicate the number of radial helices. The preceding superscript is a serial number used to identify different mesojunctions with the same number of strands and radial helices.

Du *et al.* (1992) have constructed and characterized the branched structures depicted in Fig. 1.4 containing 16 nucleotides per arm (equating to approximately 3 half turns). Gel electrophoresis indicates that the antijunctions and mesojunctions have a tendency to oligomerize into unknown structures of higher molecular weight.

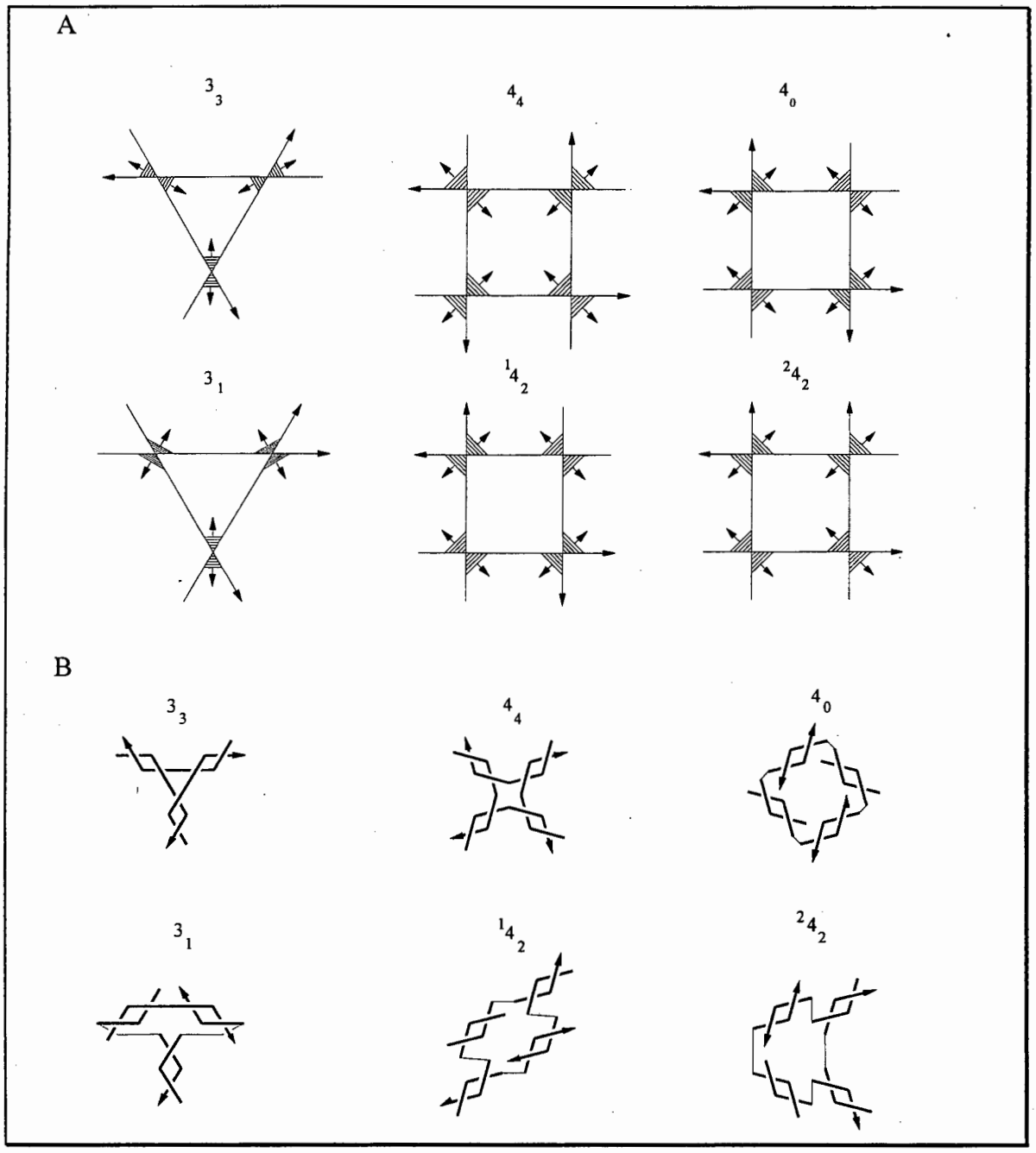

Fig. 1.4 Schematic representation of the three- and four-stranded junctions, antijunctions, and mesojunctions. (A) Each polygon is flanked by the single DNA strand, the arrow heads of which indicate the polarity of the strands. The parallel lines at the vertices of the polygon represent the base pairs and the short arrows, perpendicular to the base pairs, the helical axis of each duplex region. B represents drawings of the schematically presentation of the structures in A. Figure adapted from Wang, H., and Seeman, N. C. (1995) *Biochemistry.* **34**, 920-929.

Oligomerization can be suppressed by using conditions that do not favor polymer formation such as higher temperatures and lower oligonucleotide concentrations. It should be noted that similar results were observed for certain double cross-over junctions, which are closely related to meso- and antijunctions. Mesojuctions differ from double cross-over junctions in that only one strand crosses the helical regions at each connection site, while two strands cross in the double cross-over junctions. Hydroxyl radical footprinting studies of these new structures further suggest that double-helical regions may stack to form larger helical domains, in much the same way as the arms in three- and four-way junctions stack. Wang and Seeman (1995) have recently explored the consequences of changing the number of turns in the helical arms from an odd number of 3 half turns to an even number of 2. They have shown that the four stranded antijunction 4_0 and the three stranded mesojunction 3_1 do not form monomeric structures but aggregates of high molecular weight. All four stranded mesojunctions were on the other hand successfully constructed at low concentrations and high temperatures.

It is worth noting some properties shared between mesojunctions and double-cross-over DNA structures regarding their stacking and cross-over arrangements. The $^1 4_2$ mesojunction containing 2 half turns per helical region consists of two stacked helices, in which the continuous strands have a parallel orientation. The hydroxyl radical protection patterns suggest that this structure is similar to the double cross-over structure DPE. Increasing the number of half helical turns to 3 shifts the equilibrium in favor of a structure containing three stacked helical domains. This structure is analogous to the double cross-over structures of DPOW or DPON, depending on the

number of half helical turns. The mesojunction containing 16 nucleotides per helical domain resembles DPOW where the cross-over points are separated by a major groove. When the number of helical turns is reduced to 14 nucleotides in the circumferential helices, footprinting results suggest a structure similar to DPOD. The cross-over point in the 14 nucleotide structure generates an extra minor groove separation. It should be noted that these conformations are in equilibrium, which is usually shifted in favor of one or the other, depending on the number of bases in the circumferential helices.

The $^{2}4_{2}$ mesojunction containing three half turns per helix does not form a distinct structure, due to incompatible twists between the two stacked helical regions. In the case of the $^{2}4_{2}$ mesojuction, which contains two half turns per helix, the molecule adopts a stacked structure similar to the antiparallel double cross-over branched structure DAE, with the continuous strands oriented in opposite directions.

1.10. Three-Way Junctions

Three-way junctions have been shown to be distinctly different in many aspects to their four-armed counterparts. Electrophoretic mobility studies of three-way junctions have shown that either symmetric or asymmetric structures may be adopted in the presence of Mg^{2+} (Duckett and Lilley 1990; Guo *et al.* 1990; Lu *et al.* 1991 b). In the absence of Mg^{2+}, three-way junctions always take on a more extended symmetric conformation, with the three arms pointing to the corners of an (equilateral) triangle. The screening of repulsive electrostatic forces between the phosphate backbones in the

vicinity of the branch point may explain the effect Mg^{2+} cations have on the geometry of the junction.

Lu *et al.* (1991 a) have demonstrated that the geometry of the junction is governed primarily by the identity of the bases right at the branch point and their neighboring bases. Electrophoresis was used to establish the relative angles between the arms of the junction. The structural arrangement of the three-way junction most probably varies from one of symmetry to one of considerable asymmetry, depending on the bases and neighboring bases around the branch point (Lu *et al.* 1991 a). It has been shown that junctions containing different bases at, or one position removed from, the branch have an asymmetrical configuration. In the case where both positions are identical in all arms, mobility differences are no longer observed and the junction is believed to assume a more symmetric arrangement. Three-way junctions differ from four-way junction in this respect, as only the bases abutting on the four-armed junction have been shown to influence the pairwise stacking arrangement of the arms.

Chemical footprinting and thermodynamic studies indicate that three-way junctions are inherently conformationally strained structures. Osmium tetroxide hyper-reactivity is observed for thymine bases one helical turn from the branch point (Lu *et al.* 1991), and by a diethyl pyrocarbonate purine-specific probe (Duckett and Lilley 1990). Zhong and Kallenbach (1993) have investigated the thermodynamics of branch formation in a series of complexes by constructing a parent 16-mer DNA duplex from which a neck progressively extends. These structures allowed them to investigate how the three-way branch forms, and how the physical properties change as the neck is progressively

extended. Comparative gel electrophoresis studies of the nicked duplex, and duplexes containing a neck of increasing length, show that as the neck is extended the structure becomes progressively more bent in the presence of Mg^{2+}. Single strand specific reagents like *Neurospora* endonuclease and osmium tetroxide were also shown to promote cleavage of the continuous strands as the neck extends. It is surprising that the calorimetric dissociation enthalpy changes associated with the complexes decrease as the neck is extended, despite the increasing number of bases available for base pairing. NMR spectra of the NH protons clearly demonstrate that base pairing is perturbed, accounting for the loss of stability. These results are in accordance with ligation studies of sticky-ended three-way junctions. The small size of the closed circles observed can only be explained by assuming inherent flexibility and unwinding around the branch point (Ma *et al.* 1986).

Zhong and Kallenbach (1993) have proposed a model that describes the properties of three-way junctions. They have postulated that the junction does not take on a rigid structure, like the four-way junction, but rather a more dynamic structure that is able to switch from one conformational state to another. One or more base pairs adjacent to the junction in each arm can unfold transiently, releasing bases which are then free to intercalate into the remaining duplex. It is proposed that this process can take place in all three arms. The equilibrium between the various conformations will depend on the type of nearest neighbor interaction between stacked bases abutting on the branch. This model can account for many of the structural, thermodynamic and chemical footprinting results of three-way junctions. One is, however, left wondering if a simpler model, which relies on differential unwinding around the branch point, would not

suffice. Furthermore, it has been shown that cleavage in all the arms occurs for structures containing an *Eco*RI restriction site that spans the junction (Jensch and Kemper 1986). However, it is not known whether the enzyme confers stacking of paired arms through the junction, or whether the cleavage pattern is an inherent property of the junction where transiently stable stacked helices occur between all three arms.

1.11. Unpaired Nucleotides in Three-Way Junctions

Unpaired bases are frequently found around the branch point in naturally occurring RNA three-way junctions, possibly to release conformational strain. Ribosomal RNAs and hammerhead ribozymes, for example, contain several unpaired bases at the branch point. Leontis *et al.* (1991) have investigated the effect of adding extra unpaired nucleotides at the branch point in DNA three-way junctions. They have observed an overall increase in the thermal stability, which they attributed to the release of conformational strain. It was concluded that two pendent As or Ts stabilize the three-way DNA junction, relative to junctions containing either no unpaired bases, or junctions with three or more unpaired bases. NMR studies of three-way junctions containing either two TT or TC unpaired bases in one strand have further shown potential pairwise stacking of two helical arms across the junction, with the unpaired bases being extra-helical and exposed to the solvent (Leontis *et al.* 1993). These results are in accordance with osmium tetroxide induced degradation of the pendent unpaired bases. Based on geometrical studies, Duckett and Lilley (1990) have suggested that it would be necessary to distort the stacked helices in order to

accommodate a third arm. It seems likely that the extra unpaired bases alleviate some of the steric strain associated with the stacking of two arms.

Isothermal titration calorimetric studies (Ladbury *et al.* 1994) on three-way junctions containing additional unpaired bases show that the free energy difference between the three-way junction and the individual arms is only marginal. According to these authors, the incorporation of a branch structure into a duplex region is a thermodynamically unfavorable process. If one, however, separates the enthalpic and entropic contributions from the free energy changes, it is possible to gain additional insight into the nature of this process. The difference between the formation enthalpy of the three-way junction and the sum of the enthalpy values obtained for the individual arms amounts to -8.7 kcal/mol. This shows that the formation enthalpy for the junction is somewhat more favorable than for the isolated arms. Viewing the entropy differences in the same way one finds that the entropy favors the single arms by 56 $cal.mol^{-1}.K^{-1}$. Summing up these contributions to the free energy change, one sees that the enthalpy and entropy differences work in opposite directions.

From these results it is apparent that three-way junctions containing additional unpaired bases are conformationally less strained than three-way junctions containing unpaired bases. This allows for the base pairs within each arm to be fully hydrogen bonded and stacked, as demonstrated by NMR (Leontis *et al.* 1993).

1.12. DNA Nanostructures: Scaffolds and Geometrical Objects

One of the objectives of biotechnology and nanotechnology is to design new materials at the macromolecular level for both industrial and therapeutic applications. Branched DNA junctions with sticky ends can spontaneously assemble into hinged coordination networks. The DNA junctions in conjunction with ligation reactions provide the building blocks and tools for assembling large geometrical objects, scaffolds, and periodic substances. These junctions can be used as valence clusters in the vertices of larger molecular objects, and as attachment sites for proteins and other molecules (Seeman 1981, 1982, 1985).

The first attempts in constructing geometrical objects involved the oligometrization of three and four-way junctions containing pairs of complementary cohesive arms, into cyclic trimer and larger products (Ma *et al.* 1986; Petrillo *et al.* 1988). This suggests that the valence angle between the double-helical arms of either four or three-way junctions are relatively flexible. The number of base pairs between branch points must also be taken into consideration in order to maximize the orientation of end-to-end contact between cohesive ends. For example, it was shown that if the number of nucleotide pairs between four-way junctions was reduced from 20 to 16 nucleotides (1.5 turns), the macrocyclic ligation products start primarily with the tetramers, rather than trimers.

Chen *et al.* (1989) have constructed a quadrilateral from four three-way junctions, in which one arm is closed by a hairpin and the remaining arms terminate in a unique single-strand cohesive end. By selecting specific cohesive ends they were able to exert

greater control over the products of DNA assembly. Recently the same group (Chen and Seeman 1991) constructed a covalently closed cube-like structure containing twelve double-helical edges arranged about eight three-way vertices (cf. Fig. 1.5).

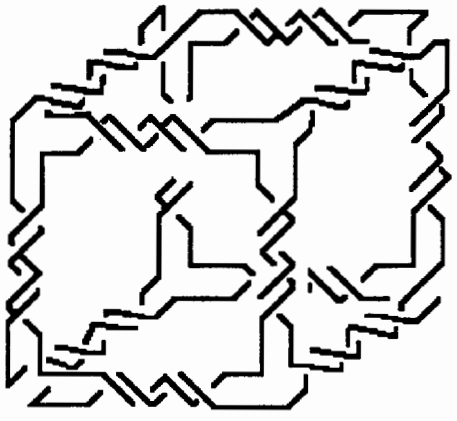

Fig. 1.5 Schematic representation of a three-dimensional double-helical cube adopted from Chen. J. H., and Seeman, N. C. (1991), *Nature (London).* **350**, 631-633.

It is easy to see that the synthesis of even larger objects would require still greater control over the intermediates formed than can be achieved by selecting unique cohesive ends alone. Zhang and Seeman (1992) have proposed a method to overcome these restrictions by using a solid-support mediated synthesis. The DNA is attached to a solid support initially with no open helical ends i.e., all DNA ends terminate in hairpins. This procedure allows the selected production of cohesive ends by the controlled removal of a given hairpin with a restriction enzyme (deblocking), followed by a ligation step to attach the next edge (coupling). The hairpins serve an additional

function by covalently closing successfully ligated products (capping) so that failure products can be readily removed by an exonuclease enzyme assay. The construction of a DNA-truncated octahedron (790 kDa) on a solid support, has shown the strength of the method (Zhang and Seeman 1994). The vertices of the octahedron were constructed from four-way junctions, so that each has an additional closed exocyclic arm. One of the exocyclic arms in the structure served as an initial attachment to the solid support.

1.13. Five and Six-Way Junctions

The construction of geometrical objects and networks with vertices of rank larger than the widely studied three- and four-way junctions, requires new structures containing five and possibly six arms. Examples of such networks have been suggested by Seeman (1982) for branched DNA structures containing 5 and 6 connections. It is evident that the study of DNA junctions of higher rank is of paramount importance for the design of more elaborate three-dimensional DNA zeolite structures. Wang *et al.* (1991) have, for example, investigated the structural properties of both five and six-way junctions using native gel electrophoresis and hydroxyl radical footprinting methods. It was shown that no junctions can be formed if the arms are 8 nucleotides long. Stable junctions were successfully assembled when the arms were extended to 16 nucleotides. The inability to construct junctions with 8 nucleotides per arm can be understood if it is assumed that the free energy cost associated with charge accumulation and/or possible steric effects around the branch point are substantial. Hydroxyl radical footprinting experiments were used to determine the structural

features of these junctions. The five-way junction seems to adopt an extended conformation, while the footprinting results for the six-way junction suggest that two helical arms are stacked. It is therefore evident that branched junctions containing more than four arms can indeed be constructed, provided their arms are sufficiently long.

1.14. Triple-Helical Three-Way Junctions

Other nucleic acid branched systems can be conjectured by extending the arms into parallel, triple, and quadruple-helices. Distifano *et al.* (1991, 1992), for example, have successfully combined triple-helices with conventional double-helices in a three-armed junction. It has become evident from my work (Hüsler and Klump 1994, 1995b) that branched three-way junctions can also accommodate triple-helices in all three arms. A number of strand arrangements are possible at the branch point. It is self-evident that the branch region of three-way junctions can only accommodate two strands from each triple-helical arm. Furthermore, triple-helices like double-helices require a particular third-strand orientation and sequence. Hence, certain constraints must be taken into consideration before designing branched DNA structures containing triple-helices. It is of great interest to see whether these structures differ from conventional double-helical three-way junctions. In the following section I shall give a brief overview of triple-helices pertinent to this thesis, before describing the finer details of triple-helical three-way junctions.

1.15. Triple-Helices

Triple-helices are readily formed when a matching poly- or oligo-pyrimidine strand binds with parallel orientation to the WC purine strand via the major groove of a target DNA or RNA duplex. The third strand is conventionally referred to as the Hoogsteen strand (HG) as it binds to the major groove by Hoogsteen hydrogen bonding (cf. Fig. 1.6). The Hoogsteen base T can bind to a complementary A of the duplex (at neutral pH) and C to G, (usually under acidic pH conditions). There are alternative HG pairs that do not conform to this mode of binding. Triple-helices have been observed containing two strands of poly dG and one strand of poly dC, and others are shown to contain two strands of poly dA and one of poly dU (Lipsett 1964; Inman 1964; Broitmann *et al.* 1987; Letai *et al.* 1988). The third strand binds to the WC purine strand via reverse Hoogsteen hydrogen bonding with an antiparallel third strand orientation (Radhakishnan *et al.* 1991, 1993b). Triple-helices containing mixed sequences of G and T bases, as the HG strand, have also been observed (Radhakishnan *et al.* 1993b). The noticeable difference between these triple-helices and those containing only pyrimidine bases in the third strand, are that they are all pH independent. For the latter, protonation of the Hoogsteen cytosines are required to supply a second hydrogen bond between C^+ and its target duplex GC pair. NMR studies on oligomeric DNA triple-helices have verified some of these structural details previously established by indirect techniques, providing a more detailed account of triplex architecture (Rajagopal and Feigon 1989; de los Santos *et al.* 1989; Pilch *et al.* 1990). The protonation of Hoogsteen cytosine bases in triple-helices is consistent with the pH dependence of their melting temperatures (Tm). Most studies have been conducted in the pH range well above the pKa of cytosine where the relationship

between 1/Tm and pH becomes linear (Plum *et al.* 1990; Xodo *et al.* 1991; Maher *et al.* 1990; Hample *et al.* 1991; Giovannangeli *et al.* 1992; Singleton and Dervan 1992; Völker *et al.* 1993). This demonstrates that triplex formation can occur in a pH range well above the pKa of cytosine. The cost of protonating the Hoogsteen C^+ bases is therefore paid for by the free energy associated with third strand binding. This result is in contrast with pH titrations studies where it was shown that the apparent pKa is considerably higher than the pKa of cytosine (pKa = 4.5). Recently it was shown (Völker *et al.* 1993, Plum and Breslauer 1995, Hüsler and Klump 1995a) that the maximum thermal stability (Tm) for a triplex containing a mixed Hoogsteen strand is observed when the pH of the solution coincides closely with the pKa of cytosine. The Tm decreases sharply on either side of the pH optimum. The nonlinear relationship between thermal stability and pH, close to the pKa value of cytosine, suggests that triple-helix stability is subject to two opposing conditions: the protonation of the Hoogsteen cytosine bases (promoted by acidic conditions) and the presence of unprotonated cytosine bases in the WC part of the triplex (favored by neutral pH conditions). Decreasing the pH below the pKa of cytosine results in the concurrent protonation of both HG and WC cytosines, effectively destabilizing both the duplex and triplex conformation. Increasing the pH above the pKa of cytosine will also result in a destabilization of the triplex because the Hoogsteen bases become deprotonated.

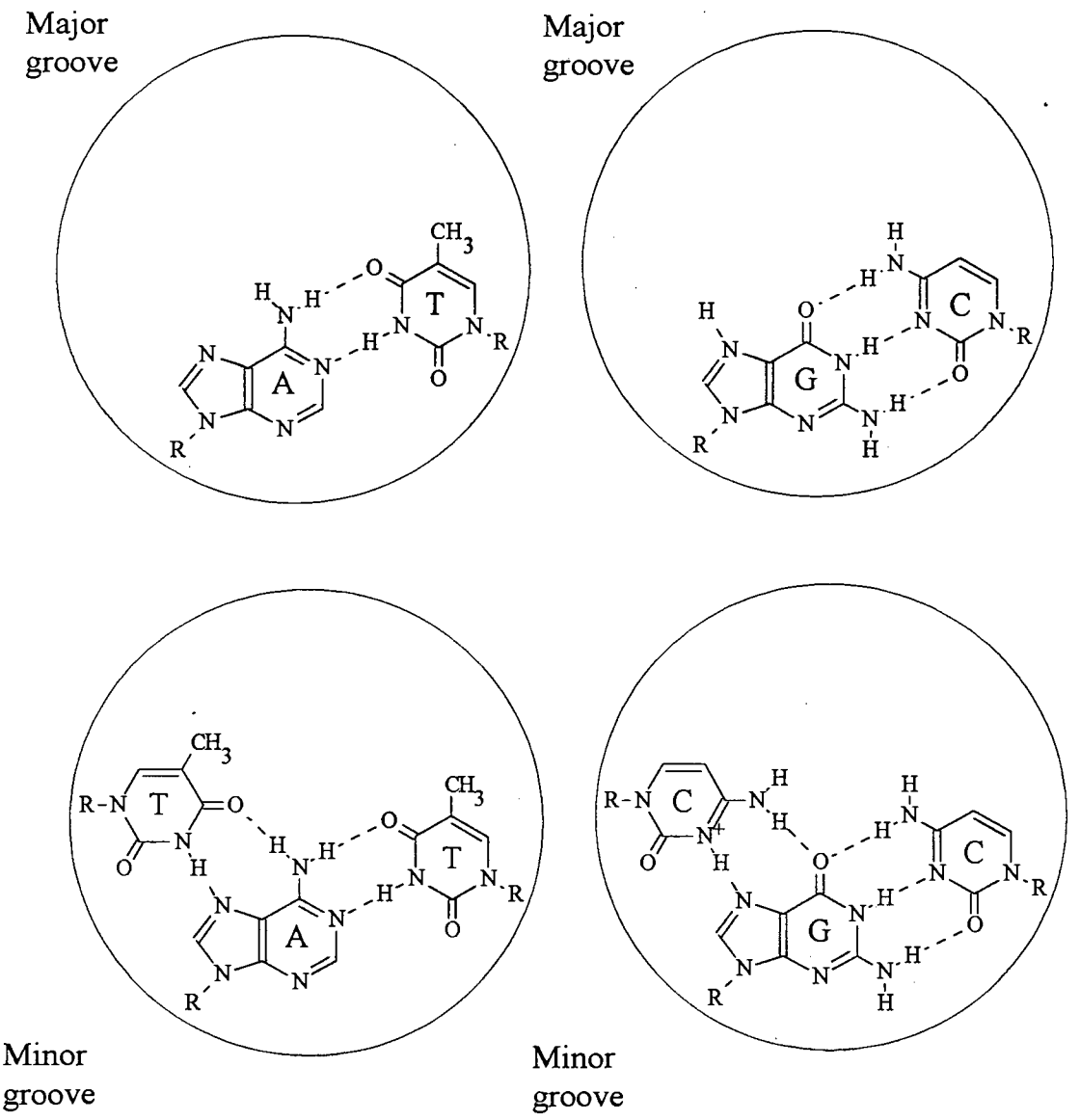

Fig. 1.6 (Top) Watson-Crick base pairs. (Bottom) triplet base pairs TAT and CGC^+. The additional pyrimidine strand is bound by Hoogsteen hydrogen bonds in the major groove of the Watson-Crick duplex.

Chapter 2.

Design of a Double-Helical Three-Way Junction and its Extension into Either a WC or HG Triple-Helical Three-Way Junction

2.1. Notation for Triple-Helical Three-Way Junctions

As mentioned in chapter 1, synthetic double-helical three-way junctions were first introduced by Guo *et al.* (1990) to study the features of double helices around the branch point. Sequences for each arm were designed primarily to minimize the number of possible alternative structures. In this thesis I have attempted to take the concept of multi-stranded DNA structures one step further, by incorporating a third strand into each arm.

It is necessary to segregate purine and pyrimidine bases onto opposite strands in the design of triple-helical DNA constructs. Thus, although is was possible for Guo *et al.* (1990) to design sequences for double helical three-way junctions by solely minimizing the possibility of alternate structures, addition of a third strand to each arm results in additional sequence constraints. Fig. 2.1 shows a schematic representation of the four different types of triple-helical three-way junctions that can be constructed from homopyrimidine and homopurine triple-helices.

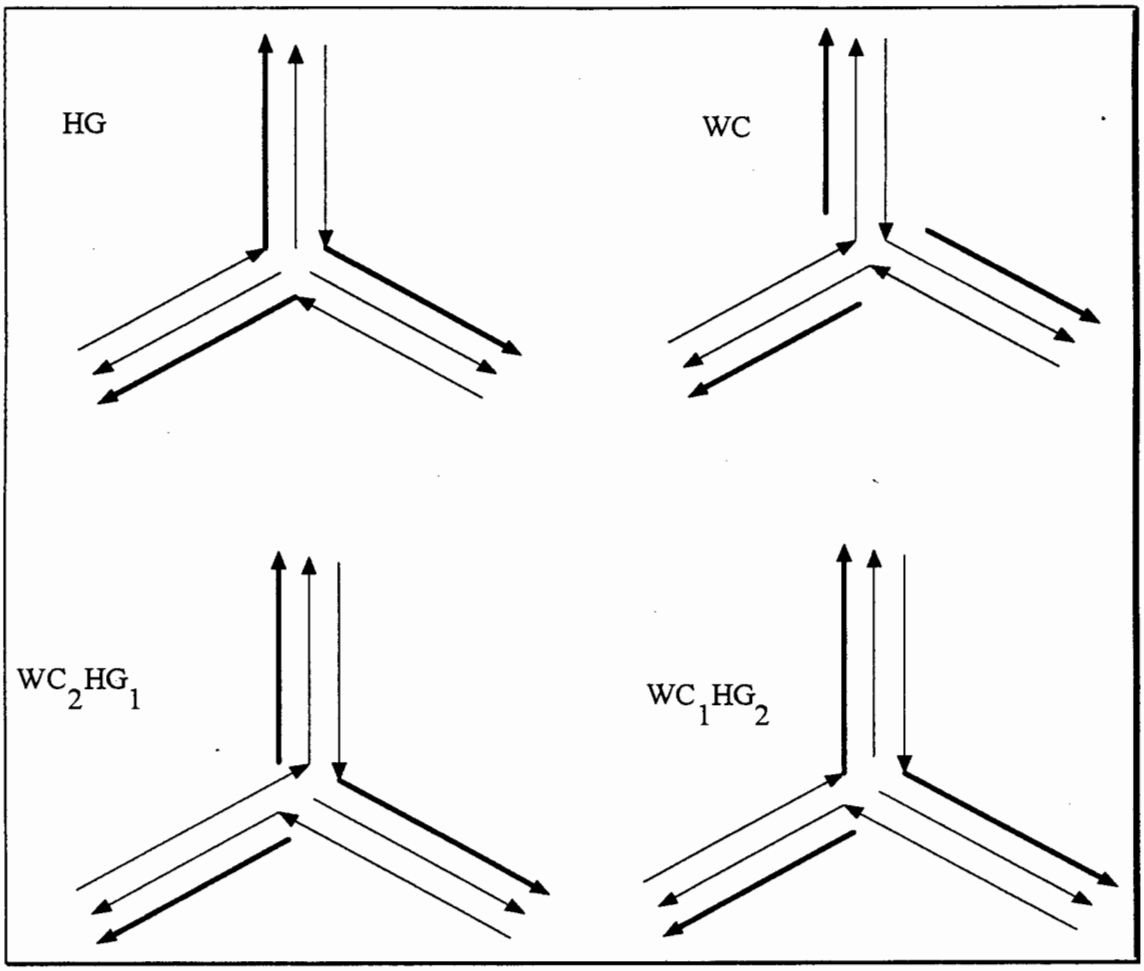

Fig. 2.1 (Top) HG and WC triple-helical three-way junctions. (Bottom) Mixed WC_2HG_1 and WC_1HG_2 triple-helical three-way junctions. Thick lines represent the Hoogsteen strands and the thin lines the Watson-Crick strands. The arrows indicate the 5'-3' polarity of the single strands. The notation refers to the type of strand connection at the branch point. HG refers to Watson-Crick and Hoogsteen connections between strands, while WC refers to connections between Watson-Crick strands only. The subscript indicates the number of HG and WC connections in mixed structures (bottom). The subscript has been dropped in structures containing only one type of connection (top).

For clarity, the three-way junction is defined according to the type of strand connection at the branch point. WC triple-helical three-way junctions designate junctions where the branch point is assembled through WC strands only. If all the arms of the junction are connected by a WC strand from one arm, and a HG strand from another arm, the junction formed is referred to as a HG triple-helical three-way junction.

It is in principle possible to assemble structures which contain a mixture of WC and HG connections. The following notation is used to distinguish between the different "mixed" triple-helical three-way junctions. WC and HG refers to the type of strand connection between arms, where the subscripts enumerate the number of HG or WC connections in mixed three-way junctions. For example, the structure WC_1HG_2 denotes a junction containing one WC connection and two HG connections (cf. Fig 2.1). Structures containing one type of connection are only referred to as HG or WC where the subscript has also been eliminated. Since mixed triple-helical three-way junctions have complex thermal unfolding pathways, which are not easily resolved, only the WC and HG triple-helical three-way junctions were investigated.

In this thesis I will concentrate only on the double-helical three-way junction, and the corresponding WC, and HG triple-helical three-way junctions. To reduce the number of strands required in the HG and WC three-way junctions from six to three, two of the strands at the ends of each arm has been joined to form a hairpin[1] or loop[1],

respectively. The same subset of sequences for the arms A, B, and C have been used in the design of all the junctions.

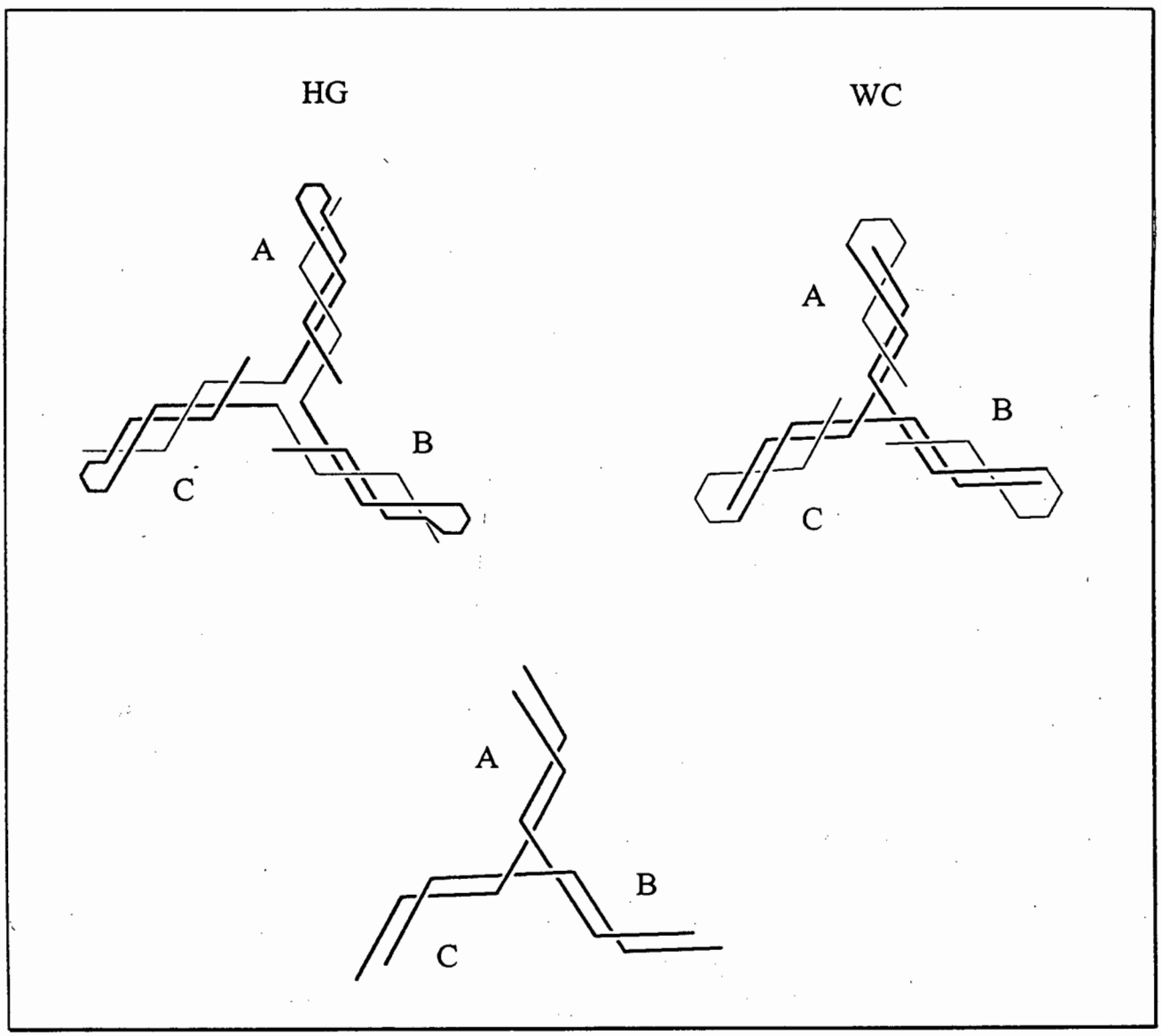

Fig. 2.2 Schematic representation of the double-helical three-way junction J, and the WC and HG triple-helical three way junctions. The thick lines represent the WC double-helical regions and the thin lines the triplex HG strands. The sequences in each of the arms A, B, and C are the same for all three junctions with only the connection at the branch point and the peripheral ends of the arms (forming the hairpin or looped regions) being different.

[1] Hairpins refer to structures containing unpaired bases that join the WC strands in double helices, while loops refer to structures containing unpaired bases that join the WC and HG pyrimidine strands in triple helices.

Chapter 3.

Theoretical Background

A great deal about the energetics and secondary structure of DNA and RNA has been learned from thermal denaturation experiments, often referred to as "melting". The process of unfolding is usually observed by following the changes in the uv absorbance (260 nm) or some other observable parameter as a function of temperature. The interpretation of melting experiments is unfortunately not straightforward, but subject to prior knowledge of the mechanism of molecular unfolding. Model-independent parameters can be obtained by using differential scanning calorimetry (DSC) where the heat (ΔCp) of unfolding can be obtained directly (Privalov, P. L. *et al.* 1986; Markey *et al.* 1987). The advantage of this procedure lies in that no prior knowledge of the unfolding process is required. Marky and Breslauer (1987) have derived the general forms of the equations required to extract thermodynamic data from either uv thermal denaturation or DSC, for systems of any molecularity. The derived equations are, however, only applicable to systems in which the process of unfolding is strictly two-state. Many unfolding processes of macromolecules can be described by a two-state model, when intermediate states do not exceed 5% (Freire, E., Biltonen, R. L. 1978a, b, c; Marky and Breslauer 1987). Under such circumstances reliable thermodynamic information can be obtained by simply monitoring the temperature-induced conformational change of the system. Unfortunately this is not the case in many macromolecules which deviate from two-state behavior. Under these

circumstances, thermodynamic data cannot be obtained directly from a single melting experiment.

Statistical thermodynamics provides an elegant and powerful means of resolving complex multi-state processes in systems where the molecularity remains constant. Freire and Biltonen (1987) have demonstrated that the partition function in conjunction with DCS data can be used to deconvolute multi-state unfolding processes. Solutions to multi-state processes in which the constraints of mass action have been relaxed are notably more complex. In most cases, no direct analytical solutions can be obtained and numerical methods are required. In this section, the solutions to multi-state processes with a molecularity of two and three are deduced using statistical thermodynamic methods. Only the solutions to models applicable to this thesis are presented.

3.1 Two-State Dissociation of Complexes with Molecularity n

The thermal unfolding process of nucleic acids can be monitored by following the change in uv absorption with increasing temperature. The shape of this profile can be analyzed to yield the value of the van't Hoff transition enthalpy and entropy. Similarly, differential scanning calorimetry can be used to determine the enthalpy and entropy changes associated with the thermal unfolding of nucleic acids by integrating the excess heat capacity curves ΔCp (where $\Delta H^\circ = \int \Delta Cp \cdot dT$ and $\Delta S^\circ = \int \Delta Cp/T \cdot dT$).

The latter procedure has the advantage that model independent enthalpy and entropy changes can be determined directly. It is sometimes convenient to express the van't Hoff and calorimetric enthalpy values as a ratio ($r = \Delta H_{VH}/\Delta H_{cal}$) to determine whether the complex undergoes two-state unfolding or not. In the case of a two-state process the ratio, r, should be equal to one. If, however, the unfolding is a multi-state process, ratios smaller than one should be obtained. Values larger than one are usually indicative of aggregation.

In general the dissociation of non-complementary sequences can be expressed as follows for an all-or-none equilibrium process:

$$A_1A_2A_3 \ldots A_n \overset{K}{\rightleftharpoons} A_1 + A_2 + A_3 + \ldots + A_n \qquad [1.1]$$

Here n is the molecularity of the reaction [1.1], and K the equilibrium constant. The equilibrium constant can be expressed in terms of the fractional amount α of the unfolded species A_1 to A_n and the molecularity n as follows:

$$\alpha = \frac{\sum_{1}^{n} A_i}{Ct} \qquad [1.2]$$

$$K = \exp - \left(\frac{\Delta H_{VH} - \Delta S_{VH} T}{RT} \right) \qquad [1.3]$$

$$K = \frac{(Ct/n)^n \alpha^n}{(1-\alpha)}. \qquad [1.4]$$

Here Ct is the total strand concentration, R the gas constant, T the temperature in Kelvin, and ΔH_{VH} and ΔS_{VH} the dissociation enthalpy and entropy changes. It is evident from equation [1.4] that α can be expressed exactly in terms of K and Ct for values of n equal to 1 or 2. For values larger than two α may be found numerically using Newton's method.

The relationship between the absorbance and temperature can be expressed in terms of α and the absorbance coefficients for the various species as follows:

$$Abs = Ct \cdot \alpha \cdot (\varepsilon_1 - \varepsilon_2) + Ct \cdot \varepsilon_2 \qquad [1.5],$$

where ε_1 is the absorbance coefficient for the species ($A_1A_2A_3$ ... A_n) and ε_2 the combined absorbance coefficient for the single stranded species A_1, A_2, A_3, ... , A_n. It is assumed that the absorbance coefficient ε_1 and ε_2 change linearly with temperature in the following way:

$$\varepsilon_1 = \beta_1 + \mu_1 \cdot T \qquad [1.6]$$

$$\varepsilon_2 = \beta_2 + \mu_2 \cdot T. \qquad [1.7],$$

where β_1 and β_2 are the extinction coefficients at 0 °C, and μ_1 and μ_2 the slope for the ordered and disordered states, respectively.

3.2. Sequential Dissociation of a Binary Complex in Three Stages with Molecularity Equal to Two

Fig. 3.1 outlines the unfolding pathway of the binary component Fa formed by combining two of oligonucleotides Fb' and Fc' which are part of the HG triple-helical three-way junction. The complex thus formed represents one of the triple-helical arms of the junction (cf. Fig 2.1), a hairpin, and a single stranded extension. The hairpins with single stranded extensions can associate when all three oligonucleotides are mixed to from the fully folded HG three-way junction.

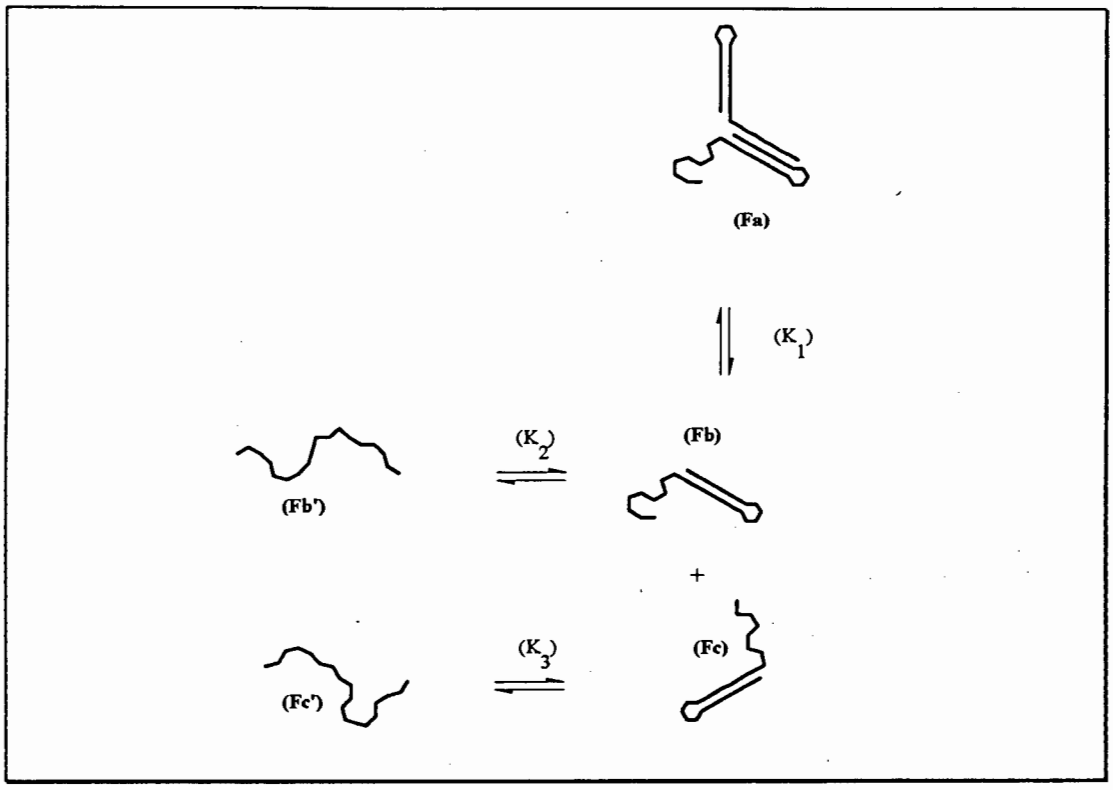

Fig. 3.1 Reaction scheme that describes the unfolding process of a binary complex with species Fa to Fc'. K_1, K_2, and K_2 represent the equilibrium constants for each step.

In this section only the dissociation pathway for the binary combinations are considered. The complex Fa can undergo thermal dissociation to form the hairpins Fb and Fc, respectively. Each of these complexes can further unfold into the random coiled single stranded species Fa' and Fb', respectively.

The following equilibria represent the individual unfolding steps for each species:

$$Fa \overset{K_1}{\leftrightarrow} Fb + Fc \qquad [2.1]$$

$$Fb \overset{K_2}{\leftrightarrow} Fb' \qquad [2.2]$$

$$Fc \overset{K_3}{\leftrightarrow} Fc' \qquad [2.3]$$

The equilibrium constants K_1, K_2, and K_3 are given as

$$K_1 = \exp-\left(\Delta H_1 - \Delta S_1 T - RT \ln Ct\right) / RT \qquad [2.4]$$

$$K_2 = \exp-\left(\Delta H_2 - \Delta S_2 T\right) / RT \qquad [2.5]$$

and

$$K_3 = \exp-\left(\Delta H_3 - \Delta S_3 T\right) / RT, \qquad [1.6]$$

where ΔH_i and ΔS_i are excess molar unfolding enthalpy and entropy terms, R is the gas constant and Ct the total molar strand concentration of the species Fb or Fc. Under normal conditions Ct of Fb and Fc are identical. It should be noted that K_1 is not a conventional equilibrium constant (Hüsler and Klump 1994), as it contains an additional term RTlnCt. This allows one to express the equilibrium constants K_1, K_1K_2, and K_1K_3 in terms of the fractional amounts of each species α_a, α_b, α_c, $\alpha_{b'}$ and $\alpha_{c'}$ rather than the molar amounts of each species, as follows:

$$K_1 = \frac{\alpha_b \alpha_c}{\alpha_a} \qquad [2.7]$$

$$K_1 \cdot K_2 = \frac{\alpha_{b'} \alpha_c}{\alpha_a} \qquad [2.8]$$

$$K_1 \cdot K_3 = \frac{\alpha_b \alpha_{c'}}{\alpha_a}. \qquad [2.9]$$

The conservation of mass is given in terms of the fractional amounts of α_a, α_b, α_c, $\alpha_{b'}$ and $\alpha_{c'}$ by equations [2.10] and [2.11].

$$1 = \alpha_a + \alpha_b + \alpha_{b'} = \alpha_a + \alpha_a(1 + K_2) \qquad [2.10]$$

$$1 = \alpha_a + \alpha_c + \alpha_{c'} = \alpha_a + \alpha_a(1 + K_3). \qquad [2.11]$$

From equations [2.10] and [2.11] one can deduce the relationship between α_b and α_c in terms of K_2 and K_3 as follows:

$$\alpha_b(1 + K_2) = \alpha_c(1 + K_3). \qquad [2.12]$$

The two partition functions Q_1 and Q_2 for the strands Fb and Fc are given by

$$Q_1 = \frac{1 + K_1(1 + K_2)}{\alpha_c} \qquad [2.13]$$

and

$$Q_2 = \frac{1 + K_1(1 + K_3)}{\alpha_b}, \qquad [2.14]$$

respectively. It follows directly from equations [2.12], [2.13] and [2.14] that $Q_1 = Q_2$.

$$Q = 1 + K_1(1 + K_2)/\alpha_c = 1 + K_1(1 + K_3)/\alpha_b \qquad [2.15]$$

The population of each species is conveniently expressed in terms of the partition function, Q as follows:

$$\alpha_a = 1/Q \qquad [2.16]$$

$$\alpha_b = K_1/\alpha_c Q \qquad [2.17]$$

$$\alpha_c = K_1/\alpha_b Q \qquad [2.18]$$

$$\alpha_{b'} = K_1 K_2/\alpha_c Q \qquad [2.19]$$

$$\alpha_{c'} = K_1 K_3 / \alpha_b Q \qquad [2.20]$$

By combining equations [2.12] and [2.16], α_c can be eliminated, yielding

$$K_1 = \frac{\alpha_b^2 (1 + K_2)}{(1 + K_3)} Q. \qquad [2.21]$$

Substituting equation [2.14] for Q into [2.21] yields α_b in terms of the three constants K_1, K_2 and K_3.

$$\alpha_b^2 \left(1 + \frac{K_1 (1 + K_3)}{\alpha_b} \right) = K_1 \frac{1 + K_3}{1 + K_2} \qquad [2.22]$$

or

$$\alpha_b{}^2 + \alpha_b K_1 (1 + K_3) - \frac{K_1 (1 + K_3)}{(1 + K_2)} = 0. \qquad [2.23]$$

The quadratic equation [2.23] can now be re-written so that α_b becomes the subject of the equation

$$\alpha_b = \frac{-K_1(1+K_3) + \sqrt{K_1^2(1+K_3)^2 + \frac{4K_1(1+K_3)}{(1+K_2)}}}{2}. \qquad [2.24]$$

The partition function Q is found directly by substituting equation [2.24] into equation [2.14]. The fraction of each species can be obtained directly by substituting the calculated values of α_b and Q into equations [2.16] to [2.20].

3.2.1. Thermal Denaturation

Ultraviolet Absorbance Changes

To determine the relationship between the absorbance and the fractional concentration of all molecular species present, we assume that the extinction coefficients change linearly with temperature

$$\varepsilon_a = \beta_a + \mu_a T \qquad [2.25]$$

$$\varepsilon_b = \beta_b + \mu_b T \qquad [2.26]$$

$$\varepsilon_c = \beta_c + \mu_c T \qquad [2.27]$$

$$\varepsilon_{b'} = \beta_{b'} + \mu_{b'} T \qquad [2.28]$$

$$\varepsilon_{c'} = \beta_{c'} + \mu_{c'} T. \qquad [2.29]$$

Here ε_i represents the extinction coefficient of each species as a function of temperature, β_i is the extinction coefficient at 0° C and μ_i the slopes for $i \in \{ a, b, c, b', c' \}$. The total absorbance at a given wavelength can thus be expressed in terms of ε_i and α_i as

$$\text{Abs} = Ct\left[\alpha_a\varepsilon_a + \alpha_b\varepsilon_b + \alpha_c\varepsilon_c + \alpha_{b'}\varepsilon_{b'} + \alpha_{c'}\varepsilon_{c'}\right]. \qquad [2.30]$$

Calorimetric Scans

The relationship between the DSC data and the above-mentioned solutions can be deduced directly from:

$$\Delta Cp = \left(\partial\alpha_b/\partial T\right)\Delta H_1 + \left(\partial\alpha_{b'}/\partial T\right)\left(\Delta H_1 + \Delta H_2\right) + \left(\partial\alpha_{c'}/\partial T\right)\Delta H_3. \quad [2.31]$$

The theoretical excess heat capacity for each reaction step corresponding to K_1-K_3 can be calculated from:

$$\Delta Cp_1 = \left(\partial\alpha_b/\partial T + \partial\alpha_{b'}/\partial T\right)\Delta H_1 \quad [2.32]$$

$$\Delta Cp_2 = \left(\partial\alpha_{b'}/\partial T\right)\Delta H_2 \quad [2.33]$$

$$\Delta Cp_3 = \left(\partial\alpha_{c'}/\partial T\right)\Delta H_3 \quad [2.34]$$

3.3. Sequential Dissociation of the WC Triple-Helical Three-Way Junction in Three Stages with Molecularity Three

Fig. 3.2 represents the unfolding pathway for the WC triple-helical three-way junction. In the first step arm A of F1 dissociates to form the intermediate species F2. It should be noted that this step is an intermolecular process. The resulting species F2 can undergo further unfolding to form the intermediate F3 and a single stranded species F4. F3 can dissociate into two single strands F5 and F6, respectively.

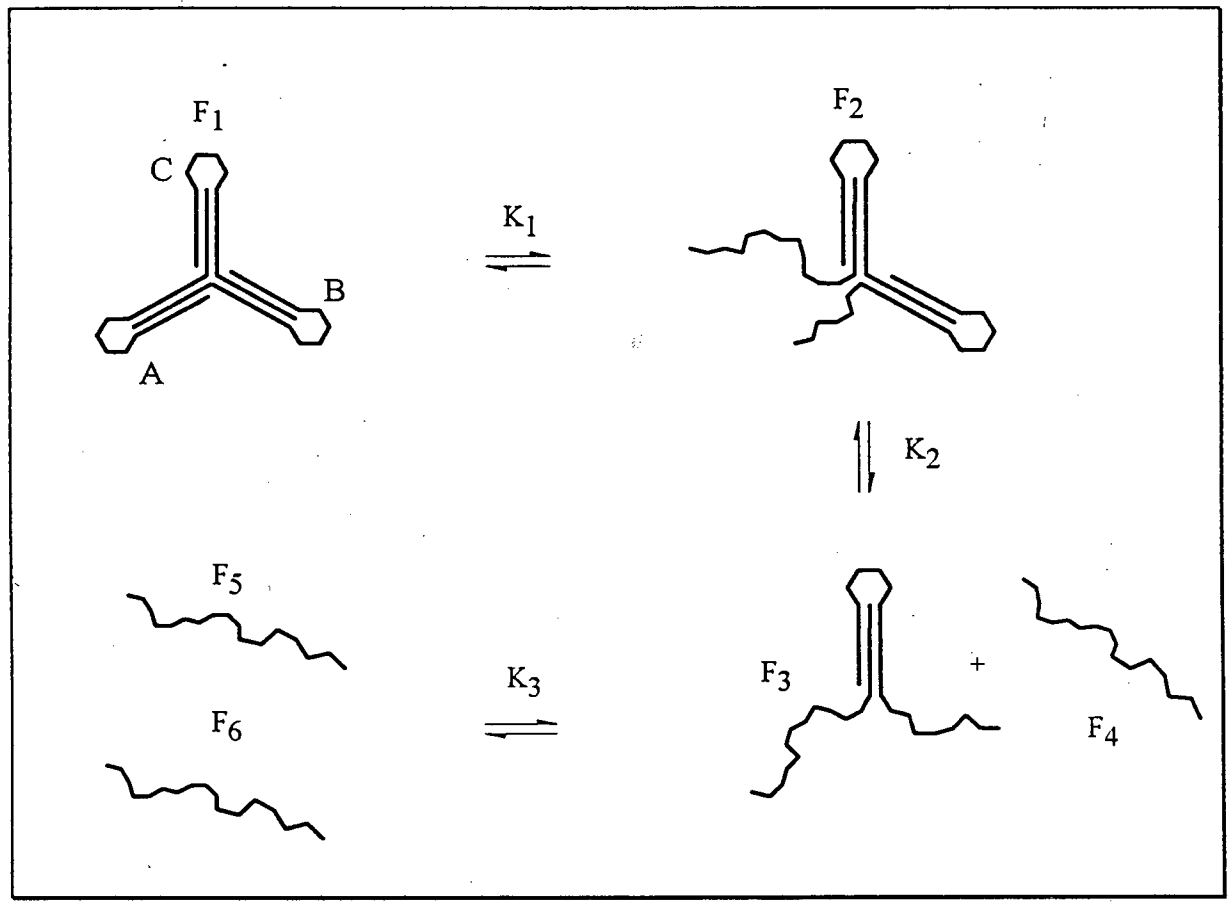

Fig. 3.2 Reaction scheme that describes the unfolding pathway for the WC triple-helical three-way junction with intermediate species F1 to F6. K_1-K_3 represent the equilibrium constants for each step.

The equilibrium constants for each of the three steps are represented by the constants K_1, K_2, and K_3, respectively.

The following equilibria represent the individual unfolding steps for each species:

$$F1 \overset{K_1}{\leftrightarrow} F2 \qquad [3.1]$$

$$F2 \overset{K_2}{\leftrightarrow} F3 + F4 \qquad [3.2]$$

$$F3 \overset{K_3}{\leftrightarrow} F5 + F6 \qquad [3.3]$$

To account for mass action the partition function is written as

$$Q = 1 + K_1 + \frac{K_1 K_2}{\alpha_4} + \frac{K_1 K_2 K_3}{\alpha_4 \alpha_6} \qquad [3.4]$$

$$[F4] = \alpha_4 Ct, \qquad [3.5]$$

$$[F6] = \alpha_6 Ct, \qquad [3.6]$$

where α_4 and α_6 are the fractional amounts of F4 and F6, respectively. The total concentration of each species F1 to F6 amounts to Ct. The concentration term Ct was

included into K_2 and K_3 (cf. Eqs. [3.8] and [3.9]). It should therefore be noted that K_2 and K_3 are not conventional equilibrium constants.

$$K_1 = \exp\frac{-(\Delta H_1 - \Delta S_1)}{RT} \qquad [3.7]$$

$$K_2 = \exp\left(\frac{-\Delta H_2}{RT} + \frac{(\Delta S_2 - R\ln Ct)}{R}\right) \qquad [3.8]$$

$$K_3 = \exp\left(\frac{-\Delta H_3}{RT} + \frac{(\Delta S_3 - R\ln Ct)}{R}\right) \qquad [3.9]$$

where ΔH_i and ΔS_i are enthalpy and entropy terms for K_i in [3.1] to [3.3] ($i = 1, 2, 3$) and T is the absolute temperature in degrees Kelvin. The fraction of each species α_i can be obtained from the partition function Q.

$$\alpha_1 = 1/Q \qquad [3.10]$$

$$\alpha_2 = K_1/Q \qquad [3.11]$$

$$\alpha_3 = (K_1K_2)/(\alpha_4 Q) \qquad [3.12]$$

$$\alpha_5 = K_1K_2K_3/(\alpha_4\alpha_6 Q) \qquad [3.13]$$

$$\alpha_4 = (Q - 1 - K_1)/Q \qquad [3.14]$$

$$\alpha_6 = \left(Q - 1 - K_1 - K_1K_2/\alpha_4\right)/Q \qquad [3.15]$$

It is convenient to express α_4 and α_6 in terms of K_1, K_2, and K_3. From Eqs. [3.14] and [3.15] one obtains

$$\alpha_4/\alpha_6 = \left(Q - 1 - K_1\right)/\left(Q - 1 - K_1 - K_1K_2/\alpha_4\right) \qquad [3.16]$$

Substituting [3.4] into [3.16] yields

$$\alpha_4/\alpha_6 = \alpha_6\left(\alpha_6 + K_3/\alpha_4\right)/K_3 \qquad [3.17]$$

or

$$\alpha_4 = \left(\alpha_6/K_3\right)\left(\alpha_6 + K_3/\alpha_4\right).$$

From Eqs. [3.4] and [3.17] α_4 is eliminated resulting in

$$Q = \frac{\alpha_6\left(\alpha_6 + K_3\right)\left(1 + K_1\right) + K_1K_2K_3\left(1 + K_3/\alpha_6\right)}{\alpha_6\left(\alpha_6 - K_3\right)} \qquad [3.18]$$

and from Eqs. [3.4] and [3.15]

$$\alpha_6 = K_1 K_2 K_3 / (\alpha_4 \alpha_6 Q). \qquad [3.19]$$

From Eqs. [3.17], [3.18], and [3.19] it follows that α_6 can be expressed in terms of K_1, K_2, and K_3:

$$(\alpha_6^4 + \alpha_6^3 K_3)(1 + K_1) + K_1 K_2 K_3 (\alpha_6^2 + \alpha_6 K_3) - K_1 K_2 K_3^2 = 0 \qquad [3.20]$$

Eq. [3.20] may be solved numerically by Newton's method, in which iterations were found to converge rapidly.

Using the values of α_6 thus obtained, α_4 can be calculated and hence Eq. [3.4] for the partition function Q.

3.3.1. Thermal Denaturation

Ultraviolet Melting Curves

To determine the relationship between absorbance and the fractional concentration of all molecular species present, it is assumed that the extinction coefficients change linearly with temperature

$$\varepsilon_1 = \beta_1 + \mu_1 T \quad [3.21]$$

$$\varepsilon_2 = \beta_2 + \mu_2 T \quad [3.22]$$

$$\varepsilon_3 = \beta_3 + \mu_3 T \quad [3.23]$$

$$\varepsilon_4 = \beta_4 + \mu_4 T \quad [3.24]$$

$$\varepsilon_5 = \beta_5 + \mu_5 T \quad [3.25]$$

$$\varepsilon_6 = \beta_6 + \mu_6 T \quad [3.26]$$

ε_j is the extinction coefficient as a function of temperature, β_j is the extinction coefficient at 0 °C, and μ_j is the slope $j \in \{1,2,3,4,5,6\}$. The total absorbance can then be expressed as

$$\text{Abs} = \text{Ct}\left[\alpha_4\varepsilon_4 + \alpha_6\varepsilon_6 + \left(\varepsilon_1 + K_1\varepsilon_2 + \frac{K_1K_2\varepsilon_3}{\alpha_4} + \frac{K_1K_2K_3\varepsilon_5}{\alpha_4 \cdot \alpha_6}\right)\Big/ Q\right]$$

or

$$Abs = Ct\left[\alpha_1\varepsilon_1 + \alpha_2\varepsilon_2 + \alpha_3\varepsilon_3 + \alpha_4\varepsilon_4 + \alpha_5\varepsilon_5 + \alpha_6\varepsilon_6\right] \quad [3.27]$$

Calorimetric Melting Curves

The relationship between DSC curves and the above-mentioned solution can be deduced from

$$\Delta Cp = \left(\partial\alpha_2/\partial T\right)\Delta H_1 + \left(\partial\alpha_3/\partial T\right)\left(\Delta H_1 + \Delta H_2\right) + \left(\partial\alpha_5/T\right)\left(\Delta H_1 + \Delta H_2 + \Delta H_3\right) \quad [3.28]$$

The theoretical excess heat capacity for each reaction step corresponding to K_1 to K_3 can be calculated from:

$$\Delta Cp_1 = \left(\partial\alpha_2/\partial T + \partial\alpha_3/\partial T + \partial\alpha_5/\partial T\right)\Delta H_1 \quad [3.29]$$

$$\Delta Cp_2 = \left(\partial\alpha_3/\partial T + \partial\alpha_5/\partial T\right)\Delta H_2 \quad [3.30]$$

$$\Delta Cp_3 = \partial\alpha_5/\partial T\,\Delta H_3 \quad [3.31]$$

3.4. Sequential Dissociation of the HG Triple-Helical Three-Way Junction in Six Stages with Molecularity Three

The unfolding pathway for the HG triple-helical three-way junction is shown in Fig. 3.3. From the unfolding pathway it is apparent that the complex F1 can unfold in six stages. In the first step, F1 unfolds into the intermediate species F2 (intermolecular process) which in turn dissociates into the species F4 (boxed region) and the hairpin F3. F4 can dissociate into two hairpin species F5 and F6, respectively. Each of the three resulting hairpins F3, F5, and F6 finally unfolds into the disordered single strands F7, F8, and F9, respectively.

The series of steps describing the unfolding pathway are represented by the following reactions:

$$F1 \overset{K_1}{\leftrightarrow} F2 \overset{K_2}{\leftrightarrow} F3 + F4 \qquad [4.1]$$

$$F3 \overset{K_3}{\leftrightarrow} F5 + F6 \qquad [4.2]$$

$$F4 \overset{K_4}{\leftrightarrow} F7 \qquad [4.3]$$

$$F5 \overset{K_5}{\leftrightarrow} F8 \qquad [4.4]$$

$$F6 \overset{K_6}{\leftrightarrow} F9 \qquad [4.5]$$

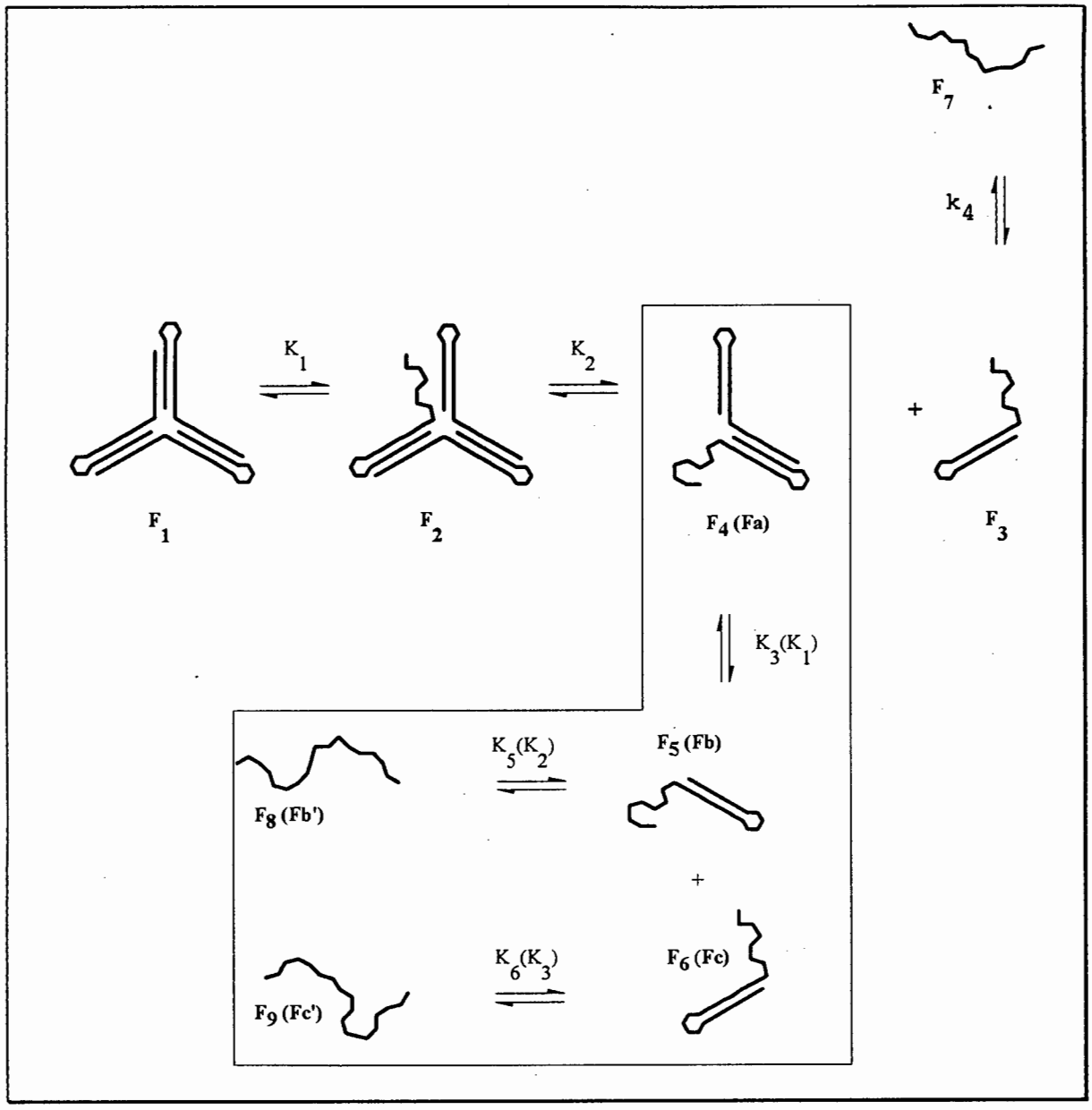

Fig. 3.3 Reaction scheme describing the unfolding pathway of the ternary complex with species F1-F9. K_1-K_6 represent the equilibrium constants for the individual steps. The boxed reaction pathway represents the unfolding steps of the binary complex. The molecular species and equilibrium constants for the boxed reaction pathway are presented in round brackets for (K_1)-(K_3), and (Fa)-(Fc'), respectively.

The equilibrium constants K_1 to K_6 are defined as follows:

$$K1 = \exp-\left(\Delta H_1 - \Delta S_1 T\right)/RT \qquad [4.7]$$

$$K_2 = \exp\left(-\Delta H_2 / RT + \left(\Delta S_2 - R\ln(Ct)\right)/R\right) \qquad [4.8]$$

$$K_3 = \exp\left(-\Delta H_3 / RT + \left(\Delta S_3 - R\ln(Ct)\right)/R\right) \qquad [4.9]$$

$$K_4 = \exp-\left(\Delta H_4 - \Delta S_4 T\right)/RT \qquad [4.10]$$

$$K_5 = \exp-\left(\Delta H_5 - \Delta S_5 T\right)/RT \qquad [4.11]$$

and

$$K_6 = \exp-\left(\Delta H_6 - \Delta S_6 T\right)/RT. \qquad [4.12]$$

Here ΔH_i and ΔS_i, for $j \in [1,2,3..6]$, are the apparent molar unfolding enthalpy and entropy terms, R is the molar gas constant, and Ct the total molar concentration of F1+F2+F3+F4+F7, F1+F2+F3+F6+F9 or F1+F2+F5+F8. It should, furthermore, be noted that K_2, and K_3 are not conventional equilibrium constants (Hüsler and Klump 1994), as they contain an additional term Rln(Ct). The relationship between the equilibrium constants and the fractional amounts of each species α_i, for $i \in \{1,2,3.... 9\}$, follows directly:

$$K_1 = \frac{\alpha_2}{\alpha_1} \qquad [4.13]$$

$$K_1K_2 = \frac{\alpha_3\alpha_4}{\alpha_1} \qquad [4.14]$$

$$K_1K_2K_3 = \frac{\alpha_4\alpha_5\alpha_6}{\alpha_1} \qquad [4.15]$$

$$K_1K_2K_4 = \frac{\alpha_5\alpha_6\alpha_7}{\alpha_1} \qquad [4.16]$$

$$K_1K_2K_3K_4 = \frac{\alpha_4\alpha_6\alpha_8}{\alpha_1} \qquad [4.17]$$

$$K_1K_2K_3K_6 = \frac{\alpha_4\alpha_5\alpha_9}{\alpha_1} \qquad [4.18]$$

The conservation of mass can also be expressed in terms of the fractional amount of each species α_i, as follows:

$$1 = \alpha_1 + \alpha_2 + \alpha_3 + \alpha_5 + \alpha_8$$

or $$1 = \alpha_1 + \alpha_2 + \alpha_3 + \alpha_5(1 + K_5) \qquad [4.19]$$

$$1 = \alpha_1 + \alpha_2 + \alpha_4 + \alpha_7$$

or $$1 = \alpha_1 + \alpha_2 + \alpha_7(1 + K_4) \qquad [4.20]$$

$$1 = \alpha_1 + \alpha_2 + \alpha_3 + \alpha_6 + \alpha_9$$

or $$1 = \alpha_1 + \alpha_2 + \alpha_3 + \alpha_6(1 + K_6). \qquad [4.21]$$

To account for mass action the partition function can be expressed in terms of the equilibrium constants and the fractional species α_4 and α_6 as follows:

$$Q = 1 + K_1 + \frac{K_1 K_2}{\alpha_4} + \frac{K_1 K_2 K_3 (1 + K_5)}{\alpha_4 \alpha_6}. \qquad [4.22]$$

It should be noted that the inclusion of the fractional quantities into the partition function has in fact produced the grand partition function for a given oligonucleotide. This is not surprising as one requires a more general assembly to handle systems in which molecular species are permitted to interact. There are in fact three partition functions for this system, one for each of the interacting oligonucleotides. The grand partition function for the entire system can be calculated from the product of all three partition functions. Even though this is in principal a more elegant approach it does not, however, simplify the mathematical solution. In this approach we only use the partition function represented by equation [4.22] in combination with equations [4.19]-[4.21] to account for the interacting systems.

The fractional amount of the species α_i can be obtained from the partition function Q

$$\alpha_1 = 1/Q \qquad [4.23]$$

$$\alpha_2 = K_1/Q \qquad [4.24]$$

$$\alpha_3 = K_1 K_2/\alpha_4 Q \qquad [4.25]$$

$$\alpha_5 = K_1K_2K_3/\alpha_4\alpha_6Q \qquad [4.26]$$

$$\alpha_6 = K_1K_2K_3(1+K_5)/(1+K_6)\alpha_4\alpha_6Q \qquad [4.27]$$

$$\alpha_4 = K_1K_2/(1+K_4)\alpha_4Q + K_1K_2K_3(1+K_5)/(1+K_4)\alpha_4\alpha_6Q \qquad [4.28]$$

Q is eliminated by combining equations [4.26] and [4.27] to obtain α_4 and α_6 in terms of the constants K_3, K_4, K_5, and K_6.

$$\alpha_5 = \frac{\alpha_6(\alpha_6 + K_3 + K_3K_5)(1+K_6)}{K_3(1+K_4)(1+K_5)} \qquad [4.29]$$

From equations [4.29] and [4.22] α_4 is eliminated yielding

$$Q = \frac{\alpha_6^2(1+K_1)(1+K_6) + K_1K_2K_3(1+K_4)(1+K_5)}{\alpha_6^2(1+K_6)} \qquad [4.30]$$

Substituting [4.29] into [4.27] yields

$$Q\alpha_6{}^3\left(1+K_6\right)^2\left(\alpha_6+K_3+K_3K_5\right)=K_1K_2K_3{}^2\left(1+K_5\right)^2\left(1+K_4\right) \qquad [4.31]$$

and by combining equations [4.30] and [4.31] it is now possible to express α_6 in terms of the equilibrium constants K_1 to K_6

$$0=\left(\alpha_6{}^4+\alpha_6{}^3K_3w_2\right)\left(1+K_1\right)w_3{}^2+K_1K_2K_3\left(\alpha_6{}^2+\alpha_6K_3w_2\right)w_1w_2w_3$$
$$-K_1K_2K_3{}^2w_1w_2{}^2 \qquad [4.32]$$

where w_1, w_2 and w_3 are defined by the following equations:

$$w_1=\left(1+K_4\right) \qquad [4.33]$$

$$w_2=\left(1+K_5\right) \qquad [4.34]$$

$$w_3=\left(1+K_6\right). \qquad [4.35]$$

Taking α_6 = 0.5 as an initial value, one can solve equation [4.32] numerically by Newton's method, in which iterations were found to converge rapidly. Using the values of α_6 obtained numerically, α_4 may be calculate from equation [4.29] and hence [4.22] can be solved for the partition function Q and α_1-α_3 and α_5. [4.23]-[4.26]. The fractional quantities of α_7, α_8 and α_9 can be obtained directly from the following equation set:

$$\alpha_7 = K_4\alpha_4 \qquad [4.36]$$

$$\alpha_8 = K_5\alpha_5 \qquad [4.37]$$

$$\alpha_9 = K_6\alpha_6 \qquad [4.38]$$

3.4.1. Thermal Denaturation

Calorimetric Thermal Denaturation (DSC)

The relationship between DSC data and the above theoretical solution can be deduced from the experimentally obtained molar transition enthalpy values ΔH_j, and the fractional amounts of the species α_i

$$\Delta C_p = (\partial\alpha_2/\partial T)\Delta H_1 + (\partial\alpha_3/\partial T)(\Delta H_1 + \Delta H_2) + (\partial\alpha_4/\partial T)(\Delta H_1 + \Delta H_2 + \Delta H_3) + (\partial\alpha_7/\partial T)(\Delta H_1 + \Delta H_2 + \Delta H_3 + \Delta H_4) + (\partial\alpha_8/\partial T)\Delta H_5 + (\partial\alpha_9/\partial T)\Delta H_6 \qquad [4.39]$$

The theoretical excess heat capacity for each reaction step corresponding to K_1-K_6 can be calculated from

$$\Delta Cp_1 = \left(\partial\alpha_2/\partial T + \partial\alpha_3/\partial T + \partial\alpha_4/\partial T + \partial\alpha_7/\partial T\right)\Delta H_1 \qquad [4.40]$$

$$\Delta Cp_2 = \left(\partial\alpha_3/\partial T + \partial\alpha_4/\partial T + \partial\alpha_7/\partial T\right)\Delta H_2 \qquad [4.41]$$

$$\Delta Cp_3 = \left(\partial\alpha_4/\partial T + \partial\alpha_7/\partial T\right)\Delta H_3 \qquad [4.42]$$

$$\Delta Cp_4 = \left(\partial\alpha_7/\partial T\right)\Delta H_4 \qquad [4.43]$$

$$\Delta Cp_5 = \left(\partial\alpha_8/\partial T\right)\Delta H_5 \qquad [4.44]$$

$$\Delta Cp_6 = \left(\partial\alpha_9/\partial T\right)\Delta H_6. \qquad [4.45]$$

3.5. pH Dependence of Triple-Helices Containing CGC^+ Triplets

3.5.1. Partition Function

I have not yet discussed the fact that the binding of C's in the Hoogsteen strand requires protonation. In the following section a model using statistical thermodynamics is developed, that describes the interdependence between thermal stability (Tm) of the triple-helices, pH, and the number of cytosines in p6p7 and p8p9 (cf. Fig. 3.4). To begin with, it is necessary to explain a number of assumptions. All cytosines in the WC, Looped, and HG sections of p6 and p8 are assumed to exchange protons with the solution only in the random coiled state. The extent of proton exchange of cytosines imbedded inside a fully folded triplex is considered to be negligible. Furthermore, it is shown that third strand binding is highly specific, as can be deduced from its intolerance to mismatches. It is thus assumed that only one of many possible proton arrangements in the pyrimidine strand will be tolerated on triplex formation. In order to determine the influence of pH on the thermal stability of triple-helices a knowledge of the population distribution of the various proton arrangements is required. It is apparent that the distribution of any given proton arrangement in the pyrimidine strands will be dependent on the pH of the solution and the sequence of the oligonucleotides. It is, furthermore, assumed that all cytosines in the WC, HG and looped regions of the unfolded single strands p6 and p8 can interact with protons. In the formalism developed here, not all HG cytosines are necessarily protonated in the triplex. The number of protonated cytosines in the triple-helices will be determined

directly from experimental data for the limiting condition where the pH is larger than the pKa of cytosine. The temperature dependence of the pKa of cytosine has been determined by Manzini *et al.* (1990). It was shown that the pKa shifts by a mere 0.13 units in the temperature range 22 to 41 °C. As the temperature dependence of the pKa of cytosine is within the experimental error margin it does not warrant a general temperature dependent treatment, and will therefore be taken as a constant in these experiments. It should non-the-less be noted that this constraint is in principal not necessary, and the formalism can handle a temperature dependence of the pKa values.

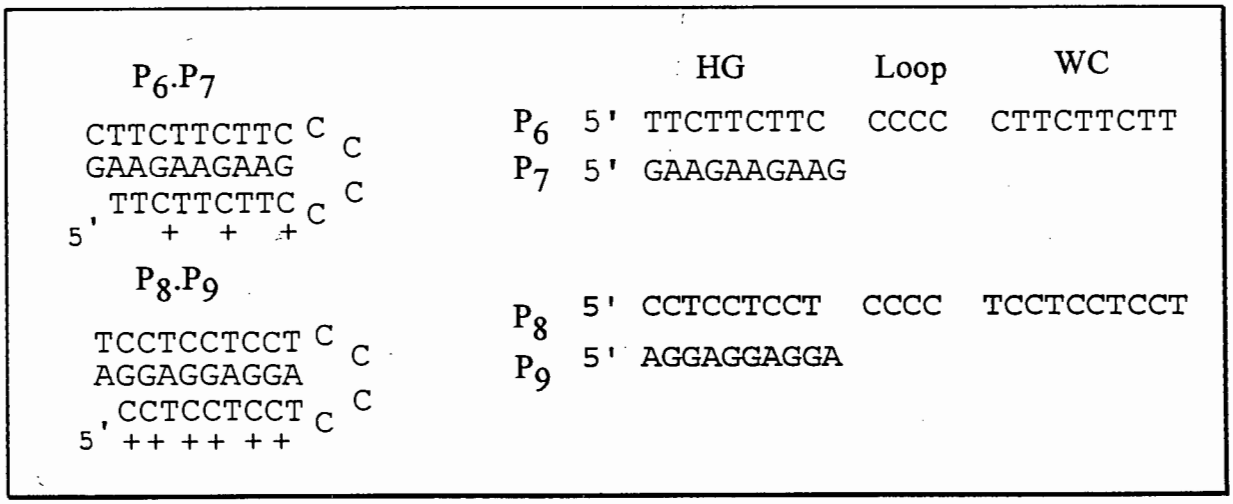

Fig. 3.4 Purine oligonucleotides p6 and p8 (A) and Pyrimidine oligonucleotides p7 and p9 used in the triple-helices p6p7 and p8p9 (B). The Hoogsteen and Watson-Crick bases are abbreviated as HG and WC respectively.

3.5.2. Polymer with Independent Proton Binding Sites

To begin with, consider a one dimensional lattice containing a number of sites, each capable of binding a proton. In general, a site that binds a proton will be treated as an occupied site. Sites that are not associated with a proton will be termed unoccupied sites.

If proton binding is to be considered independent, then from statistical thermodynamics (Poland and Scheraga. 1970), the fraction α_n of any unique arrangement of n occupied sites can be expressed in terms of the total number of sites per molecule (N), the proton concentration $[H^+]$ and the dissociation constant Ka, as follows:

$$\alpha_n = \left(\frac{[H^+]}{Ka}\right)^n \left(\frac{Ka}{[H^+]+Ka}\right)^N \qquad [5.1a]$$

As the assumption of independent sites is an over simplification I now describe a more advanced model.

3.5.3. Polymer with Interacting Proton Binding Sites

It is assumed that the binding of a proton to any given site will depend on the status of its neighboring sites and the pH of the solution. Therefore, it is of interest to determine the concentration of any given proton-binding configuration as a function of pH. I have used the grand partition function Ξ to determine the form of the matrix required to yield the desired configurations of occupied and unoccupied sites, and then adjust the number of molecules for each species with respect to λ, the absolute activity. The grand partition function can be expressed in terms of the partition functions for each species and their associated absolute activities in the following way:

$$\Xi = \sum_{m_i}^{M} \sum_{N_i}^{m_i} \prod_{i}^{n} \frac{Z_i^{N_i} . \lambda_i}{N_i !} . \qquad [5.1b]$$

Where Z_i is the partition function of a single molecule, λ_i its activity and N_i the number of molecules of type i. The absolute activity for any given molecule i with respect to the free proton concentration depends on the number of bound protons to the given species. The absolute activity λ_i for the molecular species i is, therefore, equal to the absolute activity λ_h of the free protons in solution raised to the power of the number of occupied sites. This condition defines a thermodynamic equilibrium between all molecular species and the free protons in solution. It is therefore possible to rewrite Eq. [5.1b], yielding

$$\Xi = \sum_{m_i}^{M} \sum_{N_i}^{m_i} \prod_{i}^{n} \frac{\left(Z_i \lambda_h^{k_i}\right)^{N_i}}{N_i!}, \qquad [5.2b]$$

where k_i is the number of protonated sites. It can be shown that equation [5.2b] is equivalent to

$$\Xi = \left(\sum_{ns} \prod_{s=1}^{n} \left(q_{ns} \lambda_h^{k_{ns}} \right) \right)^{M} \qquad [5.3b]$$

and

$$q_{ns} = \exp\left(-\beta \varepsilon_{ns}\right), \qquad [2.4b]$$

where s designates the position of a site in a given molecule and its state, (n is equal to 0 for an unoccupied site, and 1 if the site is occupied). k_{ns} can be either 0 for an unoccupied site (n = 0) or 1 in the case of an occupied site (n = 1). q_{ns} represents the statistical weight for a given protonated site, which is dependent on the state of its neighbors. As every DNA oligomer is assumed to be independent, the total number of molecules (M) in a system can be dropped in Eq. [5.3b]. In general the grand partition function can be written as a product of matrices as follows

$$\Xi = \mathbf{e}\prod \mathbf{W}_i\mathbf{e}^+, \qquad [5.5b]$$

where **e** is a row matrix with the first element equal to 1 and $\mathbf{e}^+$ a unit column matrix. The size of the matrices is determined by the number of neighboring interactions. For the nearest neighbor model we typically require a knowledge of two consecutive sites and therefore require a 2x2 matrix (Poland and Scheraga 1970). For homopyrimidine oligonucleotides any given site can be either occupied by a cytosine or a thymine. Thymine sites and unoccupied cytosine sites are not distinguishable. To account for the given site's identity and the state and type of its neighbor, we define 4 matrices $\mathbf{W}_{TT}$ = $\mathbf{W}_{CT}$, $\mathbf{W}_{CC}$ and $\mathbf{W}_{TC}$. The first subscript designates the nearest neighbor and the second subscript the identity of the base of interest. The four matrices are:

$$\mathbf{W}_{TT} = \mathbf{W}_{CT} = \begin{array}{c|ccc} i-1 \backslash i & 0 & + \\ \hline 0 & 1 & 0 \\ 0+ & 0 & 1 \end{array} \qquad [5.6b]$$

$$\mathbf{W}_{TC} = \begin{array}{c|ccc} i-1 \backslash i & 0 & + \\ \hline 0 & 1 & q_1\lambda_h \\ 0 & 1 & q_1\lambda_h \end{array} \qquad [5.7b]$$

$$\mathbf{W}_{CC} = \begin{array}{c|ccc} i-1 \backslash i & 0 & + \\ \hline 0 & 1 & q_1\lambda_h \\ + & 1 & q_2\lambda_h \end{array} \qquad [5.8b]$$

The table assigns the statistical weights for the *i*th site in terms of the state (occupied or unoccupied) and the identity (C, T) of its nearest neighbor at position (i-1). The statistical weights $q_1\lambda_h$ and $q_2\lambda_h$ can be expressed in terms of the dissociation constants Ka_1 and Ka_2 respectively, and the proton concentration of the solution $[H^+]$.

$$q_1\lambda_h = \frac{[H^+]}{Ka_1} \qquad [5.9b]$$

$$q_2\lambda_h = \frac{[H^+]}{Ka_2} \qquad [5.10b]$$

As we are dealing with short oligonucleotide sequences, end effects should be taken into account. For simplicity it is assumed that the first site in the oligonucleotide is similar to a buried site in the lattice with an uncharged neighbor in the position (i-1). The matrix that represents the first site is therefore equivalent to either $\mathbf{W}_{TT}$ or $\mathbf{W}_{TC}$ depending on whether the first nucleotide is a thymine or cytosine respectively. We are now in a position to calculate the grand partition function Ξ for any sequence taking nearest neighbor electrostatic interactions into account. It should be noted that the electrostatic properties of nearest neighbors has been included into q_2, as it is an

implicit value to be determined experimentally. The fraction α_i of a given sequence of protonated and unprotonated sites is then given by

$$\alpha_i = \frac{\partial \ln(\Xi)}{\partial \ln \xi_i} \qquad [5.11b]$$

Where ξ_i is the product of the statistical weight q_i for a given protonation pattern and its associated absolute activity λ_h. The grand partition function for the molecular species p6 and p8 can be expressed with the help of matrices $\mathbf{W}_{TC}$ and $\mathbf{W}_{CC}$. The matrices $\mathbf{W}_{CT}$ and $\mathbf{W}_{TT}$ are formally not included. This is justified as they are identity matrices (cf. Eq. [5.6b]).

$$\Xi_{p6} = \mathbf{e}\left(\mathbf{W}_{TC}\right)^4\left(\mathbf{W}_{CC}\right)^5\left(\mathbf{W}_{TC}\right)^2\mathbf{e}^+ \qquad [5.12b]$$

$$\Xi_{p8} = \mathbf{e}\left(\mathbf{W}_{TC}\mathbf{W}_{CC}\right)^4\left(\mathbf{W}_{CC}\right)^2\left(\mathbf{W}_{TC}\mathbf{W}_{CC}\right)^3\mathbf{e}^+ \qquad [5.13b]$$

The fraction of a given sequence of protonated sites for p6 or p8 (α_{p6} and α_{p8} respectively) can be deduced from Eq. [5.11b].

$$\alpha_{p6} = \frac{\xi_{p6}}{e(W_{TC})^4(W_{CC})^5(W_{TC})^2 e^+} \qquad [5.14b]$$

$$\alpha_{p8} = \frac{\xi_{p8}}{e(W_{TC}W_{CC})^4(W_{CC})^2(W_{TC}W_{CC})^3 e^+} \qquad [5.15b]$$

3.5.4. Thermal Stability

We are now in a position to determine the thermal stability (Tm) for a particular triplex ($pu{\cdot}pyH^+_i$) containing a short purine (pu) and a long pyrimidine (py) strand. The formation constant K for this triplex can be expressed in terms of the fraction θ of single stranded oligonucleotides (2pu/Ct or 2py/Ct), the total strand concentration of the system Ct, and α_i, the fraction of single stranded homopyrimidine oligonucleotides with a given protonation arrangement (pyH^+_i).

$$py + (1,2,3...n)H^+ \overset{K}{\leftrightarrow} pyH^+ + pyH^+_2 + ... + py\,H^+_j + ... + pyH^+_N$$

$$pyH^+_j + pu \overset{K}{\leftrightarrow} pupyH^+_i \qquad [5.16b]$$

with

$$\alpha_i = \frac{pyH^+_i}{py + \sum_{j=1}^{N} pyH^+_j}$$

$$\frac{K \cdot Ct \cdot \alpha_i}{2} = \frac{1-\theta}{\theta^2} \qquad [5.17b]$$

The denaturation temperature (Tm) where 50% of the triplex is disordered can be deduced directly from Eq. [5.17b] as follows:

$$Tm = \frac{\Delta H}{R\ln(\alpha_i) + \Delta S - R\ln 4 + R\ln(Ct)}, \qquad [5.18b]$$

where the formation enthalpy and entropy are given by ΔH and ΔS respectively and R denotes the gas constant. There are two limiting cases to be discussed with respect to Eq. [5.18b]: (1) where $[H^+]$ is considerably larger than the dissociation constants Ka_1 and Ka_2 and, (2) where $[H^+]$ is much smaller than Ka_1 and Ka_2. It can be shown that the matrices $\mathbf{W}_{CC}$ and $\mathbf{W}_{TC}$ have the following limits for $[H^+] >> Ka_1$ and Ka_2:

$$W_{TC} = \begin{pmatrix} 0 & q_1\lambda_h \\ 0 & q_1\lambda_h \end{pmatrix} \qquad [5.19b]$$

$$W_{TC} = \begin{pmatrix} 0 & q_1\lambda_h \\ 0 & q_2\lambda_h \end{pmatrix} \qquad [5.20b]$$

and hence α_i simplifies to

$$\alpha_n = [H]^{(n-N)} Ka_1^{(m-v_1)} Ka_2^{(n-m-v_2)}, \qquad [5.21b]$$

where N is the total number of sites which can bind protons in the homopyrimidine strand (i.e. cytosines) and n the number of protonated cytosines required for triplex formation. v_1 designates the number of protonated cytosines with uncharged neighbors and v_2 the number of protonated cytosines with charged neighbors ($n = v_1 + v_2$). For the case of homopyrimidine oligonucleotides with every cytosine protonated, m designates the number of cytosines with uncharged neighbors. From Eq. [5.18b] and [5.21b] one can deduce that 1/Tm vs. pH results in a linear function, as long as $[H^+] >> Ka_1$ and Ka_2:

$$\frac{1}{Tm} = \frac{R(N-n).\ln(10).pH}{\Delta H} + C, \qquad [5.22b]$$

and C is given as

$$C = \frac{\Delta S - R\ln 4 + R\ln Ct - R\ln(10)\left[(m-v_1)pKa_1\right] + (N-m-v_2)pKa_2}{\Delta H} \qquad [5.23b]$$

For $[H^+] << Ka_1$ and Ka_2 the matrices $\mathbf{W}_{CT}$ and $\mathbf{W}_{CC}$ have the following properties

$$\mathbf{W}_{TC} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \qquad [5.24b]$$

$$\mathbf{W}_{CC} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \qquad [5.25b]$$

From the matrices [5.24b] and [5.25b] it can be shown that the grand partition function Ξ approaches 1 and α_i becomes equal to

$$\alpha_i = [H]^n (Ka_1)^{-v_1} (Ka_2)^{-v_2} \qquad [5.26b]$$

Furthermore, for 1/Tm it follows that

$$\frac{1}{Tm} = \frac{-Rn\ln(10)pH}{\Delta H} + C \qquad [5.27b]$$

with

$$C = \frac{\Delta S - R\ln 4 + R\ln Ct + R\ln(10)(v_1 pKa_1 + v_2 pKa_2)}{\Delta H}. \qquad [5.28b]$$

The intersection of the linear parts of Eq. [5.22b] and [5.27b] corresponds to the average pKa value for a given oligonucleotide in its fully protonated state.

$$pH = \frac{mpKa_1 + (N-m)pKa_2}{N} \qquad [5.29b]$$

For the simpler model with independent sites (cf. [5.1a]) one can either use Eq. [5.29b] by setting $Ka_1 = Ka_2$ or by substituting Eq. [5.1a] into [5.18b]. In either case the result is the same. For the limiting case where $[H^+] >> Ka$, the relation between 1/Tm and pH becomes

$$\frac{1}{Tm} = \frac{R\ln(10)(N-n)pH}{\Delta H} + C \qquad [5.30b]$$

with

$$C = \frac{\Delta S - R\ln 4 + R\ln Ct - R\ln(10)(N-n)pKa}{\Delta H}, \qquad [5.31b]$$

and for the opposite case where $[H^+] << Ka$ it results in

$$\frac{1}{Tm} = \frac{-nR\ln(10)pH}{\Delta H} + C \qquad [5.32b]$$

with

$$C = \frac{\Delta S - R\ln 4 + R\ln Ct + nR\ln(10)pKa}{\Delta H}. \qquad [5.33b]$$

In both cases, 1/Tm is linear with respect to the pH (with opposite slope). The intersection of these two linear functions corresponds to the dissociation constant pKa.

The number of proton binding sites in the triple-helices can be deduced from the slope of 1/Tm vs pH by differentiating Eq. [5.32b]. This yields:

$$n = \frac{-\Delta H}{R \ln(10)} \frac{\partial(1/Tm)}{\partial pH}. \qquad [5.34b]$$

Chapter 4.

Methods and Materials

4.1. Chemicals and Solutions

4.1.1. Oligonucleotide Synthesis

The oligonucleotides:

p1: 5'CCTCCTCCTCCCCTCCTCCTCCTGAAGAAGAA,
p2: 5'TTTTTTTTTCCCCTTTTTTTTTTAGGAGGAGGA,
p3: 5'TTCTTCTTCCCCCCTTCTTCTTCAAAAAAAAAA,
p4: 5'TTTTTTTTTCCCCTTTTTTTTTT,
p5: 5'AAAAAAAAAA,
p6: 5'TTCTTCTTCCCCCCTTCTTCTTC,
p7: 5'GAAGAAGAAG,
p8: 5'CCTCCTCCTCCCCTCCTCCTCCT,
p9: 5'AGGAGGAGGA,
hp1: 5'AAGAAGAAGCCCCCTTCTTCTTCTTTTTTTTTT,
hp2: 5'GGAGGAGGACCCCTCCTCCTCCTCTTCTTCTTC,
hp3: 5'AAAAAAAAAACCCTTTTTTTTTTTCCTCCTCCT,
q1: 5'CCTCCTCCTGAAGAAGAAG,
q2: 5'TTTTTTTTTTAGGAGGAGGA,
q3: 5'CTTCTTCTTCAAAAAAAAAA

were synthesized on an Autogen 6500 DNA synthesizer (Milipore), using standard solid-phase cyanoethyl phosphoramidite chemistry (Beaucage and Caruthers 1981). The oligonucleotides were purified by ion exchange or reverse phase HPLC (Laughlin and Piel 1984), and shown to be pure by means of denaturing polyacrylamine gel electrophoresis (producing one band only). The concentrations of the oligonucleotides were determined in water by means of measuring the absorbance at 260 nm at 90 °C, using extinction coefficients of 7400, 8700, 15400, and 11500 $M^{-1}cm^{-1}$ for C, T, A, and G bases in the single-strand conformation, respectively, as given in the literature (Cantor *et al.* 1970). The following extinction coefficients were determined:

Oligonucleotide	Extinction coefficient, ε $[M^{-1}cm^{-1}]$
p1	$\varepsilon = 3.18 \times 10^5$
p2	$\varepsilon = 3.26 \times 10^5$
p3	$\varepsilon = 3.40 \times 10^5$
p4	$\varepsilon = 1.95 \times 10^5$
p5	$\varepsilon = 1.54 \times 10^5$
p6	$\varepsilon = 1.85 \times 10^5$
p7	$\varepsilon = 1.38 \times 10^5$
p8	$\varepsilon = 1.79 \times 10^5$
p9	$\varepsilon = 1.31 \times 10^5$
hp1	$\varepsilon = 3.25 \times 10^5$
hp2	$\varepsilon = 3.06 \times 10^5$
hp3	$\varepsilon = 3.81 \times 10^5$
q1	$\varepsilon = 2.18 \times 10^5$
q2	$\varepsilon = 2.18 \times 10^5$
q3	$\varepsilon = 2.36 \times 10^5$

4.1.2. Solutions

The oligonucleotides were dissolved in either 20 mM Na_3PO_4 or Na_2HPO_4 buffer. The sodium concentrations were adjusted by mixing a stock solution containing only buffer with a second solution containing buffer and 2 M sodium chloride, in various proportions. The pH was adjusted by adding HCl or NaOH to the stock solutions or by adjusting the pH of the sample directly.

4.2. Physical Methods

4.2.1. Ultraviolet Absorbance vs Temperature Profiles

Ultraviolet absorbance (260 nm or 280 nm) was recorded as a function of temperature using either a Hewlett-Packard HP 8450a diode array spectrophotometer equipped with a HP 89101a temperature probe and a HP 98100 temperature controller unit, or a Pye Unicam spectrophotometer equipped with a custom-made temperature controller interfaced to an analog to digital device (oasis/4; 3D Digital Design and Development Ltd., London) and IBM 386 computer. The data were sampled every 0.1 or 0.05 °C and stored on disk for further analysis. All experiments were conducted at a heating rate of 1 °C/min using a 1-cm quartz cuvette or a 0.1-cm quartz cuvette, (Helma, Mulheim, Germany).

4.2.2. Differential Scanning Calorimetry

DSC experiments were performed at a heating rate of 1 °C/min. on a DASM-4 differential scanning microcalorimeter (Mashpriborintork, Moscow) interfaced to an analog to digital device (oasis/4; 3D Digital Design and Development Ltd., London) and IBM 386 computer. Data were collected every 0.05 °C and stored on disk for further processing on an IBM 486 PC.

4.2.3. Polyacrylamide Gel Electrophoresis (PAGE)

Polyacrylamide gels were made up from solutions containing working buffer, 20% (w/v) acrylamide, 2% bisacrylamide, and 0.05-0.08% ammonium persulfate. All experiments under nondenaturing conditions were run either at room temperature or at 4°C in a cold room. Typically gels run at pH 8 contained 150 mM NaCl and TBE (89 mM Tris, 89 mM boric acid, and 0.25 mM EDTA), while gels run at pH less than 7 contained 100 mM ammonium acetate and 50 mM NaCl, or 100 mM sodium acetate and 50 mM $MgCl_2$. The pH was adjusted with either acetic acid or HCl. Five to ten micrograms of sample were loaded per lane with bromophenol blue as tracking dye. Gels were run at approximately 10 V/cm, constant voltage. The contents of the cathode and anode buffer tanks were circulated to maintain a constant pH. Gels were stained with stains-all dye in 1:1 water:formamide and photographed after having been destained in water. The staining intensity of oligonucleotides at the same concentration are not necessarily equal and depend on the oligonucleotide sequence, for example, base composition and length.

4.2.3. Temperature Gradient Gel Electrophoresis

Gels were made up from solutions containing working buffer, 20% (w/v) acrylamide, 2% (w/v) bisacrylamide, and 0.05-0.08% (w/v) ammonium persulfate. Working buffer contained either 100 mM ammonium acetate and 50 mM sodium acetate, or 100 mM sodium acetate, 50 mM $MgCl_2$ at pH 5.5. A temperature gradient perpendicular to the electric field was applied ranging from either 20-70 °C, 5-85 °C,

or 25-60 °C. Gels were run at constant voltage at approximately 10V/cm. DNA samples were loaded with bromophenol blue as tracking dye. The contents of the cathode and anode buffer tanks were circulated to maintain a constant pH. Gels were stained as described above.

4.2.5. Circular Dichroism Spectroscopy

Circular dichroism scans were carried out on a Jasco J-40 spectrophotometer. The temperature of the sample was controlled by a thermostatically regulated water bath attached directly to the quartz cuvette.

4.2.6. Software Development

All customized software interfacing with instrumentation was developed in the programming language C/C^{++}. Solutions to equations were processed numerically, provided no analytical solutions were obtainable. Equations presented in chapter 3 were fitted to experimental data by the method of least squares using either the steepest descent method (Bevington 1969), or the Levenberg-Marquard procedure (Press *et al.*). In nearly all cases the Levenberg-Marquard method was found to be superior to the steepest descent technique. Algorithms for curve fitting and solutions to equations were written in the programming language C/C^{++}. None of the source code for the various algorithms is given in this thesis due to its complexity and size.

4.3. Mathematical Analysis of Calorimetric and uv Melting Curves

4.3.1. Thermal Unfolding of the WC Triple-Helical Three-Way Junction p1p2p3 and the Double-Helical Three-Way Junction q1q2q3

Since the WC triple-helical three-way junction and the double-helical three-way junction are formed by three oligomers, p1, p2, and p3 and q1, q2, and q3, respectively, the following reversible multistate reaction scheme was used to determine the changes in enthalpy and entropy for each step (cf. Fig. 3.2):

$$F1 \overset{K_1}{\leftrightarrow} F2 \qquad [3.1]$$

$$F2 \overset{K_2}{\leftrightarrow} F3 + F4 \qquad [3.2]$$

$$F3 \overset{K_3}{\leftrightarrow} F5 + F6, \qquad [3.3]$$

where K_1, K_2, and K_3 are equilibrium constants (cf. chapter 3.3). It is apparent from the model that the first step is a conventional equilibrium constant. The two consecutive steps are bimolecular and the equilibrium constants assigned to these steps are modified to include a concentration term to facilitate the mathematical treatment. The corresponding Eqs. [3.1]-[3.3] can be solved using the method described in chapter 3.3 (equations are numbered in the order in which they appear in the chapter).

The optical density (e.g., absorbance at 260 nm) of a solution containing the three-way junctions can be predicted from the equation

$$Abs = Ct\left[\alpha_1\varepsilon_1\alpha +_2 \varepsilon_2 + \alpha_3\varepsilon_3 + \alpha_4\varepsilon_4 + \alpha_5\varepsilon_5 + \alpha_6\varepsilon_6\right] \qquad [3.27]$$

where Ct is the concentration of the individual strands (p1, p2, and p3, for the WC triple-helical junction or q1, q2, and q3 for the double-helical junction), α_4 and α_6 are the fractions of F4 and F6 in the single-stranded state, respectively, and Q is the partition function.

$$Q = 1 + K_1 + \frac{K_1K_2}{\alpha_4} + \frac{K_1K_2K_3}{\alpha_4\alpha_6} \qquad [3.4]$$

The extinction coefficients for Fj are given by ε_j where $j \in \{1, 2, 3, 4, 5, 6\}$. It is assumed that the extinction coefficients for each intermediate species changes linearly with temperature according to:

$$\varepsilon_1 = \beta_1 + \mu_1 T \qquad [3.21]$$

$$\varepsilon_2 = \beta_2 + \mu_2 T \qquad [3.22]$$

$$\varepsilon_3 = \beta_3 + \mu_3 T \qquad [3.23]$$

$$\varepsilon_4 = \beta_4 + \mu_4 T \qquad [3.24]$$

$$\varepsilon_5 = \beta_5 + \mu_5 T \qquad [3.25]$$

$$\varepsilon_6 = \beta_6 + \mu_6 T \qquad [3.26]$$

Eq. [3.27] was fitted to the uv melting data using a nonlinear regression approach, taking β_j, μ_j, ΔH_i, and ΔS_i as adjustable parameters with $i \in \{1,2,3\}$.

DSC curves can be described by the equation

$$\Delta C_p = (\partial\alpha_2/\partial T)\Delta H_1 + (\partial\alpha_3/\partial T)(\Delta H_1 + \Delta H_2) + (\partial\alpha_5/T)(\Delta H_1 + \Delta H_2 + \Delta H_3) \qquad [3.28]$$

where ΔC_p is the molar excess heat capacity. Eq. [3.28] was fitted to the experimental DSC data set, using the gradient search method of least squares (Press et al.; Bevington 1969) by treating ΔC_p, ΔH_i, and ΔS_i as adjustable parameters.

4.3.2. Thermal Unfolding of the Binary Complexes hp1hp2, hp2hp3, and hp1hp3

The unfolding pathway of the three binary complexes hp1hp2, hp2hp3, or hp1hp3 can be described by the following multi-state reaction scheme, where each step is considered to be fully reversible (cf. chapter 3.2, Fig. 3.1):

$$\text{Fa} \overset{K_1}{\leftrightarrow} \text{Fb} + \text{Fc} \qquad [1.1]$$

$$\text{Fb} \overset{K_2}{\leftrightarrow} \text{Fb'} \qquad [1.2]$$

$$\text{Fc} \overset{K_3}{\leftrightarrow} \text{Fc'}, \qquad [1.3]$$

where K_1, K_2, and K_2 are the equilibrium constants. Fa represents the binary complex formed between the hairpins Fa and the 3' pyrimidine extension of the hairpin Fc. Fb' and Fc' represent the single stranded (coiled) conformation of the hairpin oligonucleotides Fa and Fc, respectively. It should be noted that K_1 is not a conventional equilibrium constant, K_1 has been modified to include a concentration term, to facilitate the mathematical treatment of the reaction mechanism (cf. chapter 3.2). The remaining constants K_2 and K_3 are conventional equilibrium constants.

The optical absorbance (260 nm) for the three binary complexes hp1hp2, hp2hp3 and hp1hp3 can be calculated from the following equation:

$$Abs = C_t\left[\alpha_a\varepsilon_a + \alpha_b\varepsilon_b + \alpha_c\varepsilon_c + \alpha_{b'}\varepsilon_{b'} + \alpha_{c'}\varepsilon_{c'}\right] \qquad [1.30]$$

where Ct is the concentration of each of the oligonucleotides used in the reaction, α_i represents the fractional amounts of the species F_i, and ε_i their associated extinction coefficients, $i \in \{a,b,c,b',c'\}$. It is assumed that the extinction coefficients for each species changes linearly with temperature (cf. Eqs. [2.25]-[2.29], chapter 3.2.1). The calculated absorbance (260 nm) (Eq. [1.30]) was fitted to the experimental uv melting curves using the Levenberg-Marquard procedure of nonlinear regression (Press *et al.*) where ε_i, ΔH_j, and ΔS_j, $j \in \{1,2,3\}$ are adjustable parameters.

DSC curves can be described by the following equation (cf. chapter 3.2.1):

$$\Delta Cp = \left(\partial\alpha_b/\partial T\right)\Delta H_1 + \left(\partial\alpha_{b'}/\partial T\right)\left(\Delta H_1 + \Delta H_2\right) + \left(\partial\alpha_{c'}/\partial T\right)\Delta H_3. \qquad [1.31]$$

where ΔCp is the molar excess heat capacity. Equation [1.31] was fitted to the DSC experimental data set, where ΔH_j and ΔS_j were taken as adjustable parameters.

4.3.3. Thermal Unfolding of the Ternary Complexes hp1hp2hp3

The unfolding process of the ternary complex (HG triple-helical three-way junction) hp1hp2hp3 can be described by the following reaction scheme (cf. Fig. 3.3):

$$F1 \overset{K_1}{\leftrightarrow} F2 \overset{K_2}{\leftrightarrow} F3 + F4 \qquad [4.1]$$

$$F3 \overset{K_3}{\leftrightarrow} F5 + F6 \qquad [4.2]$$

$$F4 \overset{K_4}{\leftrightarrow} F7 \qquad [4.3]$$

$$F5 \overset{K_5}{\leftrightarrow} F8 \qquad [4.4]$$

$$F6 \overset{K_6}{\leftrightarrow} F9 \qquad [4.5]$$

F1 = (hp1hp2hp3)	(all bases paired and stacked)
F2 = (hp1hp2hp3)'	(one arm partially unfolded)
F3 = (hp1hp2, or hp2hp3, or hp1hp3)	(binary combination)
F4, F5, F6 = (hp1, hp2, hp3)	(hairpins, with single stranded 3' extensions)
F7, F8, F9 = (hp1, hp2, hp3)	(single strands as random coil)

K_1- K_6 are the equilibrium constants for each reaction step. From the reaction scheme (Fig. 3.3) it can be seen that the second and third steps are bimolecular. K_2 and K_3 are, therefore, not conventional equilibrium constants, as they contain the concentration terms. (cf. Eqs. [4.8] and [4.9]). All other steps are monomolecular, and hence all other equilibrium constants are taken as conventional. The fractional amounts of each species α_1-α_9 (cf. Eqs. [4.13]-[4.18]) were computed using the method described in chapter 3.4. The molar excess heat capacity ΔCp can be calculated from

$$\Delta Cp = (\partial\alpha_2/\partial T)\Delta H_1 + (\partial\alpha_3/\partial T)(\Delta H_1 + \Delta H_2) + (\partial\alpha_4/\partial T)(\Delta H_1 + \Delta H_2 + \Delta H_3) + (\partial\alpha_7/\partial T)(\Delta H_1 + \Delta H_2 + \Delta H_3 + \Delta H_4) + (\partial\alpha_8/\partial T)\Delta H_5 + (\partial\alpha_9/\partial T)\Delta H_6 \qquad [2.39]$$

The calculated excess heat capacity curves were fitted to the experimental DSC data set by adjusting the variables ΔH_1-ΔH_6 and ΔS_1-ΔS_6, using the Levenberg-Marquard procedure of nonlinear regression (Press *et al.*; Bevington 1969). As all DSC transitions occur over a relatively large temperature range (≈ 20-100 °C), values for the partial specific heat capacity values have not been included, as these are relatively small. The partial specific heat capacity values are, furthermore, within the same order of magnitude as the experimental base line uncertainties.

4.3.4. Data Analysis of Tm Dependence on pH

Tm values for the thermal denaturation of the triple helices p6p7 and p8p9 were determined by fitting the absorbance (260 nm) vs temperature profile to a two-state model using nonlinear regression. In the case of dependent proton binding in the triple helices p6p7 and p8p9, Eq. 2.18b was fitted to the 1/Tm vs pH data set:

$$Tm = \frac{\Delta H}{R\ln(\alpha_i) + \Delta S - R\ln 4 + R\ln(Ct)}. \qquad [2.18b]$$

Here R is the molar gas constant, ΔS and ΔH the formation enthalpy and entropy changes, and Ct the total strand concentration. The fractional amounts of α_{p6} and α_{p8} were obtained directly from:

$$\alpha_{p6} = \frac{\xi_{p6}}{\mathbf{e}\left(\mathbf{W}_{TC}\right)^4\left(\mathbf{W}_{CC}\right)^5\left(\mathbf{W}_{TC}\right)^2\mathbf{e}^+} \qquad [2.14b]$$

$$\alpha_{p8} = \frac{\xi_{p8}}{\mathbf{e}\left(\mathbf{W}_{TC}\mathbf{W}_{CC}\right)^4\left(\mathbf{W}_{CC}\right)^2\left(\mathbf{W}_{TC}\mathbf{W}_{CC}\right)^3\mathbf{e}^+} \qquad [2.15b]$$

where $\mathbf{W}_{CC}$ and $\mathbf{W}_{TC}$ are matrices with elements of proton concentration $[H^+]$ and the Ka_1 and Ka_2 of cytosine.

$$\mathbf{W}_{TC} = \begin{pmatrix} 1 & [H^+]/Ka_1 \\ 1 & [H^+]/Ka_1 \end{pmatrix} \qquad [5.7b]$$

$$\mathbf{W}_{CC} = \begin{pmatrix} 1 & [H^+]/Ka_1 \\ 1 & [H^+]/Ka_2 \end{pmatrix} \qquad [5.8b]$$

In the case of the simpler model of independent proton binding sites, the fractional amounts of α_{p6} and α_{p8} (α_n) were determined as follows:

$$\alpha_n = \left(\frac{[H^+]}{Ka}\right)^n \left(\frac{Ka}{[H^+] + Ka}\right)^N \qquad [5.1a]$$

N represents the total number of proton binding sites and n the number of occupied sites. The total number of protonated sites in the complexes p6p7, p8p9, hp1hp2, hp2hp3 and hp1hp2hp3 were first calculated directly from the slope of 1/Tm vs pH in the pH range where the pH is lower than the pKa of cytosine using the following equation:

$$n = \frac{\partial(1/Tm)}{\partial pH}\left(\frac{-\Delta H}{R\ln(10)}\right). \qquad [5.34b]$$

Here n represents the total number of protonated cytosines in the Hoogsteen strand, R the molar gas constant, and ΔH the formation enthalpy changes associated with the triple helices (obtained from uv thermal denaturation or DSC data).

Chapter 5.

WC Triple-Helical Three-Way Junction

The WC junction was constructed from the oligonucleotides p1, p2, and p3 (cf. Fig. 5.1). The three arms (A, B, and C) of the completely folded complex contain the same structural elements, which are joined at the branch point by an underlying double-helical three-way junction (cf. chapter 7). Each arm consists of a triple-helix made from a unique purine sequence and two complementary pyrimidine sequences which represent the Watson and Hoogsteen strands, respectively. The sequences are linked by a four-member loop consisting of cytosines. Arm A contains exclusively TAT triad bases, while arms B and C contain increasing numbers of CGC^+ triad bases. To prevent possible steric interference (crowding) at the branch point, the Hoogsteen strands were shortened by one base as compared to the underlying WC duplex. To simplify the investigation of the thermal behavior of the individual arms, it was necessary to construct the triple-helical arms independently with oligonucleotides lacking the 3' extension. The arms A, B, and C for the WC triple-helical three-way junction are represented by the binary complexes formed between the oligonucleotides p4p5, p6p7, and p8p9, respectively (cf. Fig. 5.2).

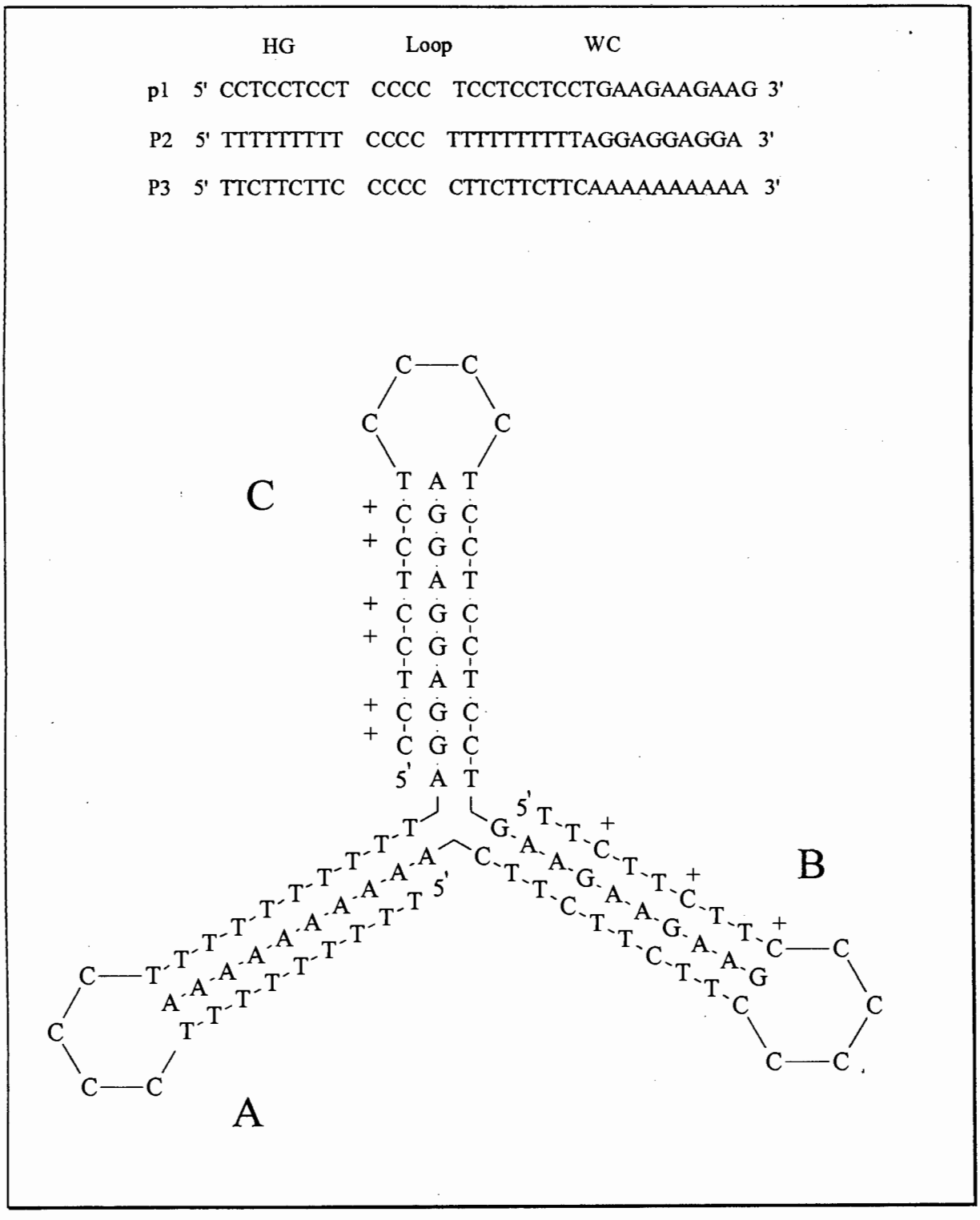

Fig. 5.1 WC triple-helical three-way junction with arms marked A, B, and C. The oligonucleotides used to construct the junction are p1, p2, and p3. The Hoogsteen and Watson-Crick bases are abbreviated as HG and WC, respectively.

Binary complexes are formed by combining one homopurine (10-mer) and one homopyrimidine oligonucleotide (23-mer) in a 1:1 ratio. The homo-pyrimidine strand consists of a 10-mer domain which forms the Crick region and a 9-mer Hoogsteen region bridged by four cytosine bases (loop). At neutral pH the Crick sequence can associate with the complementary 10-mer Watson homopurine oligonucleotide to form a duplex with a single-strand extension. The single-strand extension can be made to fold back onto the duplex by dropping the pH to 5 for p6p7 and p8p9, to form the triplex through Hoogsteen base pairs. The triplex p4p5 contains exclusively thymine Hoogsteen bases and should therefore be spontaneous and pH independent (Massoulie 1968).

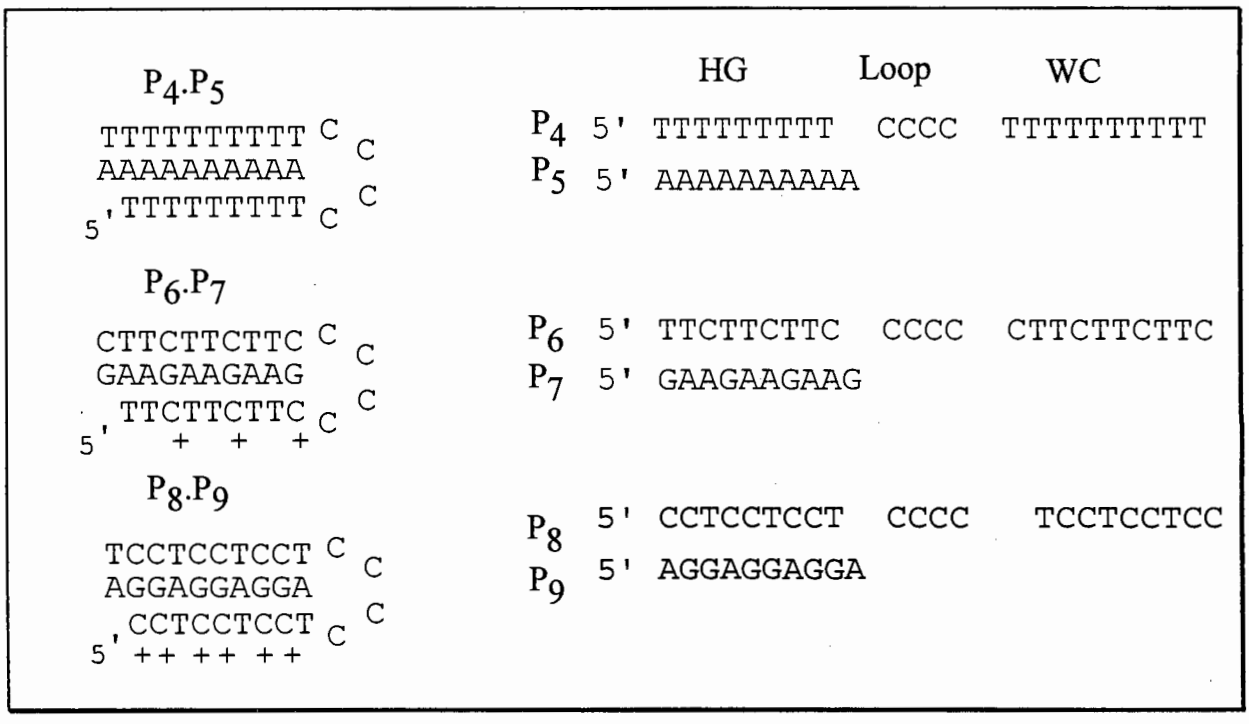

Fig. 5.2 Purine oligonucleotides p5, p7, and p9 and pyrimidine oligonucleotides p4, p6, and p8 were used to form the triple-helices p4p5, p6p7, and p8p9. The Hoogsteen and Watson-Crick bases are abbreviated as HG and WC, respectively.

5.1. Electrophoresis

Complex formation of the triple-helix p6p7 was monitored by gel electrophoresis at pH 5. Fig. 5.3 depicts the results of a gel shift experiment in a continuous variation titration experiment of p6 and p7. As shown, the molar fractional amounts of p6 increases from 0 to 1 in lane 1 to 9, and p7 in lanes 9 to 1. The lower and upper bands correspond to the 10-mer purine (p7) and 23-mer pyrimidine strand (p6), respectively. The band with intermediate mobility corresponds to a 1:1 complex formed between p6 and p7. It is apparent that only one species is formed when the stoichiometry of p6 and p7 is 1:1 (lane 2-4) with no competing structures. The low mobility bands appearing at the top right of the gel arises as a result CC^+ complex formation in p6.

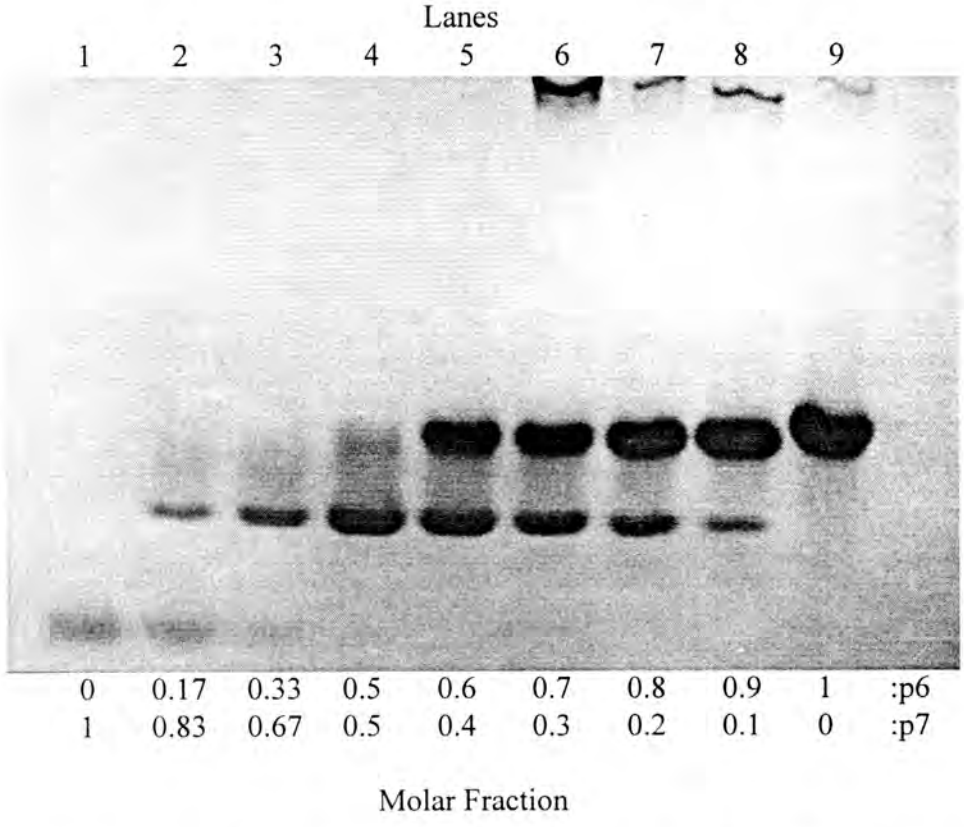

Fig. 5.3 20% native polyacrylamide gel electrophoresis at 20 °C. The values at the bottom of the gel represent the molar fractional amounts of the oligonucleotides p6 and p7, respectively. Solvents contained 100 mM sodium acetate, pH 5.

Native gel electrophoresis in 20 % polyacrylamide (Fig. 5.4) for the WC triple-helical three-way junction was investigated at pH 8. At this pH arms B and C of the WC three-way junction adopt a double-helical conformation (cf. also Fig. 5.8) while arm A remains triple-helical as it contains exclusively TAT triad bases. It is therefore of interest to determine whether the WC junction can form a well defined three-way junction with two double-helices and one triple-helical arm. The bands in lanes 1 to 3 represent the single strands p1, p2, and p3 and lanes 3 to 4 the pairwise combinations of p1p2, p2p3, and p1p3, respectively. The bands at the top of the gel in lanes 4 and 5 represent the binary complexes formed by p1p2 and p2p3, respectively. The difference in mobility between these two complexes can be attributed to p2p3 forming a triplex rather than a duplex. For lane 6 there are two species present, one of which runs with mobility similar to the single strand p3 in lane 3. It is also possible that two p1 strands can hydrogen bond to the HG and WC pyrimidine regions of p3 to form a double-helical ternary complex. The binary complexes dominate in the case of p1p2 and p2p3. The stoichiometric combination of the three single strands (lane 7) exhibit three (to four) bands. The lowest band coincides with the species in lane 3, while the intermediate band migrates with similar mobility to the binary combination in lane 5. The dominant band in lane 7 is assigned to the 1:1:1 complex. It is apparent from the extra bands that the complete junction seems to be in equilibrium with the single strands and at least one of the binary combinations. Conventional double-helical three-way junctions have been characterized and shown to form stable structures under similar conditions (cf. chapter 7; Duckett and Lilley 1990; Guo *et al.* 1990; Lu *et al.* 1991).

Fig 5.4 20 % native polyacrylamide gel electrophoresis at 4 °C. Lanes 1, 2, and 3, single-strand p1, p2, and p3, respectively; lanes 4, 5, and 6, two-strand combinations p1p2, p2p3, and p1p3, respectively. An equimolar mixture of p1p2p3 was loaded in lane 7. Solvents contained 150 mM NaCl, 89 mM Tris, and 89 mM boric acid, pH 8. The various components stain with different efficiency.

5.2. Temperature Gradient Gel Electrophoresis

The thermal unfolding of the individual triple-helical arms and the three-way junction were investigated by temperature gradient gel electrophoresis (Rosenbaum and Riesner 1987). Only the results of p6p7 are shown for the isolated arm B. Fig 5.5 represents a temperature gradient gel of the triple-helix p6p7, with a linear temperature gradient ranging from 24 to 65 °C. The mobility of the complex increases only marginally from 25 to 50 °C. Above 50 °C the triplex unfolds rapidly to produce two new bands with lower and higher mobility. The band with low mobility corresponds to the 23-mer oligonucleotide p6 and the high mobility band to the 10-mer oligonucleotide p7. No intermediate complexes appear to play a role in this process. It is evident from these

results that the triplex unfolds directly to single stands over a very narrow temperature range with no intermediate species.

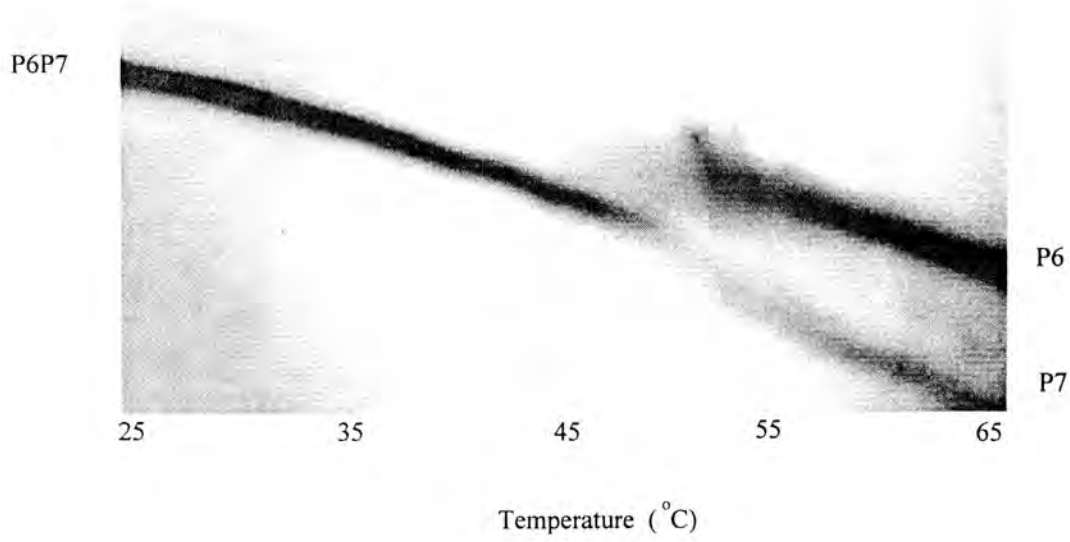

Fig 5.5 Temperature gradient gel electrophoresis of the triplex p6p7 in 100 mM ammonium acetate and 50 mM sodium acetate at about pH 5.5.

To shift the equilibrium of the 1:1:1 structure shown in Fig. 5.1 in favor of the fully folded WC triple-helical three-way junction, the pH was lowered to about 5.5. Fig. 5.6 represents a temperature gradient gel, ranging from 20 to 70 °C. In the temperature range 20 to 35 °C only one band is apparent, the mobility of which increases only marginally in this temperature range. It is evident that only one species is present, consistent with the fully folded WC triple-helical three-way junction (cf. Fig 5.19, F1). From about 40 °C the

mobility starts to increase dramatically. This increase in mobility can be attributed to the thermal unfolding of one of the arms in the junction (F2). Between temperatures 45 and 55 °C an additional band with larger mobility emerges, corresponding to the release of one strand (F4). In the temperature range 60 to 70 °C two closely spaced diffused bands are still apparent, which correspond to the mobility of the single strands F4, F5 and F6 (cf. Fig. 5.4, lanes 1-3).

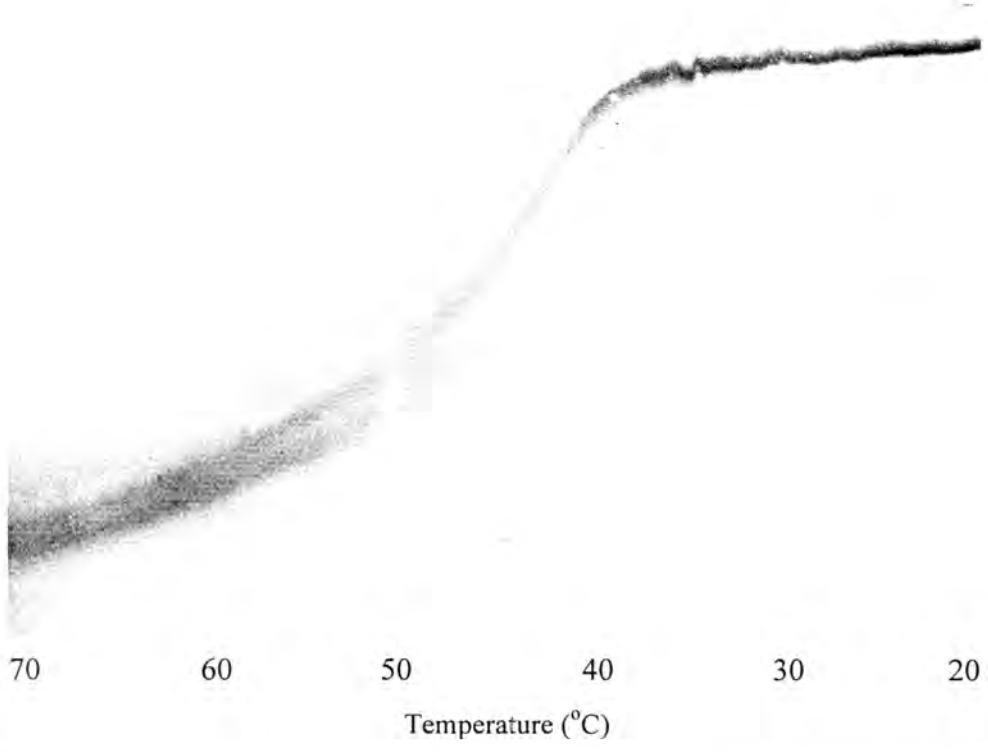

Fig 5.6. Temperature gradient gel electrophoresis of the WC triple-helical three-way junction (p1p2p3) in 100 mM ammonium acetate and 50 mM sodium acetate at about pH 5.5.

5.3. pH Titration

It is generally accepted that the binding of the third strand containing cytosines is largely due to the protonation of those cytosine resulting from CGC^+ triples (Lee *et al.* 1974; Moser and Dervan 1987; Xodo *et al.* 1990; Plum *et al.* 1990; Manzini *et al.* 1990). Lowering the pH of the solution should therefore shift the equilibrium in favor of triplex formation. To investigate the effect of pH on triplex formation in the isolated arms B and C, the absorbance change (260 nm) was measured as a function of pH, ranging from 2.5 to 9.5 at 20 °C. The pH was initially lowered to 2.5 by adding microliter quantities of 1 M HCl to the solution and then increased again stepwise by adding small quantities of 1 M NaOH. The samples were allowed to stand for approximately two minutes after each addition of NaOH before measuring the absorbance at 260 nm. Fig. 5.7 depicts the absorbance vs pH profiles for the two triple-helices p6p7 and p8p9. From these profiles it is apparent that there is hardly any difference in the absorbance vs pH profiles between the two complexes. The absorbance for both complexes remains unchanged between pH 2.5 and 5. This can be attributed to the fully folded triple-helical nature of both complexes. Increasing the pH from 5 to 7 results in a sigmoidal increase in the absorbance profile, corresponding to the unfolding of the triple-helices in favor of the duplex. Between pH 7 and 9 the absorbance remains constant indicating that triple-helices have unfolded completely in favor of the double-helices. These findings are in agreement with previously published results for triple-helices containing CGC^+ triplets (Manzini *et al.* 1990; Hample *et al.* 1991; Xodo *et al.* 1991; Völker *et al.* 1993; Hüsler and Klump 1994;). It should be noted that this type of behavior is not observed in the triplex p4p5 as expected, as it contains only thymine bases in the Hoogsteen strand (data not shown). The apparent pKa values as deduced from Fig. 5.7 for the triple-helices p6p7 and p8p9

are 6.15 and 6.19, respectively. It should be noted that these pKa values do not represent the true pKa of the cytosine bases but rather an apparent pKa as a result of triplex formation.

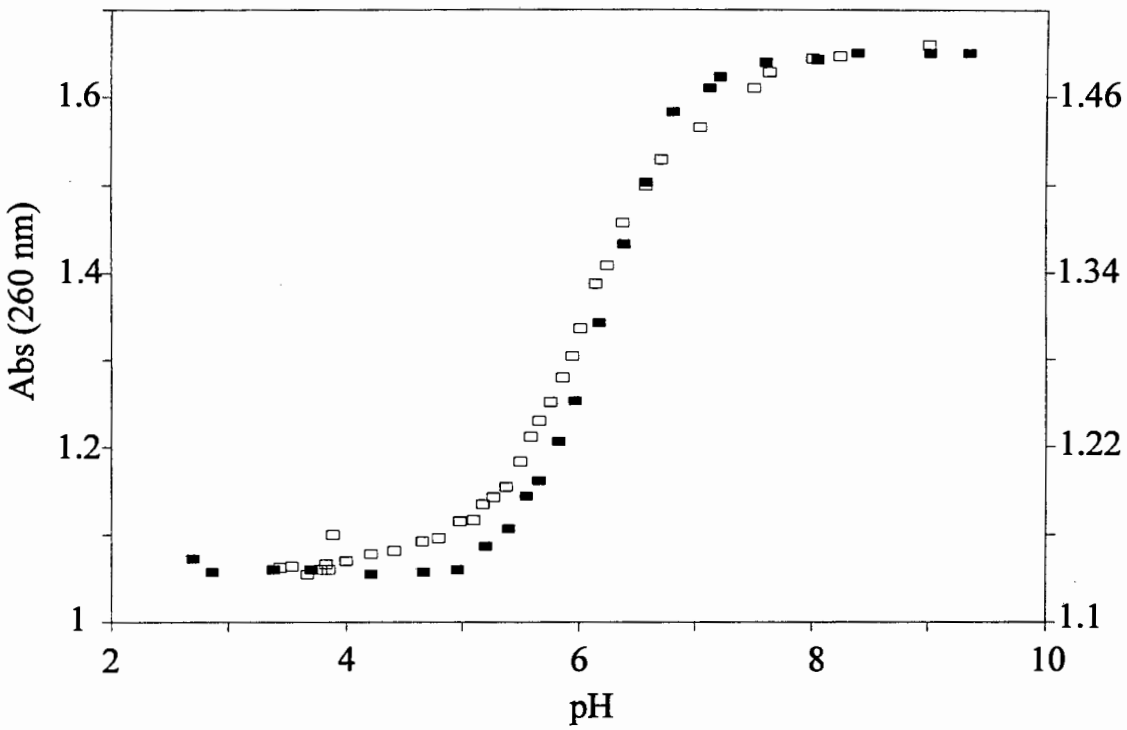

Fig. 5.7 pH titration of p6p7 (□) and p8p9 (■) in 20 mM Na_3PO_4 and 1M NaCl at 20 °C. The absorbance was measured at 260 nm by increasing the pH using microliter quantities of a 1 M NaOH solution.

To investigate the folding properties of the WC three-way junction on lowering the pH, the uv absorbance change at 260 nm (20 °C) was also measured as a function of pH between 2 and 12 (cf. Fig. 5.8). An overall high absorbance was obtained in the pH range 11-12, suggesting that the strands are completely dissociated and single stranded. On lowering the pH from 10.5 to 9.5 the absorbance decreases sharply to an intermediate plateau which declines only marginally up to about pH 7. The predominant species in this

pH interval are presumably a mixture of the partially folded three-way junction with one triple-helical arm and the binary complexes and single strands (cf. Fig. 5.4). Lowering the pH from 7 to 5.5 results in a large sigmoidal hypochromic shift in the absorbance, reaching a shallow trough between pH 5.5 and 3.5. The change in absorbance can be attributed to the formation of the fully folded triple-helical three-way junction. Between pH 3.5 and 2 the absorbance increases sharply again, indicating that a further decrease in the pH results in the complete dissociation of the junction into its single stranded oligonucleotides. It is worth noting that the apparent pKa (≈5.9) thus obtained for the WC three-way junction is only slightly smaller than for the isolated triple-helices p6p7 and p8p9.

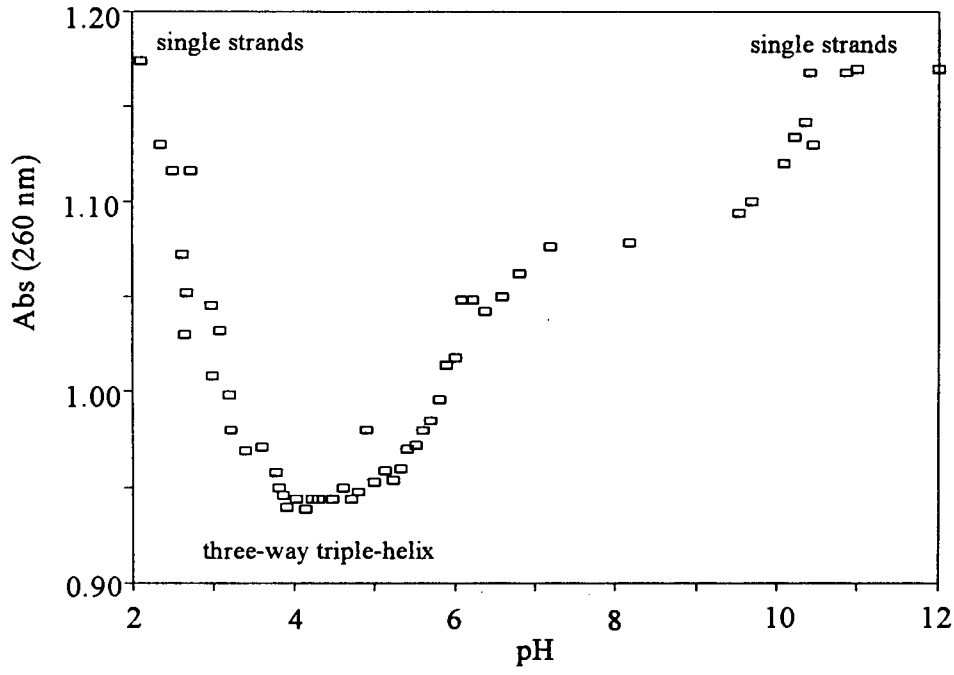

Fig. 5.8 pH titration of p1, p2, and p3 (1:1:1) in 20 mM Na_3PO_4 and 1 M NaCl at 20 °C. The absorbance was measured at 260 nm after adjusting the pH and equilibrating for 15 min

5.4. CD Spectroscopy of the WC Triple-Helical Three-Way Junction

The CD spectrum of the three-way junction was investigated at 20 and 100 °C in 20 mM Na_3PO_4 and 1 M NaCl at pH 4.5 to demonstrate that the arms of the junction are triple-helical (Fig. 5.9 B). A negative band at 215 nm and a red shift of the positive band at 280 nm was previously shown to be indicative of triple-helix formation (Lee *et al.* 1979). The complementary results of X-ray diffraction (Arnott and Selsing 1974; Arnott *et al.* 1976) and CD studies have been used to assign the negative band at 215 nm to DNA in the A-conformation (Gray *et al.* 1978; Gray and Ratliff 1975), assumed to be the preferred conformation adopted by triple-helical DNA. The spectrum at 20 °C, where the WC triple-helical three-way junction is completely formed, displays a negative band between 210 and 225 nm. This result suggests that the triple-helical arms of the junction have indeed formed. At 100 °C the spectrum is typical of single stranded DNA, characterized by the small negative and positive bands at 245 and 278 nm, respectively. As a control, the spectrum of the double-helical three-way junction, which contains the same sub-set of sequences in the arms, is shown in Fig. 5.9A. The spectrum at 20 °C resembles that of B-DNA (Bush 1974) and as expected, is distinctly different from the spectrum obtained for the WC triple-helical three-way junction.

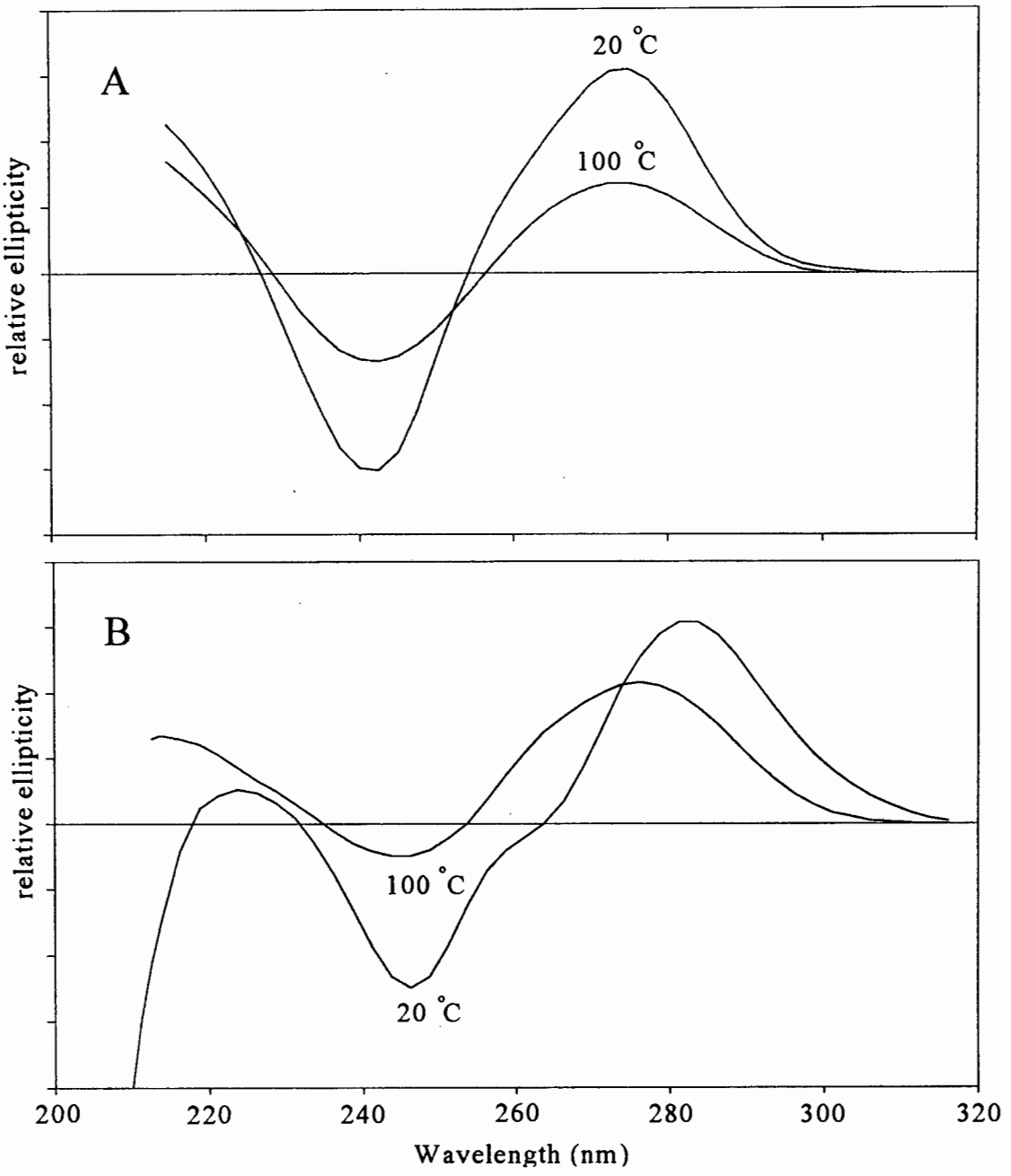

Fig. 5.9 A: CD spectrum of the double-helical three-way junction (q1q2q3) at 20 °C and 100 °C at pH 7. B: CD spectrum of the WC triple-helical three-way junction p1p2p3 at temperatures 100 °C and 20 °C at pH 4.5. Buffer: 20 mM Na_3PO_4 and 1 M NaCl.

5.5. Ultraviolet Thermal Denaturation of the Isolated Arms and the Three-Way Junction

Fig. 5.10 represents the thermal denaturation of the triplex p4p5 in 20 mM Na_3PO_4 and 1M NaCl at pH 7. The complex undergoes a simple monophasic transition in one step from the triple-helix to random coil. The thermal denaturation for the triple-helices p6p7 and p8p9 (at various pH values) are given in Figs. 5.11 and 5.12 (in 1M NaCl and 20 mM Na_3PO_4). Both binary structures undergo simple monophasic thermal denaturation. It should be noted that at pH 5.8 (cf. Fig. 5.12) the triplex p8p9 displays a biphasic transition. The first transition observed in the temperature range 35 to 40 °C is attributed to the unfolding of the Hoogsteen strand. The unfolding of the remaining WC duplex (Manzini *et al.* 1990; Pilch *et al.* 1990a, 1990b; Plum *et al.* 1990; Xodo *et al.* 1991; Völker *et al.* 1993;) takes place between 45 and 60 °C (cf. also Fig. 5.17 C). The two transitions merge completely on lowering the pH by 0.25 units. The resulting monophasic melting curve is attributed to the transition from triplex to random coil (cf. Fig. 5.12, pH 5.55). It is thus evident that the thermal stability of triple-helices representing arms B and C, is pH sensitive.

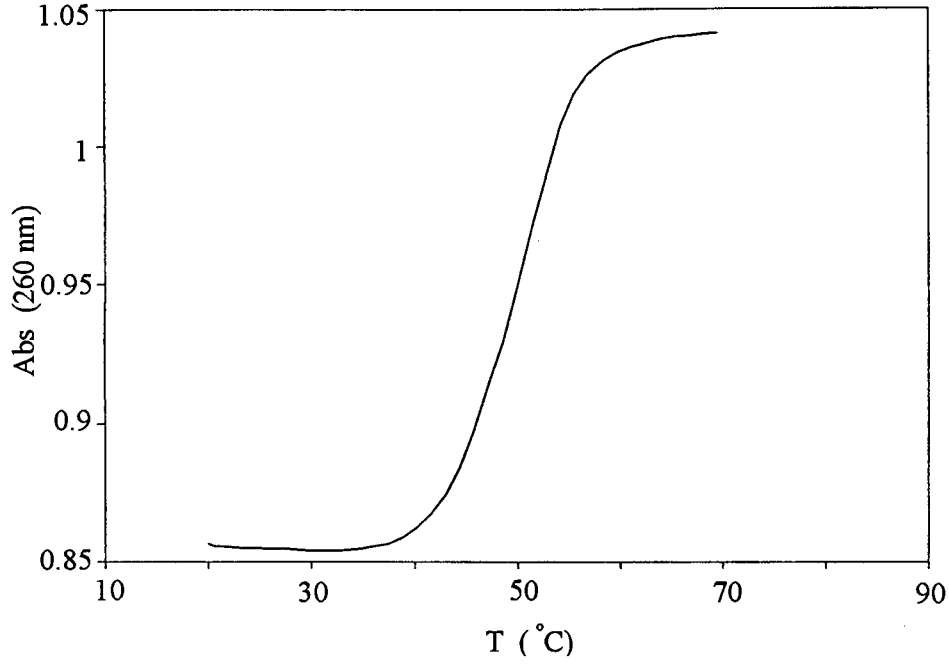

Fig 5.10 Ultraviolet thermal denaturation (260 nm) of the triplex p4p5, recorded at pH 7 in 20 mM Na_3PO_4 and 1 M NaCl.

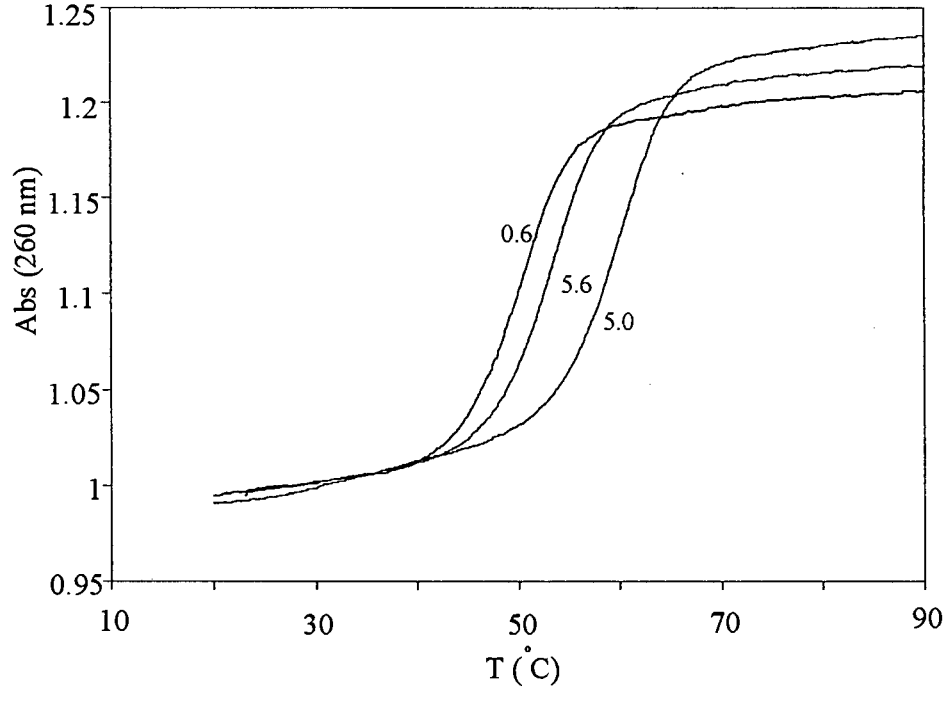

Fig. 5.11 Ultraviolet thermal denaturation (260 nm) curves for the triplex p6p7 recorded at pH 6.0, 5.6, and 5. Buffers contained 20 mM Na_3PO_4 and 1 M NaCl.

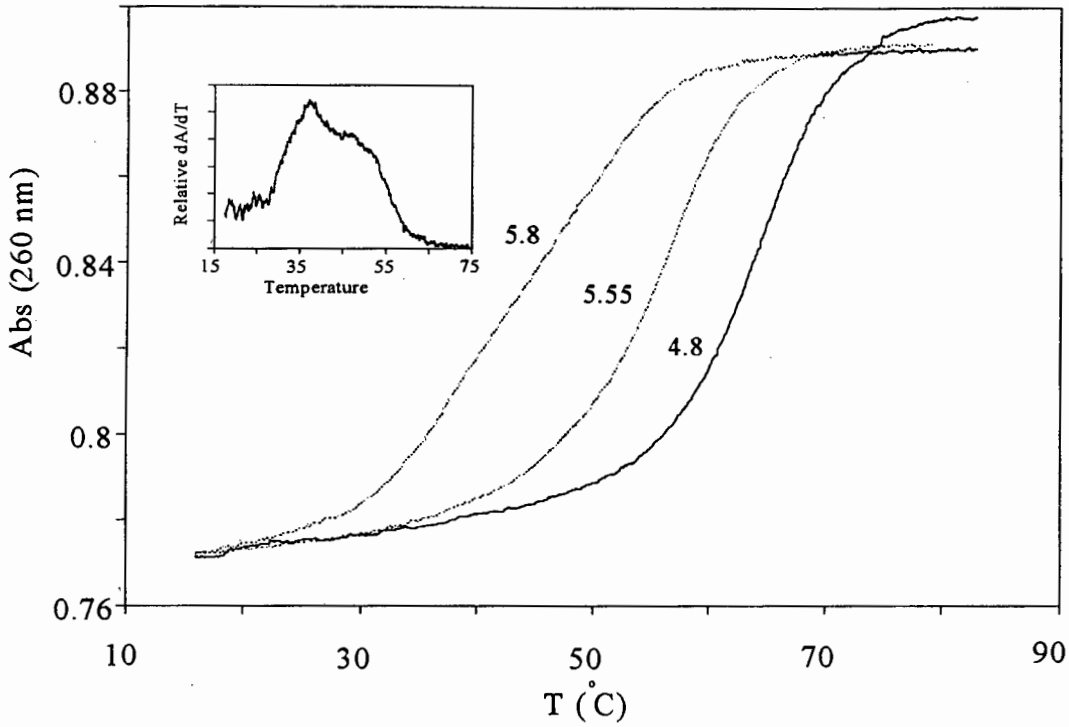

Fig. 5.12 Ultraviolet thermal denaturation (260 nm) curves for the triplex p8p9 recorded at pH 5.8, 5.55, and 4.8. Buffers contained 20 mM Na_3PO_4 and 1 M NaCl. The insert represents the first derivative of the absorbance vs temperature curve at pH 5.5.

The unfolding of the triple-helical three-way junction was also investigated at various pH values in 100 mM NaCl, 20 mM Na_3PO_4 to determine how the arms respond to pH changes once they have been integrated into a three-way junction. Fig 5.13 depicts a selection of five such derivative melting curves ranging from pH 4.5 to 6.5. Three transitions are apparent at pH 4.5. Each of these transitions arises as a result of the independent thermal unfolding of the arm A, B, and C (cf. also Fig. 5.19). The low temperature transition corresponds to the unfolding of arm A and the two upper transitions to the consecutive dissociation of the arms B and C, respectively. On increasing the pH from 5 to 5.25 the upper temperature peak (arm C) merges into the central transition peak (arm B). Increasing the pH further to 6 and then 6.5 results in the complete

amalgamation of all three transitional peaks. From this series of melting experiments it is apparent that the WC triple-helical three-way junction undergoes thermal unfolding in a multi-state manner. The pH dependence of the two peaks associated with arms B and C are attributed to the CGC^+ triad composition of these arms. It is therefore evident that the thermal unfolding properties of arms B and C can be manipulated to some degree by changing the pH of the solution.

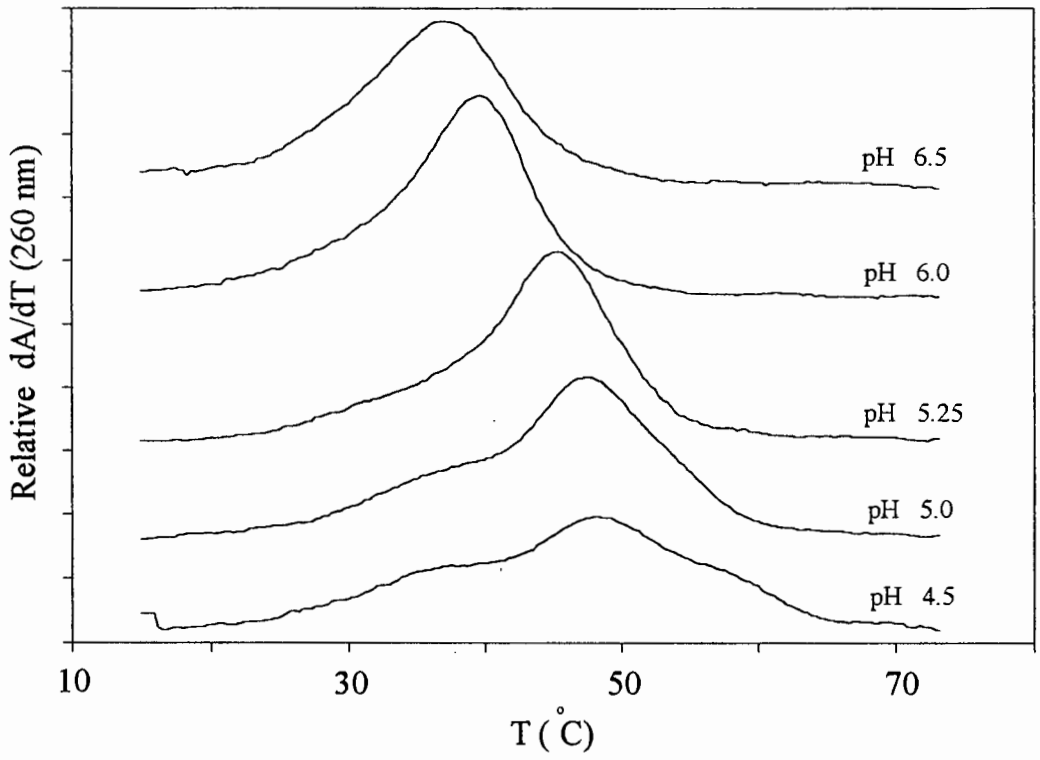

Fig 5.13 Differential uv (260 nm) melting curves for the triple-helical three-way junction (p1p2p3) as recorded under various pH conditions. Buffers contained 20 mM Na_3PO_4 and 100 mM NaCl.

5.6. pH Dependence of the Triple-Helical Arms

To investigate further the influence of pH on the thermal stability of the isolated triple-helical arms, an extensive series of thermal denaturation curves were obtained (260 nm) under various pH conditions, where only the combined transition was observed. Fig. 5.14 depicts the way in which 1/Tm, for triple-helices p6p7 (open squares) and p8p9 (solid squares), varies with pH (1M NaCl and 20 mM Na_3PO_4). In the pH range 2.8 to 3.8, 1/Tm drops off rapidly, reaching a minimum between pH 3.8 and 4.8 respectively, i.e. the pH at which the Tm is maximal. Between pH 4.8 and 6, 1/Tm increases linearly for both triple-helices (cf. also Table II). The slopes of 1/Tm vs. pH are, however, different for the two triple-helices. Similar results have been reported in the literature for the Tm dependence on pH in triple-helices containing CGC^+ triplets (Plum *et al.* 1990; Maher *et al.* 1990; Xodo *et al.* 1991; Völker *et al.* 1993). Also, it is apparent from Fig. 5.14 that the pH, where 1/Tm reaches a minimum, is lower for the triplex p8p9 than for p6p7. This can be attributed to the crowding of positively charged cytosines in the Hoogsteen strand of the triplex p8p9, as compared to the triplex p6p7, where the cytosines are evenly spaced by two thymine residues, making them independent of each other.

The formalism presented in chapter 3 section 3.5 can be used to quantify Tm dependence on pH as exemplified by the sequence of triple-helices p6p7 and p8p9. It was shown from equation [5.27b] that the slope of 1/Tm vs pH is proportional to the number of protonated cytosines (n) in the Hoogsteen strand, and inversely proportional to the formation enthalpy (ΔH).

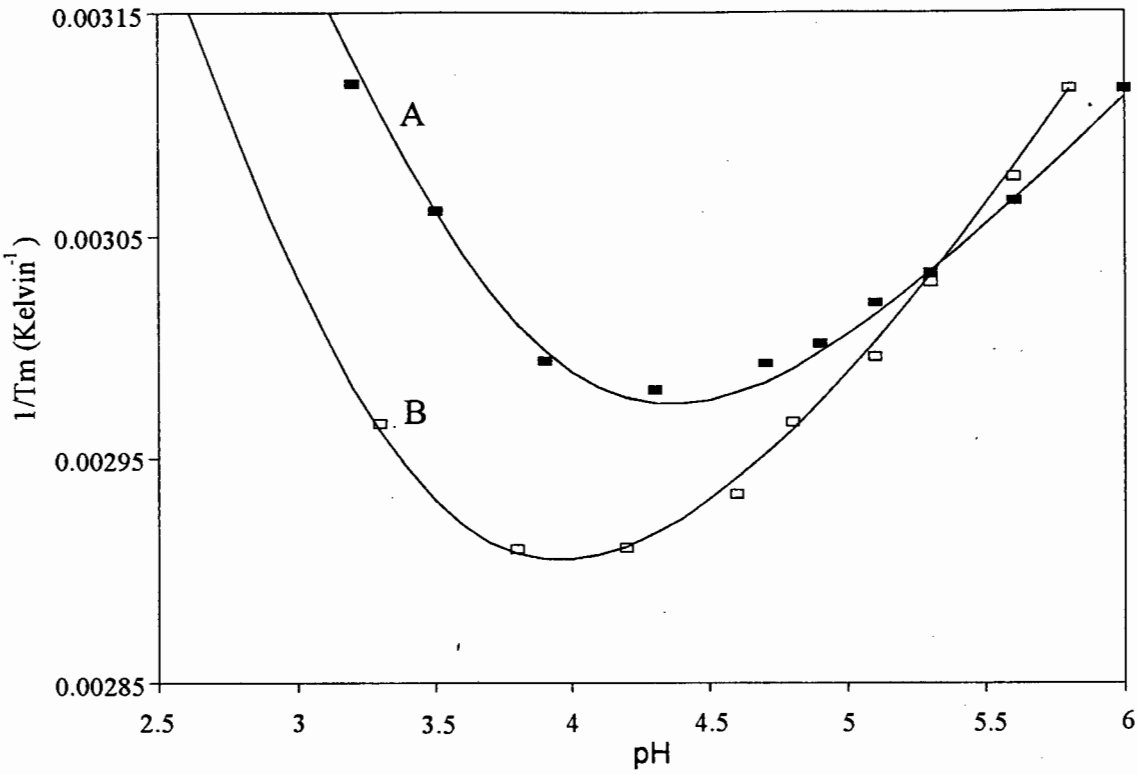

Fig. 5.14 Phase diagram representing 1/Tm (Kelvin^{-1}) vs pH for the triple-helices p6p7 ( ) and p8p9 (□). Tm values were determined from uv (260 nm) thermal denaturation curves under various pH conditions. The solid line represent the theoretical curves for the model of nearest neighbor proton binding sites. All buffers contained 20 mM Na_3PO4 and 1 M NaCl.

As the formation enthalpy (ΔH) for both complexes is similar (cf. Table I) the slope for 1/Tm vs pH will primarily reflect the fraction of protonated Hoogsteen cytosines (n) in the triplex. The slope for 1/Tm vs pH in the range pH 4.7-6 for the triple-helix p6p7 (which contains the smallest number of Hoogsteen cytosines) is smaller than the slope obtained for the triplex p8p9 (which contains the larger number of Hoogsteen cytosines). These findings are, furthermore, comparable with results obtained for the triple-helical three-way junction, which contained the triple-helices p6p7 and p8p9 as integrated arms (cf. Fig 5.15).

The nature of the integrated arms of the junction was also investigated in the pH range 4 to 5.5, where the different transitions are still discernable from one another. Fig. 5.15 shows the results for a series of uv melting experiments where 1/Tmax vs pH have been plotted for each of the three transitions. As shown in the literature (Massoulie 1968), and here for the triplex p4p5, the thermal stability of TAT triple-helices are unaffected by pH. The low-temperature transition which remains constant in the entire pH range can therefore be assigned to arm A. In contrast, the other two transitions are pH dependent, with the high-temperature transition displaying the largest dependence (largest slope), and the middle transition an intermediate slope. As shown for the individual arms p6p7 and p8p9 the dependence of Tm vs pH increases with increasing CGC^+ content. It is therefore possible to assign arms B and C to the high temperature transitions according to their CGC^+ composition and 1/Tm vs pH dependence. In this way arm C, which contains the largest number of CGC^+ triads bases, was assigned to the high temperature transition and arm B to the intermediate-temperature transition.

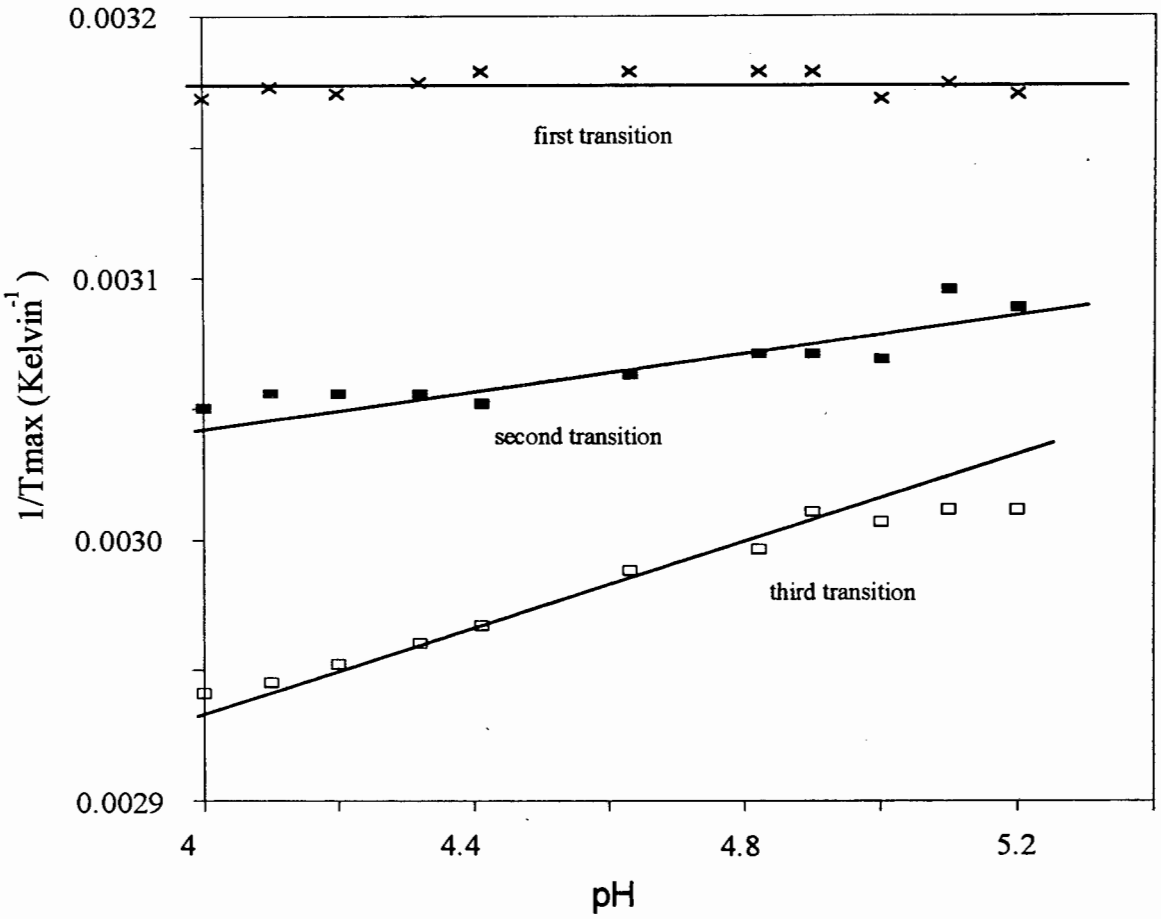

Fig. 5.15 Phase diagram representing 1/Tmax vs pH obtained from differential uv melting curves (260 nm) for the triple-helical three-way junction. Buffers contained 20 mM Na_3PO_4 and 500 mM NaCl. Tmax was determined where *dA/dT* reaches a maximum.

5.7. Ionic Strength Dependence of the Triple-Helical Arms

To substantiate the assignments of the three transitions to the arms, the thermal unfolding of the three-way junction was investigated as a function of sodium ion concentration. The pH of the solution was specifically chosen at 4.7 to ensure that the transitions remain sufficiently separated. From Fig. 5.16 it is apparent that there is a linear relationship between Tmax and log[Na^+] for each of the three transitions. It is also evident

that the Tmax dependence on log[Na^+] is different for each of the transitions. The differences in slope can readily be explained in terms of the changes in the charge density between the ordered and disordered state of each arm (Record 1967, 1978; Manning 1978; Anderson and Record 1982). As a result of the large negative charge associated with the three backbone strands in each of the arms, and the positive charge associated with the Hoogsteen cytosines in arms B and C, the overall charge density of each arm will be different. As arm A contains exclusively TAT triplets, it has the largest charge density. Arms B and C exhibit the intermediate and smallest overall charges respectively, as a result of the increasing number of positively charged CGC^+ triplets associated with each of them. Such differences in charge density are clearly reflected in the experimental results shown in Fig. 5.16. Arm C typically contains the largest number of CGC^+ triads and has consequently the highest thermal stability, but smallest Tmax dependence on sodium ion concentration (dTmax/d log[Na^+] = 1.8). Arm B contains an intermediate number of CGC^+ triads and thus has an intermediate thermal stability, but comparatively larger Tmax dependence on changes in sodium ion concentration (dTmax/d log[Na^+] = 8.3). Arm A, which exhibits the largest change in charge density upon melting, also has the largest slope (dTmax/d log[Na^+] = 11.9). Similar results have been obtained by others for polynucleotide triple-helices such as poly(UAU), with a slope dTm/$d$log[Na^+] of 26 °C and a moderate slope of 10 °C for poly(CGC^+) (Thiele *et al.* 1972; Blake *et al.* 1967). The comparatively moderate slopes obtained for the triple-helical arms compared to polymeric triple-helices can readily be explained by the end effects in the arms of the junction. The over-all tendency is, however, the same for both systems where the slopes, dTmax/d log[Na^+], for the arms A, B, and C, show a trade-off between the Na^+ ion shielding of the

triple-helical arms and the charge reduction associated with positively charged CGC^+ triads in arms B and C.

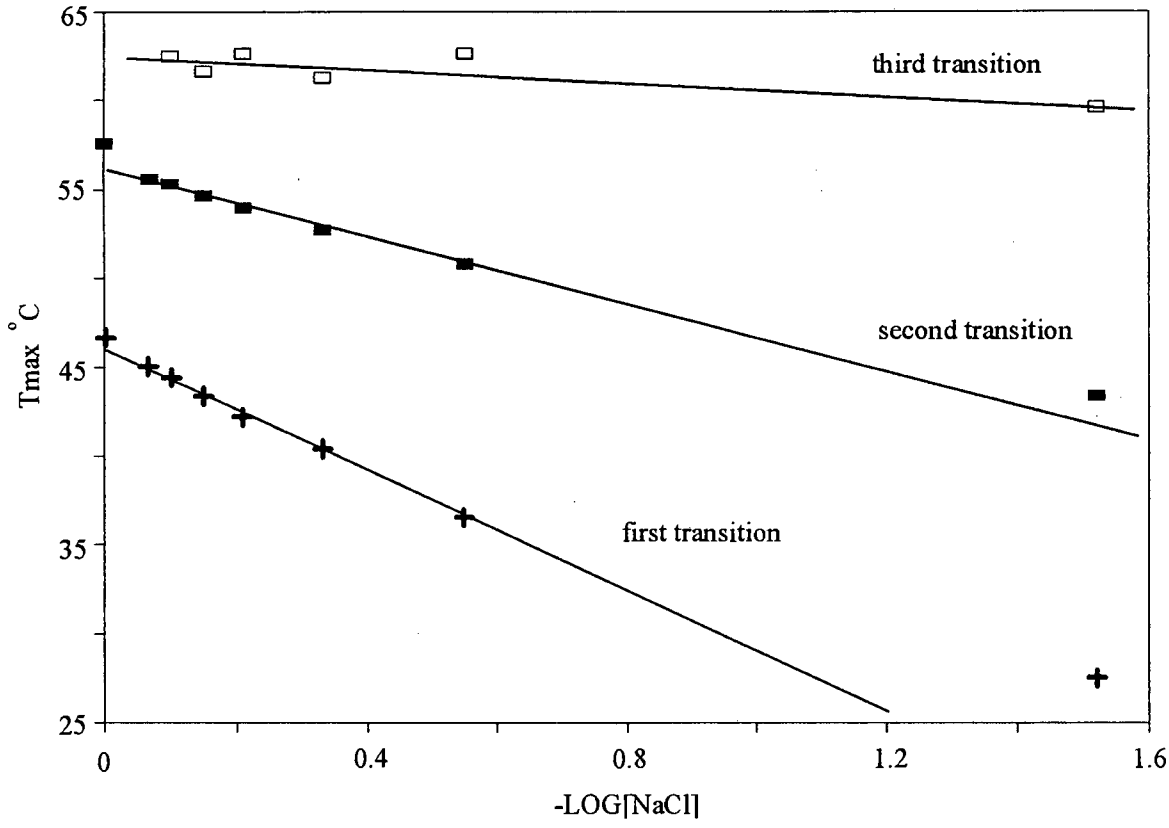

Fig. 5.16 Phase diagram representing Tmax vs Log[Na+] from differential uv melting curves (260 nm) for the three-way junction. Buffers contained 20 mM Na_3PO_4, pH 4.7.

5.8. Thermodynamics of the Isolated Arms p4p5, p6p7, and p8p9

The formation enthalpy (ΔH) for the triple-helices p4p5, p6p7 and p8p9 was determined directly by differential scanning calorimetry (cf. Table I). In the case of triplex p4p5 and p6p7, simple triplex to coil transitions were observed at pH 7 and 6, respectively (cf. Fig. 5.17 A and B). However, for the triple-helix p8p9, a biphasic

transition was observed at pH 5.3, with only one at pH 4.8 (Fig. 5.17 A, B and C). The first step (pH 5.3) in Fig. 6C coincides with the detachment of the Hoogsteen strand and the second step to the thermal unfolding of the remaining WC duplex. The transition as observed by DSC occurs at a higher temperatures than the corresponding transition observed by uv thermal denaturation. This increase in Tm is due to the large oligonucleotide concentration required for DSC. The unfolding of the Hoogsteen strand (first transition) will proceed through an intramolecular process, and hence mass action does not play a role in the triplex to duplex transition (Marky and Breaslauer 1987). The duplex to coil transition proceeds through a bimolecular process. It is therefore apparent that the Tm of the second transition (melting of the duplex) will be dependent on the oligonucleotide concentration. Larger oligonucleotide concentrations used in DSC shifts the duplex to coil transition to higher temperatures, resulting in a more pronounced separation of the two transitions. The enthalpy changes for the unfolding of the triple-helices p4p5, p6p7 and p8p9 were determined directly by integration of the DSC peaks. The results are as listed in Table I. The enthalpy values obtained are in accordance with the published data of Xodo *et al.* (1990) for the unfolding of a similar looped triplex containing four Hoogsteen cytosines (129 kcal/mol). There is a general increase in the overall enthalpy changes associated with the triple-helices as the number of cytosines increase. Differences in the enthalpy changes between the triple-helices p6p7 and p8p9 are, however, smaller than for the triple-helices p4p5 and p6p7. This may suggest that neighboring cytosines (p8p9) are less favorable than cytosines separated by thymine bases (p6p7) as a consequence of opposing electrostatic positive charges in the Hoogsteen strand.

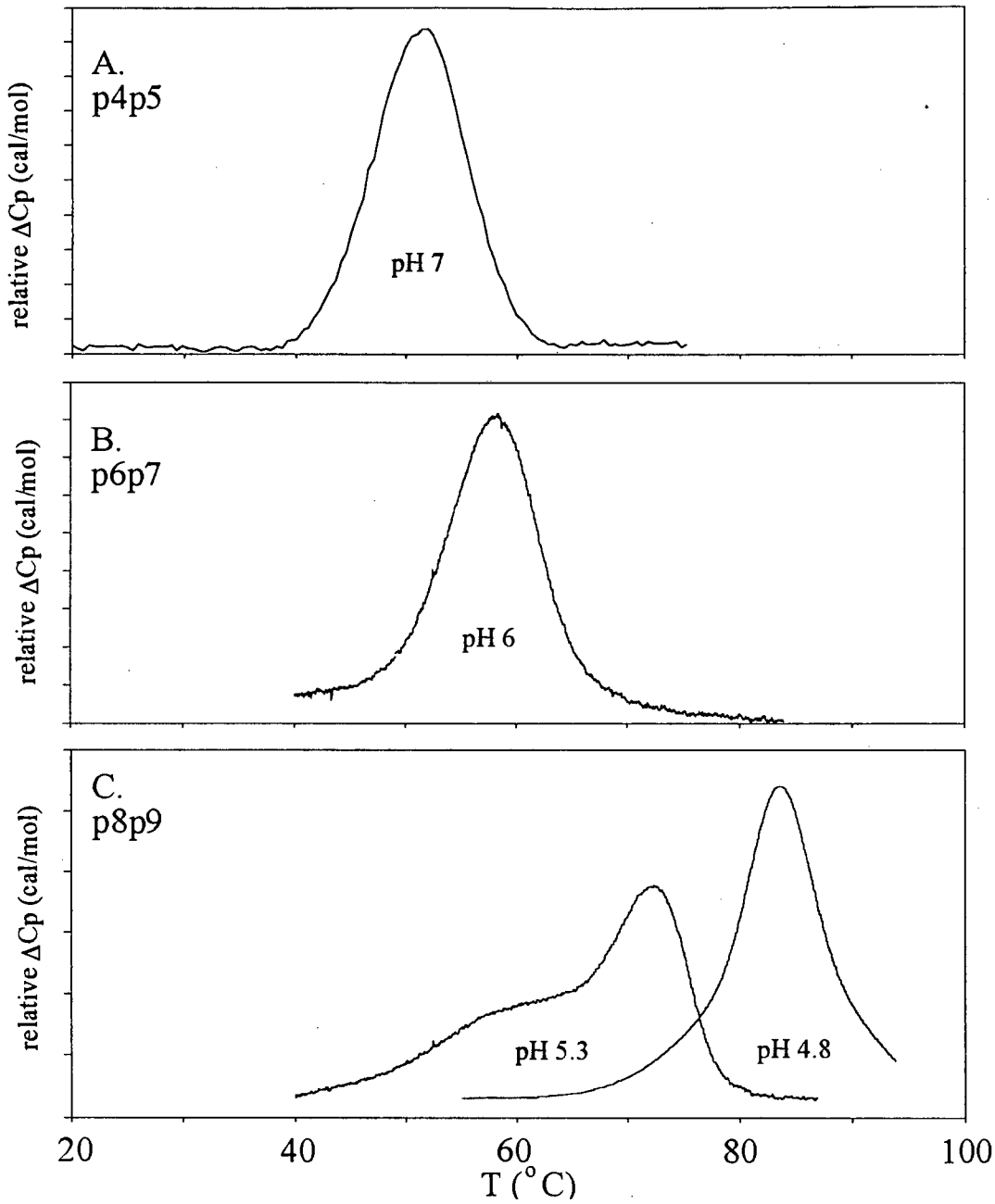

Fig. 5.17 Excess heat capacity (ΔCp) vs temperature for the triplexes p4p5 (A), p6p7 (B), and p8p9 (C). The buffers contained 20 mM Na_3PO_4 and 1M NaCl at the pH values indicated in the figure. The oligonucleotide concentration (purine + pyrimidine) were 0.515, 0.236, and 222 μM for the triple-helices p4p5, p6p7, and p8p9, respectively.

Table I.

Thermodynamic Parameters for the Unfolding of the Triple-Helices p4p5, p6p7, and p8p9.

Triplex	ΔH^a (kcal/mol)	ΔS^b (e.u.)	Tm (°C)	$\frac{\partial(1/Tm)}{\partial pH}10^{-5}$	Predicted n value	Integer value for n	Number of Hoogsteen cytosines
p4p5[e]	-97.0	-380.0	51.3	0	0	0	0
p6p7[c]	-128.7 ± 5.8	-397.0 ± 13	58.2	9.71 ± 0.5	2.7 ± 0.14	3	3
p8p9[d]	-135.9 ± 5.9	-399.4 ± 24	60.0	16.3 ± 1.0	4.8 ± 0.30	5	6

Thermodynamic parameters for the unfolding of the triple-helices p4p5, p6p7 and p8p9 are obtained by differential scanning calorimetry and the number of protonated Hoogsteen sites n was deduced from the slope.

Note. Buffer: 20 mM Na_3PO4, 1M NaCl.

[a] Units for ΔH are $kcal.mol^{-1}$ (1 cal = 4.18J).

[b] Units for ΔS are $cal.mol^{-1}.K^{-1}$

[c] pH 6

[d] pH 5.3

[e] pH 7

5.9. Determination of the Number of Protonated Cytosines in Triple-Helices p6p7 and p8p9

The number of protonated sites in the triple-helices p6p7 and p8p9 can be determined from the formalism developed in section 3.5, for the limiting case where the pH is larger than the pKa of cytosine. It can be shown by differentiating equation [5.34b] that the number of protonated sites is directly related to:

$$n = \frac{\partial(1/Tm)}{\partial(pH)}\left(\frac{-\Delta H}{R\ln(10)}\right), \qquad [5.34b]$$

where n is the total number of protonated cytosines in the triplex. Physically, n can only be represented by integers and the calculated values should be adjusted accordingly to the nearest integer. ΔH is the formation enthalpy change associated with the folding of the triplex determined directly by DSC.

The slope (cf. Table I) was obtained in the pH range 4.7 to 6 for the limiting case where 1/Tm vs pH is linear (cf. Fig 5.14). By substituting ΔH (obtained by DSC) and the slope (obtained from uv thermal denaturation) into equation [5.34b], n can be calculated directly. For the triplex p6p7 the number of protonated cytosines matches precisely the number of Hoogsteen cytosines. In the case of the triple-helix p8p9, only five of the six Hoogsteen cytosines appear to be protonated. The position of this unprotonated Hoogsteen cytosine cannot be deduced directly from the formalism described in this thesis. It is, however, most likely that the terminal 3' cytosine in the pyrimidine strand (p8) of the triplex p8p9, is partially unstacked and uncharged as a result of end effects. Hence its contribution to the overall stability of the triplex will be comparatively small. There is a striking similarity to the results obtained by Xodo *et al.* (1991) for a triplex where three of six Hoogsteen cytosines were protonated.

5.10. Nearest Neighbor Proton Binding Effects in the Isolated Arms p6p7 and p8p9

To determine the extent of nearest neighbor interaction in the triple-helices p6p7 and p8p9, Tm vs pH curves were fitted assuming either interacting proton binding sites or independent proton binding sites. In the case of independent proton binding sites; equations [5.1a] and [5.18b] were used to fit the adjustable variables pKa, ΔH and ΔS.

$$\alpha_n = \left(\frac{[H^+]}{K_a}\right)^n \left(\frac{K_a}{[H^+]+K_a}\right)^N, \qquad [5.1a]$$

$$T_m = \frac{\Delta H}{R\ln(\alpha_n) + \Delta S - R\ln(C_t)} \qquad [5.18b]$$

The constants n and N (from equation [5.1b]) were set to 3 and 11 respectively for the triplex p6p7, and 5 and 16 for the triplex p8p9. N represents the total number of cytosines in any portion of the sequence as a potential proton binding site and n the number of protonated sites in the Hoogsteen strand. The variables were adjusted using nonlinear regression to obtain the best fitting curve (cf. methods, chapter 4). The results are shown in Table II. The same numerical procedure was used for the model which is based on interacting proton binding sites. In the case of the triplex p6p7, equations [5.14b] and [5.18b] were used and for the triplex p8p9, equations [5.15b] and [5.18b] are applicable.

$$\alpha_{p6} = \frac{\xi_{p6}}{\mathbf{e}\left(\mathbf{W}_{TC}\right)^4\left(\mathbf{W}_{CC}\right)^5\left(\mathbf{W}_{TC}\right)^2\mathbf{e}^+} \qquad [5.14b]$$

$$\alpha_{p8} = \frac{\xi_{p8}}{\mathbf{e}\left(\mathbf{W}_{TC}\mathbf{W}_{CC}\right)^4\left(\mathbf{W}_{CC}\right)^2\left(\mathbf{W}_{TC}\mathbf{W}_{CC}\right)^3\mathbf{e}^+} \qquad [5.15b]$$

$$T_m = \frac{\Delta H}{R\ln(\alpha_i) + \Delta S - R\ln(C_t)}, \qquad [5.18b]$$

where the matrices $\mathbf{W}_{TC}$ and $\mathbf{W}_{CC}$ represent the matrices required to account for nearest neighbor protonation interactions in the triple-helices p6p7 and p8p9 (cf. chapter 3.5). The results obtained for interacting sites are also listed in Table II. The quality of the fit for interacting proton binding sites can be judged from Fig. 5.14 (solid lines).

Table II.

Thermodynamic Parameters for the Unfolding of the Triple-Helices p6p7 and p8p9

	Correlation of 2 (interacting sites)				Correlation of 1 (independent sites)		
Triplex	ΔH^a (kcal/mol)	ΔS^b (e.u.)	pKa_1	pka_2	ΔH^a (kcal/mol)	ΔS^b (e.u.)	pKa
p6p7	122.5	348.8	3.88	3.90	119.4	339.6.0	3.91
p8p9	135.3	368.0	3.79	3.36	144.6	397.1	3.63

Thermodynamic parameters for the unfolding of the triple-helices p6p7 and p8p9 were calculated from equation 2.18, 2.14 and 2.15 respectively. For the assumed model of independent proton binding sites, equations 1 and 2.18 were used. The Levenberg-Marquardt algorithm of least squares (Press *et al.*) was used to fit the variables ΔH, ΔS`, pKa_1, pKa_2 and pKa to the data set depicted in Fig. 5.14.

Note. Buffer: 20 mM Na_3PO_4, 1M NaCl. The errors margins for ΔH and ΔS were not greater than 4%. Error margins for pKa, pKa_1 and pKa_2 were less than 0.5%.
[a] Units for ΔH are $kcal.mol^{-1}$ (1 cal = 4.18J)
[b] Units for ΔS are $cal.mol^{-1}.K^{-1}$

The formation enthalpy for triple-helices p6p7 and p8p9 from the numerical fitting procedure, as listed in Table II, are in good agreement with the results obtained by calorimetry (cf. Table I). The formation enthalpy obtained for the triplex p6p7 is insensitive to the model, as the cytosines are sufficiently spaced. There is, however, a discrepancy between the formation enthalpies for p8p9 assuming either interacting or independent proton binding models. This can be attributed to the large number of positively charged nearest neighbor cytosines in the triplex p8p9.

The extent of nearest neighbor effects in the triplex p6p7 and p8p9 can be correlated with pKa_1 and pKa_2. The pKa values obtained for the triplex p6p7, assuming either

independent or interacting proton binding sites, are approximately the same with an average pKa value of approximately 3.9. This is consistent with the arrangement of cytosines in p6, where the cytosines are separated by two thymine bases. Only in the loop are there four consecutively arranged cytosines.

In the case of the triplex p8p9, pKa_1 and pKa_2 differ considerably. The observed difference between pKa_1 and pKa_2 confirms our previous assumption that interacting charged neighbors will lower the pKa of cytosine. The pKa for a cytosine with a positively charged neighbor is approximately 0.5 pH units smaller than the pKa of a cytosine with uncharged neighbors (pKa_2 =3.36). It should be noted that a similar analysis of the integrated arms in the three-way junction is not possible as the latter transitions are always partially overlapping. In other words, the unfolding of a neighboring arm will invoke some influencing factor on the thermal unfolding of the arms. The results for Tm vs pH curves for the triple-helical three-way junction are therefore only treated qualitatively. A further matter that may complicate the quantitative analysis of proton binding in the three-way junction is the fact that the arms seem to undergo partial unfolding around the branch point (see section 5.12).

5.11. Thermodynamics of Unfolding in the WC Triple-Helical Three-Way Junction

To obtain reliable data on the thermal unfolding of the three-way junction, DSC experiments were performed at 1M NaCl, 20 mM Na_3PO_4, at either pH 5 or pH 4.5. The curves in Fig. 5.18 A and B represent the excess heat capacity associated with the stepwise thermal unfolding of the three-way junction into the single strands. The enthalpy changes were obtained from the area under these curves. In this way it is possible to obtain a model independent estimate of the enthalpy and entropy values associated with the unfolding of the three-way junction. Unfortunately, it is not possible to extract the contributions for the individual transitions unambiguously as they overlap. For this reason it was necessary to use a thermodynamic model to extract the exact contributions of each arm.

The proposed pathway of unfolding is shown in Fig. 5.19 (cf. also chapter 3.3). The first transition corresponds to the unfolding of arm A in the triple-helical three-way junction (F1), resulting in the intermediate F2. F2 corresponds to a structure with two connected triple-helical arms (B and C) and two partially unfolded single strands.

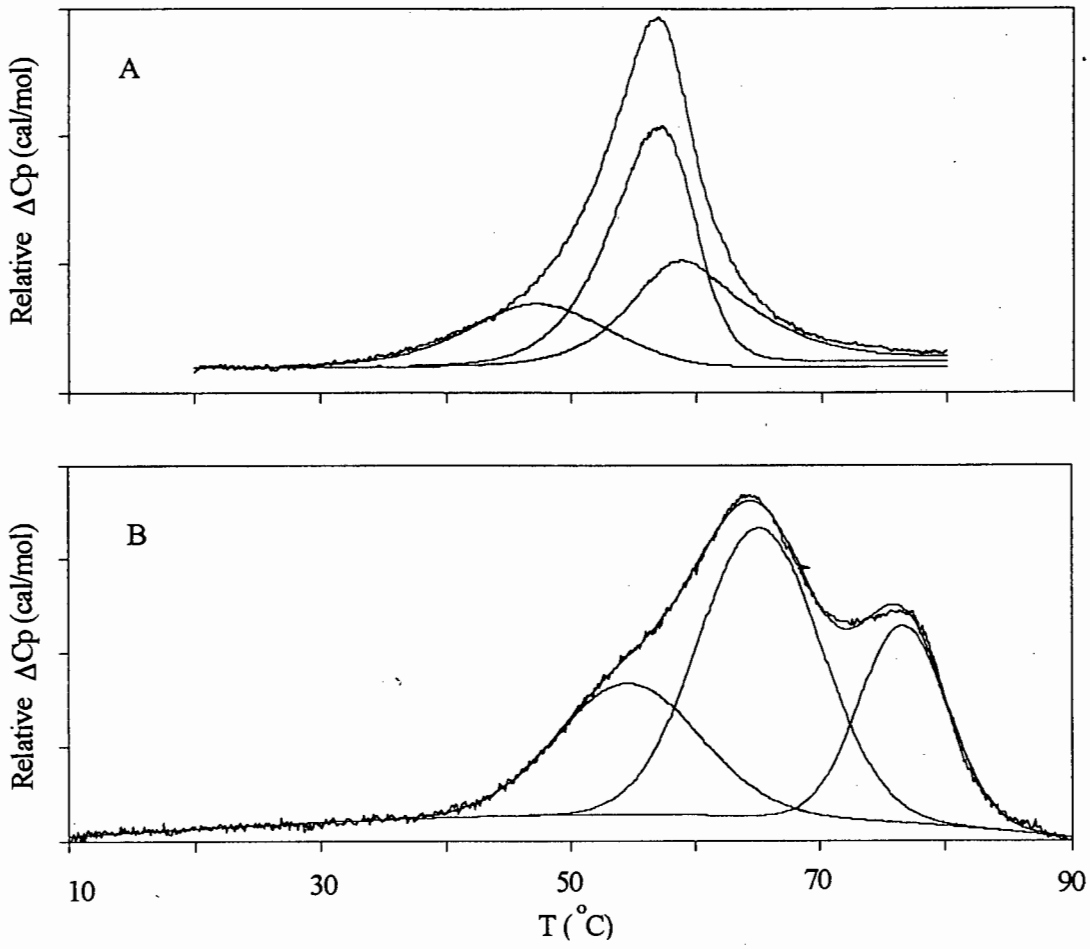

Fig. 5.18. Excess heat capacity (ΔCp) vs temperature curves for the denaturation for the three-way junction and deconvoluted enthalpy contributions of arms A, B, and C at (A) pH 5 and (B) pH 4.5. Buffers contained 20 mM Na_3PO_4 and 1 M NaCl. The theoretical curve in A is left out as it is superimposed completely on the experimental curve.

This assumption is supported by temperature gradient gel electrophoresis (cf. Figs. 5.6) and by the results of the concentration independence of the first transition as deduced from uv thermal denaturation curves (dA_{260}/dT) depicted in Fig. 5.20. The concentration independence of the first transition is also consistent with an intramolecular unfolding process. The species F2 can undergo unfolding of arm B at intermediate temperatures. This process will be accompanied by the release of one of the oligonucleotide strands (F4)

and the formation of arm C (F3). The high-temperature transition is attributed to the thermal unfolding of arm C, resulting in the formation of the two single strands F5 and F6, respectively. The intermolecular nature of the last two transitions is clearly manifested by the oligonucleotide concentration dependence of the latter two transitions (cf. Fig. 5.20). This can be formulated more precisely by considering the equilibrium constants as defined by equations [3.7], [3.8], and [3.9]. The equilibrium constant K_1 in Eq. [3.7] is a conventional equilibrium constant which depends solely on the Gibbs free energy, while K_2 and K_3 (Eqs. [3.8] and [3.9]) are dependent on the Gibbs free energy and on the total strand concentration Ct. To calculate the transition temperatures it is necessary that the subtransitions be partially resolved. It is evident that the Tm for the first transition is concentration independent as described by equation [3.7], where the $Tm = \Delta H/\Delta S$. The later two transitions described by Eqs. [3.8] and [3.9] are dependent on Ct as can be deduced from $Tm = \Delta H/[\Delta S - R\ln(Ct)]$. From model studies it can be shown that the last two transitions are dependent on the strand concentration Ct, even if they are not completely separated. The initial peak (Fig. 5.20B) has a pronounced shoulder on the low-temperature side which corresponds to the low temperature transition, while the two high temperature transitions are partially superimposed under the main peak. At a 10-fold higher strand concentration a more detailed transition curve is observed (Fig. 5.20A). The low temperature shoulder remains unchanged (transition 1). The main peak (transition 2) shifts to a slightly higher temperature, and the new shoulder (transition 3) appears at the higher temperature side of the main peak. The slight upward shift of the transition is in good agreement with the theory as outlined in section 3.3, discussing the thermodynamic solutions.

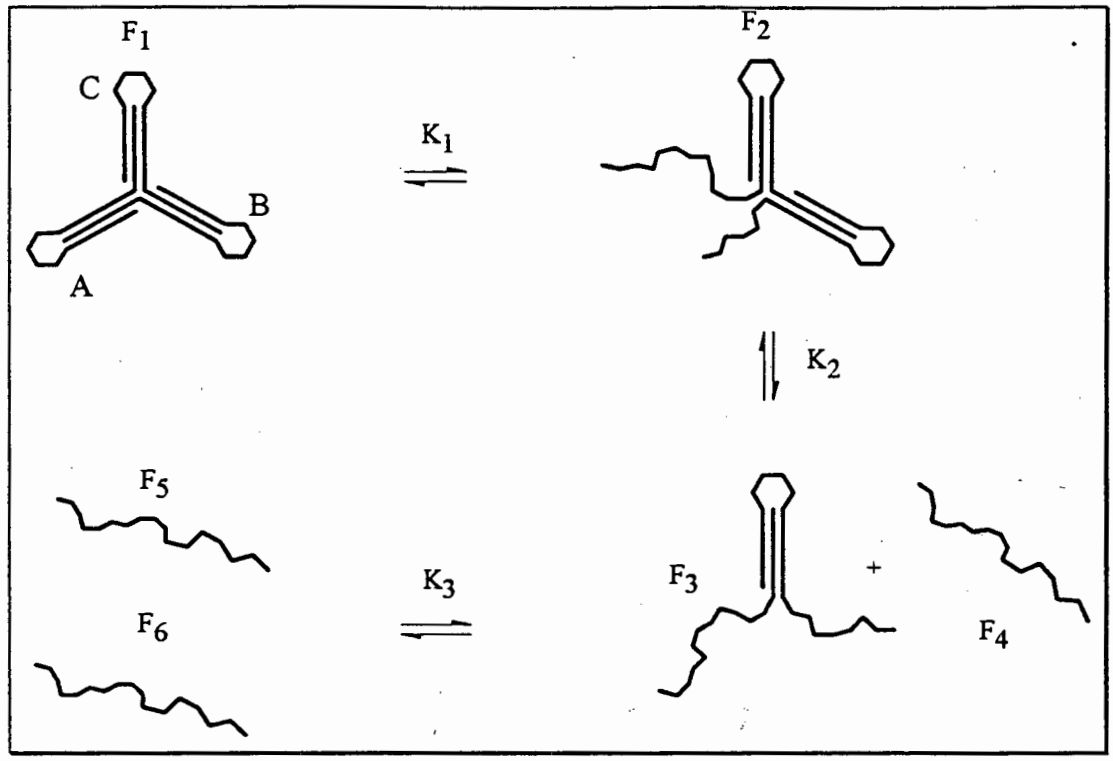

Fig. 5.19 Reaction scheme that describes the unfolding pathway of the WC triple-helical three-way junction with species F1 to F6. K_1, K_2, and K_3 are the equilibrium constants for each step.

The fitted DSC deconvoluted curves at pH 5.5 and 4,5 in Figs 5.18 A and B demonstrate that the first transition seems to be pH sensitive. At first this seems contrary to the above results, but in fact they are completely in accordance with the proposed mathematical model. From computer simulations it can be shown that the low-temperature transition will remain unaffected by pH, provided the higher-temperature transitions occur at temperatures well above the first transition.

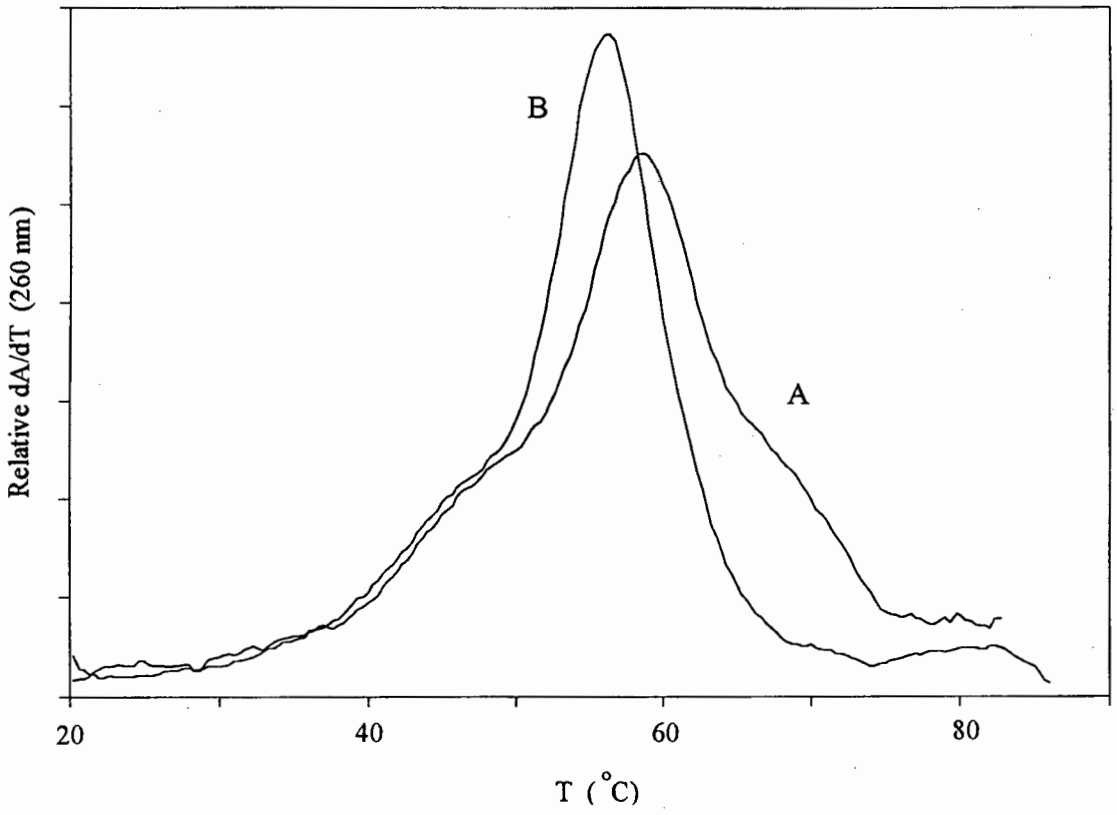

Fig. 5.20 Differential uv melting curves of the WC triple-helical three-way junction. The concentration of the oligonucleotides p1, p2, and p3 were (A) 14.25μM and (B) 1.14 μM. Buffers contained 20 mM Na_3PO_4 and 1M NaCl at pH 4.5.

If the upper transitions overlap or occur below the thermal stability of the first transition it will effect the later transition by shifting it to lower temperatures. This can clearly be seen from the derivative uv melting curves in Fig. 5.13 where all three transitions have merged at pH 6.5 and 6 and all three arms melt almost simultaneously, implying an apparent pH dependence of the first transition.

5.12. Model Calculations

The thermodynamic parameters for the unfolding of the WC triple-helical three-way junction have been characterized by nonlinear regression procedures in which the Eqs. [3.27] and [3.28] were fitted to uv melting and DSC data sets, respectively.

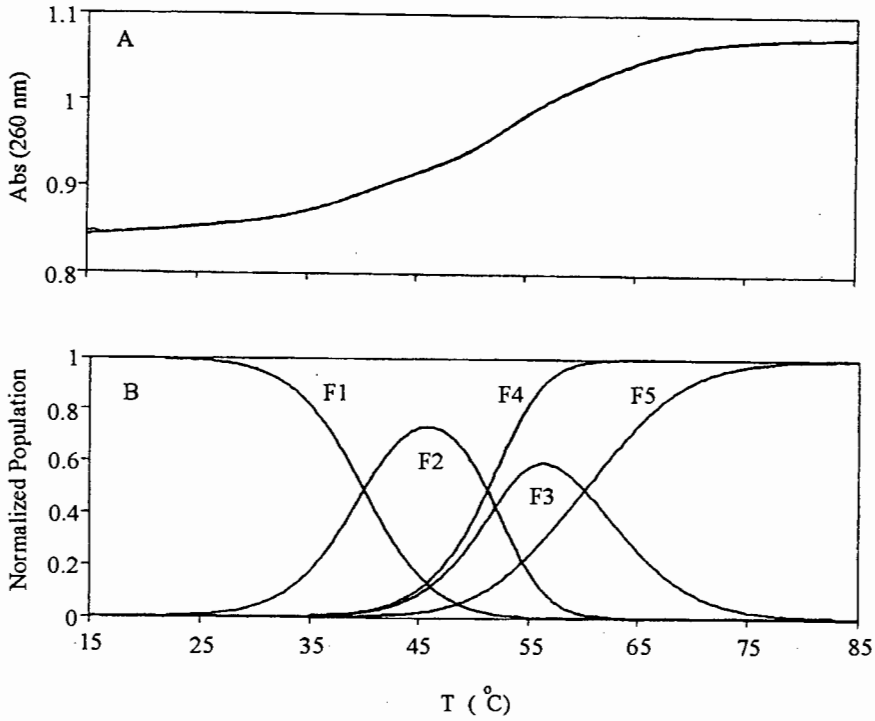

Fig. 5.21 (A) Absorbance (260 nm) vs temperature profile for the three-way junction and the fitted curves in 20 mM Na_3PO_4 and 1 M NaCl, pH 4.5. (B) Normalized populations of F1 to F6 as a function of temperature.

The fitted enthalpy and entropy values are listed in tables III and IV. The enthalpy changes obtained from DSC scans are in reasonable agreement with the enthalpy changes calculated from uv thermal denaturation. There is, however, a small difference in the DSC results for enthalpy changes at pH 5.5 and 4.5. This discrepancy can be attributed to the relatively large temperature range over which the three transitions occur, making them

somewhat susceptible to base line correction errors. Fig. 5.18 A and B show the result of the fitted DSC and uv thermal denaturation curves. Also given in Fig. 5.21 B are the deconvoluted subtransitions for the unfolding of arms A, B and C. The population distribution for the various molecular species F1 to F6 is depicted (Fig 5.21 B).

TABLE III

Thermodynamic Parameters for the Unfolding of the WC Triple-Helical Three-Way Junction in Sodium Phosphate Buffer at pH 4.5 and 5.5, Based on DSC.

pH	ΔH_1[a]	ΔH_2[a]	ΔH_3[a]	Tm_1[b]	Tm_2[b]	Tm_3[b]
4.5	69.4	130.1	64.0	54.1	65.4	76.7
5.5	59.0	130.0	70.0	54.5	58.0	60.0

pH_1[a] = 4.5 $\quad \Delta H_1$[a] + ΔH_2[a] + ΔH_3[a] = 263.5

pH_1[a] = 5.5 $\quad \Delta H_1$[a] + ΔH_2[a] + ΔH_3[a] = 255.0

Average

ΔH_1[a] = 62.7 $\quad \Delta H_2$[a] = 130.1 $\quad \Delta H_3$[a] = 67.0

ΔS_1[c] = 192 $\quad \Delta S_2$[c] = 389 $\quad \Delta S_3$[c] = 197

Note. Buffers: 20 mM Na3PO4, 1 M NaCl. Error margins were normally not greater than 5 %. Only the error for ΔH_1 is 11 %.

[a] Units for ΔH are kcal.mol^{-1} (1 cal = 4.18 J)

[b] Temperature in °C.

[c] Units for ΔS are cal.mol^{-1}.K^{-1}.

TABLE IV

Thermodynamic Parameters for the Unfolding of the WC Triple-Helical Three-Way Junction, Based on Ultraviolet Absorbance Changes vs Temperature..

pH	ΔH_1[a]	ΔH_2[a]	ΔH_3[a]	Tm_1[b]	Tm_2[b]	Tm_3[b]
4.00	55.7	110	66.8	38.2	54.4	64.4
4.25	56.0	110	67.2	39.8	54.4	63.0
4.50	59.0	110	67.2	39.2	53.4	63.0
4.75	59.0	110	67.0	42.5	54.4	62.0
5.00	58.8	110	66.5	41.4	55.3	59.5
5.25	59.8	109	65.9	38.4	51.4	53.2
5.5	59.2	108	65.2	35.3	51.3	49.7
Average						
ΔH_1[a] = 58.1	ΔH_2[a] = 109.6	ΔH_3[a] = 66.5				
ΔS_1[c] = 186	ΔS_2[c] = 336	ΔS_3[c] = 200				

Note. Buffers: 20 mM Na3PO4, 0.1 M NaCl, The parameters were determined from uv melting curves using a least squares fitting procedure. Error margins were not greater than 5 %.
[a] Units for ΔH are kcal.mol^{-1} (1 cal = 4.18 J)
[b] Temperature in °C.
[c] Units for ΔS are cal.mol^{-1}.K^{-1}.

The triple-helical three-way junction undergoes thermal denaturation in three stages, each transition corresponding to the consecutive unfolding of the arms A, B and C. The integrated arms of the triple-helical three-way junction and the isolated triple-helices representing these arms show considerable differences between their calorimetric enthalpy

changes. The enthalpy change associated with arms A and C is considerably smaller than for the triplex p4p5, and p8p9, respectively. The dissociation enthalpy change for the triple-helix p4p5 amounts to 97 kcal/mol. The discrepancy between this value and the enthalpy changes associated with arm A (62.7 kcal/mol) can only be understood if it is assumed that there is a considerable degree of unstacking in arm A around the branch point of the junction. This assumption is also supported by a lower than expected entropy change associated with this arm. A major contribution to the entropy change arises from the conformational entropy gained by the single strands during thermal denaturation (De Voe and Tinoco 1962). Similar considerations hold for arms B and C. Surprisingly the enthalpy and entropy changes associated with the triplex p6p7 (128.7 kcal/mol, 397 $cal.mol^{-1}.k^{-1}$) are similar to the calculated values obtained for the integrated arm B (130.1 kcal/mol, 389 $cal.mol^{-1}.k^{-1}$). This result suggests that the adjoining two triple-helices (arms A and C) do not have a noticeable destabilizing effect on the integrated arm B. The calculated enthalpy changes for arm C (67 kcal/mol, 197 $cal.mol^{-1}.k^{-1}$) are also considerably smaller than for the triplex p8p9 (135.9 kcal/mol, 399.4 $cal.mol^{-1}.k^{-1}$). This can be explained by the same basic assumption as that for arm A. It is therefore apparent from the thermodynamic parameters that arms A and C are partially unstacked around the branch point, while arm B remains completely folded. The results bare some resemblance to double-helical three-way junctions, where it has been shown that the bases around the branch point may be unstacked. Zhong and Kallenbach (1993) have demonstrated that the dissociation enthalpy associated with Duplex DNA decreases considerably as an adjoining neck is progressively extended from the center of the duplex, despite the larger number of base pairs available for base pairing. The degree of unfolding can be estimated by taking the ratio of the calculated enthalpy changes for each of the arms of the three-way junction

and the isolated triple helices representing the arms. In the case of arm A, only 65 % of the bases seem to be hydrogen bonded, while 50 % is hydrogen bonded in the case of arm C. In the same way it can be shown that arm B is completely folded (100.3 %).

It was shown that the thermal stability of isolated and integrated triple-helices containing Hoogsteen cytosines are dependent on the pH and both the number and position of the cytosines within the sequence of the triplex. The proposed theory used to describe the way in which the Tm changes with pH is, furthermore, in accordance with the experimental results. The prediction of the enthalpy and entropy changes calculated from the dependence of Tm on pH are in good agreement with the results obtained directly from differential scanning calorimetry. However, it was not possible to quantitate proton binding in the integrated arms of the WC three-way junction as a result of the partially overlapping transitions. Nevertheless, it was still possible to interpret the results qualitatively, as arms B and C have the repeating sequences CTT and CCT. The formation enthalpy and the number of protonated cytosines should therefore change proportionately in these arms as unwinding takes place. As a result, the slope $d(1/\mathrm{Tm})/d(\mathrm{pH})$ should remain approximately constant, whether the arms are partially unstacked and unwound or completely folded. This can be understood in terms of the relationship between 1/Tm and the pH, where the slope was shown to be directly proportional to the number of bound protons and inversely proportional to the formation enthalpy.

Chapter 6.

HG Triple-Helical Three-Way Junction

The HG triple-helical three-way junction is based on the same subset of sequences used for the three arms (A, B, and C) of the WC triple-helical three-way junction. The polarity of the sequences has been inverted and the position of the loops changed. The junction was constructed from three 33-mer oligonucleotides hp1, hp2, and hp3. Each oligonucleotide sequence can form a Watson-Crick homo-purine (9-mer) homopyrimidine (10-mer) hairpin, linked by a four membered cytosine loop. Each of the hairpins contains a homopyrimidine 10-mer single strand extension. The extensions can mutually associate with one of the other hairpins through Hoogsteen hydrogen bonding, resulting in the assemblage of the HG triple-helical three-way junction with a HG branch point and triple-helical arms. For example, the WC hairpin formed by the 5' purine sequence and the central pyrimidine sequence of the oligonucleotide hp3 constitute the backbone of arm A, while the 3' pyrimidine extension constitutes the Hoogsteen strand of arm C. As in the WC triple-helical three-way junction, arm A contains exclusively TAT triad bases, while arms B and C contain an increasing number of CGC^+ triads, respectively. To counteract possible crowding effects at the branch point, the 5' purine sequences as part of the hairpins were shortened by one base. The arms of the HG triple-helical three-way junction have been characterized separately by binary combination of the oligonucleotides hp1+hp2, hp1+hp3, and hp2+hp3.

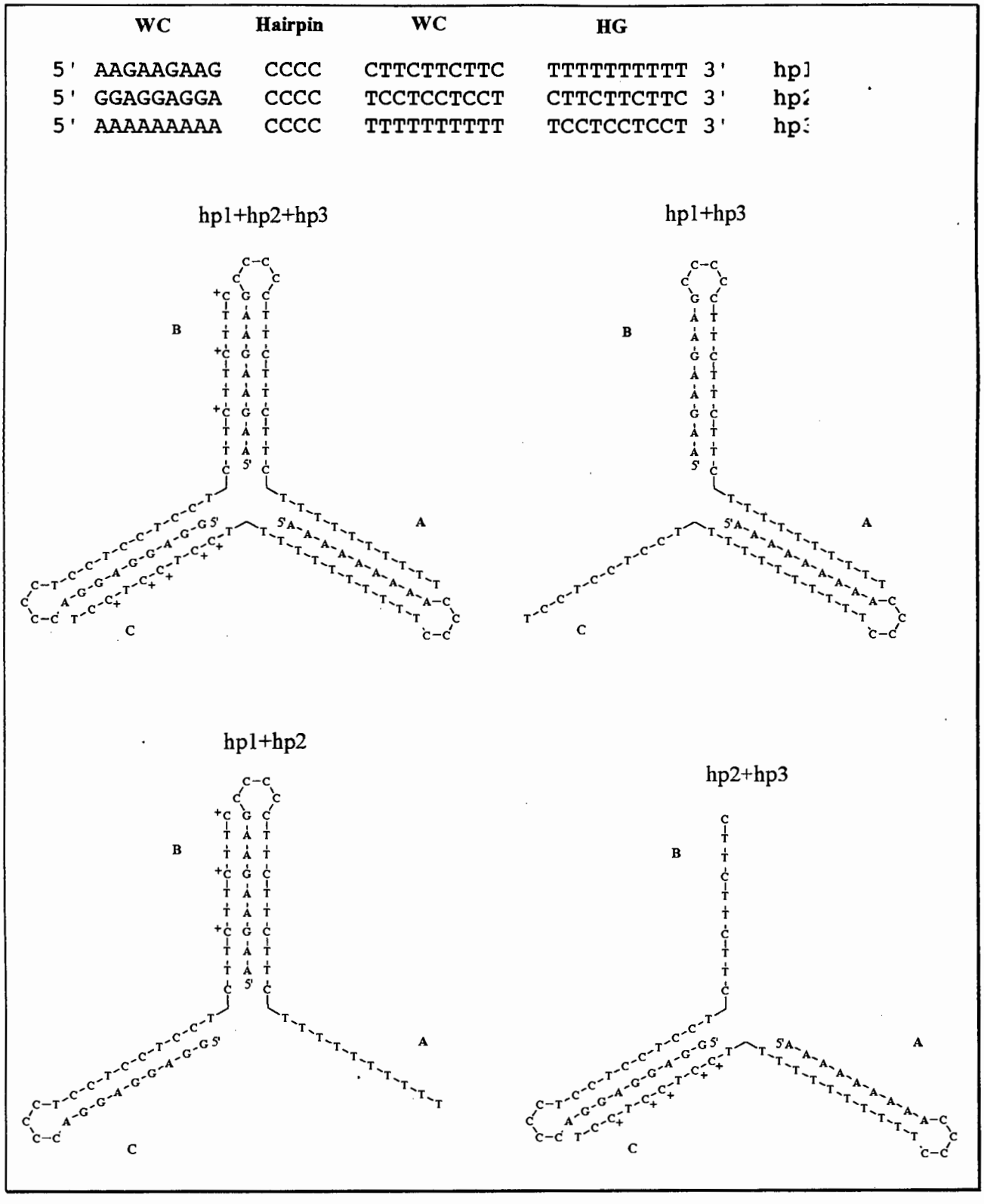

Fig. 6.1 The HG triple-helical three-way junction (hp1hp2hp3) represents the ternary complex with arms marked as A, B, and C. The complexes hp1p2, hp2p3, and hp1hp3 depict the binary complexes. The binary complexes contain three regions namely: a single stranded extension, a hairpin, and triple helix. The triple-helical region of the binary combinations corresponds to the separate arms of the ternary complex. The Watson-Crick and Hoogsteen bases are abbreviated as WC and HG, respectively.

In this chapter it will be shown that the HG triple-helical three-way junction does not unfold in a two-state manner. Below 50 °C the three arms of the junction unfold through loss of Hoogsteen hydrogen bonding, resulting in three separate WC hairpins. Above 50 °C the three hairpins denature independently according to their GC content. The calorimetric transition enthalpy for the low temperature transitions is equal to the sum of the triplex to hairpin transition enthalpies obtained for the three separate arms. These results suggest that the local geometry at the branch point does not perturb or influence the structure of the arms. This is an unexpected result, distinctly different from that obtained for the WC triple-helical three-way junction and the ordinary three-way junction with similar core sequences in the arms (chapter 5 and 7).

6.1. Electrophoresis

Electrophoresis was used to establish the formation of the HG triple-helical three-way junction under non denaturing conditions at 4 °C, pH 5.5, 100 mM sodium acetate and 50 mM $MgCl_2$ in 20 % polyacrylamide. The results are shown in Fig. 6.2 for the single strands, binary combinations, and the ternary combination of the three strands hp1, hp2, and hp3. The lanes 5-7 represent the isolated oligonucleotides hp1, hp2, and hp3, respectively. Under these conditions the oligonucleotides should fold into a hairpin structure with an extended single strand at the 3' end. The single strand extension of hp1 contains exclusively thymine bases, while hp2 and hp3 contain an increasing fraction of cytosines. The fuzziness of the band in lane 6 which corresponds to hp2 and the lower mobility of the band in lane 5 corresponding to hp3, most probably arise as a result of

dimerization of the single stranded extensions through CC^+ self complex formation. The dimerization is, in addition, more pronounced in the case of hp3 which contains three repeats of CCT. Judging from the fuzziness of the dimeric complex formed by hp2 (lane 6) it appears that this complex is in rapid equilibrium with the monomer. This can possibly explain the intermediate mobility of the band in lane 6, compared to the higher mobility of the strictly monomeric hairpin hp1 (lane 7), and the relatively small mobility of hp3 which most probably forms a self complex (lane 5).

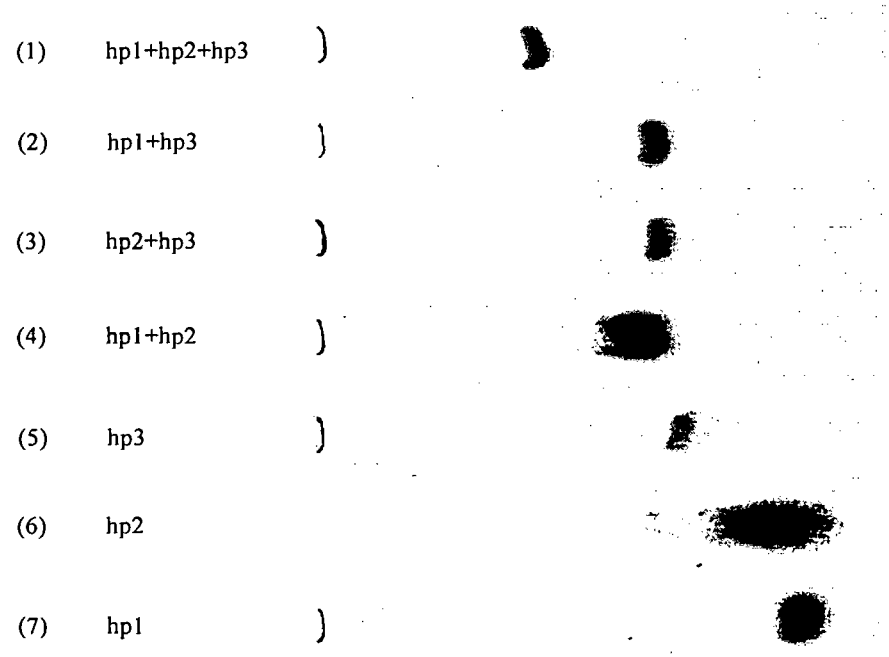

Fig. 6.2 20% native polyacrylamide gel electrophoresis at 4 °C, pH 5.5. Lanes 7, 6, and 5, oligonucleotides hp1, hp2, hp3, respectively; Lanes 4, 3, and 2, binary combinations hp1hp2, hp2hp3, and hp1hp3, respectively. An equal molar mixture of hp1, hp2, and hp3 was loaded in lane 1. Solvents contained 100 mM Sodium acetate, 50 mM $MgCl_2$. The various components stain with different intensities.

The binary complexes formed by mixing equal molar amounts of the oligonucleotides hp1, hp2, and hp3 in different combinations, are shown in lanes 2-4. The mobility of the three complexes are approximately identical, exhibiting, on average, a lower mobility than the monomeric oligonucleotides in lanes 5-7. Furthermore, it is apparent from the presence of a single band, that only one molecular species is formed for the binary combinations. The fully folded HG triple-helical three-way junction was formed by mixing stoichiometric amounts of hp1, hp2, and hp3 (lane 1). Again it is apparent from the single band in lane 1 that only one species is present. Also worth noting is the low mobility of the three-way junction as compared to the individual oligonucleotides in lanes 5-7 and the binary combinations in lanes 2-4.

6.2. Temperature Gradient Gel Electrophoresis

Temperature gradient gel electrophoresis was used to investigate further the thermal unfolding properties of the HG triple-helical three-way junction. Fig. 6.3 depicts the result of a temperature gradient gel in 20 % polyacrylamide, 100 mM sodium acetate, 50 mM magnesium acetate at pH 5.5. The temperature gradient ranges from 10 to 85 °C with its orientation perpendicular to the electrical field. In the temperature range 5 to 35 °C, where the fully folded junction is stable, a single band with only minimal increasing mobility is observed. Between the temperatures 35 and 55 °C the mobility of the band increases rapidly, which can be attributed to the unfolding of the three-way junction into the constituent hairpins hp1, hp2, and hp3, by the loss of Hoogsteen hydrogen bonding (cf. also Fig. 6.7). From the results of uv thermal denaturation (cf. Fig. 6.5) and DSC (cf. Fig. 6.10) it can be shown that the hairpins undergo thermal unfolding in the temperature range

range 55 and 90 °C. It is therefore apparent from the broad band in the temperature range 55-90 °C, that the thermal unfolding of the hairpins cannot be resolved sufficiently using temperature gradient gel electrophoresis. Similar results have also been observed in other hairpin structures (unpublished).

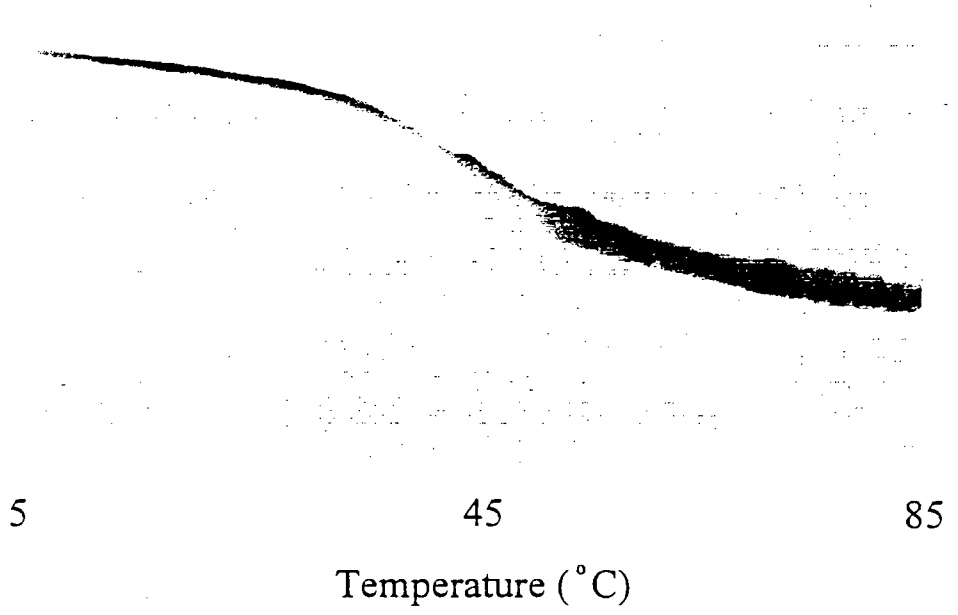

Fig. 6.3 Temperature gradient gel electrophoresis of the ternary complex (hp1hp2hp3) in 100 mM Sodium acetate and 50 mM $MgCl_2$ at pH 5.5.

6.3. pH Titration

As the formation of the HG triple-helical three-way junctions requires protonation of the cytosine bases in the Hoogsteen strand, it is convenient to follow this process by monitoring the change in the absorbance at 260 nm as a function of pH. Fig. 6.4

represents the results of a titration experiment ranging from pH 2 to 12 (Lee *et al.* 1974; Moser and Dervan 1987; Xodo *et al.* 1990; Plum *et al.* 1990; Manzini *et al.* 1990). The pH was initially changed from 7 to 2 by adding microliter quantities of concentrated HCl to a solution containing equal molar amounts of hp1, hp2, and hp3. In the same way, the pH was raised from 7 to 12 by adding similar amounts of concentrated NaOH to an aliquot of the same solution. In the pH range 12-11 the oligonucleotides are in the single stranded coiled form, as inferred by the overall high absorbance in this pH range. Decreasing the pH from 11 to 10 results in a sigmoidal decrease in absorbance as a consequence of WC hairpin formation in addition to the formation of a binary complex between hp1 and hp3. Take note that the binary complex formed between hp1 and hp3 is stable in the pH range 2.5-10 as it does not require any protonated cytosines for its formation. Between pH 11 and 6 the absorbance remains almost constant. On lowering the pH from 6 to 4 the absorbance drops off rapidly again reaching a minimum trough between pH 4 and 3. The decrease in absorbance in this pH range can be attributed to the formation of the fully folded HG triple-helical three-way junction between the hairpins hp1, hp2 and hp3. This result is also consistent with protonation of the Hoogsteen cytosines required to promote branch formation (Hüsler and Klump 1994). Decreasing the pH a little further results in the unfolding of the three-way junction, as inferred by an increase in the absorbance between pH 3.5 and 2.

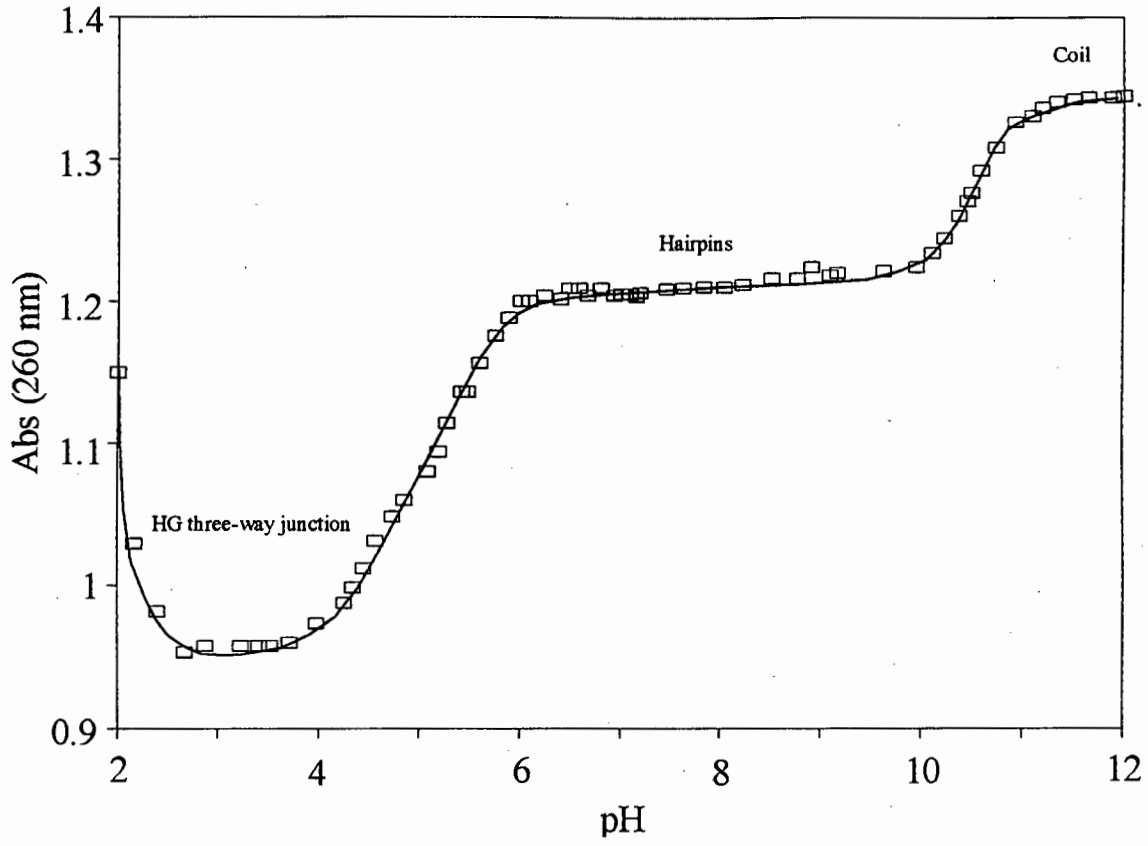

Fig. 6.4 pH titration of an equal molar mixture of hp1, hp2, and hp3 representing the ternary complex in 1 M NaCl and 20 mM Na_2HPO_4 at 20 °C. The absorbance was measured at 260 nm after adjusting the pH.

6.4. Absorbance vs Temperature Scans at 260 nm of the Ternary Complexes and the Three Binary Complexes

The thermal unfolding of the ternary complex (HG triple-helical three-way junction) and the three binary complexes have been established by recording the changes in absorbance as a function of temperature. Fig. 6.5 depicts the differential uv (260 nm) melting profiles of the three binary complexes hp1hp2, hp2hp3, and hp1hp3 and the

completely folded three-way junction hp1hp2hp3. For the three binary complexes the low-temperature transition occurring between 15 and 50 °C corresponds to the temperature induced release of the Hoogsteen strand, resulting in two isolated hairpins. The corresponding high temperature transitions arise as a result of the thermal denaturation of the hairpins. The high temperature transitions for the binary complex hp2hp3 displays two partially overlapping peaks at 67.5 and 81 °C, respectively. The observation of these two transitions is possibly due to the large difference in GC content of the hairpins hp2 and hp3. Typically the hairpin formed by hp2 contains six GC pairs while hp3 contains exclusively AT base pairs. Accordingly their melting temperatures should be quite different. The differences in GC content of the other pairwise combinations (hp1hp2 and hp1hp3) is insufficient to give good separation of their unfolding transitions (Fig. 6.5).

The HG triple-helical three-way junction's thermal unfolding is split into two temperature regions in much the same way as the binary complexes. The low-temperature transition between 15-50 °C is attributed to the disintegration of the three-way junction into separate haipins due to the loss of Hoogsteen hydrogen bonding. The high-temperature transition corresponds to the consecutive unfolding of the three remaining hairpins hp1, hp2, and hp3 in the temperature range 55-95 °C.

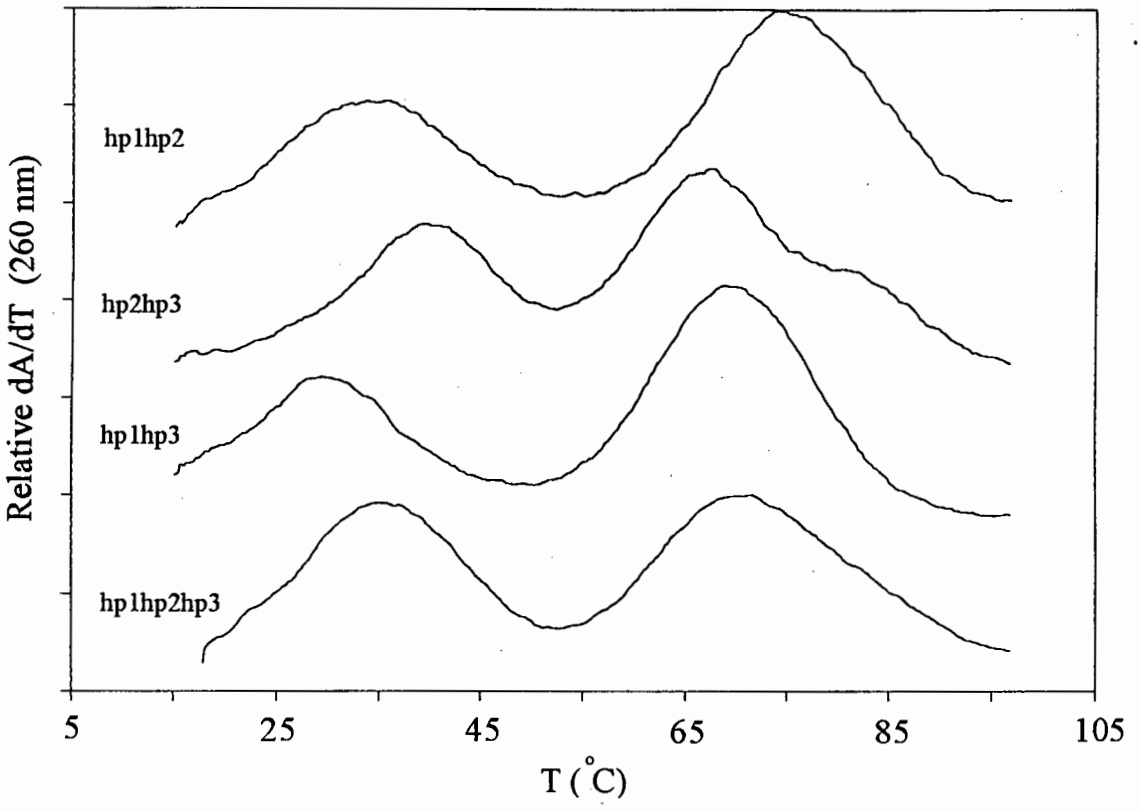

Fig. 6.5 Differential uv (260 nm) thermal denaturation curves of the binary complexes hp1hp2, hp2hp3, hp1hp3, and the ternary complex hp1hp2hp3. Buffers contained 100 mM sodium acetate, 50 mM magnesium acetate at pH 5.5.

6.5. CD Spectroscopy of the HG Triple-Helical Three-Way Junction

The CD spectrum of the ternary complex was determined at different temperatures to establish the identity of the low and high temperature transitions of the three-way junction, respectively. The spectrum of the ternary complex is given in Fig. 6.6 at 10, 55, and 100 °C in 1 M NaCl, 20 mM Na_2HPO_4 and pH 5.5. At 100 °C the spectrum is typical of single

stranded oligonucleotides, characterized by the small negative and positive bands at approximately 245 nm and 278 nm, respectively.

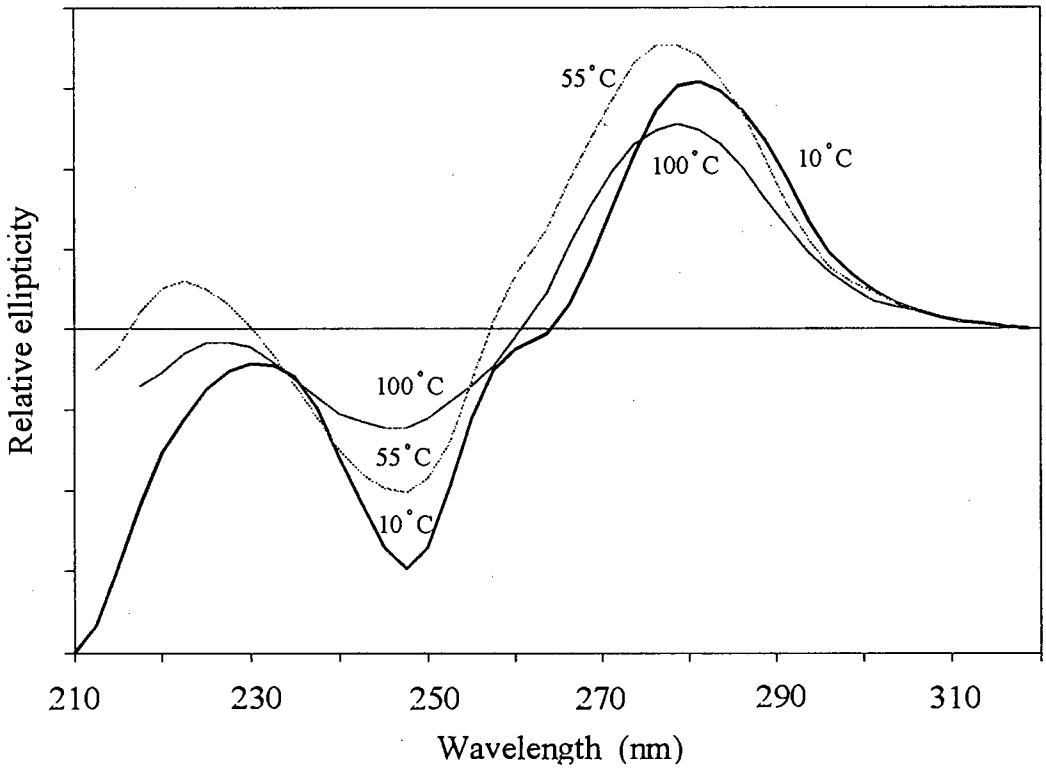

Fig. 6.6 CD spectrum of the ternary complex hp1hp2hp3 at temperatures 100, 55, and 10 °C in 1 M NaCl and 20 mM Na_2HPO_4 at pH 5.5.

The spectrum obtained at 55 °C is typical of B-DNA, indicating that only the hairpins are stable in this temperature range. Dropping the temperature to 10 °C, which is below the dissociation temperature of the Hoogsteen strands (cf. also Fig. 6.5 and 6.8), results in a red shift of the positive band (280 nm), and a strong negative band between 210-225 nm. This band suggest that the arms of the three-way junction adopt an A-like conformation

which has been proposed as indicative of triple helix formation (Lee *et al.* 1978). Furthermore, the red shift of the positive band at 280 nm suggests that the Hoogsteen cytosine bases have undergone protonation.

6.6. Calorimetric and van't Hoft Enthalpy and Entropy Determinations for the Unfolding of the Binary Complexes hp1hp2, hp2hp3, and hp1hp3

The boxed section in Fig. 6.7 describes schematically the unfolding pathway of the three binary complexes. The fully folded binary complex formed by mixing stoichiometric amounts of two oligonucleotides is represented by the molecular species Fa. In the first step Fa dissociates into the two hairpin species Fb and Fc. The hairpins Fb and Fc will undergo further unfolding to form the single stranded coiled species Fb' and Fc', respectively. The equilibrium constants (K_1, K_2, and K_3) associated with these steps are placed in round brackets. The enthalpy and entropy changes were determined by fitting equations [2.30] and [2.31] to the experimental uv melting and DSC data sets respectively, using the Levenberg-Marquardt method of least squares. The results are listed in Tables V and VI. Note that the respective enthalpy and entropy changes (ΔH_1-ΔH_6 and ΔS_1-ΔS_6), as listed in Tables V and VI, refer to the blocked reaction scheme in Fig 6.7.

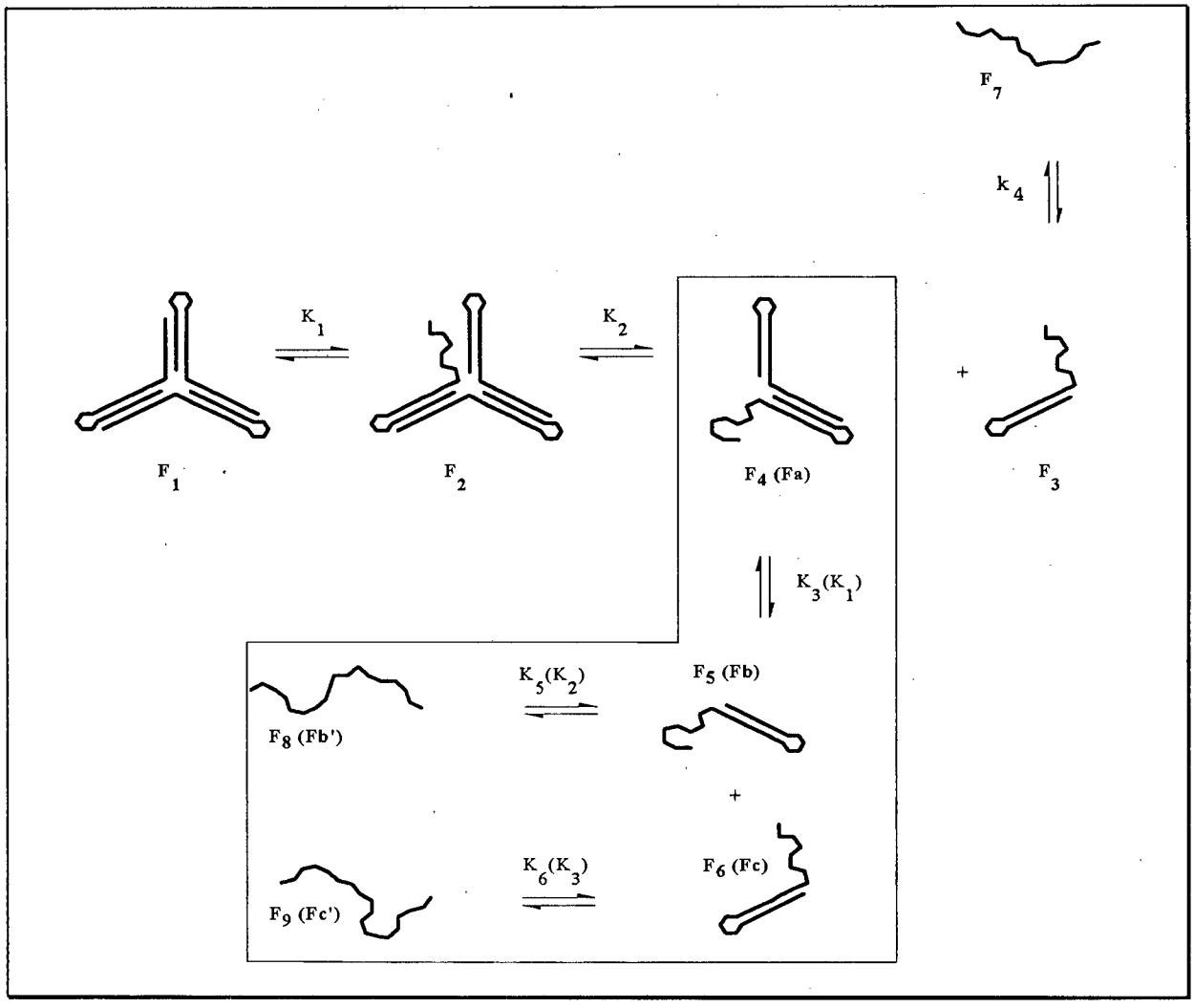

Fig. 6.7 Reaction scheme describing the unfolding pathway of the ternary complex with species F1-F9. K_1-K_6 represent the equilibrium constants for each step. The boxed reaction pathway represents the unfolding steps of the binary complex. The molecular intermediates and equilibrium constants for the boxed reaction pathway of the binary complex are presented in round brackets for (K_1)-(K_3), and (F_a)-($F_{c'}$), respectively.

The enthalpy changes determined directly by DSC, for the binary complexes are in reasonable agreement with the enthalpy changes obtained by curve fitting from uv melting data. Fig. 6.8 A to C depict the heat capacity profiles for the binary complexes and the corresponding hairpin to coil transitions. Also shown are the deconvoluted sub-transitions for each of the steps. The low temperature peak (Cp1) represents the molar excess heat

capacity associated with the unfolding of the binary complex (triplex), while the high temperature peaks, Cp2 and Cp3, represent the molar excess heat capacities associated with the unfolding of the two hairpins.

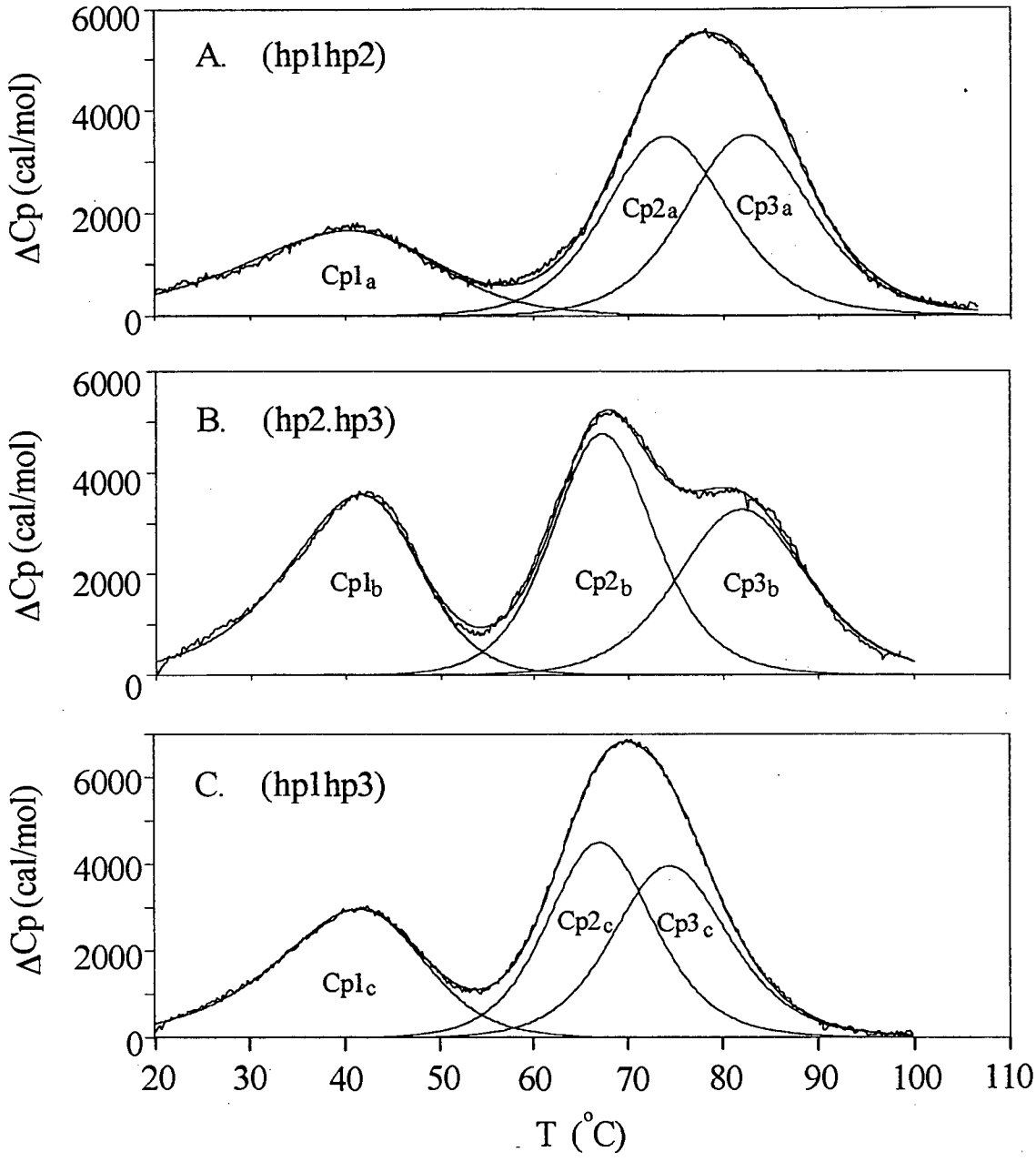

Fig. 6.8. Excess heat capacity (ΔCp) vs temperature profiles for the thermal unfolding of the binary complexes A, hp1hp2; hp2hp3, B; and C, hp1hp3. The envelope curve represents the experimental DSC data curve and the overlapping curve, the fitted theoretical curve (cf. also chapter 3). The underlying peaks Cp_1-Cp_3 represent the enthalpy contributions for each reaction step K_1 to K_3 (cf. boxed section in Fig. 6). Buffers contained 1 M NaCl, 20 mM Na_2HPO_4 at pH 5.5.

TABLE V

Thermodynamic Parameters for the Unfolding of the Binary Complexes and the HG Three-Way Junction Obtained by DSC

	Triple Helices						Hairpin Double Helices					
Complex	hp1+hp2		hp1+hp3		hp2+hp3		hp1		hp2		hp3	
	ΔH_1	ΔS_1	ΔH_2	ΔS_2	ΔH_3	ΔS_3	ΔH_4	ΔS_4	ΔH_5	ΔS_5	ΔH_6	ΔS_6
hp1+hp2	45.0	142.9	—	—	—	—	59.6	171.8	62.0	174.6	—	—
hp1+hp3	—	—	60.8	192.6	—	—	61.5	177.1	—	—	64.0	188.2
hp2+hp3	—	—	—	—	64.8	205.0	—	—	57.9	162.9	64.6	190.1
hp1+hp2+hp3	42.1	137.4	57.4	177.7	63.7	203.1	65.5	193.8	61.8	174.2	70.9	205.4
Average	43.6	140.2	59.1	185.2	64.3	204.1	62.2	180.9	60.6	170.6	66.5	194.6

Triple Helices (hp1.hp2, hp1+hp3 & hp2+hp3):

$\Delta H_1 + \Delta H_2 + \Delta H_3 = 170.6$ $\Delta S_1 + \Delta S_2 + \Delta S_3 = 540.5$

Three Way Junction:

$\Delta H_1 + \Delta H_2 + \Delta H_3 = 163.2$ $\Delta S_1 + \Delta S_2 + \Delta S_3 = 518.2$

Thermodynamic parameters for the unfolding of the binary complexes hp1hp2, hp1hp3, hp2hp3 and the HG three-way junction hp1hp2hp3, were obtained by DSC. Eq [3.31] was fitted to the DSC data set for the complexes hp1hp2, hp1hp3 and hp2hp3, and equation [4.39] was fitted to the DSC data set obtained for the HG three-way junction. The Levenberg-Marquardt algorithm of least squares was used to fit the enthalpy and entropy parameters.

Note. Buffers: 20 mM Na_2HPO_4, 1M NaCl.
ΔH has units of $kcal.mol^{-1}$ (1 cal = 4.18J).
ΔS has units of $cal.mol^{-1}.K^{-1}$

Table VI.

Thermodynamic Parameters for the Unfolding of the Binary Complexes calculated from uv Thermal Denaturation

		Triple Helices						Hairpin Double Helices					
		hp1+hp2		hp1+hp3		hp2+hp3		hp1		hp2		hp3	
Complex	pH	ΔH_{1VH}	ΔS_{1VH}	ΔH_{2VH}	ΔS_{2VH}	$\Delta H_{3\,VH}$	$\Delta S_{3\,VH}$	ΔH_{4VH}	ΔS_{4VH}	ΔH_{5VH}	ΔS_{5VH}	ΔH_{6VH}	$\Delta S_{6\,VH}$
hp1+hp2	5.5	50.9	172.0	—	—	—	—	65.4	189.9	65.1	183.1	—	—
	5.2	50.7	168.3	—	—	—	—	65.0	188.8	60.2	170.3	—	—
	5	50.8	166.0	—	—	—	—	65.6	189.4	59.6	167.8	—	—
Average:		50.8						65.3	189.4	61.6	173.7		
hp1+hp3	6	—	—	63.5	212.1	—	—	63.3	181.1	—	—	62.4	183.5
	5.5	—	—	63.6	211.7	—	—	63.3	180.9	—	—	62.4	183.2
	5.2	—	—	63.6	211.4	—	—	63.3	180.9	—	—	62.4	183.3
Average:				63.6				63.3	180.9			62.4	183.3
hp2+hp3	5.6	—	—	—	—	64.8	215.6	—	—	62.3	174.3	64.9	192.1
	5.4	—	—	—	—	64.8	211.6	—	—	62.3	174.4	65.0	192.6
	5.2	—	—	—	—	64.8	216.5	—	—	62.3	172.7	65.1	192.1
Average:						64.8				62.3	173.8	65.0	192.3

Average Hairpin Enthalpies: $\Delta H_{4VH} = 64.3$ $\Delta H_{5VH} = 61.1$ $\Delta H_{6VH} = 63.7$

Average Hairpin Entropies: $\Delta S_{4VH} = 185.2$ $\Delta S_{5VH} = 173.8$ $\Delta S_{6VH} = 187.8$

$\Delta H_1 + \Delta H_2 + \Delta H_3 = 179.2$ $\Delta S_1 + \Delta S_2 + \Delta S_3 = 546.8$

Thermodynamic parameters for the unfolding of the binary complexes hp1hp2, hp1hp3, and hp2hp3 were calculated from uv thermal denaturation data set using equations [3.30]. The Levenberg-Marquardt algorithm of least squares was used to fit the variables ΔH and ΔS.

Note. Buffers: 20 mM Na_2HPO_4, 1M NaCl.
ΔH_{VH} has units of $kcal.mol^{-1}$ (1 cal = 4.18J).
ΔS_{VH} has units of $cal.mol^{-1}.K^{-1}$

6.7. pH Effects on the Thermal Stability of the Binary Complexes and HG Triple-Helical Three-Way Junction

A series of melting experiments was performed at different pH values (pH 4.8 to 6) to determine the effect of pH on the thermal stability of the binary and ternary complexes. Fig. 6.9 represents a plot of 1/Tm vs pH for the low-temperature transitions of the three binary complexes (Arnott *et al.* 1976) and the ternary complex. It is apparent from this plot that the Tms for the binary complex hp1hp3 are pH independent, while the low temperature transitions for the complexes hp1hp2 and hp2hp3 are sensitive to changes in pH. The independence of Tm from pH for hp1hp3 is consistent with the TAT triad composition of the triple-helical region in this complex. In contrast, the binary complexes hp1hp2 and hp2hp3 show different slopes with respect to 1/Tm vs pH. The steeper slope obtained for the binary complex hp1hp2 is consistent with the larger number of HG cytosines in the triplex region as compared to hp2hp3. These results are in good agreement with the results obtained for the isolated arms (cf. Chapter 5) of the WC triple-helical three-way junction and with published data (Plum and Breslauer 1995). The high-temperature transitions in each of the binary complexes is pH independent in the pH range studied here. The results presented above suggest the loss of Hoogsteen hydrogen bonding in the low temperature regions and the dissociation of the hairpins in the high temperature region.

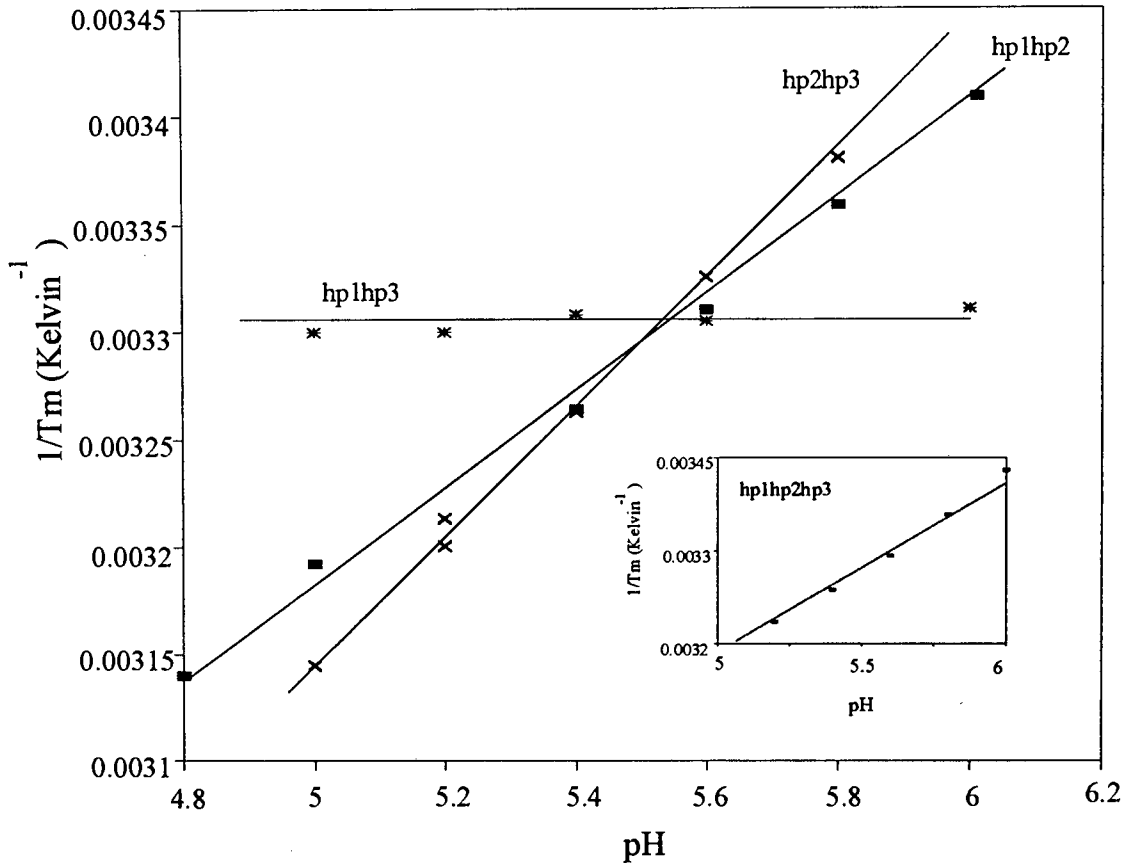

Fig. 6.9. Phase diagram representing 1/Tm vs pH as deduced from uv thermal denaturation for the low temperature transition of the binary complexes hp1hp2, hp2hp3, and hp1hp3 and the ternary complex hp1hp2hp3 (insert). Buffers contained 100 mM sodium acetate and 50 mM magnesium acetate.

The pH dependence of the low temperature transition of the HG triple-helical three-way junction was also investigated (insert of Fig. 6.9). The plot shows a linear relationship between 1/Tm and pH, the slope ($1.8X10^{-4}$) of which is within the same range as exhibited by the three binary complexes.

6.8. Proton Binding in the Triple-Helical Regions of the Binary Complexes and the HG Triple-Helical Three-Way Junction

The number of protonated cytosines in the Hoogsteen strand have been determined directly from the 1/Tm vs pH profiles in the limiting range where the pH is higher than the pKa of cytosine. As was demonstrated in chapter 3.5, the number of protonated sites can be determined directly from the slope of 1/Tm vs pH by:

$$n = \frac{\partial(1/Tm)}{\partial(pH)}\left(\frac{-\Delta H}{R \ln 10}\right), \qquad [5.34b]$$

where n is the total number of protonated cytosines in the Hoogsteen strand, R is the molar gas constant and ΔH the formation enthalpy of the triplex. By substituting ΔH (obtained from uv thermal denaturation or DSC) into equation [5.34b], one can calculate n directly.

The number of protonated sites for the binary complexes hp1hp2 and hp2hp3 is summarized in Table VII at either 100 mM NaCl and 50 mM $MgCl_2$ or 1 M NaCl. For the complex formed between hp1 and hp2, the number of protonated sites was found to fit the number of Hoogsteen cytosines. In the case of the complex hp2hp3 only 4 of 6 Hoogsteen cytosines appear to be protonated. Similar results were obtained for the triplex p8p9 where only 5 of 6 Hoogsteen cytosines are protonated (chapter 5). Both these complexes have a similar sequence. It seems likely that incomplete protonation may arise as a result of nearest neighbor electrostatic interactions between positively charged cytosines. The enthalpy contribution for Hoogsteen strand binding (low-temperature

transition) in the HG triple-helical three-way junction was determined indirectly from the slope of 1/Tm vs pH, under the assumption that the total number of protonated cytosines in the three-way junction is simply the sum of the number of protonated cytosines estimated for the three binary complexes. It should be noted that this can only be done if it is assumed that the three arms (A, B, and C) dissociate in a two-state manner or that the unfolding of the arms is strongly coupled i.e., one observes highly overlapping transitions. It will be shown shortly that the transitions do not actually proceed in a two-state manner. The transitions are indeed strongly coupled as demonstrated by the single melting profile for the low-temperature transition (cf. Figs. 6.5 and 6.10). An enthalpy change of 172.8 kcal/mol was obtained from the slope of 1/Tm vs pH assuming that the number of protonated cytosines amount to 7 (cf. Table VII). The calculated enthalpy change is in surprisingly good agreement with the enthalpy changes obtained directly by DSC (cf. Table VI), and is approximately equal to the sum of the enthalpy changes associated with the binary complexes. The corresponding van't Hoft enthalpy changes were also determined from a shape analysis of the uv thermal denaturation curves, assuming a two state model (Marky and Breslauer 1987). The ΔH_{VH} enthalpy changes thus obtained are notably smaller than the enthalpy changes obtained from either the Tm dependence on pH (cf. Table VII) or from DSC data directly (cf. Table VI). The ratio (r) between the van't Hoft enthalpy and the calorimetric enthalpy, suggest that the unfolding proceeds through multiple steps. Typically, ratios of $r = 0.36$ were obtained at pH 5 and $r = 0.52$ at pH 5.4. A ratio less than one usually implies that intermediate states are present.

Table VII.

Thermodynamic Parameters for the Triple-Helices hp1hp2, hp2hp3 and the HG Three-Way Junction hp1hp2hp3 Obtained From the 1/Tm vs pH.

Triplex Complex	ΔH^c	ΔS^d	$\delta(1/Tm)/\delta pH$	n
hp1hp2a	47.5	148.6	2.91×10^{-4}	3 (3.2)
hp1hp2b	58.4	178.8	2.36×10^{-4}	3 (2.6)
hp2hp3a	63.3	209.9	2.71×10^{-4}	4 (3.8)
hp2hp3b	65.2	196.2	2.82×10^{-4}	4 (3.9)
hp1hp2hp3a	176.2	551.2	1.83×10^{-4}	7
hp1hp2hp3b	172.8	535.6	1.87×10^{-4}	7

Thermodynamic parameters for the unfolding of the triple helices hp1hp2, hp2hp3 and the Hoogsteen three-way junction hp1hp2hp3 were obtained from the 1/Tm dependence on pH; $\delta(1/Tm)/\delta pH$. Tm values were determined by uv thermal denaturation.

n: Number of protonated Hoogsteen sites, Actual calculated values are presented in parentheses.
[a] 1 M NaCl, 20 mM Na_2HPO_4..
[b] 100 mM Na acetate, 50 mM magnesium acetate.
[c] Units for ΔH are $kcal.mol^{-1}$ (1 cal = 4.18J).
[d] Units for ΔS are $cal.mol^{-1}.K^{-1}$.

6.9. Following the Unfolding of the Three-Way Junction by DSC

Since the unfolding of the complex does not follow a simple two-state model, it is impossible to extract the thermodynamic parameters directly from the analysis of uv thermal denaturation curves as the transitions are not sufficiently resolved. Therefore, DSC was used to measure the integral transition enthalpy of the low temperature transitions of the three-way junction directly. To extract the partial contribution of the arms A, B, and C from the dissociation of all Hoogsteen strands and consecutively the denaturation of the resulting hairpins, it was necessary to follow the sequential unfolding

process as depicted in Fig. 6.7. The relationship between conformers which are in equilibrium with one another are given by equations [4.23] to [4.28].

Fig. 6.10 depicts the DSC scan for the thermal unfolding of the three-way junction in 1 M NaCl and 20 mM Na_2HPO_4, at pH 5.5. The envelope curve represents the experimental DSC data and the smooth curve passing through the experimental data represents the fitted curve, as defined by equation [4.39] (cf. section 3.4, chapter 3). The underlying peaks Cp1-Cp6 represent the deconvoluted molar excess heat capacities for each of the unfolding steps (cf. Eqs [4.40]-[4.45]), as modeled by the reaction pathway in Fig 6.7. From the close match of the experimental and calculated curves it is obvious that the choice of the fitting parameters is adequate. The low temperature transition can be sub-divided into three sub-transitions (peaks Cp1-Cp3) which represent the dissociation of the Hoogsteens strands from their corresponding hairpin loops. The sub-transitions Cp1-Cp3 cannot be assigned to individual triple-helical arms by simply comparing the transition temperatures (Tm or Tmax) to $Cp1_a$-$Cp1_c$ in Figs. 6.8 A to C. This is due to the coupling of the three sub-transitions, Cp1-Cp3. It follows from the model of the thermal unfolding (Fig. 6.7), that the first sub-transition Cp1 is associated with an intramolecular step, which is concentration independent, while the corresponding transitions for the binary complex involves intermolecular steps, which are concentration dependent. To assign the subtransitions to individual steps of the unfolding pathway, the fitted data were compared with the enthalpy and entropy values obtained by DSC for the corresponding low temperature transition in the binary complexes hp1hp2, hp2hp3, and hp1hp3.

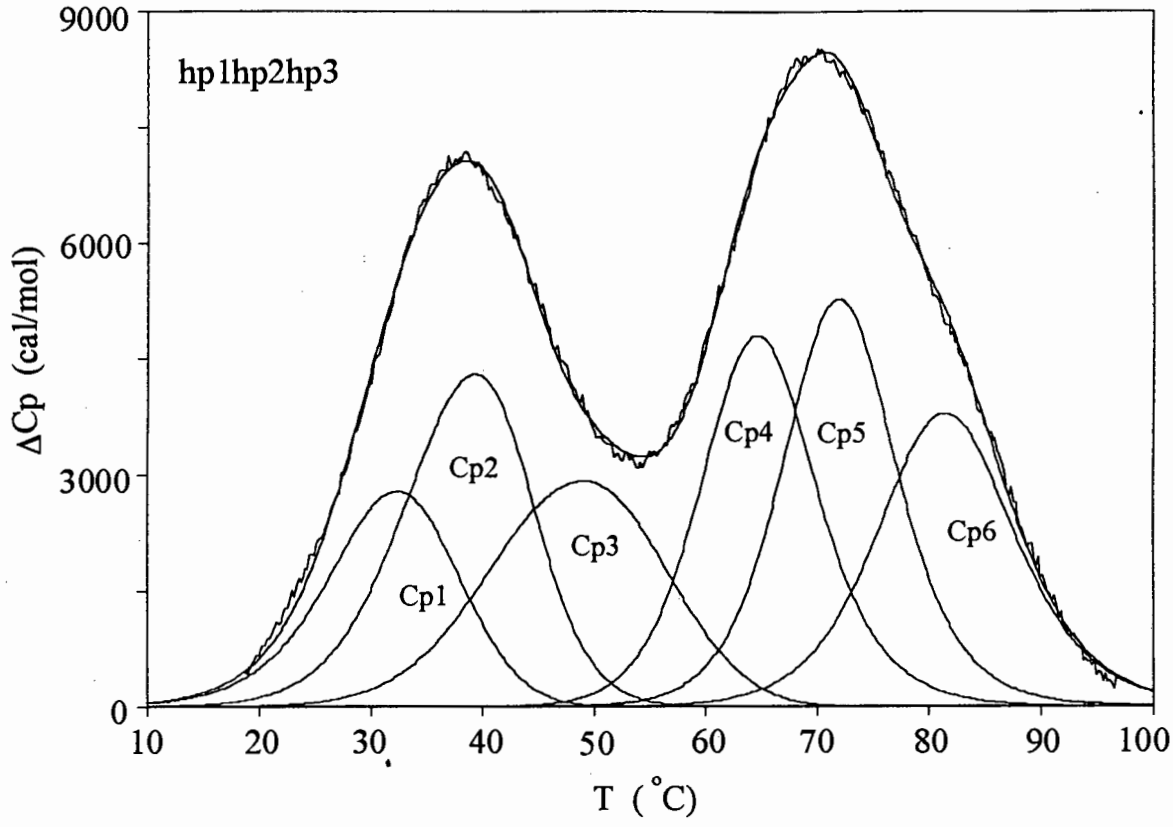

Fig. 6.10 Experimental excess heat capacity (ΔCp) vs temperature profiles for the thermal unfolding of the ternary complex hp1hp2hp3 and the calculated curve (cf. chapter 3.4). The underlying peaks Cp_1-Cp_6 represents the deconvoluted enthalpy contribution for each of the reaction step K_1-K_6 (cf. also Fig. 6.7). Buffers contained 1 M NaCl and 20 mM Na_2HPO_4 at pH 5.5.

Accordingly, the sub-transitions Cp1, Cp2, and Cp3 correspond to the dissociation of the Hoogsteen strand in arms B (hp1hp2), A (hp1hp3), and C (hp2hp3), respectively. The high temperature DSC peak can also be broken down into three sub-transitions. Each subtransition (Cp4-Cp6) corresponds to the thermal denaturation of the hairpins hp1, hp2 and hp3, respectively. As the the hairpin structures unfold independently, the correlation between Cp4-Cp6 and hp1, hp2, and hp3 is straight forwardly derived by comparing the melting temperatures of the isolated hairpin transitions with the deconvoluted steps of

unfolding for the HG three-way junction. The peaks Cp4, Cp5, and Cp6 can thus be assigned to the unfolding of the hairpins hp3, hp1, and hp2, respectively. The close match of the sum of the thermodynamic parameters of the separate entities with the parameters obtained for the subtransitions of the HG three-way junction, suggests that the geometry of all helices around the branch point in the three-way junction is not perturbed. This is in contrast to the results obtained for the WC triple-helical three-way junction (Hüsler and Klump 1994; chapter 5). In that case it was shown that two of the three arms are partially unstacked, implying that the structure is conformationally strained. Similar results have been reported when additional base pairs were introduced into the nick of a 16-mer DNA duplex (Zhong and Kallenbach 1991), used to investigate how a double helical three-way branch forms, and how its properties change as the third arm is extended (cf. chapter 1.10).

6.10. Influence of Sodium and Magnesium Ion Concentration on the Formation of the HG Three-Way Junction

To investigate the influence of sodium ion concentration on the transition temperature of either the low or high temperature transitions of the HG three-way junction, a series of six melting experiments was performed. The [Na^+] concentration was varied from 0-600 mM, in 20 mM Na_2HPO_4. The results are depicted in Fig. 6.11.

With no NaCl added only the high temperature transition, due to the unfolding of the hairpin structures, can be observed between 30-70 °C. Increasing the sodium chloride concentration to 100 mM results in the appearance of an additional transition in the

temperature range 0-40 °C. A further stepwise increase in the Na^+ concentration to 300 and finally 600 mM is accompanied by an increase in the area under the low temperature transition peak.

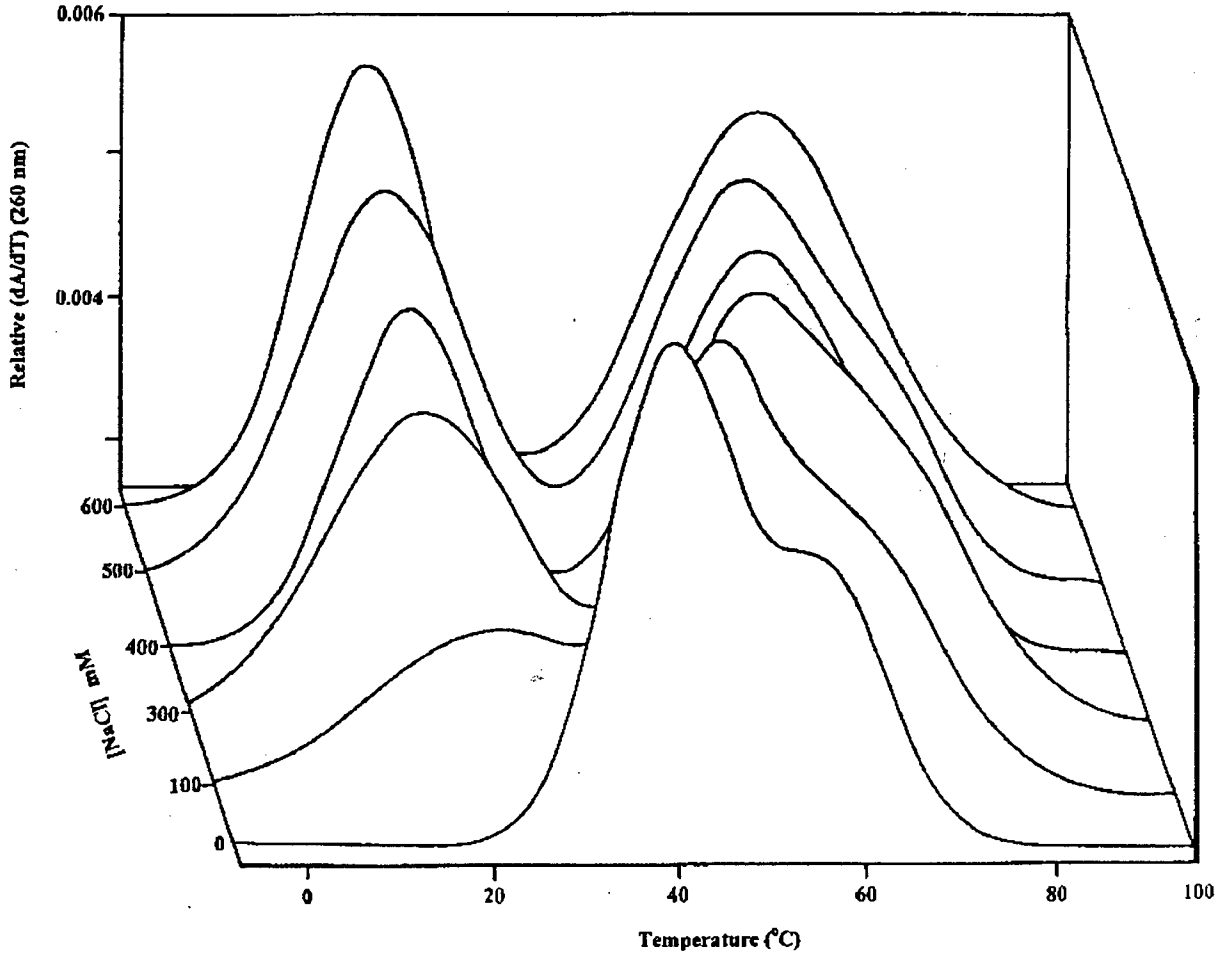

Fig. 6.11 Differential uv (260 nm) thermal denaturation curves for the ternary complex hp1hp2hp3 at various NaCl concentrations. Buffers contained 20 mM Na_2HPO_4 at pH 5.5.

The extent of HG three-way junction formation must increase with increasing NaCl concentration. This result deviates from those obtained previously for a WC triple-helical

three-way junction, where it was shown that the junction was already stable at low salt concentrations (Chapter 5).

The effect of Mg^{2+} ions on the stability of the ternary and the three binary combinations representing the arms of the three-way junction, was also investigated. It was found that the three-way junction forms a stable complex in 100 mM Na^{+} and 30-60 mM Mg^{2+}. To characterize the effect of the Mg^{2+} ion concentration on the formation of the HG three-way junction and the binary complexes, changes in absorbance at 260 nm were recorded as a function of Mg^{2+} concentration (0-200 mM). All solutions, containing 100 mM sodium acetate at pH 5.5, were titrated by adding µl quantities of 1M Mg^{2+} to adjust the magnesium concentration of the solutions (25 °C). Absorbance readings were corrected for dilution effects. Fig. 6.12 depicts the normalized absorbance vs Mg^{2+} concentration plot for the ternary and the three binary complexes. In the case of the hp1hp2 complex, the absorbance decreases between 0-90 mM Mg^{2+}, reaching an absorbance minimum between 90 and 120 mM Mg^{2+}. The decrease in absorbance reflects the formation of the binary complex between hp1 and hp2 due to the association of the extension of hp1 with the hairpin helix of hp3. Increasing the Mg^{2+} concentration further from 120 to 200 mM results in the dissociation of the complex. Thus, it was found that Mg^{2+} ion concentration has a different effect on the other two binary complexes.

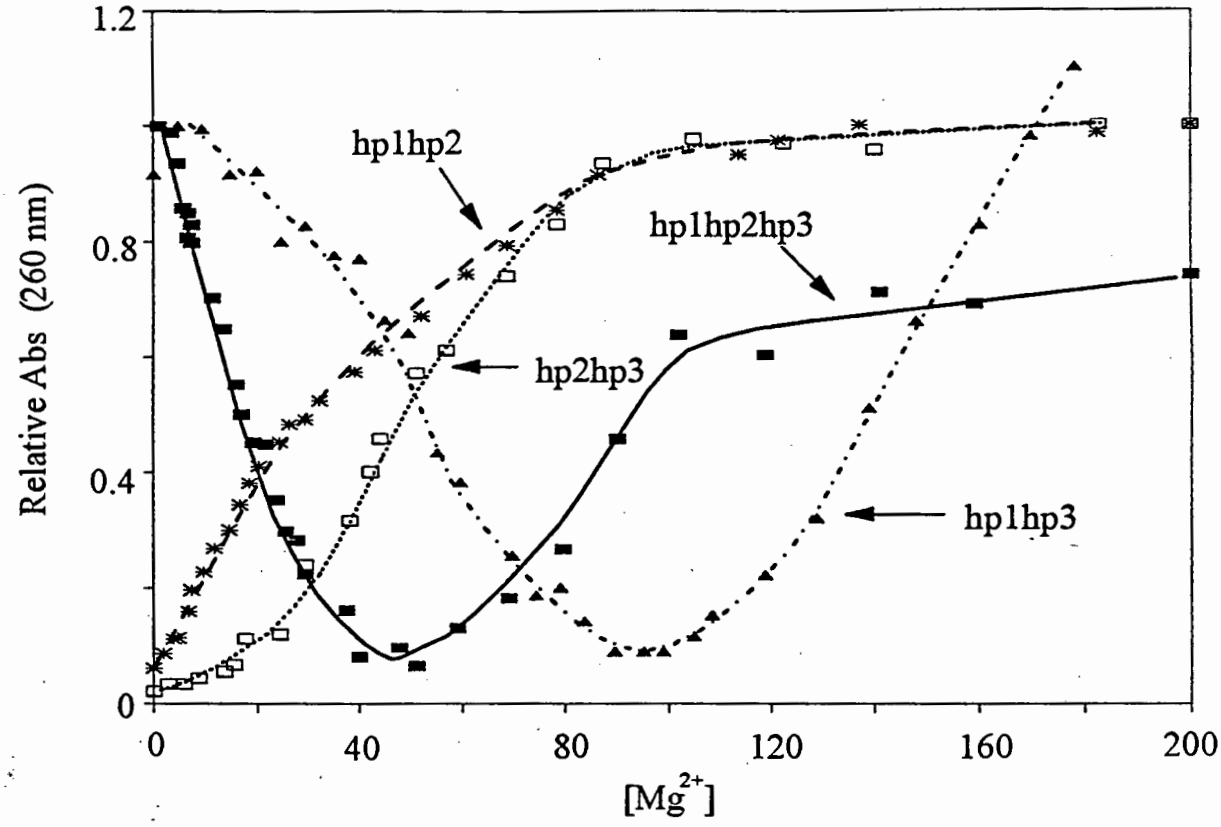

Fig. 6.12 Absorbance (260 nm) vs Mg^{2+} ion concentration for the three binary combinations hp1hp2, hp2hp3, and hp1hp3 and the ternary complex hp1hp2hp3 at 25 °C. Buffers contain100 mM sodium acetate at pH 5.5.

For the complex hp2hp3 the absorbance is minimal between 1 and 20 mM Mg^{2+} and increases sharply between 20 and 100 mM, indicating that addition of Mg^{2+} destabilizes the complex. Similarly, Mg^{2+} ions (0-100 mM) have a destabilizing effect in the case of the complex hp1hp2. From these results it is apparent that Mg^{2+} ions are required to stabilize the triplex formed between hp1 and hp3, This can be attributed to the need to shield the large number of negatively charged phosphates along the backbone which arises from the close proximity of the Watson-Crick and Hoogsteen strands. Mg^{2+} ions are, however, destabilizing in the case of the triple helices hp2hp3 and hp1hp2. Both these triple-helices

have a mixed cytosine and thymine Hoogsteen strand in common. The absorbance of the three-way junction hp1hp2hp3 decreases between 0-40 mM Mg^{2+}, reaching a shallow minimum between 40-60 mM Mg^{2+}, corresponding to the formation of the fully folded three-way junction. A further increase in the Mg^{2+} concentration between 60-100 mM is accompanied by a sharp increase in absorbance. Thus, it is apparent from the Mg^{2+} titration results that the ternary complex shows characteristics of all three individual binary complexes.

Chapter 7.

Double Helical Three-Way Junction

This three-way junction is a conventional double-helical junction based on the same sub-set of sequences in the arms A, B, and C as was used for the WC and HG triple-helical three-way junctions. Three (20-mer) oligonucleotides q1, q2, and q3, were used which, when mixed in a 1:1:1 ratio, mutually associate to form a fully folded double-helical three-way junction (cf. Fig. 7.1). Arm A contains exclusively AT base pairs and arms B and C contain an increasing number of GC base pairs (cf. Fig. 7.1).

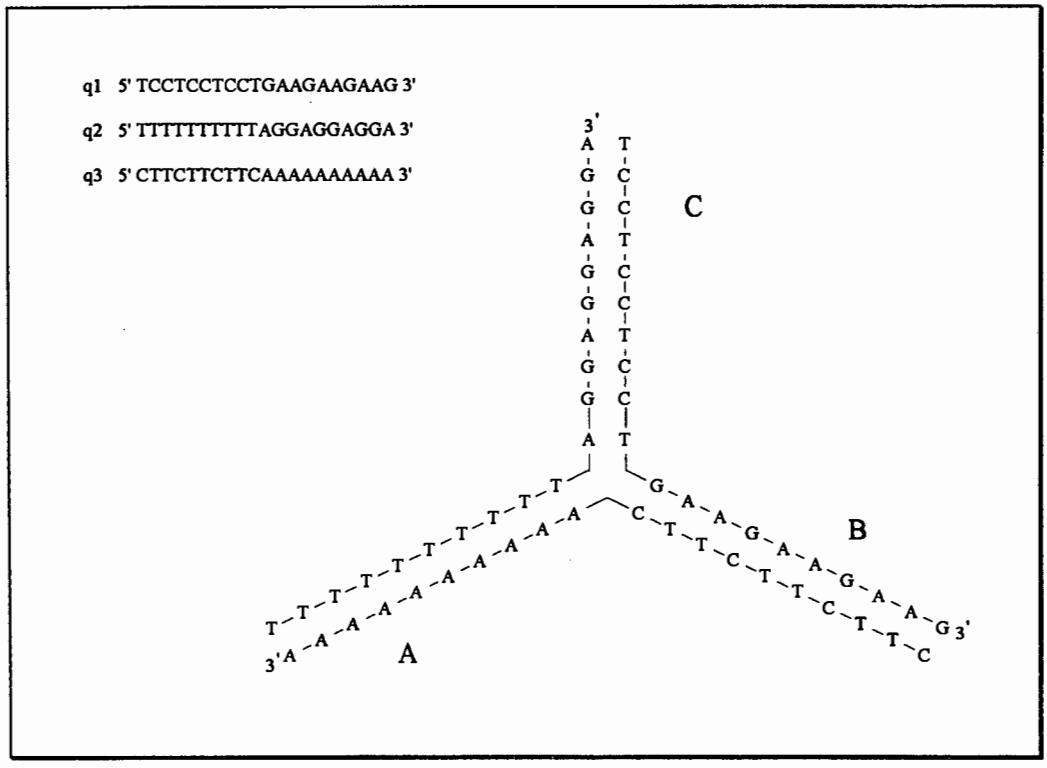

Fig. 7.1. Double-helical three-way junction with arms marked A, B, and C. The oligonucleotides q1, q2, and q3 were used to construct the junction.

7.1. Gel Electrophoresis

The results of native gel electrophoresis in 20 % polyacrylamide, pH 8 are shown in Fig. 7.2. Lanes 1-3 correspond to the single strands, where only a single band for each of the oligonucleotides, q1, q2, and q3 is apparent. Two of the oligonucleotides q1 and q3 (lanes 1 and 3) have similar mobilities, while the mobility of q2 (lane 2) is somewhat lower. The reason for this is not clear but it may arise as a result of secondary structure in q2. The pair-wise combinations of the binary complexes q1q2, q2q3, and q1q3 are depicted in lanes 4-6, respectively. It is clear from the single bands in lanes 4-6 that only one molecular species is present in each of these complexes. Lane 7 exhibits the ternary complex (three-way junction), formed by mixing an equal molar amount of q1, q2, and q3.

Fig 7.2 20 % native polyacrylamide gel electrophoresis at 4 °C. Lanes 1, 2, and 3, single strands q1, q2, and q3, respectively; Lanes 4, 5, and 6 two-strand combinations q1q2, q2q3, and q1q3, respectively. An equilmolar mixture of q1, q2, and q3 was loaded on lanes 7. Solvents contained 150 mM NaCl, 89 mM Tris, and 89 mM boric acid, pH 8.

The three-way junction exhibits a single band with anomalous mobility relative to the single strands (lanes 1-3) as well as the binary complexes (lanes 4-6).

7.2 Thermal Denaturation of the Double-Helical Three-Way Junction

The thermal unfolding of the three-way junction (ternary complex) and its isolated arms (binary complexes) was obtained by recording the uv (260 nm) absorbance changes as a function of temperatures in 20 mM Na_2HPO_4 and 1 M NaCl, pH 7. Fig. 7.3 A-C represents the thermal denaturation of all four complexes. The binary complexes undergo simple monophasic thermal unfolding at similar temperatures (A-C). Similarly, the ternary complex unfolds in a monophasic way, but at a lower temperature. The unfolding of the isolated arms take place in the range 60-80 °C, while the three-way junction undergoes thermal unfolding between 30 and 50 °C. The lower overall thermal stability of the three-way junction, relative to the isolated arms, can be explained by assuming sizable unstacking of base pairs in the arms at the branch point. Zhong and Kallenbach (1993) have shown by calorimetry that the dissociation enthalpy changes associated with a duplex containing a short neck, which represents the three-way junction, decreases as the neck is extended despite the potential formation of additional base pairs. Chemical footprinting experiments have given extra evidence that conformational strain around the branch point may lead to the loss of hydrogen bonding in the arms (Lu et al. 1991; Duckett and Lilley 1990).

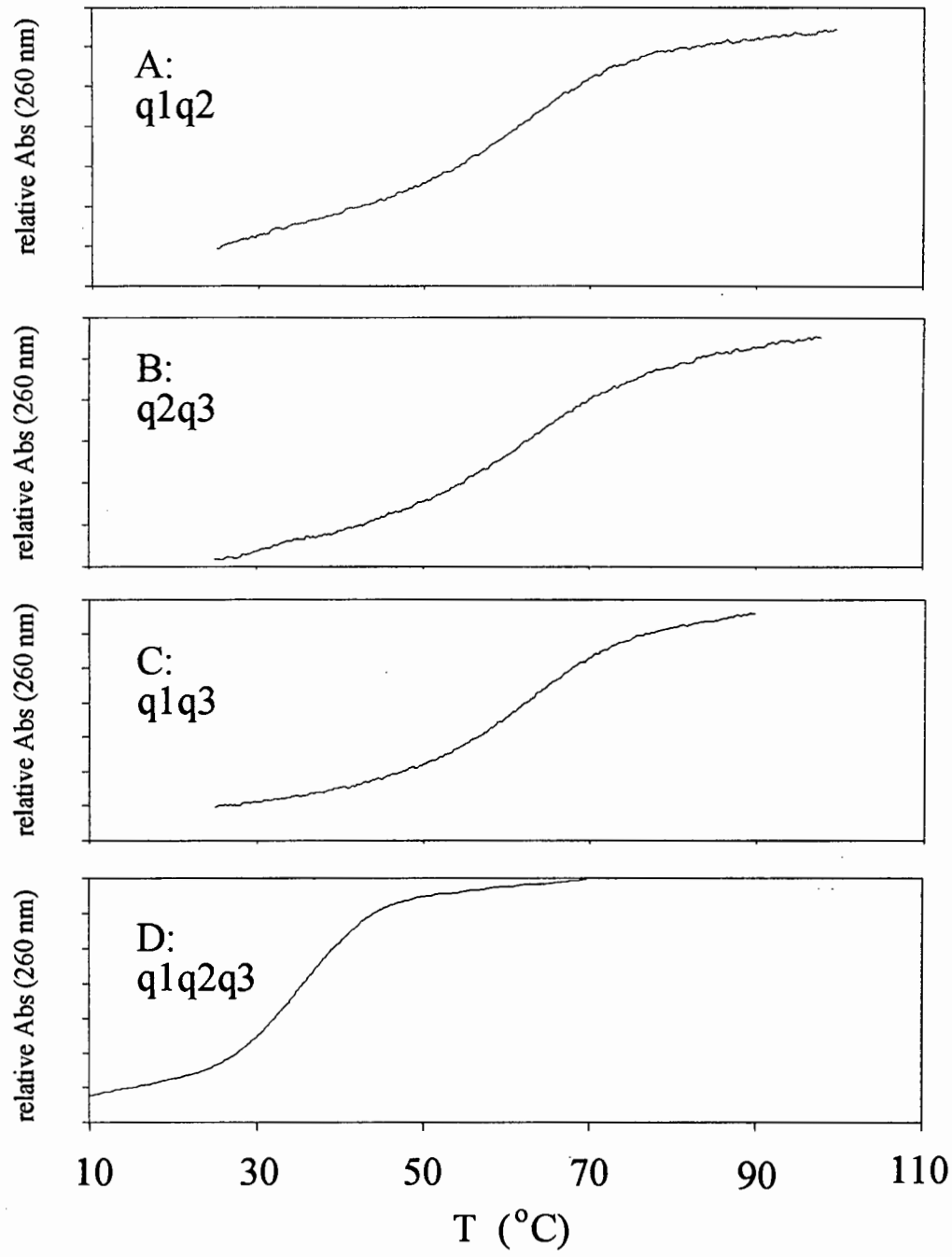

Fig. 7.3 uv (260 nm) thermal denaturation curves of the binary complexes A: q1q2, B: q2q3, C: q1q3 which represent the isolated arms of the three-way junction. The thermal denaturation of the three-way junction q1q2q3 is depicted in D. Buffers contained 20 mM Na_2HPO_4 and 1M NaCl, pH7.

Table VIII

Thermodynamic Parameters for the Double-Helical Three-Way Junction and Binary Complexes

Complex	ΔH_1 [a]	ΔS_1 [b]	Tm_1 [c]	ΔH_2 [a]	ΔS_2 [b]	Tm_2 [c]	ΔH_3 [a]	ΔS_3 [b]	Tm_3 [c]
q1.q2.q3 [d]	35.4	112.0	316.2	56.0	178.8	333.4	62.8	205.7	322.2
q1.q2.q3 [d]	30.8	96.5	319.3	59.1	189.1	331.5	69.3	225.4	323.0
Average:	**33.1**	**104.2**	**317.8**	**57.6**	**184.0**	**332.5**	**66.1**	**215.6**	**322.6**
q1q2 [e]	59.3	175.5							
q2q3 [e]							62.4	190.7	
q1q3 [e]				61.74	182.8				

van't Hoft enthalpy and entropy for the three-way junction
[e] $\Delta H_{VH} = 118$ [e] $\Delta S_{VH} = 372.9$ [e] $Tm = 317.9$

Binary combinations	Three-way junction
$\Delta H_1 + \Delta H_2 + \Delta H_3 = 183.4$	$\Delta H_1 + \Delta H_2 + \Delta H_3 = 156.8$
$\Delta S_1 + \Delta S_2 + \Delta S_3 = 549.0$	$\Delta S_1 + \Delta S_2 + \Delta S_3 = 503.8$

Thermodynamic parameters for the three-way junction q1q2q3 and the binary complexes were obtained by DSC and uv thermal denaturation. Enthalpy and entropy values (ΔH_i and ΔS_i) for the three-way junction were calculated by fitting Eq [3.27] to DSC data, and van't Hoft enthalpy and entropy changes from uv melting, assuming a two-state model. Enthalpy and entropy changes associated with the binary complexes were calculated directly from uv melting curves assuming a two-state model

Note: buffers contained 1M NaCl, 20 mM Na_2HPO_4, pH 7. Error margins were between 5 and 10%.
[a] Units for ΔH are $kcal.mol^{-1}$ (1 cal = 4.18 J).
[b] Units for ΔS are $cal.mol^{-1}.K^{-1}$.
[c] $Tm = \Delta H/(\Delta S - \ln Ct)$.
[d] Values determined by DSC.
[e] van't Hoft values determined by uv thermal denaturation.

To determine the extent of unwinding at the branch point, van't Hoft enthalpy and entropy changes for the three-way junction were compared to the sum of the enthalpy changes obtained for the binary complexes (isolated arms). The enthalpy and entropy changes associated with the thermal unfolding of the three-way junction and the isolated arms were calculated by fitting a two state model to uv melting data (cf. chapter 3.1). The results of the fitted variables are listed in Table VIII. From the listed values it is evident that the sum of the van't Hoft enthalpy change (ΔH = 183 kcal/mol) for the three binary complexes is much bigger than the enthalpy change (ΔH = 118 kcal/mol) obtained for the three-way junction.

To determine whether the thermal unfolding of the three-way junction deviates from a two-state process, model independent calorimetric DSC data were obtained (cf. Fig. 7.6). The overall enthalpy change deduced directly from DSC is, surprisingly, 38.8 kcal/mol larger than the van't Hoft enthalpy. Conventionally, differences between the van't Hoft and calorimetric enthalpy changes are expressed as a ratio $r = \Delta H_{VH}/\Delta H_{cal}$. The ratio r, for the double helical three-way junction, amounts to 0.75. This suggests that the unfolding process of the three-way junction is not a two-state process. The overall calorimetric enthalpy change associated with the unfolding of the three-way junction is, however, still smaller than the sum of the enthalpy changes associated with the isolated arms, in line with the assumption of partial unstacking of the arms around the branch point.

To investigate further the extent of non two-state behavior, absorbance vs temperature profiles for the three-way junction were determined at two different wavelengths. In the case of true two-state thermal denaturation, normalized melting curves should produce identical profiles. However, in the case of non two-state thermal unfolding, normalized melting profiles usually result in curves that are not superimposable.

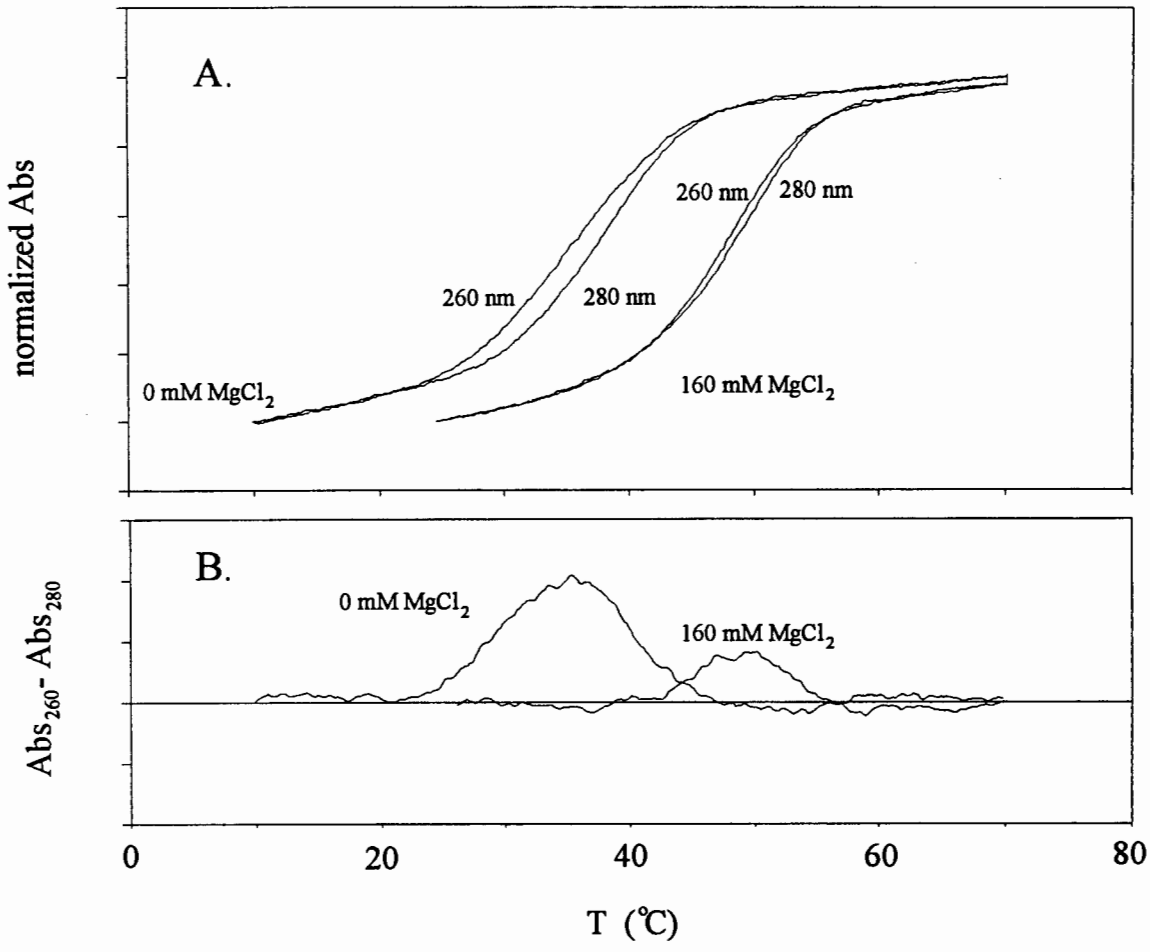

Fig. 7.4 A: uv thermal denaturation curves of the double-helical three-way junction at 260 and 280 nm in 1 M NaCl, 20 mM sodium acetate in either 0 or 160 mM Mg^{2+}. B: depicts the difference curves (Abs 260nm - Abs 280 nm) at 0 and 160 mM Mg^{2+}.

This can be attributed to the presence of intermediate species that contribute differently to the overall normalized absorbance. Fig. 7.4 A depicts normal uv thermal denaturation curves for the three-way junction in either the absence of magnesium, or in 160 mM Mg^{2+} (at 260 and 280 nm). At 0 mM Mg^{2+}, the 260 and 280 nm uv melting profiles are not superimposable. This is clearly illustrated by the difference between the curves in Fig. 7.4 B. On increasing the Mg^{2+} concentration to 160 mM, differences between the melting profiles at 260 and 280 nm are much smaller, suggesting that the complex approaches two-state unfolding behavior (cf. also difference curves in Fig. 7.4 B).

7.3 Model Calculations for the Thermal Unfolding of the Three-Way Junction

To deduce characteristics of the unfolding process, a theoretical model is proposed which describes the unfolding pathway of the three-way junction. It should be noted that the proposed model is based on the formalism developed for the WC triple-helical three-way junction in chapter 5. Fig. 7.5 shows a schematic representation of the unfolding pathway of the double-helical three-way junction. F1 represents the fully folded three-way junction in which one of the arms unfolds, forming the intermediate species F2. Subsequently a second arm in F2 can then dissociate to form the intermediate F3 and a single-stranded oligonucleotide (F4). The molecular species F3 finally dissociates into its constituent single strands F5 and F6. The unfolding steps are represented by the equilibrium constants K1, K2, and K3, respectively.

To extract the exact enthalpy and entropy changes associated with each step, Eq. [3.28] was fitted to the excess heat capacity data set (DSC), by adjusting the enthalpy and entropy parameters using the Levenberg-Marquart method of least squares.

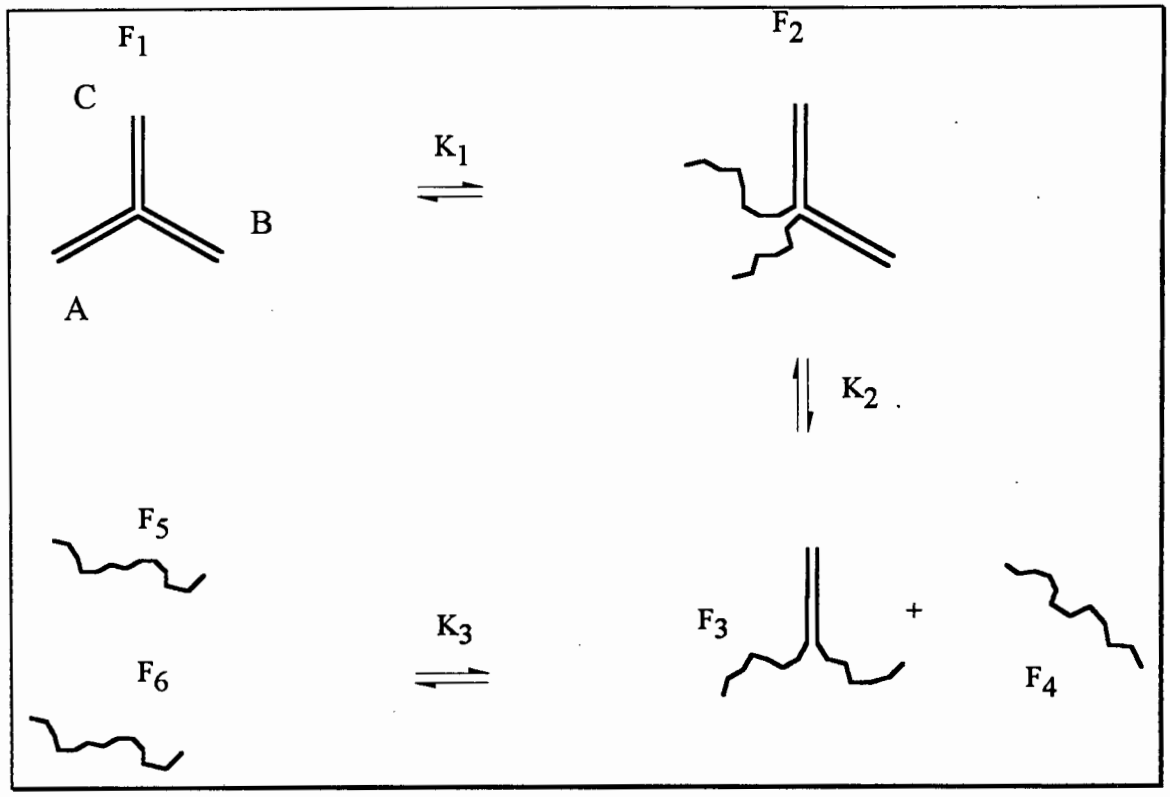

Fig. 7.5 Reaction scheme that describes the unfolding pathway of the WC double-helical three-way junction with species F2 to F6. K1-K3 represent the equilibrium constants for each step.

The envelope curve in Fig 7.6 represents the experimental data, while the smooth curve passing through the experimental data, represents the fitted curve. The underlying peaks demonstrate the deconvoluted sub-transitions associated with the equilibrium constants K1, K2, and K3, respectively (cf. Eqs. [3.29], [3.30], and [3.31]). It is apparent from the results that the three-way junction undergoes thermal unfolding in at least two stages. The

first step takes place at approximately 50 °C, while the two other transitions take place at approximately 55 °C. It should be noted that the upper temperature transitions have identical profiles, differing only in the area under these peaks.

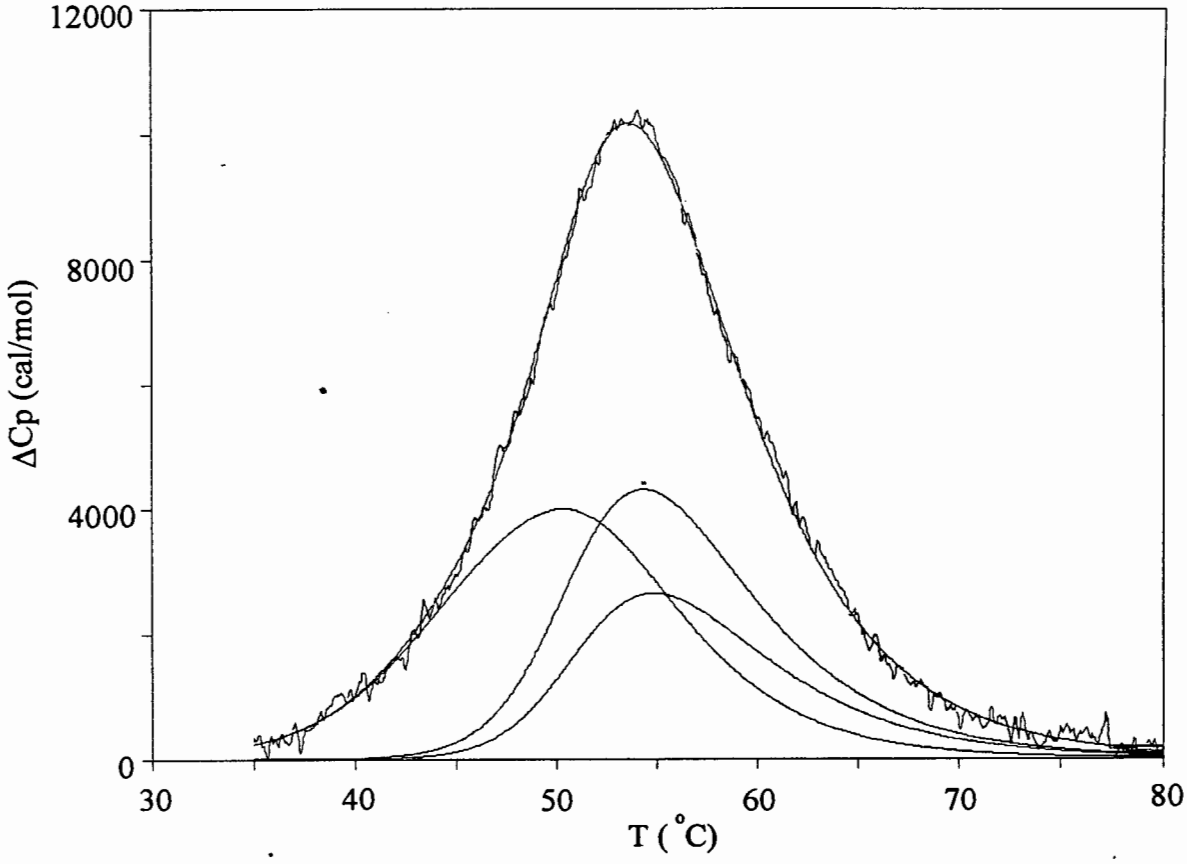

Fig. 7.6 Excess heat capacity (ΔCp) vs temperature curves for the denaturation of the three-way junction and deconvoluted enthalpy contributions for the arms A, B, and C at pH 7. Buffers contained 20 mM Na_2HPO_4 and 1 M NaCl.

This further suggests that the latter two transitions may in fact melt in a cooperative fashion. Comparing the enthalpy changes associated with the isolated arms (binary combinations) and the deconvoluted values obtained by calorimetry for the junction, it seems that at 1 M NaCl one arm in the junction is partially unstacked. This is shown by the

finding that the unfolding enthalpy for each isolated arm amounts to about 60 kcal/mol on average, while deconvolution of DSC curves gives 33.1, 57.6, and 66.1 kcal/mol for the arms A, B, and C of the three-way junction, respectively. It is evident that the unfolding of two of the arms has similar enthalpy values to the isolated arms, while the enthalpy change of the third arms is considerably smaller. It was not possible to assign the sub-transition to the arms of the three-way junction because of the overlapping transitions.

7.4. Thermal Unfolding Enthalpy ΔH_{VH} is Dependent on Mg^{2+} Concentration

It was shown that the thermal unfolding of the double helical three-way junction approaches a two-state process in the presence of Mg^{2+} ions (cf. Fig. 7.4 B). To investigate this effect in more detail a series of uv melting experiments was performed between 0 and 90 mM Mg^{2+}. Fig. 7.7 depicts the calculated van't Hoft enthalpy changes assuming a two-state model. It is apparent that the enthalpy increases initially with increasing Mg^{2+} concentration between 0 and 25 mM Mg^{2+}, and levels off between 40 and 100 mM. The increase in ΔH_{VH} from 0 to 100 mM Mg^{2+} amounts to about 50 kcal/mol. At 100 mM the van't Hoft enthalpy change has reached 170 kcal/mol, which closely approximates the sum of the enthalpy changes obtained for the isolated arms (ΔH = 183 kcal/mol). This suggests that the addition of magnesium ions to the solution (at relatively large concentrations) reduces unfolding in at least one of the arms.

To sum up, the double-helical three-way junction seems to be a conformationally strained structure with partial unfolding of at least one of the arms. Furthermore, thermal unfolding of the three-way junction, in the absence of Mg^{2+}, cannot be described by a two-state model. Thus caution should be exercised when trying to characterize these structures thermodynamically.

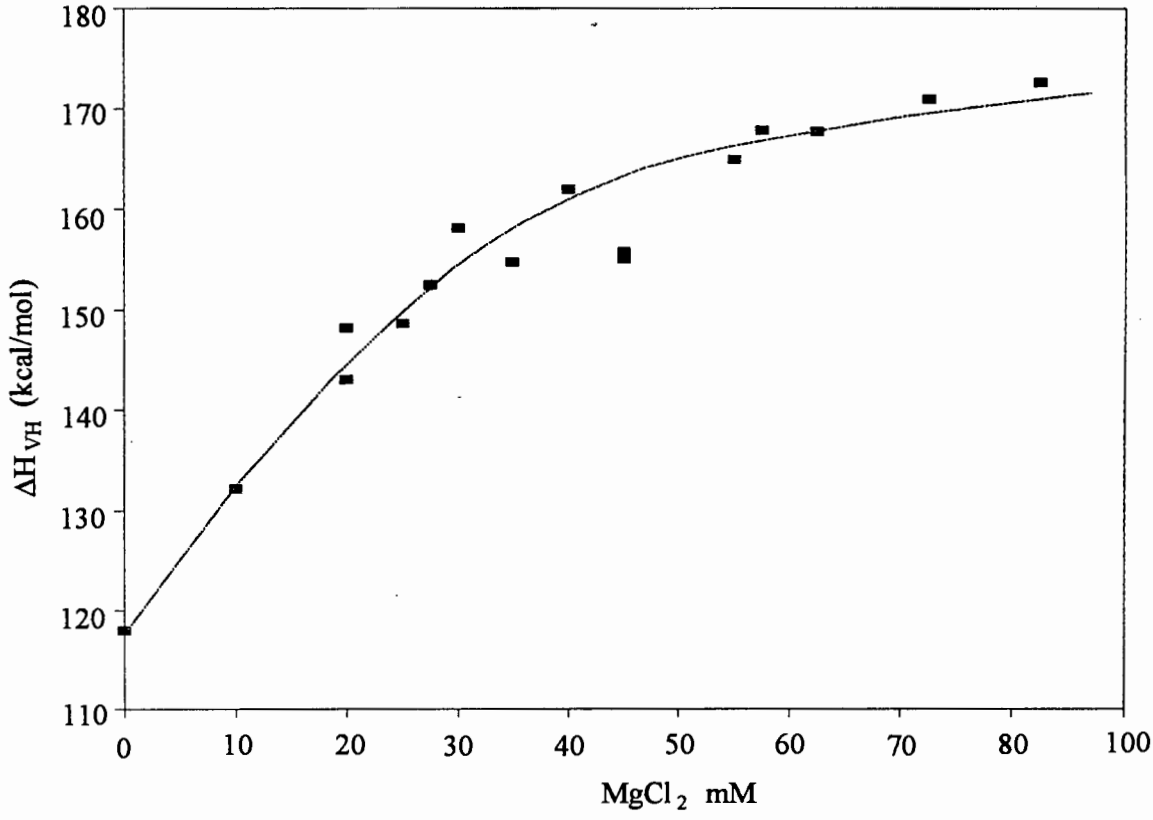

Fig. 7.7 Plot of the van't Hoft enthalpy changes associated with the unfolding of the three-way junction. The van't Hoft enthalpy changes were determined from uv thermal denaturation curves (cf. chapter 3.1).

Chapter 8.

Summary

In this thesis it is shown that triple-helices, like double-helices, can be incorporated into branched DNA structures. Two types of triple-helical three-way junctions are characterized thermodynamically and shown to form stable complexes under conditions of varying pH. In the case of the WC triple-helical three-way junction the branch point is formed through the WC strands, while in the HG triple-helical three-way junction the branch point is formed by the Hoogsteen strands. The sequences that constitute the arms are selected so that the CGC^+ triad composition increases from arms A to C. This makes them increasingly pH dependent. Unlike double-helical three-way junctions that have been studied previously, triple-helical three-way junctions require acidic conditions to promote complete folding. These structures are characterized by a number of physical and biochemical techniques, such as PAGE, temperature gradient gel electrophoresis, CD, uv-spectroscopy, and differential scanning calorimitry.

Thermal unfolding of the WC and HG triple-helical three-way junctions are shown to proceed through a number of intermediate states. In the case of the WC three-way junction, each of the arms unfolds independently, resulting in three sub-transitions. As expected from the sequence composition, arms B and C exhibit a pH-dependent thermal unfolding, while arm A exhibits pH-independent unfolding. In contrast to the results obtained for the WC junction, the HG triple-helical three-way junction unfolds in two temperature regions. At low temperatures the arms unfold due to the dissociation of the

Hoogsteen strands, while the resulting hairpins unfold in the high temperature region. Only the low temperature transition for the HG three-way junction is pH-dependent. It is shown that the unfolding of the HG junction can be brocken down into six stages. Three transitions occur in the low temperature region with sequential unfolding of the triple-helical arms, and three in the high temperature region attributed to the unfolding of the hairpins. These results, in conjunction with CD spectroscopy and pH titration, support the proposed unfolding pathway for both the WC and HG triple-helical three-way junctions.

Models using statistical thermodynamics are proposed, which describe the unfolding pathway for the WC, HG and double helical three-way junctions. The proposed models are used to extract the exact enthalpic and entropic contributions for each of the sub-transitions by means of nonlinear regression. The deconvoluted enthalpy changes for the integrated arms of the WC triple-helical three-way junction and the isolated triple-helices representing the arms, show a systematic difference between the different sets of enthalpy changes. The experimentally obtained enthalpy changes associated with arms A and C are considerably smaller than the corresponding changes obtained for the isolated arms, indicating that the arms are partially unstacked around the branch point. In contrast, the enthalpy change associated with arm B is similar to the value obtained for the corresponding isolated arm, suggesting that the two adjoining triple-helices (arms A and C) do not have a noticeable destabilizing effect on arm B. In the case of the double-helical three-way junction, only one of the arms is shown to be partially unstacked. This result is in accordance with the work done by other groups, where it was shown that double-helical three-way junctions are inherently strained structures.

The enthalpy changes associated with each of the integrated arms of the HG three-way junction match the enthalpy changes associated with the isolated triple-helical arms. This suggests that the arms of the HG three-way junction are completely folded and unperturbed around the branch point.

Finally, it is shown that the WC triple-helical and double helical three-way junctions differ considerably from the HG triple-helical three-way junction. A further result here is that the double-helical and WC three-way junctions are inherently strained structures, resulting in partial unwinding around the branch point. This is in contrast to the situation in the HG triple-helical three-way junction which does not seem to be conformationally strained and the arms remain completely folded.

Tables and Figures

Figures

Chapter 1.

Fig. 1.1 Schematic representation of an immobile junction showing the extended structure (A) and the two stacked isomers corresponding to the antiparallel and parallel structures (B and C, respectively). The arrows indicate the 5' to 3' orientation of the oligonucleotides 1-4. Each of the oligonucleotides participate in forming the double-helical arms of the junction. In this scheme only two of the four possible stacked isomers are shown, where arm I is stacked onto arm II, and arm III onto arm IV. (6)

Fig. 1.2 Top and side views of the antiparallel stacked X like four-way junction, adopted from Murchie, A. I. H., Portugal, J., and Lilley, D. M. J. (1991) *EMBO.* **10**, 713-718. (8)

Fig. 1.3 Schematic drawing of the five double crossover DNA structures. The structures are named according to their basic characteristics. All the abbreviations start with the letter D for double crossover. The letter that follows indicates the orientation of the double-helices, A for antiparallel and P for parallel. The third letter refers to the number of half helical turns between the double crossover regions, O for odd and E for even. In parallel double crossover structures with an odd number of half helical turns two structures are possible. The last letter refers to a major half helical turn separation (W for wide) or a minor half helical separation (N for narrow). Figure adopted from Fu, T., and Seeman, N. C. (1993) *Biochemistry.* **32**, 3211-3220. (15)

Fig. 1.4 Schematic representation of the three- and four-stranded junctions, antijunctions, and mesojunctions. (A) Each polygon is flanked by the single DNA strand, the arrow heads of which indicate the polarity of the strands. The parallel lines at the vertices of the polygon represent the base pairs and the short arrows, perpendicular to the base pairs, the helical axis of each duplex region. B represents drawings of the schematically presentation of the structures in A. Figure adopted from Wang, H., and Seeman, N. C. (1995) *Biochemistry.* **34**, 920-929. (18)

Fig. 1.5 Schematic representation of a three-dimensional double-helical cube adopted from Chen. J. H., and Seeman, N. C. (1991) *Nature (London).* **350**, 631-633. (26)

Fig. 1.6 (Top) Watson-Crick base pairs. (Bottom) triplet base pairs TAT and CGC^+. The additional pyrimidine strand is bound by Hoogsteen hydrogen bonds in the major groove of the Watson-Crick duplex. (31)

Chapter 2.

Fig. 2.1 (Top) HG and WC triple-helical three-way junctions. (Bottom) Mixed WC_2HG_1 and WC_2HG_1 triple-helical three-way junctions. Thick lines represent the Hoogsteen strands and the thin lines the Watson-Crick strands. The arrows indicate the 5'-3' polarity of the single strands. The notation refers to the type of strand connection at the branch point. HG refers to Watson-Crick and Hoogsteen connections between strands, while WC refers to connections between Watson-Crick strands only. The subscript indicates the number of HG and WC connections in mixed structures (bottom). The subscript has been dropped in structures containing only one type of connection (top). (34)

Fig. 2.2 Schematically representation of the double-helical three-way junction J, and the WC and HG triple-helical three way junctions. The thick lines represent the WC double-helical regions and the thin lines the triplex HG strands. The sequences in each of the arms A, B, and C are the same for all three junctions with only the connection at the branch point and the peripheral ends of the arms (forming the hairpin or looped regions) being different. (36)

Chapter 3.

Fig. 3.1 Reaction scheme that describes the unfolding process of a binary complex with species Fa to Fc'. K_1, K_2, and K_2 represent the equilibrium constants for each step. (42)

Fig. 3.2 Reaction scheme that describes the unfolding pathway for the WC triple-helical three-way junction with intermediate species F1 to F6. K_1-K_3 represent the equilibrium constants for each step. (50)

Fig. 3.3 Reaction scheme describing the unfolding pathway of the ternary complex with species F1-F9. K_1-K_6 represent the equilibrium constants for the individual steps. The boxed reaction pathway represents the unfolding steps of the binary complex. The molecular species and equilibrium constants for the boxed reaction pathway are presented in round brackets for (K_1)-(K_3), and (Fa)-(Fc'), respectively. (58)

Fig. 3.4 Purine oligonucleotides p6 and p8 (A) and Pyrimidine oligonucleotides p7 and p9 used in the triple-helices p6p7 and p8p9 (B). The Hoogsteen and Watson-Crick bases are abbreviated as HG and WC respectively. (67)

Chapter 5.

Fig. 5.1 WC triple-helical three-way junction with arms marked A, B, and C. The oligonucleotides used to construct the junction are p1, p2, and p3. The Hoogsteen and Watson-Crick bases are abbreviated as HG and WC, respectively. (98)

Fig. 5.2 Purine oligonucleotides p5, p7, and p9 and pyrimidine oligonucleotides p4, p6, and p8 were used to form the triple-helices p4p5, p6p7, and p8p9. The Hoogsteen and Watson-Crick bases are abbreviated as HG and WC, respectively. (99)

Fig. 5.3 20% native polyacrylamide gel electrophoresis at 20 °C. The values at the bottom of the gel represent the molar fractional amounts of the oligonucleotides p6 and p7, respectively. Solvents contained 100 mM sodium acetate, pH 5. (100)

Fig. 5.4 20 % native polyacrylamide gel electrophoresis at 4 °C. Lanes 1, 2, and 3, single-strand p1, p2, and p3, respectively; lanes 4, 5, and 6, two-strand combinations p1p2, p2p3, and p1p3, respectively. An equimolar mixture of p1p2p3 was loaded in lane 7. Solvents contained 150 mM NaCl, 89 mM Tris, and 89 mM boric acid, pH 8. The various components stain with different efficiency. (102)

Fig. 5.5 Temperature gradient gel electrophoresis of the triplex p6p7 in 100 mM ammonium acetate and 50 mM sodium acetate at about pH 5.5. (103)

Fig. 5.6 Temperature gradient gel electrophoresis of the WC triple-helical three-way junction (p1p2p3) in 100 mM ammonium acetate and 50 mM sodium acetate at about pH 5.5. (104)

Fig. 5.7 pH titration of p6p7 (□) and p8p9 (■) in 20 mM Na_3PO_4 and 1M NaCl at 20 °C. The absorbance was measured at 260 nm by increasing the pH using microliter quantities of a 1 M NaOH solution. (106)

Fig. 5.8 pH titration of p1, p2, and p3 (1:1:1) in 20 mM Na_3PO_4 and 1 M NaCl at 20 °C. The absorbance was measured at 260 nm after adjusting the pH and equilibrating for 15 min. (107)

Fig. 5.9 CD spectrum of the WC triple-helical three-way junction p1p2p3 at temperatures 100 °C and 20 °C in 20 mM Na_3PO_4 and 1 M NaCl, at pH 4.5. (109)

Fig 5.10 Ultraviolet thermal denaturation (260 nm) of the triplex p4p5, recorded at pH 7 in 20 mM Na_3PO_4 and 1 M NaCl. (111)

Fig. 5.11 Ultraviolet thermal denaturation (260 nm) curves for the triplex p6p7 recorded at pH 6.0, 5.6, and 5. Buffers contained 20 mM Na_3PO_4 and 1 M NaCl. (111)

Fig. 5.12 Ultraviolet thermal denaturation (260 nm) curves for the triplex p8p9 recorded at pH 5.8, 5.55, and 4.8. Buffers contained 20 mM Na_3PO_4 and 1 M NaCl. The insert represents the first derivative of the absorbance vs temperature curve at pH 5.5. (112)

Fig 5.13 Differential uv (260 nm) melting curves for the triple-helical three-way junction (p1p2p3) as recorded under various pH conditions. Buffers contained 20 mM Na_3PO_4 and 100 mM NaCl. (113)

Fig. 5.14 Phase diagram representing 1/Tm ($Kelvin^{-1}$) vs pH for the triple-helices p6p7 (■) and p8p9 (□). Tm values were determined from uv (260 nm) thermal denaturation curves under various pH conditions. The solid line represent the theoretical curves for the model of nearest neighbor proton binding sites. All buffers contained 20 mM Na_3PO4 and 1 M NaCl. (115)

Fig. 5.15 Phase diagram representing 1/Tmax vs pH obtained from differential uv melting curves (260 nm) for the triple-helical three-way junction. Buffers contained 20 mM Na_3PO_4 and 500 mM NaCl. Tmax was determined where *dA/dT* reaches a maximum. (117)

Fig. 5.16 Phase diagram representing Tmax vs Log[Na+] from differential uv melting curves (260 nm) for the three-way junction. Buffers contained 20 mM Na_3PO_4, pH 4.7. (119)

Fig. 5.17 Excess heat capacity (ΔCp) vs temperature for the triplexes p4p5 (A), p6p7 (B), and p8p9 (C). The buffers contained 20 mM Na_3PO_4 and 1M NaCl at the pH values indicated in the figure. The oligonucleotide concentration (purine + pyrimidine) were 0.515, 0.236, and 222 μM for the triple-helices p4p5, p6p7, and p8p9, respectively. (121)

Fig. 5.18. Excess heat capacity (ΔCp) vs temperature curves for the denaturation for the three-way junction and deconvoluted enthalpy contributions of arms A, B, and C at (A) pH 5 and (B) pH 4.5. Buffers contained 20 mM Na_3PO_4 and 1 M NaCl. The theoretical curve in A is left out as it is superimposed completely on the experimental curve. (129)

Fig. 5.19 Reaction scheme that describes the unfolding pathway of the WC triple-helical three-way junction with species F1 to F6. K_1, K_2, and K_3 are the equilibrium constants for each step. (131)

Fig. 5.20 Differential uv melting curves of the WC triple-helical three-way junction. The concentration of the oligonucleotides p1, p2, and p3 were (A) 14.25μM and (B) 1.14 μM. Buffers contained 20 mM Na_3PO_4 and 1M NaCl at pH 4.5. (132)

Fig. 5.21 (A) Absorbance (260 nm) vs temperature profile for the three-way junction and the fitted curves in 20 mM Na_3PO_4 and 1 M NaCl, pH 4.5. (B) Normalized populations of F1 to F6 as a function of temperature. (133)

Chapter 6.

Fig. 6.1 The HG triple-helical three-way junction (hp1hp2hp3) represents the ternary complex with arms marked as A, B, and C. The complexes hp1p2, hp2p3, and hp1hp3 depict the binary complexes. The binary complexes contain three regions namely: a single stranded extension, a hairpin, and triple helix. The triple-helical region of the binary combinations corresponds to the separate arms of the ternary complex. The Watson-Crick and Hoogsteen bases are abbreviated as WC and HG, respectively. (140)

Fig. 6.2 20% native polyacrylamide gel electrophoresis at 4 °C, pH 5.5. Lanes 7, 6, and 5, oligonucleotides hp1, hp2, hp3, respectively; Lanes 4, 3, and 2, binary combinations hp1hp2, hp2hp3, and hp1hp3, respectively. An equal molar mixture of hp1, hp2, and hp3 was loaded in lane 1. Solvents contained 100 mM Sodium acetate, 50 mM $MgCl_2$. The various components stain with different intensities. (142)

Fig. 6.3 Temperature gradient gel electrophoresis of the ternary complex (hp1hp2hp3) in 100 mM Sodium acetate and 50 mM $MgCl_2$ at pH 5.5. (144)

Fig. 6.4 pH titration of an equal molar mixture of hp1, hp2, and hp3 representing the ternary complex in 1 M NaCl and 20 mM Na_2HPO_4 at 20 °C. The absorbance was measured at 260 nm after adjusting the pH. (146)

Fig. 6.5 Differential uv (260 nm) thermal denaturation curves of the binary complexes hp1hp2, hp2hp3, hp1hp3, and the ternary complex hp1hp2hp3. Buffers contained 100 mM sodium acetate, 50 mM magnesium acetate at pH 5.5. (148)

Fig. 7.1. Double-helical three-way junction with arms marked A, B, and C. The oligonucleotides q1, q2, and q3 were used to construct the junction. (169)

Fig. 7.2 20 % native polyacrylamide gel electrophoresis at 4 °C. Lanes 1, 2, and 3, single strands q1, q2, and q3, respectively; Lanes 4, 5, and 6 two-strand combinations q1q2, q2q3, and q1q3, respectively. An equilmolar mixture of q1, q2, and q3 was loaded on lanes 7. Solvents contained 150 mM NaCl, 89 mM Tris, and 89 mM boric acid, pH 8. (170)

Fig. 7.3 uv (260 nm) thermal denaturation curves of the binary complexes A: q1q2, B: q2q3, C: q1q3 which represent the isolated arms of the three-way junction. The thermal denaturation of the three-way junction q1q2q3 is depicted in D. Buffers contained 20 mM Na_2HPO_4 and 1M NaCl, pH7. (172)

Fig. 7.4 A: uv thermal denaturation curves of the double-helical three-way junction at 260 and 280 nm in 1 M NaCl, 20 mM sodium acetate in either 0 or 160 mM Mg^{2+}. B: depicts the difference curves (Abs 260nm - Abs 280 nm) at 0 and 160 mM Mg^{2+}. (175)

Fig. 7.5 Reaction scheme that describes the unfolding pathway of the WC double-helical three-way junction with species F2 to F6. K1-K3 represent the equilibrium constants for each step. (177)

Fig. 7.6 Excess heat capacity (ΔCp) vs temperature curves for the denaturation of the three-way junction and deconvoluted enthalpy contributions for the arms A, B, and C at pH 7. Buffers contained 20 mM Na2HPO4 and 1 M NaCl. (178)

Fig. 7.7 Plot of the van't Hoft enthalpy changes associated with the unfolding of the three-way junction. The van't Hoft enthalpy changes were determined from uv thermal denaturation curves (cf. chapter 3.1). (180)

Tables

Chapter 5.

Chapter 6.

Table VIII.
Thermodynamic Parameters for the Double-Helical Three-Way Junction and Binary Complexes. (173)

References

Anderson, F. C., and record, M. T., Jr. (1982)
Polyelectrolyte theories and their application to DNA.
Annu. Rev. Phys. Chem. **33**, 191-222.

Arnott, S., Bond, P. J., Selsing, E. and Smith, P. J. C. (1976)
Model of triple-stranded polynucleotides with optimized stereochemistry.
Nucleic Acids. Res. **3**, 2459-2470.

Arnott, S., and Selsing, E. (1974)
Structure for the polynucleotide complex ploy(dA) · poly(dT) and ploy(dA)·poly(dT)·ploy(dT)
J. Mol. Biol. **88,** 509-521.

Arnott, S., and Bond, P. J. (1973)
Structure for poly(U)·Poly(A)·poly(U) triple stranded polynucleotides
Nature New Biol. **224**,244, 99-101.

Beaucage, S. L., and Caruthers, M. H. (1981)
Deoxynucleoside phosphoramidites - A new class of key intermediates for deoxypolynuceotide synthesis.
Tetrahedron Lett. **22**, 1859-1862.

Bevington, R. P., (1969)
in Data Reduction and Error Analysis for the Physical Sciences.
McGraw-Hill, New York.

Blake, R. D., Massoulie, J., and Fresco, R. J. (1967)
Polynucleotides VIII. a spectral approach to the equilibria between polyadenylate and ployribouridylate and their complexes.
J. Mol. Biol. **30**, 291-308.

Broitmann, S. L., Im, D. D., and Fresco. J. R. (1987)
Formation of the triple stranded polynucleotide helix, poly (AAU)
Proc. Natl. Acad. Sci. USA. **84**, 5120-5124.

Bush, C. A. (1974)
Ultraviolet spectroscopy, Circular dichorism, and optical rotatory dispersion. In basic Principles
in Nucleic Acid Chemistry (Ts'o, P. O. P., ed.) vol. 2, pp. 91-169, Academic Press, NY.

Cantor, C. R., Warshaw, M. M., and Shapito, H. (1970)
oligonucleotide interaction. III. circular dichroism studies of the conformation of deoxyoligonucleotides.
Biopolymers. **9**, 1059-1077.

Chen, J. H., Churchill, M. E. A., Tullius, T. D., Kallenbach, N. R., and Seeman, N. C. (1988)
Construction and analysis of monomobile DNA junctions.
Biochemistry. **27**, 6032-6038.

Chen, J. H., Kallenbach, N. R., and Seeman, N. C. (1989)
A specific quadrilateral synthesized from DNA branched junctions.
J. Am. Chem. Soc. **111**, 6402-6407.

Chen, J. H., and Seeman, N. C. (1991)
Synthesis from DNA of a molecule with the connectivity of a cube.
Nature (London). **350**, 631-633.

Chen, S. M., Heffron, F., Leupin, W., and Chazin, W. J. (1991)
Two-dimentional H NMR studies of synthetic immobile holliday junctions.
Biochemistry. **30**, 766-771.

Clegg, R. M., Murchie, A. I. H., Zechel, A., Carlberg, C., Diekmann, S., and Lilley, D. M. J. (1992)
Fluorescence resonance energy transfer analysis of the structure of the four-way DNA junction.
Biochemistry. **31**, 4846-4856.

Clery, D. (1995)
DNA goes electric
Science. **267**, 1270

Cooper, J. P., and Hagerman, P. J. (1987)
Gel Electrophoretic analysis of the geometry of a DNA four-way junction.
J. Mol. Biol. **189**, 711-719.

Cooper, J. P., and Hagerman, P. J. (1989)
Geometry of a branched DNA structure in solution.
Proc. Natl. Acad. Sci. USA. **86**, 7336-7340.

Cooper, J. P., and Hagerman, P. J. (1990)
Analysis of flourescence energy transfer in duplex and branched DNA molecules.
Biochemistry. **29**, 9261-9268.

Churchill, M. E. A., Tullius, T. D., Kallenbach, N. R., and Seeman, N. C. (1988)
A Holliday recombination intermediate is twofold symmetric.
Proc. Natl. Acad. Sci. **85**, 4653-4656.

Distefano, M. D., Shin, J. A., and Dervan, P. B. (1991)
Cooperative binding of oligonucleotides to DNA by triple-helix formation: Dimerization via Watson-Crick hydrogen bonding.
J. Am. Chem. Soc. **113**, 5901-5902.

Distefano, M. D., Shin, J. A., and Dervan, P. B. (1992)
Ligand-promoted dimerizarion of oligonucleotides binding cooperatively to DNA.
J. Am. Chem. Soc. **114**, 11006-11007.

Du, S. M., Zhang, S., and Seeman N. C. (1992)
DNA junctionsa, antijunctions, and mesojunctions.
Biochemistry. **31**, 10955-10963.

Duckett, D. R., and Lilley, D. M. J. (1990)
The three-way DNA junction is a Y-shaped molecule in which there is no helix-helix stacking.
EMBO. **9**, 1659-1664.

Duckett, D. R., and Lilley, D. M. J. (1991)
Effects of base mismatches on the structure of the four-way DNA junction.
J. Mol. Biol. **221**, 147-161.

Duckett, D. R., Murchie, A. I. H., Diekmann, S., von Kitzing, E., Kemper, B., and Lilley, D. M. J.(1988)
The structure of the Holliday junction, and its resolution.
Cell. **55**, 79-89.

Freire, E. (1994)
Statistical thermodynamic analysis of differential scanning calorimetry data: Structural deconvolution of heat capacity function of proteins
Methods in Enzymology. **240**, 502-530.

Freire, E., and Biltonen, R. L. (1978)
Statistical mechanical deconvolution of thermal transitions in macromolecules. I. Theory and application to homogeneous systems.
Biopolymers. **17**, 463-479.

Freire, E., and Biltonen, R. L. (1978)
Statistical mechanical deconvolution of thermal transitions in macromolecules. II. General treatment of cooperative phenomena.
Biopolymers. **17**, 481-496.

Freire, E., and Biltonen, R. L. (1987)
Statistical mechanical deconvolution of thermal transitions in macromolecules. III. Application to double-stranded to single-stranded transitions of nucleic acids.
Biopolymers. **17**, 497-510.

Fu, T. J., and Seeman, N. C. (1993)
DNA double-crossover molecules.
Biochemistry. **32**, 3211-3220.

Gou, Q., Lu M., Churchill. M. E. A., Tullius, T. D., and Kallenbach, N. R. (1990)
Asymmetric structure of a three-arm DNA junction.
Biochemistry. **29**, 10927-10934.

Gough, G. W., and Lilley, D. M. J. (1985)
DNA bending induced by cruciform formation.
Nature (London). **313**, 154-156.

Giovannangeli, C., Rougee, M., Garestier, T., Thuong, N. T., and Helene, C. (1992)
Triple-helix formation by oligonucleotides containing the three bases thymine, cytosine, and guanine.
Proc. Natl. Acad. Sci. U.S.A. **89**, 8631-8635.

Gray, D. M., Morgan, A. R., and Ratliff, R. L. (1978)
A comparison of the circular dichroism spectra of synthetic DNA sequences of the homopurine.homopyrimidine and Mixed purine-pyrimidine types.
Nucleic. Acids. Res. **5**, 3679-3695.

Gray, D. M., and Ratliff, R. L. (1975)
Circular-dichroism spectra of poly[d(AC)-d(GT)], poly[r(AC)-r(GU)], and hybrids poly [d(AC)-r(GU)], and poly[r(AC)-d(GT)] in the presence of ethanol.
Biopolymers. **14**, 487-498.

Holliday, R. (1964)
A mechanism for gen conversion in fungi.
Genet. Res. **5**, 282-304.

Hüsler, P. L., Klump, H. H. (1994)
Unfolding of a branched double-helical DNA three-way junction with triple-helical ends.
Arch. Biochem. Biophys. **312**, 29-38.

Hüsler, P. L., Klump, H. H. (1995)
Prediction of pH-dependent properties of DNA triple helices.
Arch. Biochem. Biophys. **317**, 46-56.

Hüsler, P. L., Klump, H. H. (1995)
Thermodynamic characterization of a DNA triple-helical three-way junction joined by a Hoogsteen branch point.
Arch. Biochem. Biophys. in Press.

Hample, K. J., Crossen, p., and Lee, J. S. (1991)
Polyamines favor DNA triplex formation at neutral pH.
Biochemistry. **30**, 4455-4459.

Inman, R. B. (1964)
Multistranded DNA homopolymer interactions.
J. Mol. Biol. **10,** 137-146.

Jayasena, S. D. and Johnston, B. H. (1992a)
Oligonucleotide directed triple-helix formation at adjacent oligopurine and oligopyrimidine DNA tracts by alternate strand recognition.
Nucleic. Acids. Res. **20**, 5279-5288.

Jayasena, S. D. and Johnston, B. H. (1992b)
Intramolecular triple-helix formation at $(Pu_nPy_n)\cdot(Pu_nPy_n)$ tracts: Recognition of alternate strands via Pu·PuPy and Py·PuPy base triplets.
Biochemistry. **31**, 320-327.

Jensch, F., and Kemper, B. (1986)
Endonuclease VII resolve Y-junctions in branched DNA in vitro.
EMBO. J. **5**, 181-189

Kallenbach, N. R., Ma, R.I., and Seeman, N. C.(1983)
An immobile nucleic acid junction constructed from oligonucleotides.
Nature (London). **305**, 829-831.

Kimball, A., Guo, Q., Lu, M., Cunningham, R. P., Kallenbach, N. R., Seeman, N. C., Tullius, T. D. (1990)
Construction and analysis of parallel and antiparallel Holliday junctions.
J. Biol. Chem. **265**, 6544-6547.

von Kitzing, E., Lilley, D. M. J., and Diekman, S. (1990)
The stereochemistry of a four-way junction: a theoretical study.
Nucleic. Acids. Res. **18**, 2671-2683.

Ladbury, J. E., Sturtevant, J. M., Leontis, N. B. (1994)
The thermodynamicsof formation of a three-strand, DNA three-way junction complex.
Biochemistry. **33**, 6828-6833.

Lee, J. S., Johnon, D. A., and Morgan, R. A. (1979)
Complexes formed by $(pyrimidine)_n \cdot (purine)_n$ DNA's on lowering the pH are three-stranded.
Nucleic. Acids. Res. **6**, 3073-3091.

Leontis, N. B., Hills, M. T., Piotto, M., Malhotra, A., Nussbaum, J., and Gorenstein, D. G. (1993)
A model for the solution structure of a branched, three-strand DNA complex.
J. Biomol. Struct. Dyn. **11**, 215-223.

Leontis, N. B., Kwok, W., and Newman, J. S. (1991)
Stability and structure of three-way junctions containing unpaired nucleotides.
Nucleic. Acids. Res. **19**, 759-766.

Letai, A. G., Pallidino, M. A., Fromm, E., Rizzo, V. and Fresco, J. R. (1988)
Specificity in formation of triple-stranded nucleic acid helical complexes: studies with agarose-linked polyribonucleotide affinity columns.
Biochemistry. **27**, 9108-9112.

Lipsett, M. (1964)
Complex formation between polycytidylic acid and guanine oligonucleotides.
J. Mol. Chem. **239**, 1256-1260.

Lu, M., Guo, Q., and Kallenbach, N. R. (1991 a)
Effect of sequence on the structure of three-arm DNA junctions.
Biochemistry. **30**, 5815-5820.

Lu, M., Guo, Q., Marky, L. A., Seeman, N. C., and Kallenbach, N. R.(1992)
Thermodynamics of DNA branching.
J. Mol. Biol. **223**, 781-789.

Lu, M., Guo, Q.,Mueller, J. E., Kemper, B., Studier, F. W., Seeman, N. C., and Kallenbach, N. R. (1990a)
Characterization of a bimobile DNA junction.
J. Biol. Chem. **265**, 16778-16785.

Lu, M. Guo, Q., Seeman, N. C., and Kallenbach, N. R. (1991b)
Parallel and antiparallel Holliday junctions differ in structure and stability.
J. Mol. Biol. **221**, 1419-1432.

Lu, M., Guo, Q., Wink, D. J., and Kallenbach, N. R. (1990b)
Charge dependence of Fe(II)-catalyzed DNA cleavage.
Nucleic. Acids. Res. **18**, 3333-3337.

Ma, R. I., Kallenbach, N. R., Sheardy, R. D., Petrillo, M. L., and Seeman, N. C. (1986)
Three-arm nucleic acid junctions are flexible.
Nucleic. Acids. Res. **14** , 9745-9753.

Maher, L. J., Dervan, P. B., Wold, B. J. (1990)
Kinetic analysis of oligodeoxyribonuceitide-directed tripl-helix formation on DNA.
Biochemsitry. **29**, 8820-8826.

Manning, G. S., (1978)
The Molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides.
Quart. Rev. Biophys. **11**, 179-246.

Manzini, G., Xodo, L. E., Gasparotto, D., Quadrifoglio, F., van der Marel, G. A., and van Boom, J. H. (1990)
Triple-helix formation by oligpurine-oligopyrimidine DNA fragments.
J. Mol. Biol. **213**, 833-843.

Marky, L. A., Kallenbach, N. R., McDonough, K. A., Seeman, N, C., and Breslaure, K. J. (1987)
The melting behavoir of a DNA junction structure: a clorimetric and spectroscopic study.
Biopolymers. **26**, 1621-1634.

Marky. L. A., and Breslauer, K. J. (1987)
Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves.
Biopolymers. **26**, 1601-1620.

Massoulie, J., (1968)
Associations de poly A et ploy U en milieu acide.
Eur. J. Biochem. **3**, 439-447.

Me Laughlin, L. W., and Piel, N. (1984)
in Oligonucleotide Synthesis: A practical approach.
(Gait, M. J., Ed.) pp. 117-133, IRL Press Limited, Oxford.

Moser, H. E., and Dervan, P. B. (1987)
Sequence-specific cleavage of double-helical DNA by triple-helix formation.
Science. **238**, 645-650.

Murchie, A. I. H., Carter, W. A., Portugal, J., and Lilley, D. M. J.(1990)
The tertiary structure of the four-way DNA junction affords protection against DNase I cleavage.
Nucleic. Acids. Res. **18**, 2599-2606.

Murchie, A. I. H., Portugal, J., and Lilley, D. M. J. (1991)
Cleavage of a four-way DNA junction by a restriction enzyme spanning the point of strand exchange.
EMBO. **10**, 713-718.

Murchie, A. I. H., Clegg, R. M., von Kitzing, E., Duckett, D. R., Diekmann, S., and Lilley, D. M. J. (1989)
Fluorescence energy transfer shows that the four-way DNA junction is a right-handed corss of antiparallel molecules.
Nature (London). **341**, 763-766.

Olmsted. M. C., and Hagerman, P. J. (1994)
Excess counterion accumulation around branched nucleic acids.
J. Mol. Biol. **243**, 919-920.

Petrillo, M. L., Newton, C. J., Cunningham, R. P., Ma, R. I., Kallenbach, n. R., and Seeman, N. C. (1988)
Ligation and flexibility of four-arm junctions.
Biopolymers. **27**, 1337-1352.

Pilch, D. S., Brousseau, R., and Shafer, R. H. (1990)
Thermodynamics of triple-helix formation: spectroscopic studies on the $d(A)_{10}{\cdot}d(T)_{10}$ and $d(C^{+}_{3}T_4C^{+}_{3}){\cdot}d(G_3A_4G_3){\cdot}d(C_3T_4C_3)$ triple-helices.
Nucleic. Acids. Res. **18**, 5743-5750.

Pilch, D. S., Levenson, C., and Shafer, R. H. (1990)
Structural analysis of the $(dA)_{10}{\cdot}2(dT)_{10}$ triplex.
Proc. Natl. Acad. Sci. USA. **87**. 1942-1946.

Plum, G. E., Park, Y., Singleton, S. F., Dervan, P. B., and Breslauer, K. J. (1990)
Thermodynamic characterization of the stability of the melting behaviour of a DNA triplex: A spectroscopic and calorimetric study.
Proc. Natl. Acads. Sci. USA. **87**, 9436-9440.

Plum, G. E., and Breslauer, K. J. (1995)
Thermodynamics of an intramoleular DNA triple helix: A calorimetric and spectroscopic study oh the pH and salt dependence of thermally induced structural transitions
J. Mol. Biol. **248**, 679-695

Poland, D., and Scheraga, H. A. (1970)
in Theory of Helix-Coil Transitions in Biopoymers.
Academic Press, New York and London.

Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterking W. T.,
in Numerical Recipes in C: The Art of Scientific Computating, pp 542-547.
Cambridge.

Privalov, P. L., and Potekhin, S. A. (1975)
Scanning microcalorimetry in studing temperature-induced changes in proteins.
Methods in Ezymology. **131**, 4-50.

Guo, Q., Lu, M., Churchill, M. E. A., Tullius, T. D., and Kallenbach, N. R. (1990)
Asymmetric structure of a three-arm DNA junction.
Biochemistry. **29**, 10927-10934.

Record, M. T., Jr., and Lohman, T.M., (1978)
A semiempirical extention of polyelectrolyte theory to the treatment of oligoelectrolytes: Application to oligonucleotide helix-coil transitions
Biopolymers. **17**, 159-166.

Record, M. T., Jr. (1967)
Electrostatic effects on polynucleotide transitions. II. behaviour of titrated systems.
Biopolymers. **5**, 993-1008

Rajagopal, P., and Feigon, J.(1989)
NMR studies of triple-strand formation from the homopurine-homopyrimidine deoxyribonucleotide $d(GA)_4$ and $d(TC)_4$.
Biochemistry. **28**, 7859-7870.

Rajagopal, P., and Feigon, J.(1989)
Triple strand formation in the homopurine-homopyrimidine DNA oligonucleotides $d(G\text{-}A)_4$ and $d(T\text{-}C)_4$.
Nature (London). **339**, 637-640.

Radhakrishnan I, de los Santos, C., Patel, D. J. (1991)
Nuclear Magnetic resonance structural studies of intramolecular purine.purine.pyrimidine DNA triplexes in solution.
J. Mol. Biol. **221**, 1403-1418.

Radhakrishnan I, de los Santos, C., Patel, D. J. (1993a)
Solution structure of a purine.purine.pyrimidine DNA-triplex containing G.GC and T.AT triplets.
Structure. **1**, 17-32.

Radhakrishnan I, de los Santos, C., Patel, D. J. (1993b)
Nuclear Magnetic resonance structural studies of A.AT base triple alignment in intramolecular purine.purine.pyrimidine DNA triplexes in solution.
J. Mol. Biol. **234**, 188-197.

Rosenbaum, V., and Riesner, D. (1987)
Electrophoresis-thermodynamic analysis of nucleic-acids and proteins in purified form and in celluar-extracts.
Biophys. Chem. **26**, 235-246.

de los Santos, C., Rosen, M., and Patel, D. (1989)
NMR studies of DNA $(R^+)_n \cdot (Y^-)_n \cdot (Y^+)_n$ triple helices in solution: Imino and amino proton markers of $T \cdot A \cdot T$ and $C \cdot G \cdot C^+$ base-triple formation.
Biochemistry. **28**, 7282-7289.

Seeman, N. C. (1985)
Macromolecular design, nucleic acid junctions, and crystal formation.
J. Biomol. Struct. Dyns. **3**, 11-34

Seeman, N. C. (1981)
in Nucleic acid junctions: building blocks for genetic engineering in three dimensions. Proceeding of the Second SUNYA Convention in the Discipline Biomolecular Stereodynamics. (Ramaswamy H. Sarma., ED.) **1**, pp.269, Adenine Press, New York.

Seeman, N. C. (1982)
Nucleic acid junctions and lattices.
J. Theor. Biol. **99**, 237-247.

Seeman, N. C. (1993)
The construction of a trefoil knot from DNA branched junction motif.
Biopolymers. in press

Seeman, N. C. (1992)
The design of single-stranded nucleic aciud knots.
Mol. Eng. **2**, 297-307.

Seeman, N. C., Chen, J. H., Kallenbach, N. R. (1989)
Gel electrophoretic analysis of branched junctions.
Electrophoresis. **10**, 345-354.

Seeman, N. C., and Kallenbach, N. R. (1983)
Design of immobile nucleic acid junctions.
Biophys. J. **44**, 201-209.

Seeman, N. C., and Kallenbach, N. R. (1994)
DNA Branched junctions.
in Annual Reviews of Biophysics and Biomolecular structure. (ed Stroud, R. M., Cantor, C. R., and Pollard, T. D.) **23**, 53-86.

Seeman, N. C. (1990)
De Novo design of sequences for nucleic acid structural engineering.
J. Biomol. Str. Dyns. **8**, 573-581.

Sigal N., and Alberts B. (1972)
Genetic recombination: the nature of a crossed strand-exchange between two homologous DNA molecules.
J. Mol. Biol. **71**, 789-793.

Singleton, S. F., and Dervan, P. B. (1992)
Influence of pH on the equilibrium association constants for oligodeoxyribonucleotide-directed triple-helix formation at single DNA sites.
Biochemistry. **31**, 10995-11003.

Timsit, Y., and Moras, D. (1991)
Groove-backbone interaction in B-DNA, Implication of DNA condensation and recombination.
J. Mol. Biol. **221**, 919-940.

Thiele, D., Guschlbauer, W., and Favre, A. (1972)
Protonated polynucleotide structure X. Optical properties of poly(I).ploy(C) and its disproportionation complexes.
BBA. **272**, 22-26

De Voe, H., and Tinoco, I., Jr. (1962)
The hypochromism of helical polynucleotides.
J. Mol. Biol. **4**, 500-517.

Völker, J., Botes, D. P., Lindsey, G. G., and Klump, H. H. (1993)
Energetics of a stable intramolecular DNA triplex formation.
J. Mol. Biol. **230**, 1278-1290.

Watson, J. D., and Crick, F. H. C. (1953a)
Genetical implications of the structure of deoxyribonucleic acids.
Nature (London). **171**, 964-967.

Watson, J. D., and Crick, F. H. C. (1953b)
Molecular structure of nucleic acids
Nature (London). **171**, 737-740.

Wang, H., and Seeman, N. C. (1995)
Structural domains of DNA mesojunctions.
Biochemistry. **34**, 920-929.

Wang, Y., Mueller, J. E., Kemper, B., and Seeman, N. C. (1991)
Assembly and Characterization of five-arm and six-arm DNA branched junctions.
Biochemistry. **30**, 5667-5674.

Wemmer, D. E., Joshua Wand A., Seeman, N. C., and Kallenbach, N. R. (1985)
NMR analysis of DNA junctions: imino proton NMR studies of individual arms and intact junction.
Biochemistry. **14**, 5745-5749.

Xodo, E. L., Manzini, G., and Quadrifoglio, F. (1990)
Spectroscopic and calorimetric investigation on DNA triplex formed by d(CTCTTCTTTCTTTTCTTTCTTCTC) and d(GAGAAGAAAGA) at acidic pH.
Nucleic. Acids. Res. **18**, 3557-3564.

Xodo, E. L., Manzini, G., Quadrifoglio, F., van der Marel, G. A., and van Boom, J. H. (1991)
Effect of 5-methylcytosine on the stability of triple-stranded DNA - a thermodynamic study.
Nucleic Acids Res. **19**, 5625-5631.

Zhang, Y., and Seeman, N. C. (1992)
A solid-support methodology for the construction of geometrical objects from DNA.
J. Am. Chem. Soc. **144**, 2656-2663.

Zhang, Y., and Seeman, N. C. (1994)
Construction of a DNA-Truncated Octahedron.
J. Am. Chem. Soc. **116**, 1661-1669.

Zhong, M., and Kallenbach, N. R. (1993)
Conformation and thermodynamics of DNA 'Necks" models of three-arm branch formation in a duplex.
J. Mol. Biol. **230**, 766-778.