Orientation Preferences of Backbone Secondary Amide Functional Groups in Peptide Nucleic Acid Complexes: Quantum Chemical Calculations Reveal an Intrinsic Preference of Cationic D-Amino Acid-Based Chiral PNA Analogues for the *P-form*

Christopher M. Topham* and Jeremy C. Smith†‡

*Institut de Pharmacologie et de Biologie Structurale, Centre National de la Recherche Scientifique UMR 5089, Toulouse, France; †Computational Molecular Biophysics, IWR, Universität Heidelberg, Heidelberg, Germany; and ‡Oak Ridge National Laboratory/University of Tennessee Center for Molecular Biophysics, Oak Ridge National Laboratory, Oak Ridge, Tennessee

ABSTRACT Geometric descriptions of nonideal interresidue hydrogen bonding and backbone-base water bridging in the minor groove are established in terms of polyamide backbone carbonyl group orientation from analyses of residue junction conformers in experimentally determined peptide nucleic acid (PNA) complexes. Two types of interresidue hydrogen bonding are identified in PNA conformers in heteroduplexes with nucleic acids that adopt *A-like* basepair stacking. Quantum chemical calculations on the binding of a water molecule to an O2 base atom in glycine-based PNA thymine dimers indicate that junctions modeled with *P-form* backbone conformations are lower in energy than a dimer comprising the predominant conformation observed in *A-like* helices. It is further shown in model systems that PNA analogs based on D-lysine are better able to preorganize in a conformation exclusive to *P-form* helices than is glycine-based PNA. An intrinsic preference for this conformation is also exhibited by positively charged chiral PNA dimers carrying 3-amino-D-alanine or 4-aza-D-leucine residue units that provide for additional rigidity by side-chain hydrogen bonding to the backbone carbonyl oxygen. Structural modifications stabilizing *P-form* helices may obviate the need for large heterocycles to target DNA pyrimidine bases via PNA·DNA-PNA triplex formation. Quantum chemical modeling methods are used to propose candidate PNA Hoogsteen strand designs.

INTRODUCTION

More than a decade ago, Nielsen and co-workers described an electrostatically neutral chimera between nucleic acids (the nucleobases) and (pseudo-) peptides (the backbone), termed ''polyamide nucleic acids'', or PNAs (1,2). These molecules, which are resistant to both nuclease and proteinase attack (3), comprise a backbone that is structurally homomorphous to the deoxyribose phosphate backbone, containing achiral N-(2-aminoethyl) glycine (*aeg*) units to which the nucleobase is attached via a methylene carbonyl linker (Fig. 1). PNAs form specific and highly thermally stable complexes with complementary single-stranded DNA or RNA, mediated by Watson-Crick hydrogen bonding (4). Unique among oligonucleotide analogs, PNAs are additionally able to strand invade double-stranded DNA (5,6). Their remarkable strand invasion properties, alongside demonstrations that modular PNA-conjugate constructs can cross cellular (7–9) and nuclear (10,11) membrane barriers, have made PNAs promising lead candidate molecules for the therapeutic control of gene expression (12–16).

In practice, the efficient targeting of double-stranded DNA via (PNA·DNA-PNA) triplex formation remains essentially limited to homopurine (pu) DNA tracts (17). The use of pseudocomplementary (pc) PNA oligomers, comprising 2, 4-diaminopurine and 4-thiouracil base replacements for adenine and thymine to disfavor unwanted self association of the pcPNA strands, allows mixed (pu/py) DNA sequences with a minimum 50% AT content to be targeted by a double duplex invasion mechanism (18). The future incorporation of pseudocomplementary guanine and cytosine nucleobase analogs is expected not only to increase the number of DNA sequences that can be targeted, but should also permit the covalent tethering of pcPNA oligomers as bis-pcPNAs so as to reduce the molecularity of the strand displacement process (14).

The amenability of the *aeg* PNA structural framework to chemical modification affords considerable advantages in the development of a pharmacologically efficacious antigene agent (19,20). The incorporation of charge and chirality, and manipulation of PNA backbone flexibility provide opportunities to increase solubility and bioavailability, and to improve selective binding to DNA targets. PNA binding to double-stranded DNA is effectively kinetically controlled (21), and the decrease in the magnitudes of association rate constants with increase in salt concentration (22–24) makes binding inhibition at physiological ionic strength a potential obstacle to in vivo applications. Marked improvements in binding rates at elevated ionic strengths obtained with bis-PNAs carrying positive charge either within the interchain linker or in the N- and C-terminal peptide tail sections (25,26) suggest that the judicious placement of charged

Submitted January 13, 2006, and accepted for publication October 5, 2006.

Address reprint requests to Christopher M. Topham, E-mail: christopher.topham@novaleads.com.

Christopher M. Topham's present address is Novaleads, Centre de Bioinformatique de la Haute Garonne, CEEI Théogone, 10 avenue de l'Europe, 31520 Ramonville-Saint-Agne, France.

© 2007 by the Biophysical Society

0006-3495/07/02/769/18 $2.00

doi: 10.1529/biophysj.105.079723

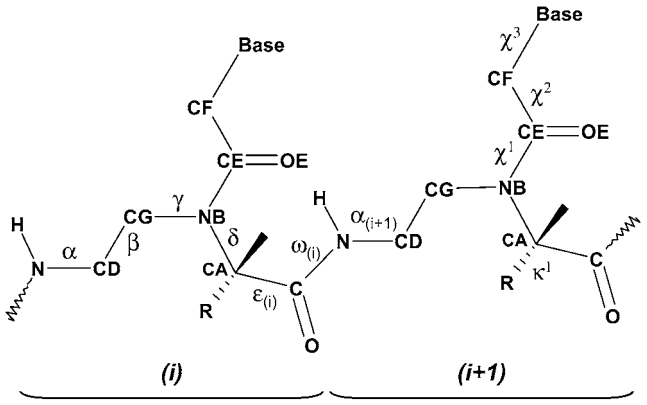

FIGURE 1 Schematic representation of a PNA residue junction with atom and dihedral angle nomenclature. Chiral PNA analogs based on D-aminoethylamino acid units carry side-chain (*R*) replacements of the *pro*-R hydrogen in the glycine moiety of the prototype (*aeg*) design.

groups on the *aeg* PNA backbone proper could similarly increase binding rates under physiological conditions without loss of sequence selectivity. The inherent flexibility of the prototype *aeg* PNA design (27–30), and the attendant entropy losses incurred upon binding to DNA, have at the same time prompted the search for conformationally constrained backbone skeletons that are preorganized so as to favor complex formation. Many strategies have focused on the covalent integration of (chiral) cyclic ring systems in designs that conform to the guiding principle of maintaining the base separation from and along the backbone by the same numbers of bonds as in DNA (see 20,31–33 for review and recent advances). In principle, rigidity may also be conferred without the creation of covalent cyclic structures by the attachment of one or more chemical substituents at pro-chiral carbon centers in the *aeg* PNA backbone (see Fig. 1). Much interest has been shown in the improved solubility and modified hybridization properties toward single-stranded DNA of PNA variants carrying positively charged chiral N-(2-aminoethyl) D-lysine units (34–37). A recent crystallographic study of an anti-parallel mixed-sequence PNA-DNA decamer heteroduplex, comprising a three-residue unit D-lysine ''chiral box'', provides evidence that increased polyamide backbone conformational rigidity can be obtained through the introduction of chiral centers (38).

The availability of experimentally determined structures of PNA complexes provides a rich knowledge base for the rational design of conformationally restricted chiral PNA analogs bearing functional groups. Experimental studies of PNA complexes have thus far revealed the existence of two distinct morphological helix forms. Crystal structures of the D-lysine-based PNA-DNA heteroduplex (38), a homopyrimidine PNA·DNA-PNA (py·pu-py) triplex (39), and four mixed (pu/py) sequence self-complementary PNA-PNA duplexes (40–43) all possess an unusual low-twist angle helical morphology, known as the *P-form*. These underwound structures are characterized by 16-fold (or 18-fold) helical repeats, according to whether (or not) a partner Watson-Crick DNA strand is present, and a pronounced displacement of the bases from the helix axis. In contrast to the *P-form* helices, which have average (local) twist angle values in the range 19–23°, antiparallel mixed-sequence heteroduplex *aeg* PNA-DNA octamer (44) and *aeg* PNA-RNA hexamer (27) NMR solution structures have significantly higher average (local) helix twist angles of 28° and 30°, respectively. Although the C2′-*endo* sugar puckering observed in the *aeg* PNA-DNA duplex is more typical of a B-form helix, both heteroduplexes adopt similar *A-like* helical topologies, with displacement of inclined Watson-Crick basepairs toward the minor groove.

We have previously reported that helix morphology exerts a significant influence on PNA backbone conformation and flexibility (45). The largest conformational variation is observed experimentally in rotations around the two bonds flanking the backbone secondary amide at junctions connecting residues at positions (*i*) and (i + 1) in the PNA chain, described by the $\varepsilon_{(i)}$ and $\alpha_{(i+1)}$ dihedral angles (see Fig. 1). Values of $\varepsilon_{(i)}$ and $\alpha_{(i+1)}$ were shown to be highly correlated over certain ranges, providing the basis for a PNA backbone conformation classification scheme. Most notably from a design perspective, it was found that preferred Watson-Crick PNA backbone conformations in *P-form* helical structures differ from those in *A-like* helices, possibly due in part to differential solvation effects in the minor groove. These findings suggest that, in addition to restricting conformational flexibility, the incorporation of bulky and/or charged groups to the PNA backbone could be used to control the basepair stacking pattern through the selective stabilization of backbone conformers associated with a particular helical form. This thesis is supported by the existence of the D-lysine based PNA-DNA duplex as a *P-form* helix, rather than an *A-like* helical structure, although the possibility of the difference in helix morphology compared to the *aeg* PNA-DNA duplex studied by NMR (44) being the result of crystal packing, strand length difference and/or base sequence effects cannot as yet be completely ruled out (38). Conversely, a L-arginine based PNA T_{10} decamer was observed to bind a complementary DNA A_{10} sequence with 1:1 stoichiometry, rather than as a 2:1 (PNA·DNA-PNA) triplex as does its *aeg* PNA T_{10} counterpart (46). Notwithstanding the role of electrostatic effects or steric hindrance of Hoogsteen strand binding by the L-arginine side chain, a plausible explanation for this change in binding stoichiometry may be an inability of the 2-aminoethyl-L-arginine backbone to adopt conformations compatible with a *P-form* helix.

It is with this design philosophy in mind that we present here a structural analysis of the spatial orientation preferences of the backbone amide −NH and >C=O groups at residue junctions in experimental structures of PNA complexes, classified according to backbone conformation. Ambiguities concerning the controversial participation of (i + 1) backbone −NH groups in hydrogen bonding

interactions with the carbonyl oxygen of the backbone-base linker at the preceding residue position (*i*), as proposed by Bruice and co-workers (47,48) but otherwise widely refuted (see, for example, 27,41,44,49), have been clarified using an operational definition of a hydrogen bond routinely employed in the classification of protein secondary structure (50,51). We have investigated the energetic and the structural changes involved in the binding of a water molecule to an O2 base atom in a model symmetry-restrained PNA thymine dimer (pTT) system using quantum chemical methods. Collectively the results are consistent with bound water molecules in the minor groove playing a structurally more important role in the stabilization of Watson-Crick PNA backbone conformation in *aeg* PNA *P-form* structures than in *A-like* hybrid double helices, where a tendency to form weak $(i + 1)/(i)$ interresidue hydrogen bonds of less than ideal geometry is evident in at least two PNA conformation classes. Our findings shed light on conflicting results obtained in molecular dynamics simulations of solvated *aeg* PNA-DNA duplexes (49,52).

In another series of quantum chemical calculations on a structurally modified pTT junctions, in which the *pro-R* hydrogen of the *aeg* PNA CA atom is replaced by aliphatic chains carrying positive charge (see Fig. 1), we exploit the symmetry-restrained dimer as a model with which to investigate local backbone preorganization. In contrast to the *aeg* PNA junction, the *P-form* backbone conformation common to the D-lysine ''chiral box'' and the PNA·DNA-PNA triplex can be identified at minima on the chiral dimer potential energy surfaces. Shortening of the D-lysine aliphatic chain allows direct electrostatic interaction with the PNA backbone carbonyl oxygen in this conformation, suggesting that the introduction of chiral PNA analogs based on either 3-amino-D-alanine or 4-aza-D-leucine should also favor *P-form* helix formation.

We conclude with a proposal that conformationally restricted structural modifications preferentially stabilizing *P-form* helices could be exploited to promote PNA·DNA-PNA triplex formation, and in particular to target mixed-sequence double-stranded DNA via triplex invasion mechanisms using small bases and other functional groups, rather than large heterocycles, attached to Hoogsteen strand carbonyl linkers to recognize DNA pyrimidine bases in the major groove. Quantum chemical modeling methods are used in support of candidate PNA Hoogsteen strand designs.

COMPUTATIONAL METHODS

PNA residue junction conformer database

The database, comprising a total of 207 nonredundant PNA residue junction atom sets, was compiled from eight experimental structures of PNA complexes in the Protein Data Bank (PDB) (53). Of these, 50 and 56 atom sets were, respectively, extracted from NMR solution structures of *A-like* mixed (py-pu) sequence PNA-RNA (PDB code 176D; (27)) and PNA-DNA (PDB code 1PDT; (44)) heterodimers. For comparative purposes, we also included data sets obtained by constrained energy minimization of the NMR models with PNA base atoms and all atoms of the nucleic acid strands held fixed (see Topham and Smith (45)). The remaining atom sets were culled from PNA strands in *P-form* x-ray crystal structures of a PNA·DNA-PNA triplex (PDB code, 1PNN (32 sets); (39)), a mixed-sequence PNA-DNA heteroduplex, comprising a three-residue unit D-lysine ''chiral box'' (PDB code, 1NR8 (9 sets); (38)), and four self-complementary right-handed homoduplexes (PDB codes, 1PUP (16 sets), (40); 1QPY (28 sets), (41); 1HZS (10 sets), (42); 1RRU (6 sets), (43)). Hydrogen atoms were added to the Protein Data Bank crystal structures using the HBUILD (54) facility in CHARMM (55), before energy minimization with all heavy atoms held fixed. PNA force field parameters were abstracted from the distributed all-atom CHARMM22 nucleic acid (56) and protein (57) sets, supplemented as previously reported (45).

Structural analysis of PNA residue junctions

Analysis of (α, β, γ, δ, ε, and ω) backbone and (χ^1, χ^2, and χ^3) backbone-base linker dihedral angles in the residue junction database was carried out according to the contiguous atom quartet definitions given previously (45). A $\{æ_{(i)}, \beta_{(i)}, \gamma_{(i)}, \delta_{(i)}, \beta_{(i+1)}, \gamma_{(i+1)}, \delta_{(i+1)}\}$ dihedral domain vector was defined at each junction between adjacent residue positions (i) and ($i + 1$) in the PNA chains, where $æ_{(i)}$ is given by $\alpha_{(i+1)} + \varepsilon_{(i)}$. We refer to æ as the coupling constant. Negative-signed values of æ in the range ($-150° > æ \geq 0°$) are assigned to the $\{æ^-\}$ domain, positive-signed values in the range ($0° > æ \geq 150°$) are classed as $\{æ^+\}$, and values in the range ($150° > æ \geq 210°$) are assigned to the {æ-*trans*} domain. β-domains are, respectively, categorized as $\{g^+\}$, {*trans*} or $\{g^-\}$ for angles in the range ($0° \geq \beta < 120°$), ($120° \geq \beta < 210°$) or ($210° \geq \beta < 360°$). The γ and δ dihedrals are either assigned to the $\{+90°\}$ domain for angle values falling in the range ($0° \geq (\gamma, \delta) < 180°$), or the $\{-90°\}$ domain for angles in the range ($180° \geq (\gamma, \delta) < 360°$). The orientation of the carbonyl group of the PNA backbone secondary amide bond was measured using the pseudodihedral angle (ν) introduced by the Orozco and Laughton groups (49). Recast in terms of the atom nomenclature used previously (39,45), and retained here (see Fig. 1), ν is defined at the i^{th} residue position ($\nu_{(i)}$) by the atom quartet, $CF_{(i)}$–$NB_{(i)}$–$C_{(i)}$–$O_{(i)}$.

Our PNA residue junction classification scheme allows for the coarse classification of conformers according to the {æ} domain occupied, and assignment to a particular subset or class according to the flanking $\{\beta, \gamma, \delta\}$ torsion angle domain combination (45). Hierarchic flanking angle domain pattern searches are now more robustly conducted by the ordered examination of the four-component $\{\gamma_{(i)}, \delta_{(i)}, \beta_{(i+1)}, \gamma_{(i+1)}\}$ and $\{\delta_{(i)}, \beta_{(i+1)}, \gamma_{(i+1)}, \delta_{(i+1)}\}$ vectors, before consideration of the $\{\beta_{(i)}, \gamma_{(i)}, \delta_{(i)}\}$ and $\{\delta_{(i+1)}, \beta_{(i+1)}, \gamma_{(i+1)}\}$ three-component vectors. Seven conformational classes are currently recognized (Table 1). It should be noted that all values of γ and δ were inadvertently systematically interchanged in our earlier analysis of PNA conformational preferences (45), and the $\{æ, \beta, \gamma, \delta\}$ vector defining the ($æ^+$)-β-*trans* conformational class is accordingly redefined in Table 1 as $\{æ^+, trans, +90°, -90°\}$. Three residue conformers in the 1PDT data set, previously assigned as ($æ^+$)-β-g^-, are now reclassified as ($æ^+$)-β-*trans*, and one residue junction in the 1PUP data set, previously annotated ($æ^-$)-P, is more appropriately considered as a member of a newly recognized class, (æ-*trans*)-P, uniquely observed in PNA-PNA duplexes. Structurally, these conformers resemble ($æ^-$)-P conformers, with compensatory shifts in values β, δ, and $\varepsilon_{(i)}$ dihedral angles permitting minor groove water bridging between the base and the backbone $-NH_{(i+1)}$ group to be maintained. A statistical analysis of backbone dihedral angle values in all seven conformational classes is presented in Table 2.

Electrostatic interaction energies between the backbone-base linker carbonyl group ($>CE_{(i)} = OE_{(i)}$) at residue position i and the backbone $-NH_{(i+1)}$ group at residue position $i + 1$ were calculated according to the Kabsch and Sander interatomic distance formula (50). An interresidue hydrogen bond was considered to exist when the interaction energy was < -0.5 kcal mol^{-1}. The generous cutoff proposed by the authors allows for bifurcated hydrogen bonds and errors in coordinates.

TABLE 1 Geometric criteria for PNA backbone conformation classification

Helix Morphology	Conformational Class	æ	β	γ	δ	Other Conditions
P-form	(æ$^-$)-P	{æ$^-$}	g^+	+90°	+90°	$180° \geq \alpha_{(i+1)} < 360°$
P-form	(æ$^-$)-P_{minor}	{æ$^-$}	g^+	+90°	+90°	$0° \geq \alpha_{(i+1)} < 180°$
P-form	(æ-*trans*)-P	{æ-*trans*}	g^+	+90°	+90°	
A-like	(æ$^-$)-β-g^+	{æ$^-$}	g^+	+90°	+90°	
A-like	(æ-trans)-β-g^+	{æ-*trans*}	g^+	+90°	+90°	
A-like	(æ$^+$)-β-*trans*	{æ$^+$}	*trans*	+90°	−90°	
A-like	(æ$^+$)-β-g^-	{æ$^+$}	g^-	−90°	−90°	

Ab initio quantum chemical calculations on model PNA thymine dimers

Hartree-Fock (HF) calculations on a model PNA thymine dimer (pTT) system, with hydrogen atoms in place of the terminal −NH and >C=O groups, were performed using GAUSSIAN 94 (58) or 98 (59). Geometry optimizations were carried out using either the 6-31G* or 3-21G* basis set as reported in Table 3. HF/6-31G* provides a sufficiently high level of quantum chemical theory appropriate for the calculation of hydrogen bonding properties and water binding interaction energies in small stable systems (56,60,61). Input geometries were specified as a Z-matrix, and symmetry restraints applied to chemically equivalent bond lengths, valence bond angles, and dihedral angles in the two residue units. Dimer configurations existing at (local) energy minima in regions of dihedral angle space characteristic of the main *P-form* and *A-like* conformer classes were identified from geometry optimizations in the absence of constraints using starting internal coordinate values obtained from analyses of experimental structures. Other conformers in a given class were then generated by the application of a constraint to fix α to a selected value within the experimentally observed range. Initial values for ε were calculated from the æ coupling constant. In cases such as the (æ$^-$)-P reference model 1.2 (Table 3), where it was necessary to ensure the integrity of α/ε dihedral angle coupling, a constraint was applied on the pseudodihedral angle ν without directly fixing either α or ε. High-occupancy solvent binding in the minor groove was modeled in the dimer systems by optimization of the interaction of a water molecule with the O2 thymine base atom of the first (N-terminal) residue unit. In these calculations, water molecule (HW-OW) bond length and (HW-OW-HW) bond angle values were constrained to TIP3P reference values (62). Dimer geometries were either held fixed to their optimized configurations, determined in the absence of the water molecule, or a full optimization of the dimer-water complex performed (with dimer symmetry restraints) to explore induced conformational changes. Default convergence criteria were satisfied in all cases. All reported energy differences were determined from (Møller-Plesset) MP2 single point energy calculations on the HF/6-31G* optimized geometries, allowing for a more accurate treatment of electron correlation effects than afforded by Hartree-Fock theory.

Comparison of base stacking patterns in modeled residue junctions and experimental structures

Intrastrand base stacking patterns in four representative PNA-containing experimental structures, the standard A80 and B80 DNA fiber conformations, and the pTT model dimers were compared in a pairwise manner using (averaged) root mean square distance (RMSD) values, calculated over the (12) heavy ring atoms of stacked py/py bases (see Supplementary Material, Table S1). A total of 20 stacked py/py base atom sets were culled from (Watson-Crick) PNA strands in the 176D (chain A, 10 models) Protein Data Bank (PDB) structure, and 16 from each from the 1PDT (chain B, 8 models) and 1PNN (chains A and C) coordinate sets. None of the four experimental PNA-PNA homoduplex structures contains stacked py/py bases, and 10 stacked py/py base atom sets were generated by reconstruction of chains A and B in the 1PUP PDB structure as polypyrimidine strands using JUMNA (63,64). Helicoidal parameters were calculated using a CURVES (65,66) analysis of the original coordinates. Canonical A80 and B80 AT duplex structures were generated using JUMNA and the helicoidal parameter sets

TABLE 2 Statistical analysis of selected (pseudo-) dihedral angle values in seven conformational classes of PNA residue junction in experimental structures

Helix morphology	Conformational class	Median æ (°)	Average $\alpha_{(i+1)}$ (°)	Average $\varepsilon_{(i)}$ (°)	Average $\nu_{(i)}$ (°)	Median $\alpha_{(i+1)}+\nu_{(i)}$ (°)	Median β (°)	Median γ (°)	Median δ (°)	n	(%)
P-form	(æ$^-$)-P	−116	−109 ± 9.0	−8 ± 10.7	+118 ± 15.7	+11	+69	+69	+90	67	66.3
	(æ$^-$)-P_{minor}	−127	+75 ± 10.1	+156 ± 11.6	−56 ± 11.7	+19	+67	+69	+89	25	24.8
	(æ-*trans*)-P	−162	−109 ± 11.2	−54 ± 14.1	+81 ± 8.3	−29	+95	+74	+120	7	6.9
	Others									2	2.0
	Total									101	100
A-like	(æ$^-$)-β-g^+	−114	+168 ± 14.5	+79 ± 10.4	−131 ± 18.4	+37	+69	+83	+75	35	33.0
	(æ-*trans*)-β-g^+	−163	−139 ± 14.2	−41 ± 38.7	+114 ± 43.1	−14	+113	+88	+129	5	4.7
	(æ$^+$)-β-*trans*	+117	+116 ± 65.5	−4 ± 80.0	−166 ± 84.5	−41	+148	+91	−143	20	18.9
	(æ$^+$)-β-g^-	+121	+150 ± 54.2	−42 ± 66.0	−176 ± 69.8	−36	−138	−77	−118	8	7.6
	Unclassified									38	35.8
	Total									106	100

Conformational classes were assigned to residue junctions in the database according to the {æ} domain occupied and hierarchic searches of flanking backbone {β, γ, δ} torsion angle domain vector patterns as described in the Computational Methods section. The total number of residue junctions in each class is indicated by n. Analysis of the *A-like* heteroduplexes was performed on the original NMR solution data. Contiguous bonded atom quartet definitions of α, β, γ, δ, and ε dihedral angles (see Fig. 1) are as given previously (45). The (i/i + 1) residue junction pseudodihedral angle (æ) is defined as the sum of the $\varepsilon_{(i)}$ and $\alpha_{(i+1)}$ dihedral angle values. The $\nu_{(i)}$ pseudodihedral angle, introduced by Soliva et al. (49) to describe the orientation of the carbonyl group of the PNA backbone secondary amide bond, is defined at the i^{th} residue position by the atom quartet, $CF_{(}$–$NB_{(i)}$–$C_{(i)}$–$O_{(i)}$.

TABLE 3 Structural analysis of geometry-optimized quantum chemical PNA thymine dimer models

Model	R*	Basis set	$H_2O^{\dagger}$	æ (°)	α (°)	ε (°)	ν (°)	ω (°)	β (°)	γ (°)	δ (°)	χ^1 (°)	χ^2 (°)	χ^3 (°)	κ^1 (°)
1.1	−H	6-31G*	−	−116.4	−97.5	−18.8	108.7‡	−166.6	71.3	72.4	108.1	−2.3	−164.7	78.5	–
1.2	−H	6-31G*	+	−117.5	−99.3	−18.2	108.7‡	−167.0	71.8	73.6	101.4	3.3	−174.2	85.4	–
1.3	−H	6-31G*	+	−89.6	−110.8	21.2	158.7	−178.9	58.0	71.6	72.1	2.9	−179.3	85.5	–
1.4	−H	6-31G*	−	−79.4	−104.6	25.2	173.9	177.5	58.5	76.4	82.8	−2.2	−166.5	74.7	–
2.1	−H	6-31G*	−	−108.2	91.7	160.1	−57.9	172.9	60.8	76.0	85.6	8.6	179.0	80.0	–
2.2	−H	6-31G*	+	−105.0	101.7	153.3	−59.1	176.4	60.3	76.4	101.4	3.9	−178.7	82.9	–
2.3	−H	6-31G*	+	−99.8	107.2	153.0	−60.7	176.8	57.7	77.5	92.5	5.3	179.6	80.7	–
2.4	−H	6-31G*	−	−106.6	101.8	151.6	−61.6	178.3	61.0	77.8	93.6	7.0	179.2	84.3	
3.1	−H	6-31G*	−	−113.3	145.0‡	101.7	−100.6	−174.9	64.8	77.0	83.2	4.3	−179.4	81.4	–
3.2	−H	6-31G*	−	−111.3	155.9	92.8	−106.7	−174.4	63.6	76.5	81.9	5.0	−179.9	82.1	–
3.3	−H	6-31G*	+	−115.0	155.9‡	89.1	−113.1	−169.7	62.8	75.0	77.0	11.1	173.6	88.5	–
3.4	−H	6-31G*	−	−107.4	168.2‡	84.4	−113.1	−175.0	61.8	75.8	80.9	5.6	179.3	82.1	–
3.5	−H	6-31G*	+	−109.0	168.2‡	82.8	−119.7	−170.6	57.7	71.2	73.0	13.3	171.8	86.3	–
3.6	−H	6-31G*	−	−102.4	180.0‡	77.6	−119.2	−176.0	59.9	75.1	80.6	5.4	178.9	81.2	–
3.7	−H	6-31G*	−	−96.6	−167.0‡	70.4	−126.6	−177.4	58.4	74.7	81.2	4.5	179.4	80.3	–
3.8	−H	6-31G*	−	−91.2	−154.0‡	62.8	−136.2	−178.6	58.2	74.9	82.6	2.7	178.7	79.2	–
4.1	$-(CH_2)_4NH_3^+$	3-21G*	−	−131.2	−111.3	−19.9	119.4	−161.5	82.9	100.2	91.7	20.2	−134.5	60.3	59.8
4.2	$-(CH_2)_4NH_3^+$	3-21G*	+	−136.2	−115.0	−21.2	120.9	−163.8	86.5	91.2	102.6	5.9	−177.8	82.8	58.5
4.3	$-(CH_2)_4NH_3^+$	3-21G*	−	−113.8	168.2‡	78.0	−104.7	−163.4	70.5	86.8	86.1	−13.9	178.1	69.7	53.3
5.1	$-CH_2NH_3^+$	3-21G*	−	−102.5	−99.9	−2.6	157.5	−167.8	69.5	92.3	102.3	−9.1	−160.5	65.9	65.1
5.2	$-CH_2NH_3^+$	3-21G*	+	−125.4	−94.4	−30.1	113.1	−168.4	80.4	83.7	103.4	9.4	−174.0	81.8	62.5
5.3	$-CH_2NH_3^+$	3-21G*	−	−108.3	−103.1	−5.1	151.9	−168.2	71.8	81.2	114.4	−16.5	−158.9	65.2	168.1
5.4	$-CH_2NH_3^+$	3-21G*	+	−124.2	−97.2	−27.0	119.0	−169.0	80.0	73.2	113.4	0.1	−174.0	75.0	175.1
5.5	$-CH_2NH_3^+$	6-31G*	+	−130.8	−103.7	−27.0	115.8	−164.0	81.9	78.1	110.3	4.3	−171.5	79.6	170.7
6.1	$-CH_2(CH_3)_2NH^+$	3-21G*	+	−125.7	−97.3	−28.4	114.4	−167.8	80.0	84.9	100.8	10.3	−174.7	83.3	58.9
6.2	$-CH_2(CH_3)_2NH^+$	3-21G*	+	−125.0	−98.2	−26.8	119.5	−168.7	80.8	73.6	113.3	0.3	−174.4	75.2	177.3

Geometry optimizations of symmetry-restrained pTT dimer systems were performed using the indicated basis set as described in the Computational Methods section. Atom quartet definitions of (α, β, γ, δ, ε, and ω) backbone and (χ^1, χ^2, and χ^3) backbone-base linker dihedral angles (see Fig. 1) are those given previously (45). Pseudodihedral angles (æ) and ν are defined in the legend to Table 2. The κ^1 side-chain dihedral angle refers to rotation around the CA – C^{β} bond in cationic D-amino acid-based chiral pTT dimer analogues, and is defined by the atom quartet $NB_{(i)} - CA_{(i)} - C^{\beta}_{(i)} - X^{\gamma}_{(i)}$, where X is carbon in analogues in which the prototype glycine moiety is replaced by D-lysine, and nitrogen in analogues based on 3-amino-D-alanine or 4-aza-D-leucine.
**pro*-R substituent at the CA atom (see Fig. 1).
†Geometry optimizations carried out in the presence (+) or absence (−) of a water molecule targeted to the O2 atom of the N-terminal thymine base.
‡Values of α or ν held fixed during geometry optimization.

tabulated by Lavery et al. (64). Base stacking patterns were compared using principal coordinates analysis (classical scaling) techniques (67). RMSD values were loaded into a ($n \times n$) distance (dissimilarity) matrix, and the elements normalized in a driver routine before analysis using the CMDS Fortran77 subroutine (http://astro.u-strasbg.fr/~fmurtagh//mda-sw/). The program outputs the ($n - 1$) eigenvalues, and projections of the base stacking pattern on the first seven principal components for plotting.

Quantum chemical modeling of interactions between Hoogsteen strand PNA analog derivative model compounds and pyrimidine bases

Geometric interactions of proposed functional groups with DNA pyrimidine bases in the major groove (see Fig. 2) were investigated using quantum chemical techniques and spatial constraints derived from the 1PNN (py·pu-py) PNA·DNA-PNA triplex x-ray crystal structure (39). Calculations were performed using the (May, 2004) GAMESS (68) implementation of the B3LYP hybrid Hartree-Fock/density functional method and the 6–311++G(d,p) basis set. The model systems (Fig. 9) comprised a pyrimidine base and *N*,*N*-dimethyl substituted alkyl amide derivatives of the functional group: (*A*) isopropyl ($-CH(CH_3)_2$) to target the thymine 5-methyl atom, and (*B*) imidazolyl (attached at the N1 position) or (*C*) isoxazolyl (attached at the C5 position) to hydrogen bond with cytosine. Six spatial constraints (three dihedral angles, two angles, and one distance) were imposed between the C5 and C6 atoms of the target pyrimidine base and atom positions in the model compounds analogous to the PNA backbone CA, NB, and CG atoms (Fig. 2). Pyrimidine base coordinates were obtained by conversion of the 1PNN polypurine DNA chains using JUMNA (63,64). Helicoidal parameters were calculated from a CURVES (65,66) analysis of the heterotriplex. Spatial constraints were derived from centrally positioned representative pT20·dA5 and pT18·dA3 PNA·DNA duplets (chains C and D), remote from the out-swinging Hoogsteen pC16 base in PNA chain A of the other triplex in the asymmetric unit. To take co-ordinate error into account, the CA, NB, and CG atom positions were not taken directly from the original 1PNN data, but from coordinates obtained by superposition of all heavy atoms in a HF/6-31G* geometry-optimized *N*,*N*-dimethyl thymine-1-acetamide model structure, built in the same conformation as the experimental PNA residue unit using dihedral angle constraints.

Strain energy in the model compounds in the pyrimidine base complexes was calculated from energy differences with respect to the fully geometry-optimized conformation of the isolated molecule in the absence of constraints. Calculations of atomic solvent-accessible surface area were carried out using NACCESS version 2.1.1 (69), a probe size of 1.4 Å, and the program default van der Waals radii of Chothia and co-workers for common chemical atom types. Basepair propeller and buckle parameters in the complexes of the azolyl derivatives with cytosine were calculated using CURVES (65,66) with a guanine base in place of the azole bases. Superposition was performed using the following ring atom mappings to guanine N9, C4, C5, N7, and C8 positions: N1, C2, N3, C4, C5 (imidazole); C5, O1, N2, C3, C4 (isoxazole).

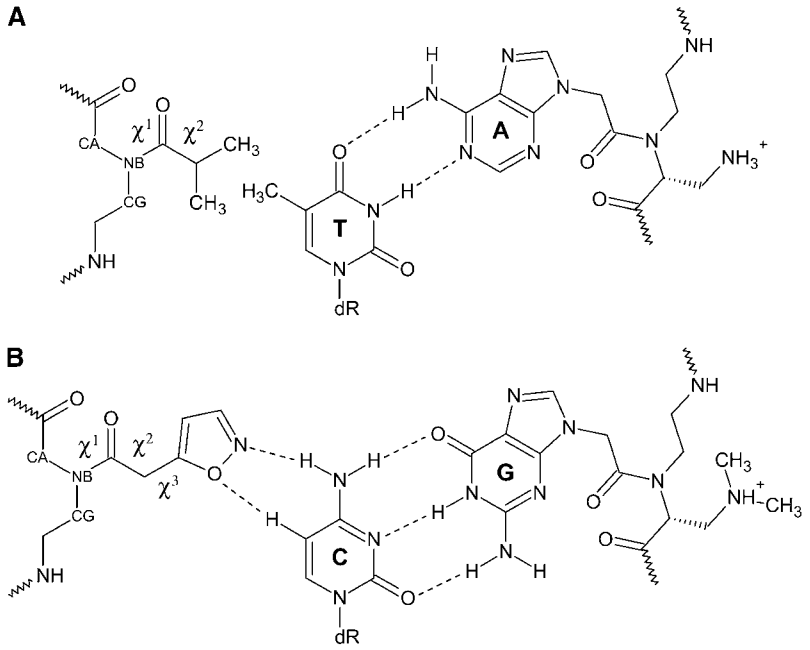

FIGURE 2 Proposed tandem use of conformationally restricted PNA and small Hoogsteen strand functionalities to target DNA pyrimidine bases via triplex formation. (*A*) Solvent shielding of thymine 5-methyl group by an isopropyl group. (*B*) Major groove recognition of cytosine by an isoxazole base. Watson-Crick PNA strand chiral PNA analogs based on 3-amino-D-alanine (*A*) or 4-aza-D-leucine (*B*) may promote *P-form* helix formation.

RESULTS

Influence of α/ε dihedral angle coupling on PNA backbone carbonyl group orientation

Our previous work drew attention to the existence of coupled rotations around the two bonds flanking the backbone secondary amide at junctions connecting residues at positions (i) and ($i + 1$) in the PNA chain, described by the $\varepsilon_{(i)}$ and $\alpha_{(i+1)}$ dihedral angles (45). Coupling of these dihedral angles operates over relatively short ranges of values characteristic of the conformer class. Fig. 3 shows that values of the pseudodihedral angle $\nu_{(i)}$, describing the orientation of the carbonyl group of the PNA backbone secondary amide bond at residue positions (i), correlate closely with values of $\alpha_{(i+1)}$ in the seven conformational classes identified in experimental structures. The correlations indicate that coupled change in $\alpha_{(i+1)}$ and $\varepsilon_{(i)}$ exerts a direct effect on the orientation of the secondary amide carbonyl group.

Backbone carbonyl group orientation accommodates minor groove water binding in P-form structures

The (æ$^-$)-P conformer is present in all PNA chains of *P-form* crystal structures. All the experimental (æ$^-$)-P residue junction $\nu_{(i)}$ data fall within the so-called ''forward'' ν domain (Fig. 3), defined by Soliva et al. (49) for PNA residue units with backbone carbonyl groups pointing toward the helix C terminus, corresponding to values of ν in the range $120 \pm 90°$. This geometric arrangement allows the backbone −NH group at the next ($i^{th} + 1$) residue position of Watson-Crick PNA chains to interact with pyrimidine $O2_{(i)}$ or purine $N3_{(i)}$ base atoms via the intermediary of a bound water molecule in the minor groove. The modeled (æ$^-$)-P pTT dimer junction (1.2) shown in Fig. 4 *A* was minimized in the presence of a bridging water molecule with ν held fixed at the average value of 108.7° (±11.8°) for (16) residue junctions in Watson-Crick PNA chains of the PNA·DNA-PNA triplex x-ray diffraction structure (39). Base stacking in the quantum chemical model is similar to that in the homopyrimidine Watson-Crick chains of the 1PNN coordinate set, with an average RMSD value of 0.35 Å for the ring atoms.

In contrast to the (æ$^-$)-P conformation, the backbone carbonyl group points in the opposite direction in (æ$^-$)-P_{minor} residue junctions. Experimental $\nu_{(i)}$ data for (æ$^-$)-P_{minor} residue junctions all lie within the ''*backward*'' domain, covering the ν angle range $-60 \pm 90°$. This type of *P-form* residue junction geometry is found in the four PNA-PNA duplexes and the first two junctions at the N-terminus of the mixed-sequence PNA-DNA decamer, carrying a centrally positioned three-residue unit D-lysine ''chiral box'' (38). The orientation of the carbonyl group in the experimental structures supports minor groove water bridging with pyrimidine O2 or purine N3 base atoms in the same PNA residue. The geometry-optimized reference model (2.1) of an (æ$^-$)-P_{minor} pTT residue junction (Fig. 4 *C*) exists at an energy minimum on the pTT energy surface. The value of ν in the dimer model is $-57.9°$, in close agreement with an average value of $-56°$ for experimental data. Base stacking in model 2.1 most closely resembles stacking in the 1PUP PNA-PNA data set, the average RMSD being 0.40 Å for the ring atoms.

Backbone carbonyl group orientation in A-like heteroduplexes associated with nonideal interresidue hydrogen bonding

The (æ$^-$)-β-g^+ class is the most populated of the four conformational classes observed in *A-like* PNA-DNA and

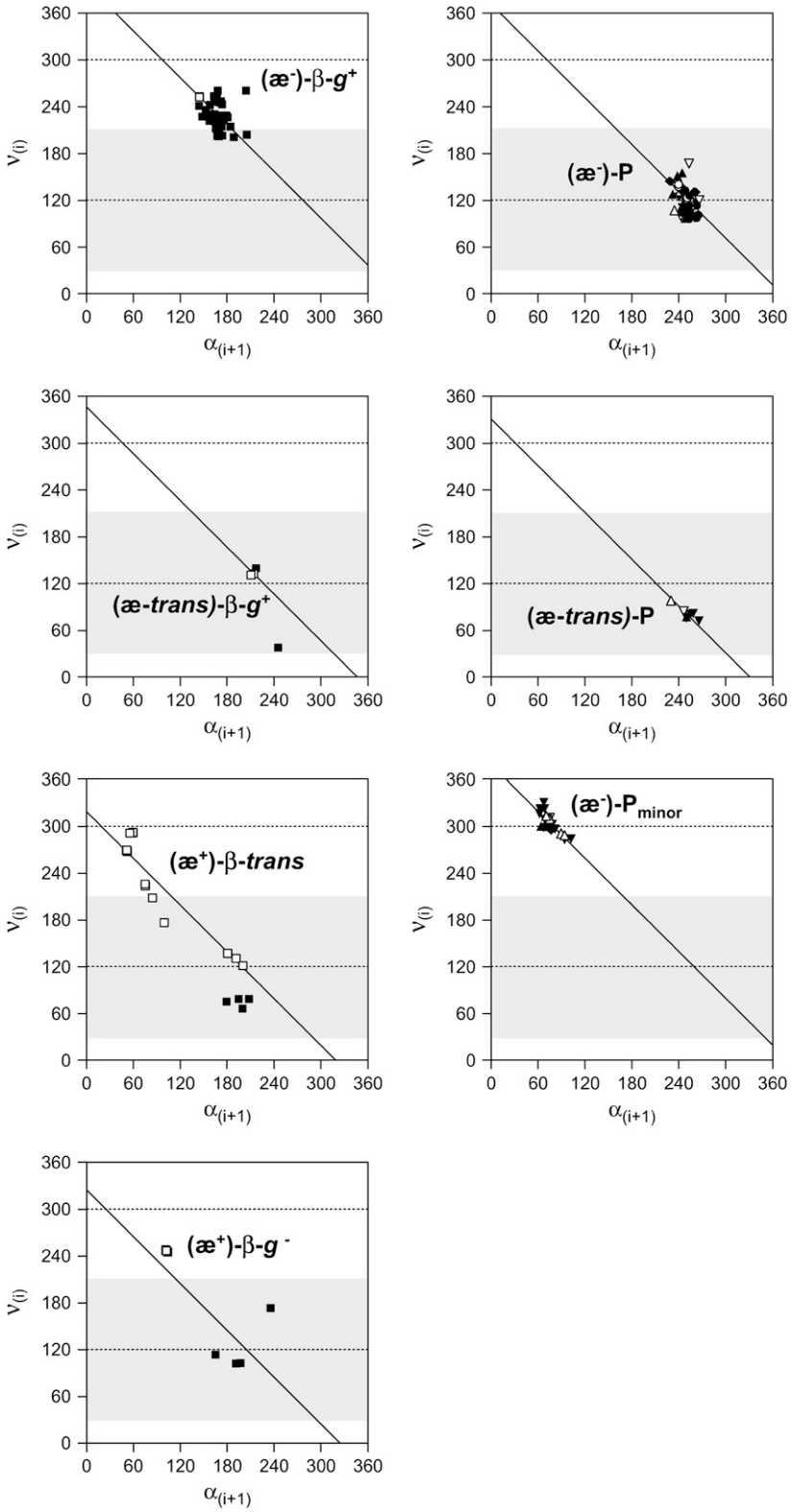

FIGURE 3 Dependence of PNA polyamide backbone carbonyl group orientation on the ∝ dihedral angle at residue junctions in experimental structures. Plots of the pseudodihedral angle $\nu_{(i)}$ at the i^{th} PNA residue unit versus the $\alpha_{(i+1)}$ dihedral angle at the following residue position are shown. Dihedral angle pairs from four conformers identified in PNA strands in *A-like aeg* PNA-RNA (PDB code 176D, ■) and *aeg* PNA-DNA (1PDT, □) helical structures are plotted in panels on the left-hand side. Data for three PNA conformer classes in *P-form* structures are shown in panels on the right-hand side: 1PNN homopyrimidine PNA·DNA-PNA triplex (Watson Crick, ●, and Hoogsteen strands, O); 1PUP (Δ), 1HZS (▲), 1QPY (▼), and 1RRU (∇) self-complementary PNA-PNA homoduplexes; 1NR8 (♦) PNA-DNA heteroduplex comprising a three-residue D-lysine ''chiral box''. Further details concerning the origins of the data are given in the Computational Methods section. The ''forward'' and ''backward'' pseudodihedral angle domains, describing 180° rotations of the PNA backbone carbonyl oxygen (O) atom around the NB-C vector with respect to the CF atom, defined by Soliva et al. (49), are, respectively, indicated as gray shaded and unshaded areas, centered at values of ν of 120° and −60° (300°). Solid line fits to the angle data were constructed from median estimates of $(\alpha_{(i+1)} + \nu_{(i)})$ given in Table 2. Regression analysis of $\nu_{(i)}$ on $\alpha_{(i+1)}$ yielded a r^2 (coefficient of determination) value of 0.95 ($n = 125$) for combined (classified) data in the $\{æ^-\}$ domain, and respective r^2 values of 0.71 ($n = 12$) and 0.83 ($n = 28$) for data in the {æ-*trans*} and $\{æ^+\}$ domains. Corresponding r^2 values obtained from regression analysis of $\varepsilon_{(i)}$ on $\alpha_{(i+1)}$ are 0.97 ($n = 127$), 0.43 ($n = 12$), and 0.91 ($n = 28$), respectively.

PNA-RNA heteroduplex helices studied by NMR. Average values of $\nu_{(i)}$, $\alpha_{(i+1)}$, and $\varepsilon_{(i)}$ for this class are ~90° out of phase with respect to corresponding values in $(æ^-)$-P and $(æ^-)$-P_{minor} residue junctions. The range of values of $\nu_{(i)}$ in the experimental structures (200–260°) straddles the boundary at 210° (−150°) separating the ''*forward*'' and ''*backward*'' ν domains. The backbone carbonyl group in these conformers points toward the solvent, allowing the

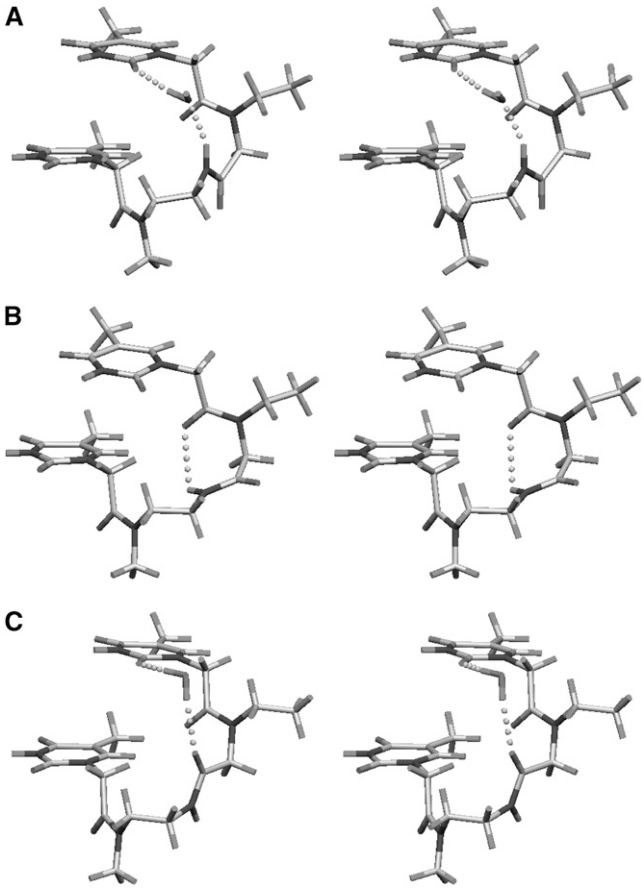

FIGURE 4 HF/6-31G* geometry-optimized *aeg* pTT dimer models of reference PNA conformers. (*A*) (æ$^-$)-P model 1.2 in the presence of a water molecule. (*B*) (æ$^-$)-β-g^+ model 3.2. (*C*) (æ$^-$)-P_{minor} model 2.1 showing the optimized interaction geometry of a bridging water molecule, the water oxygen lying approximately in the plane of the first thymine base. Detailed structural and energetic analysis of individual models can be found in Tables 3–5. The N- to C-terminal direction is from right to left. Hydrogen bond interactions are represented as spheres. Stereo graphics images were prepared using SETOR (84).

backbone −NH group of the next residue unit in the chain to interact with the backbone-base carbonyl (OE) oxygen, as exemplified in the geometry-optimized model (3.2) of an (æ$^-$)-β-g^+ junction shown in Fig. 4 *B*. The dimer exists in a local energy minimum, with an α-value of 156° and a ν-value of −107° (253°), as compared to respective average $\alpha_{(i+1)}$ and $\nu_{(i)}$ values of 168° and −131° (229°) for (æ$^-$)-β-g^+ conformers in *A-like* experimental structures. The $OE_{(i)}$–$HN_{(i+1)}$ distance (2.59 Å) and $OE_{(i)}$–$HN_{(i+1)}$–$N_{(i+1)}$ angle (116.0°) in this model are close to respective average experimental values of 2.81 ± 0.39 Å and 127 ± 9° for (æ$^-$)-β-g^+ conformers. While base stacking in model 3.2 is not particularly *A-like*, it being most similar to that of the 1PNN structure with an average RMSD value of 0.41 Å, a shift away from the 1PUP PNA-PNA homoduplex stacking pattern, relative to stacking in the 1.2 and 2.1 *P-form* reference model residue junctions, is nevertheless evident from the principal coordinates analysis in Fig. 5 *A*.

The chemical shifts and solvent exchange properties of $HN_{(i+1)}$ amide protons in the PNA-RNA (27) and PNA-DNA (44) duplexes appear to contradict the existence of interresidue hydrogen bonds (70). Indeed, respective average $OE_{(i)}$–$HN_{(i+1)}$ distances of 2.81 (±0.39) Å and 2.69 (±0.51) Å for (æ$^-$)-β-g^+ and pooled (æ$^+$)-β-*trans* and (æ$^+$)-β-g^- residue junctions in the NMR solution structures are longer than the 2.5-Å limit often employed in standard definitions of hydrogen bonding (71,72). An analysis of the dependence of the electrostatic interaction energy (E_{HB}^{elec}) of the $>CE_{(i)}=OE_{(i)}$ and $-NH_{(i+1)}$ groups, as a convenient single-parameter description of hydrogen bond quality, on the $\nu_{(i)}$ pseudodihedral angle in the (æ$^-$)-β-g^+ and pooled (æ$^+$) residue junction sets is shown in Fig. 6. The results show sharp minima in E_{HB}^{elec} as $\nu_{(i)}$ approaches ~240° (−120°) in (æ$^-$)-β-g^+ conformer junctions, or alternatively as $\nu_{(i)}$ tends to ~120° in (æ$^+$)-β-*trans* and (æ$^+$)-β-g^- conformers that posses δ dihedral angles in the −90° domain. According to the Kabsch and Sander (50) formula, a good hydrogen bond should have an electrostatic interaction energy of ~−3 kcal mol^{-1}. Only in data obtained by constrained in vacuo molecular mechanics energy minimization (panels *B* and *D*) do E_{HB}^{elec} values approach this limit. The raw solution data (panels *A* and *C*) do not support the existence of ideal interresidue hydrogen bonding.

The energetics of interresidue hydrogen bonding was further investigated in (æ$^-$)-β-g^+ residue junctions in a series of quantum chemical calculations on the pTT dimer by varying α over the experimentally observed $\alpha_{(i+1)}$ range from 145° to 206° (−154°). Dihedral angle values in the six geometry-optimized models (3.1, 3.2, 3.4, and 3.6–3.8) are reported in Table 3. Values of the æ coupling constant remain close to the median experimental value of −114°, at least up to α-values of 180°. Inspection of the relative total energies (ΔE in Table 4) confirms our earlier conclusions, based on molecular mechanics calculations (45), that α/ε dihedral angle coupling in PNA chains involves little net energy change up to limiting α-values of ~180°. The hydrogen bonding geometry and the calculated electrostatic energy component for the ($-NH_{(i+1)}$–$O_{(i)}=C_{(i)}<$) interaction vary as a function of α (and consequently ν) across the series (Table 4). In accordance with the analysis of the raw (æ$^-$)-β-g^+ NMR data, the strength of the most favorable calculated electrostatic interaction energy is appreciably <−3 kcal mol^{-1} (−1.61 kcal mol^{-1}). This is observed for model 3.6 (α = 180°) with a ν-value of 241° (−119°). Regression analysis of ΔE_{HB}^{elec} on ΔE, with model 3.2 serving as the reference geometry, yielded a r^2 value close to zero (0.05), implying that the weak hydrogen bonds do not contribute significantly to the stabilization of the (æ$^-$)-β-g^+ conformation.

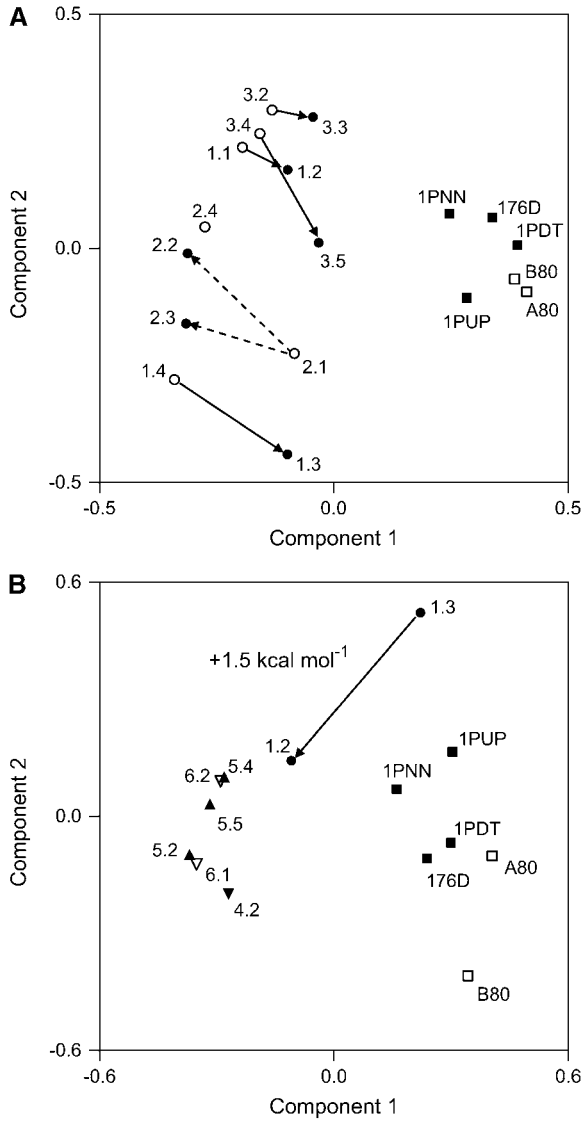

FIGURE 5 Principal coordinates analysis of base stacking patterns in modeled and experimental PNA residue junctions. Scatter plots show projections of the first two principal components (*1* and *2*) for py/py base-step RMSD data in Table S1: Panel *A* shows structural shifts in *aeg* pTT dimer stacking patterns (*open symbols*) toward (*solid arrows*) and away from (*dashed arrows*) experimental stacking patterns (■) induced by water molecule binding to the O2 atom of the N-terminal thymine base (*solid symbols*); (*B*) Imposition of a pseudodihedral angle constraint to clamp the orientation of the backbone secondary amide groups in the *aeg* pTT dimer results in reorganization of base-step stacking in the presence of a backbone-base bridging water (*solid arrow*) closer to patterns in the *P-form* 1PNN triplex crystal structure and in fully geometry-optimized cationic D-amino acid-based chiral pTT dimer analogs. The first two dimensions account for 48.2% of the total variation in panel *A* and 52.4% in panel *B*. Data for standard A80 and B80 DNA fiber conformations (□) are included for reference. Geometric descriptions of annotated symmetry-restrained quantum chemical dimer models are given in Table 3.

Structural and energetic analysis of base-water molecule binding in the *aeg*-pTT dimer

Minor groove high-occupancy water sites and spines of hydration have been observed experimentally in *P-form* crystal structures (38–43) and in molecular dynamics simulations of PNA-DNA and PNA-RNA heteroduplexes (49). Here we have used the symmetry-restrained pTT dimer as a model system to probe the structural and energetic influence of the binding of a water molecule in the minor groove by quantum chemical methods. Energy data for water molecule binding to the O2 atom of the N-terminal base as a function of the backbone conformation are summarized in Table 5. Variations in base-stacking patterns are represented two-dimensionally in Fig. 5 as projections on the first two components produced by principal coordinates analysis of pairwise RMSD data.

The (æ$^-$)-P conformer permits water bridging with the backbone $-NH_{(i+1)}$ group, whereas neither of the $-NH_{(i+1)}$ and $>C_{(i)}=O_{(i)}$ backbone functional groups is available in the (æ$^-$)-β-g^+ conformer. As expected, the calculated water interaction energy (E_{H2O}) is higher for the (æ$^-$)-P model 1.2 (-18.5 kcal mol^{-1}) compared to that for the (æ$^-$)-β-g^+ model 3.2 conformer (-11.2 kcal mol^{-1}). Although each system nominally comprises a total of two hydrogen bonds in the presence of a water molecule, the *P-form* reference (æ$^-$)-P conformer is -2.9 kcal mol^{-1} lower in energy than the (æ$^-$)-β-g^+ junction. In contrast, the isolated (æ$^-$)-P pTT dimer (model 1.2) is $+4.3$ kcal mol^{-1} higher in energy following removal of the water molecule than the (æ$^-$)-β-g^+ conformer model 3.2, geometry-optimized in the absence of a water molecule.

Geometry optimization of (æ$^-$)-β-g^+ model 3.2 in the presence of a water molecule targeted to the O2 atom yielded structure 1.3. Binding incurred the rupture of the weak (i + 1)/(i) interresidue hydrogen bond and the formation of a water bridge with the backbone $-NH_{(i+1)}$, accompanied by a marked transition in α from 156° to 249° (-111°). Holding α fixed at 156° or 168°, the average $\alpha_{(i+1)}$ value for (æ$^-$)-β-g^+ conformers in *A-like* experimental structures, permitted the retention of the interresidue hydrogen bond in structures 3.3 and 3.5, respectively, obtained by optimization of starting geometries 3.2 and 3.4 in the presence of a water molecule. In both cases water molecule binding induced shifts in base-step stacking toward patterns in the four PNA experimental reference structures, and in particular the *P-form* 1PNN and 1PUP structures (Fig. 5 *A*).

The high-end α-value of 249° in model 1.3 is more characteristic of *P-form* (æ$^-$)-P or (æ-*trans*)-P conformations than *A-like* (æ$^-$)-β-g^+ conformers. However, model 1.3 displays signs of α/ε dihedral angle decoupling (æ = -90°) and has a ν-value of 156° that is notably higher than respective average $\nu_{(i)}$ values of 118° or 81° in (æ$^-$)-P or (æ-*trans*)-P experimental junctions. Removal of the bound water in model 1.3 followed by a second round of optimization led to

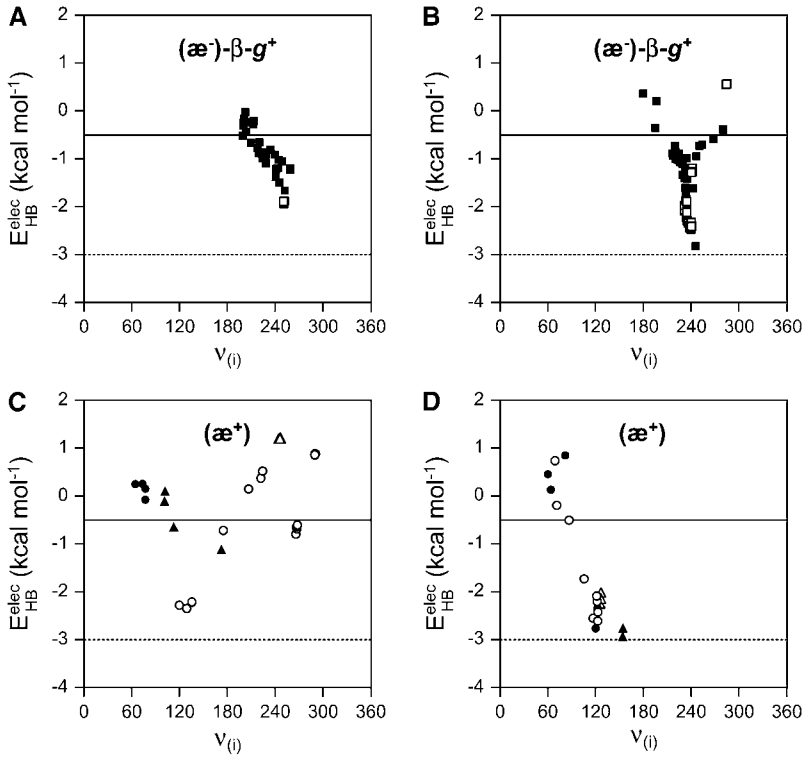

FIGURE 6 Interresidue hydrogen bond electrostatic energy as a function of backbone carbonyl group orientation in hybrid duplexes determined by NMR. Plots of the electrostatic energy of interaction between the $-NH_{(i+1)}$ and $>CE_{(i)}=OE_{(i)}$ groups (E_{HB}^{elec}) versus the pseudodihedral angle $\nu_{(i)}$ are shown for pooled (æ$^-$)-β-g^+ (*A* and *B*) and (æ$^+$)-β-*trans* and (æ$^+$)-β-g^- (*C* and *D*) conformer sets identified in the 176D (■) PNA-RNA and 1PDT (□) PNA-DNA PDB coordinate sets. Raw solution data are plotted in panels *A* and *C*. Panels *B* and *D* show data for conformers obtained by constrained energy minimization with the PNA bases and all atoms of the nucleic acid strands held fixed. The Kabsch and Sander (50) hydrogen bond cutoff energy of -0.5 kcal mol^{-1} is indicated by a horizontal solid line. According to this definition, $(i + 1)/(i)$ interresidue hydrogen bonds can be assigned to 77% (27 of 35) (æ$^-$)-β-g^+ residue junctions in *A*, and 86% (51/59) in *B*. Hydrogen bonds are found in 39% (11/28) of the pooled nonminimized (**æ$^+$**) residue junctions (*C*), rising to 75% (15/20) following energy minimization (*D*). No hydrogen bonds could be attributed using the same cut off in control calculations on 99 classified experimental *P-form* conformers (Table 2). The dotted reference line corresponds to an ideal hydrogen bond with an electrostatic interaction energy of -3.0 kcal mol^{-1}.

structure 1.4, which shows even more pronounced α/ε decoupling (æ = $-79°$) and a further increase in ν to 174°. The structural change for water binding (i.e., 1.4→1.3) again involves a shift in base stacking closer to stacking patterns in the *P-form* experimental reference structures, as does the analogous 1.1→1.2 water-binding induced structural transition with ν fixed to 108.7° to provide for α/ε dihedral angle coupling in better agreement with experimental data for triplex (æ$^-$)-P junctions (see Fig. 5 *A*). Structural changes resulting from the imposition of the pseudodihedral angle constraint in the pTT dimer, respectively, incur +1.5 kcal mol^{-1} (1.3→1.2) and +2.3 kcal mol^{-1} (1.4→1.1) increases in the total energy in the presence or absence of a backbone-base bridging water molecule. The 1.3→1.2 shift in base stacking relative to stacking variation in the PNA experimental structures and canonical A80 and B80 forms is highlighted in Fig. 5 *B*. The constraint on ν exerts a more pronounced effect on the ε and δ dihedral angles than on α (Table 3). It is of note that reorientation of the backbone amide functional groups or shifts in base stacking closer in agreement with experimental (æ$^-$)-P junction stacking could not be obtained by constraining either ε or α individually.

Backbone-base water bridging in *P-form* (æ$^-$)-P_{minor} conformers is with the $>C_{(i)}=O_{(i)}$ group. As in the case of the (æ$^-$)-P conformer, calculated water interaction energies of -15.3 or -14.4 kcal mol^{-1}, depending on the solvent molecule-binding mode, are higher for the (æ$^-$)-P_{minor} reference dimer (model 2.1) than for the (æ$^-$)-β-g^+ conformer ($E_{H2O} = -11.2$ kcal mol^{-1}, model 3.2). Similarly, the isolated

TABLE 4 Interresidue hydrogen bond geometry and energy analysis of HF/6-31G* geometry-optimized *aeg* pTT dimer models with (æ$^-$)-β-g^+ backbone conformations

Model	α (°)	ν (°)	$OE_{(i)}\cdots HN_{(i+1)}$ (Å)	$OE_{(i)}\cdots N_{(i+1)}$ (Å)	$OE_{(i)}$–$HN_{(i+1)}$–$N_{(i+1)}$ (°)	$CE_{(i)}$–$OE_{(i)}$–$HN_{(i+1)}$ (°)	E_{HB}^{elec} (kcal mol^{-1})	ΔE_{HB}^{elec} (kcal mol^{-1})	ΔE (kcal mol^{-1})
3.1	145.0	259.4	2.76	3.23	109.5	109.3	−1.104	0.216	0.137
3.2	155.9	253.3	2.59	3.15	116.0	107.8	−1.320	0.000	0.000
3.4	168.2	246.9	2.44	3.10	123.7	105.3	−1.515	−0.195	0.426
3.6	180.0	240.8	2.33	3.07	131.3	101.9	−1.608	−0.288	1.173
3.7	193.0	233.4	2.25	3.06	138.4	97.6	−1.573	−0.253	2.059
3.8	206.0	223.8	2.24	3.09	142.0	91.8	−1.275	0.045	2.884

Electrostatic energies of interaction between the $-NH_{(i+1)}$ and $>CE_{(i)}=OE_{(i)}$ groups (E_{HB}^{elec}) were calculated according to Kabsch and Sander (50). Relative electrostatic interaction energies (ΔE_{HB}^{elec}) and MP2/6-31G*//HF/6-31G* quantum chemical energy differences (ΔE) are reported with respect to model 3.2, which lies at a local minimum on the energy surface. Further details of the model structures and geometry optimization protocols can be found in Table 3 and Computational Methods.

TABLE 5 Water molecule interactions with thymine base O2 atom in symmetry-restrained *aeg* pTT dimer model systems

Model	H_2O*	$\Delta E^{\dagger}$ pTT (kcal mol^{-1})	$\Delta E^{\ddagger}$ pTT + H_2O (kcal mol^{-1})	θ § N1–C2–O2–OW (°)	E_{H2O}¶ (kcal mol^{-1})
1.1	–	2.828	−1.902	−4.9	−15.960
1.2	+	4.339	−2.934	−19.4	−18.503
1.3	+	3.605	−4.432	−20.8	−19.267
1.4	–	0.504	–	–	–
2.1	–	3.282	−0.759	4.9	−15.272
2.1	–	3.282	0.072	−90.8	−14.441
2.2	+	3.713	−1.352	−1.0	−16.295
2.3	+	3.984	−0.840	−95.4	−16.054
2.4	–	3.396	–	–	–
3.2	–	0.000	0.000	−9.7	−11.230
3.3	+	0.985	−1.122	−28.0	−13.338
3.4	–	0.426	0.141	−7.1	−11.515
3.5	+	1.975	−0.210	−23.9	−13.415

Full HF/6-31G geometry optimization carried out in the presence of a water molecule (+), or partial geometry optimization of the water interaction with the *aeg* pTT dimer held fixed in its optimized configuration determined in the absence of the water molecule (−).

†MP2/6-31G*//HF/6-31G* energy differences for isolated pTT dimers (ΔE) calculated with respect to model system 3.2.

‡MP2/6-31G*//HF/6-31G* energy differences for pTT-water complexes (ΔE) calculated relative to the total energy of model 3.2 in the presence of a water molecule targeted to the O2 atom of the N-terminal thymine base.

§(OW) water oxygen atom binding mode described by θ dihedral angle.

¶MP2/6-31G*//HF/6-31G* water interaction energies (E_{H2O}) with *aeg* pTT dimer.

(æ$^-$)-P_{minor} conformer is of significantly higher energy than (æ$^-$)-β-g^+ model 3.2 ($\Delta E = +3.3$ kcal mol^{-1}). Interaction with a water molecule reduces this deficit to +0.08 kcal mol^{-1} for oxygen atom binding below the plane of the target base ($\theta = \sim -90°$), or reverses it for more favorable ($\theta = \sim 0°$) in-plane binding ($\Delta E = -0.76$ kcal mol^{-1}). Full geometry optimization of the (æ$^-$)-P_{minor} reference model 2.1 in the presence of a water molecule yielded structures 2.2 and 2.3. Water molecule binding increases base-step rise and twist, with the result that stacking in models 2.2 and 2.3 resembles experimental patterns less than the stacking in model 2.1, as shown in Fig. 5 *A*. Relative to water-bound states of model 2.1 in which binding interaction geometries were optimized with the pTT dimer held fixed, the energy gains provided by full geometry optimization are modest: −0.59 kcal mol^{-1} for in-base-plane binding (2.1 → 2.2) and −0.77 kcal mol^{-1} for below base-plane binding (2.1 → 2.3). Models 2.2 and 2.3 are most similar to model 2.4 that exists in an unbound state at a local minimum on the energy surface close in α/ε dihedral space to model 2.1 with similar energy ($\Delta E = +0.11$ kcal mol^{-1}).

Together these results indicate that minor groove water interactions favor conformational transitions from the (æ$^-$)-β-g^+ to the (æ$^-$)-P or (æ$^-$)-P_{minor} states through induced rotation of the $\varepsilon_{(i)}$ and $\alpha_{(i+1)}$ dihedral angles and changes in base stacking. However, they also underline the inherent flexibility of the prototype *aeg* PNA design, and in particular, the limited ability of *aeg* PNA to preorganize in the (æ$^-$)-P conformation.

Intrinsic preference of D-lysine-based pTT dimer for the P-form

Menchise et al. recently described the crystal structure of a *P-form* antiparallel PNA-DNA decamer heteroduplex containing a three-residue unit D-lysine ''chiral box'' in the PNA strand (38). The ''chiral box'' $(i + 1)/(i)$ residue junctions, carrying a 4-aminobutyl side chain in place of the *pro-R* hydrogen of the *aeg* PNA glycine α-carbon (CA) atom at positions ($i = 5$–7), exist in the (æ$^-$)-P conformation. This prompted us to examine whether the preference of D-lysine based PNA for the (æ$^-$)-P conformation could be observed in a modified dimer system (*D-Lys*-pTT) comprising the same 4-aminobutyl side-chain rotamer combination (i.e., $\kappa^1 \sim +60°$; κ^2, κ^3, $\kappa^4 \sim 180°$) as in the 1NR8 x-ray structure. Using *aeg*-pTT models 1.3 and 1.4 as starting configuration templates, geometry optimization of the *D-Lys*-pTT dimer in the presence or absence of a ($-NH_{(i+1)}$–$O2_{(i)}$) backbone-base bridging water molecule led to models 4.2 and 4.1, respectively. Both structures exist at energy minima with ν values of ~120° in close accord with the average of 118° for (æ$^-$)-P conformers in experimental structures. While changes in β, γ, and δ dihedral angles are evident, reorientation of the backbone secondary amide appears to be most directly related to shifts of 42° and 45° in ε (see Table 3). Base-stacking patterns in models 4.2 and 4.1 are also more similar to stacking in the 1PNN triplex crystal structure than are respective patterns in the parent starting *aeg*-pTT models 1.3 and 1.4. *D-Lys*-pTT model 4.2, which is shown in Fig. 7 *A*, has an average RMSD value of 0.556 Å with respect to ring atoms in stacked Watson-Crick strand pyrimidine bases in the 1PNN coordinate set (c.f. 0.646 Å for model 1.3), while that for model 4.1 is 0.720 Å in the absence of a backbone-base bridging water (c.f. 0.748 Å for model 1.4).

Unlike the *aeg*-pTT model system, no minimum on the potential energy surface could be found for *D-Lys*-pTT conformers comprising α-values of ~160° that allow interaction between the $-NH_{(i+1)}$ and $>CE_{(i)}=OE_{(i)}$ groups in (æ$^-$)-β-g^+ residue junctions. Fig. 7 *B* shows *D-Lys*-pTT model 4.3, built using *aeg*-pTT model 3.4 as an initial template before geometry optimization with α held fixed at 168°. Large negative base-step roll distorts stacking compared to experimental structures. Removal of the constraint on α caused the rupture of the weak $(i + 1)/(i)$ interresidue hydrogen bond, and resulted in model 4.1 of substantially lower energy ($\Delta E = -19.2$ kcal mol^{-1}).

Intrinsic preferences of hydrogen bond-locked cationic D-amino acid-based pTT dimers for the P-form

Graphical inspection of the Menchise et al. 1NR8 x-ray structure (38) reveals that the backbone $>C_{(i)}=O_{(i)}$ group

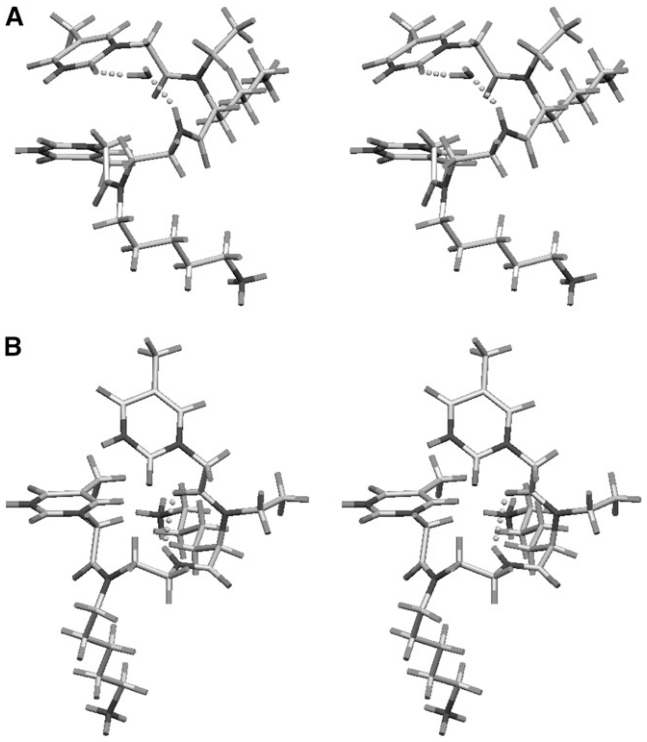

FIGURE 7 HF/3-21G* geometry-optimized *D-Lys*-pTT dimers. (*A*) Model system 4.2 exists at an energy minimum in the presence of a water molecule with a ν (pseudo-) dihedral angle of 121°, characteristic of (æ$^-$)-P conformers in experimental structures. (*B*) Model system 4.3 comprises an $(i + 1)/(i)$ interresidue hydrogen bond formed between the -$NH_{(i+1)}$ and $>CE_{(i)}$=$OE_{(i)}$ groups, maintained by holding the α dihedral angle held fixed at 168°. Other details of the model geometries are given in Tables 3 and S1. Hydrogen bonds are depicted as spheres. Stereo graphics images were prepared using SETOR (84).

oxygen atom makes close contact with the D-lysine side-chain γ carbon atom at each of the following three (i^{th} + 1) residue positions, the average interatomic separation being 3.17 ± 0.10 Å. This suggested to us the possibility of shortening the aliphatic D-lysine side chain by the removal of three methylene carbons allowing the protonated $(i + 1)$ side-chain amino group in a PNA design based on 3-amino-D-alanine (D-2,3-diaminopropionic acid) to hydrogen bond with the polyamide backbone $>C_{(i)}$=$O_{(i)}$ carbonyl. The interresidue $(-N^{\gamma}H^{+}_{3(i+1)} - -O_{(i)} = C_{(i)}<)$ D side-chain-backbone hydrogen bonding is illustrated in Fig. 8 *A*, which shows a symmetry-restrained *3-amino-D-Ala*-pTT model (5.2) dimer, geometry-optimized in the presence of a backbone-base bridging water molecule. Values of æ (−125°) and ν (113°) in model 5.2 (Table 3) are close to respective median (−116°) and average (118°) values for (æ$^-$)-P conformers in experimental structures. The average stacked pyrimidine base RMSD value of 0.503 Å with respect to the 1PNN coordinate set is lower than that for the analogous *D-Lys*-pTT model 4.2. Model 5.2 comprises a κ^1 side-chain rotamer of ~+60°, defined at the i^{th} residue position by the atom quartet $NB_{(i)} - CA_{(i)} - C^{\beta}{}_{(i)} - N^{\gamma}{}_{(i)}$. Rotation around the CA – C^{β} bond in model 5.2 to give the *trans* rotamer (i.e., $\kappa^1 \sim 180°$), followed by geometry optimization yields model 5.4, shown in Fig. 8 *B*. Model system 5.4 contains an intraresidue $(-N^{\gamma}H^{+}_{3(i)} - -O_{(i)} = C_{(i)}<)$ D side-chain-backbone carbonyl hydrogen bond, and is only slightly higher in energy than model 5.2 ($\Delta E = +0.88$ kcal mol^{-1}). Values of æ (−124°) and ν (119°) in *3-amino-D-Ala*-pTT model 5.4 are again compatible with experimental (æ$^-$)-P residue junction geometry, and the average stacked pyrimidine base RMSD with respect to 1PNN structure stacking is further reduced to 0.426 Å. In a control calculation using the 6-31G* basis set, model 5.4 geometry remained essentially unchanged (c.f. model 5.5; Table 3). Geometry optimization of models 5.2 and 5.4 following removal of the backbone-base bridging water molecules, respectively, led to structures 5.1 and 5.3. In the absence of water bridging, values of ν rise respectively to 158° and 152°, and the average base-step RMSD values with respect to the 1PNN stacking show prominent increases to 0.806 and 0.784 Å, respectively. The *3-amino-D-Ala*-pTT model system thus appears to be more sensitive than the *D-Lys*-pTT dimer to the presence of a bridging water molecule in its ability to adopt the (æ$^-$)-P conformation.

Replacement of the 3-amino group by the bulkier charge-delocalized dimethylamino group, in an alternative PNA design based on 4-aza-D-leucine, does not hinder the formation

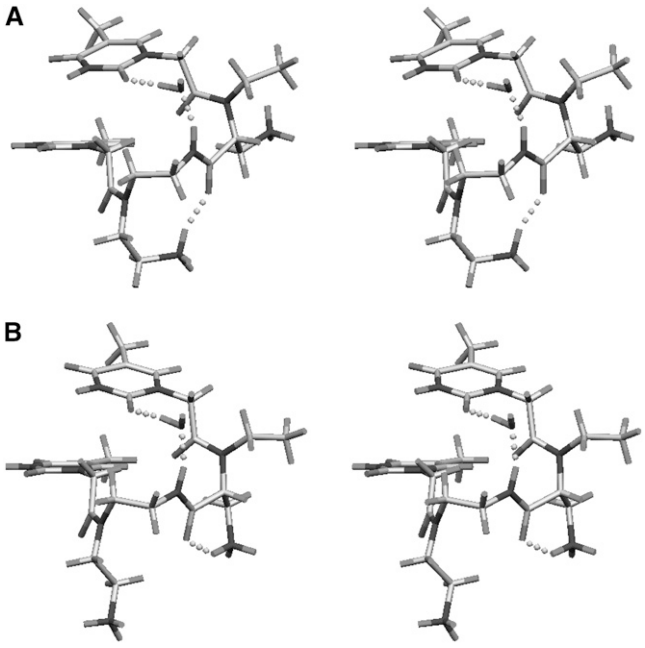

FIGURE 8 HF/3-21G* geometry-optimized *3-amino-D-Ala*-pTT dimers in the presence of a backbone-base bridging water molecule. (*A*) Model system 5.2 comprises an $(i + 1)/(i)$ interresidue $(-N^{\gamma}H^{+}_{3(i+1)} - -O_{(i)} = C_{(i)}<)$ D side-chain-backbone hydrogen bond to constrain the backbone. (*B*) Model system 5.4 possesses an analogous intra-residue hydrogen bond. Further geometric descriptions of these models are given in Tables 3 and S1. Hydrogen bond interactions are represented as spheres. Stereo graphics images were prepared using SETOR (84).

of $(i + 1)/(i)$ interresidue or $(i)/(i)$ intraresidue D side-chain-backbone carbonyl hydrogen bonding in geometry-optimized *4-aza-D-Leu*-pTT model dimers 6.1 and 6.2, which take up similar backbone conformations and base stacking patterns as models 5.2 and 5.4. As in the case of the *3-amino-D-Ala*-pTT dimer, the *trans* κ^1 rotamer (6.2), which provides for intraresidue hydrogen bonding, is of higher energy ($\Delta E = +1.3$ kcal mol^{-1}) than the $+60°$ κ^1 rotamer (6.1).

Geometric and energetic analysis of Hoogsteen strand PNA analog derivative model compound interactions with pyrimidine bases

The use of small bases or abasic functional groups, rather than large heterocycles, attached to Hoogsteen strand carbonyl linkers to recognize DNA pyrimidine bases in the major groove (Fig. 2) has been investigated in model systems using spatial constraints derived from the PNA·DNA-PNA triplex x-ray crystal structure (39). B3LYP/6-311++G(d,p) optimized interaction geometries of model compounds with pyrimidine bases are shown in Fig. 9. Structural comparisons of the model systems with the experimental py·pu-py triplex are summarized in Table 6.

The choice of an isopropyl group to target the thymine 5-methyl group (Fig. 9 *A*) was influenced by well-documented observations of the way in which valine side chains make specific van der Waals contacts with thymine 5-methyl groups at protein-DNA interfaces (73–75). Solvent shielding of the 5-methyl group rather than van der Waals interactions largely accounts for selectivity (76). The C5M carbon is the only thymine base atom to undergo a decrease in solvent-accessible surface area (of 66.5%) in the complex with the *N*,*N*-dimethyl-2-methylpropanamide. The respective approach distances of the *pro-R* and *pro-S* isopropyl methyl carbons to the C5M atom are 3.83 and 3.94 Å. The strain energy in the *N*,*N*-dimethyl-2-methylpropanamide molecule is $+1.23$ kcal mol^{-1}.

Imidazole- and other azole-2′-deoxyribonucleosides have been previously proposed as base analogs for binding to pyrimidine interruptions within antiparallel py·pu-py DNA triplexes (77–79). Third strand binding of PNA in py·pu-py heterotriplexes is in the opposite sense, with the Hoogsteen PNA amino terminus facing the 5′-end of the Watson-Crick oligonucleotide. The imidazole base N3 atom is able to accept a hydrogen bond from the cytosine N4 in the model complex with 2-imidazol-1-yl-*N*,*N*-dimethylacetamide (Fig. 9 *B*). The (cytosine) 2HN4–N3 interatomic separation is 2.32 Å, and the N4–2HN4–N3 angle is 162°. The molecular interaction energy was calculated as -5.9 kcal mol^{-1} with a strain energy of $+0.56$ kcal mol^{-1}. The proximity of the imidazole C2 and cytosine C5 atoms, imposed by the application of the spatial constraints, forces the imidazole ring out of the cytosine base plane, leaving the C2 and C5 atoms 3.62 Å apart. Differences of ~20° in basepair buckle (κ) and propeller twist (ω) angles relative to average values in the heterotriplex x-ray crystal structure are observed, and ring movement is accompanied by an increase in the χ^3 dihedral angle from a local minimum value of 98° to 115°. Closer agreement of the κ, ω, and χ^3 angles to experimental values (Table 6) is found for the complex of 2-isoxazol-5-yl-*N*,*N*-dimethylacetamide with cytosine (Fig. 9 *C*). Replacement of the imidazole C2 carbon by an (O1) oxygen in isoxazole allows a closer (cytosine) C5–O1 approach distance of 3.50 Å, facilitated by the formation of a weak −CH–O hydrogen bond. Hydrogen bonding with the cytosine N4 amine group is maintained: the 2HN4–N2 interatomic separation is 2.40 Å, and the N4–2HN4–N2 angle is 168°. The interaction energy for the isoxazolyl model compound is -6.1 kcal mol^{-1}, and the strain energy is reduced to $+0.19$ kcal mol^{-1}.

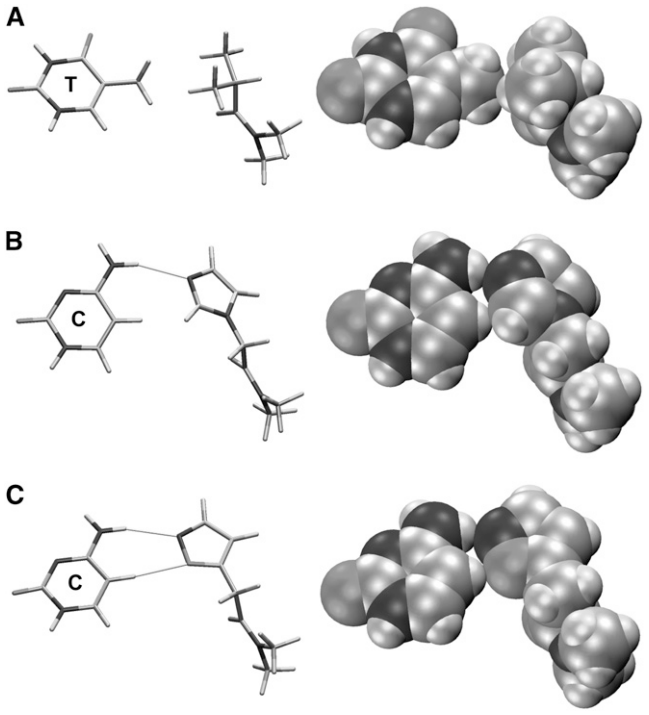

FIGURE 9 Spatially constrained B3LYP/6-311++G(d,p) optimized interaction geometries of *N*,*N*-dimethyl substituted amide model compounds with pyrimidine bases. (*A*) *N*,*N*-dimethyl-2-methylpropanamide/thymine; (*B*) 2-imidazol-1-yl-*N*,*N*-dimethylacetamide/cytosine; (*C*) 2-isoxazol-5-yl-*N*,*N*-dimethylacetamide/cytosine. Hydrogen bond interactions are represented as thin rods. Capped-stick and space-fill graphics representations were prepared using VMD (85).

DISCUSSION AND CONCLUSIONS

Previously reported differences in *aeg* PNA backbone conformations in *P-form* and *A-like* helical structures (45) may provide a basis for the rational design of more rigid chiral PNA analogs with intrinsic preferences for conformations associated with a particular helical form. Backbone modifications preferentially stabilizing *P-form* helices are of particular interest since they might be exploited to promote PNA·DNA-PNA heterotriplex formation.

TABLE 6 Structural comparison of model and experimental Hoogsteen·Watson-Crick PNA·DNA systems

Hoogsteen PNA	Watson-Crick DNA	χ^1 (°)	χ^2 (°)	χ^3 (°)	Buckle κ (°)	Propeller twist (ω −180°)
N,*N*-dimethyl-2-methylpropanamide	Thymine	−6.3	+153.2 (−82.8)†	–	–	–
N,*N*-dimethyl-2-methylpropanamide	–	−4.7	+149.4 (−88.2)†	–	–	–
2-imidazol-1-yl-*N*,*N*-dimethylacetamide	Cytosine	−2.7	+176.4	+115.2	+31.7	+25.0
2-imidazol-1-yl-*N*,*N*-dimethylacetamide	–	−7.2	−176.7	+97.5	–	–
2-isoxazol-5-yl-*N*,*N*-dimethylacetamide	Cytosine	−2.6	+171.2	+92.6	+9.6	+2.0
2-isoxazol-5-yl-*N*,*N*-dimethylacetamide	–	−9.5	−179.0	+91.2	–	–
1PNN (py) PNA*	1PNN (pu) DNA	+1.5 ± 1.5	−175.9 ± 3.6	+101.9 ± 1.6	+10.8 ± 3.8	+1.8 ± 4.1

Definitions of χ^1, χ^2, and χ^3 dihedral angles are given in Figs. 1 and 2. Basepair propeller (ω) and buckle (κ) parameters were calculated using CURVES (65,66). Further details of the *N*,*N*-dimethyl amide model structures and B3LYP/6–311++G(d,p) geometry optimization protocols can be found in Computational Methods.

*($n = 14$) for χ-dihedral angles; ($n = 17$) for κ- and ω-helicoidal parameters.

†χ^2 calculated with respect to *pro-S* (or *pro-R*) methyl group.

To better understand the origins of conformational differences in *P-form* and *A-like* structures, we have combined structural database analysis with a quantum chemical study of model *aeg* PNA and cationic D-amino acid-based chiral PNA analog (pTT) thymine dimer systems. The application of symmetry restraints allows the dimer system to be considered as a model of local preorganization in which energetic contributions from interactions with bases in the partner strand or neighboring residue units of the same strand are ignored. Comparison of base stacking in geometry-optimized quantum chemical models with that in experimentally determined structures of PNA complexes permits an assessment of the relative abilities of *aeg* PNA and chiral analog designs to preorganize. The extent to which the $\varepsilon_{(i)}$ and $\alpha_{(i+1)}$ dihedral angles that describe rotations around the two bonds flanking the backbone amide at PNA residue junctions are coupled provides an additional measure of preorganization. This can be conveniently quantified by departures of the coupling constant angle æ (defined as: $\alpha_{(i+1)} + \varepsilon_{(i)}$) from experimental values. Analysis presented here of seven conformational classes in experimental structures shows that coupled change in $\alpha_{(i+1)}$ and $\varepsilon_{(i)}$ correlates with change in the orientation of the secondary amide carbonyl group, described by the pseudodihedral angle, $\nu_{(i)}$ (49). Since coupled $\alpha_{(i+1)}/\varepsilon_{(i)}$ dihedral angle behavior operates over specific ranges characteristic of a given conformational class, departures of ν from experimental values afford a further means of assessing preorganization.

NMR solution structures of antiparallel mixed-sequence *aeg* PNA-DNA and *aeg* PNA-RNA duplexes display significant polyamide backbone conformational heterogeneity, but possess similar *A-like* helical morphologies. This is evident from the principal components scatter plot shown in Fig. 5 *B* in which base-step stacking patterns in the 1PDT and 176D data sets cluster closer to the canonical A80 stacking pattern than to stacking patterns for either the B80 DNA fiber conformation or experimental *P-form* (1PNN and 1PUP) PNA structures. The most populated of the four conformational classes observed in *A-like* heteroduplex residue junctions, we denote (æ$^-$)-β-g^+. The results of a molecular dynamics simulation by Sen and Nilsson of an explicitly solvated antiparallel PNA-DNA decamer (52) are entirely consistent with the preferred adoption of the (æ$^-$)-β-g^+ conformation in PNA-DNA systems (45).

A notable structural feature of the (æ$^-$)-β-g^+ conformation is an interresidue interaction between the backbone secondary amide $-NH_{(i+1)}$ group and the ($OE_{(i)}$) carbonyl oxygen of the backbone-base linker. Due to the flexible nature of the PNA backbone, the geometric quality of the $(i + 1)/(i)$ interresidue hydrogen bonding varies according to coupled change in the $\alpha_{(i+1)}$ and $\varepsilon_{(i)}$ dihedral angles that determines the orientation of the backbone amide functional groups, but is poor on average. Reasonable, but less than full strength, hydrogen bonding in (æ$^-$)-β-g^+ conformers is attained only over a relatively narrow pseudodihedral angle range as $\nu_{(i)}$ approaches ~240° (−120°). Nonideal hydrogen bonding interactions between the backbone-base linker $>CE_{(i)} = OE_{(i)}$ and backbone $-NH_{(i+1)}$ groups are also found in heteroduplex (æ$^+$)-β-*trans* and (æ$^+$)-β-g^- residue junctions, where limiting values of the electrostatic interaction energy are reached as $\nu_{(i)}$ tends to ~120°. In accord with our earlier molecular mechanics calculations on model heteroduplexes with fixed nucleobases (45), a quantum chemical treatment of an *aeg* pTT dimer junction in the (æ$^-$)-β-g^+ conformation with no constraints on the base atom positions indicated that weak hydrogen bonding is not a primary determinant of (æ$^-$)-β-g^+ conformer stability: no correlation could be found between relative dimer energies and changes in electrostatic interaction energy as a function of α and ν. This conclusion is at variance with the claims of Bruice and co-workers that the PNA backbone is stabilized by $(i + 1)/(i)$ interresidue hydrogen bonding (47,48). Further weakening of the hydrogen bond strengths might be expected from solvent electrostatic screening interactions, which are not taken into account in this quantum chemical analysis. Our results are more consistent with a dynamic spectrum of $(i + 1)/(i)$ interresidue electrostatic interactions of varying strength in (æ$^-$)-β-g^+ conformers. Since the range of backbone amide functional group orientations providing good interaction geometries is narrow, occupancy

of hydrogen bonded states may be correspondingly low, or vanishingly small in the limit. This sensitivity to PNA backbone flexibility could explain the apparent incompatibility of amide proton chemical shifts and solvent exchange properties in *A-like* heterodimers studied by NMR with the existence of interresidue hydrogen bonds (27,44,70).

The binding of backbone-base bridging water molecules in the minor groove may play an important role in PNA conformer stability and/or preorganization. Neither the $>C_{(i)}=O_{(i)}$ nor $-NH_{(i+1)}$ functional groups are available in (æ$^-$)-β-g^+ residue junctions for backbone-base water bridging of the type observed in *P-form* crystal structures (38–43). Consistent with this, no indication of the presence of significant water density in the minor groove was found by Sen and Nilsson in their simulation of a PNA-DNA heteroduplex (52), which retained an average helix twist angle value of 26.3 ± 3.9° close to that (28°) for the PNA-DNA octamer determined by NMR (44). Geometry optimization of our (æ$^-$)-β-g^+ pTT reference model 3.2 in the presence of a base-targeted water molecule induced the rupture of the weak hydrogen bonding interaction between the $>CE_{(i)}=OE_{(i)}$ and $-NH_{(i+1)}$ groups, and led to the formation of a water-mediated bridge between the -$NH_{(i+1)}$ backbone group and the base $O2_{(i)}$ atom (model 1.3). The structural transition involved a marked change of >90° in ν, and provided a net energy gain of −4.4 kcal mol^{-1}. A net energy gain of −1.4 kcal mol^{-1} was recorded for a 48° shift in ν in the opposite direction, accompanying the 3.2→2.2 transition and the formation of an in-base-plane water bridge between the backbone $>C_{(i)}=O_{(i)}$ carbonyl group and the $O2_{(i)}$ base atom. High occupancy minor groove water molecule binding might be linked to helix unwinding (and hence *A-like*→*P-form* transitions) through shifts in $\nu_{(i)}$ to values outside of the (æ$^-$)-β-g^+ residue junction range (200–260°). This is suggested by observations of a spine of hydration in the minor groove in a molecular simulation of the PNA-DNA octamer determined by NMR, where time-averaged duplex structures were reported to be underwound compared to experiment (49). The discrepancies between the decamer (52) and octomer (49) PNA-DNA simulations may relate to the different force fields, parameter sets, or molecular dynamics protocols employed, or to a greater tendency for the shorter heteroduplex to unwind. The magnitudes of the PNA partial atomic charges used by Soliva et al. (49), documented in Shields et al. (80), are notably larger than those of Sen and Nilsson (30,52), which may increase water residence times.

Evidence that the flexible *aeg* PNA design is less than optimally disposed for preorganization in the *P-form* is provided by the finding that the symmetry-restrained (æ$^-$)-P and (æ$^-$)-P_{minor} reference *aeg* pTT dimer models (1.2 and 2.1) do not lie in (local) energy minima in the presence of a bridging water molecule. A reorganization energy of +1.5 kcal mol^{-1}, corresponding to the 1.3→1.2 transition shown in Fig. 5 *B*, can be calculated for the (æ$^-$)-P conformer. This is consistent with the notion that water binding in the minor groove of a Watson-Crick DNA-*aeg* PNA heteroduplex does not in itself provide sufficient stabilizing energy to support topological transition from an *A-like*→*P-form* helix: this transition requires the additional binding of a Hoogsteen PNA strand in the major groove to distort the DNA structure during triplex invasion. The existence of a PNA-DNA decamer heteroduplex containing a three-residue unit PNA strand D-lysine ''chiral box'' as a *P-f*orm helix (38) demonstrates that structural change in DNA can be otherwise induced through modification of the PNA backbone. It was therefore anticipated that minimum energy structures of modified dimers bearing a 4-aminobutyl side-chain replacement of the *aeg* PNA glycine α-carbon *pro-R* hydrogen would exhibit greater conformational similarity to experimental *P-f*orm residue junctions than their *aeg* pTT counterparts. This proved to be the case both in the presence and in the absence of a bridging water molecule. Of further note was the failure to identify an (æ$^-$)-β-g^+ conformation at a local energy minimum in α dihedral angle space in the *D-Lys*-pTT dimer system. Together these results indicate that the more rigid D-lysine-based PNA analog has a stronger intrinsic preference for the (æ$^-$)-P conformation compared to prototype glycine-based PNA. Rigidity appears to be primarily conferred by a more restricted rotation around the CA-C bond, described by the $\varepsilon_{(i)}$ dihedral angle.

Intrinsic preferences for the (æ$^-$)-P conformation were also revealed in positively charged chiral dimers comprising 3-amino-D-alanine or 4-aza-D-leucine residue units. These designs provide for additional rigidity by side-chain hydrogen bonding to the backbone carbonyl oxygen, in principle, in either of two ways as shown in Fig. 8. The average stacked pyrimidine base RMSD with respect to stacking in the 1PNN triplex structure is lower in fully geometry-optimized *3-amino-D-Ala*-pTT and *4-aza-D-Leu*-pTT models with a bound water molecule than in the *D-Lys*-pTT model dimer. This suggests that PNA constructs comprising chiral analogs based on 3-amino-D-alanine or 4-aza-D-leucine should not only adopt *P-form* helices, but that they may be even better adapted to this purpose than the D-lysine-based PNA analog. In addition, these cationic D-amino acid-based chiral PNA analog designs possess shorter aliphatic side chains compared to D-lysine-based PNA, and consequently may be less prone to engage in nonspecific electrostatic interactions with nucleotide phosphate groups in other DNA molecules.

The design of nucleobases able to recognize pyrimidine bases in the major groove continues to hinder the development of antigene agents to target double-stranded DNA via triple-helix formation (81–83). The difficulties relate both to the presence of just one major groove hydrogen bonding site in the pyrimidine component of Watson-Crick C-(G) and T-(A) basepairs, and to steric obstruction by the 5-methyl group of thymine, compounded by a need to maximize π-stacking interactions to increase triplex stability. However, if binding energy gains resulting from PNA strand

backbone modification can be effectively harnessed so as to lower energy barriers for *P-form* triplex formation, base-stacking energetic contributions afforded by large heterocycles may become less critical provided that target base selectivity can still be maintained. Thus it might be possible to expand the PNA nucleobase sequence recognition alphabet through backbone modification and the tandem deployment of small bases or other functional groups attached to side-chain carbonyl linkers in the Hoogsteen PNA strand to recognize DNA pyrimidine bases (Fig. 2). Since the Watson-Crick and Hoogsteen PNA chains adopt the same (æ$^{-}$)-P conformation in the PNA·DNA-PNA triplex x-ray diffraction structure, backbone modifications in both PNA strands might be useful in overcoming difficulties presented by kinetically (preferred) ordered strand invasion mechanisms (5).

Our quantum chemical modeling studies show how an isopropyl group could be used to target the thymine 5-methyl group, mimicking the way in which valine side chains shield thymine 5-methyl groups from solvent at protein-DNA interfaces. The N2 atom of an isoxazole base attached to the prototype PNA methylene carbonyl side-chain linker via the C5 ring position is suitably positioned to accept a hydrogen bond from the N4 amine group of cytosine. Cytosine-isoxazole basepair propeller and buckle parameters are in good agreement with average values for pu-(py) pairs in the 1PNN triplex, as are side-chain linker dihedral angle values. The imidazol-1-yl analog appears less suitable due to the slightly unfavorable interaction geometry between the C2 carbon and the cytosine C5 imposed by the *P-form* helix, which forces the imidazole ring out of the cytosine base plane. Steric interference is relieved by the presence of the O1 oxygen at the equivalent position in the isoxazole ring, which is able to act as an acceptor of a weak hydrogen bond.

SUPPLEMENTARY MATERIAL

An online supplement to this article can be found by visiting BJ Online at http://www.biophysj.org.

We thank Dr. Richard Lavery for making the CURVES and JUMNA computer programs available to us. The help and support of Prof. François Amalric and Dr. Jean-Claude Meunier (CNRS, Toulouse, France) is gratefully acknowledged.

REFERENCES

1. Nielsen, P. E., M. Egholm, R. H. Berg, and O. Buchardt. 1991. Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science*. 254:1497–1500.
2. Dueholm, K. L., and P. E. Nielsen. 1997. Chemistry, properties and applications of PNA (peptide nucleic acid). *New J. Chem.* 21:19–31.
3. Demidov, V. V., V. N. Potaman, M. Frank-Kamenetskii, M. Egholm, O. Buchardt, S. H. Sönnichsen, and P. E. Nielsen. 1994. Stability of peptide nucleic acids in human serum and cellular extracts. *Biochem. Pharmacol.* 48:1310–1313.
4. Egholm, M., O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Nordén, and P. E. Nielsen. 1993. PNA hybridises to complementary oligonucleotides obeying the Watson-Crick hydrogen bonding rules. *Nature*. 365:566–568.
5. Nielsen, P. E. 2001. Targeting double stranded DNA with peptide nucleic acid (PNA). *Curr. Med. Chem.* 8:545–550.
6. Nielsen, P. E. 2001. Peptide nucleic acid targeting of double-stranded DNA. *Methods Enzymol.* 340:329–340.
7. Aldrian-Herrada, G., M. G. Desarménien, H. Orcel, L. Boissin-Agasse, J. Méry, J. Brugidou, and A. Rabié. 1998. A peptide nucleic acid (PNA) is more rapidly internalised in culture neurons when coupled to a *retro-inverso* delivery peptide. The antisense activity depresses the target mRNA and protein in magnocellular oxytocin neurons. *Nucleic Acids Res.* 26:4910–4916.
8. Pooga, M., U. Soomets, M. Hällbrink, A. Valkna, K. Saar, K. Rezaei, U. Kahl, J.-X. Hao, X.-J. Xu, Z. Wiessenfeld-Hallin, T. Hökfelt, T. Bartfai, and Ü. Langel. 1998. Cell penetrating PNA constructs regulate galanin receptor levels and modify pain transmission in vivo. *Nat. Biotechnol.* 16:857–861.
9. Mier, W., R. Eritja, A. Mohammed, U. Haberkorn, and M. Eisenhut. 2003. Peptide-PNA conjugates: targeted transport of antisense therapeutics into tumors. *Angew. Chem. Int. Ed. Engl.* 42:1968–1971.
10. Cutrona, G., E. M. Carpaneto, M. Ulivi, S. Roncella, O. Landt, M. Ferrarini, and L. C. Boffa. 2000. Effects in live cells of a c-*myc* antigene PNA linked to a nuclear localization signal. *Nat. Biotechnol.* 18:300–303.
11. Rebuffat, A. G., A. R. Nawrocki, P. E. Nielsen, A. G. Bernasconi, E. Bernal-Mendez, B. M. Frey, and F. J. Frey. 2002. Gene delivery by a steroid-peptide nucleic acid conjugate. *FASEB J.* 16:1426–1428.
12. Nielsen, P. E. 2000. Peptide nucleic acids: on the road to new gene therapeutic drugs. *Pharmacol. Toxicol.* 86:3–7.
13. Gambari, R. 2001. Peptide-nucleic acids (PNAs): a tool for the development of gene expression modifiers. *Curr. Pharm. Des.* 7:1839–1862.
14. Demidov, V. V. 2002. PNA comes of age: from infancy to maturity. *Drug Discov. Today.* 7:153–155.
15. Marin, V. L., S. Roy, and B. A. Armitage. 2004. Recent advances in the development of peptide nucleic acid as a gene-targeted drug. *Expert Opin. Biol. Ther.* 4:337–348.
16. Kaihatsu, K., B. A. Janowski, and D. R. Corey. 2004. Recognition of chromosomal DNA by PNAs. *Chem. Biol.* 11:749–758.
17. Demidov, V. V., and M. D. Frank-Kamenetskii. 2001. Sequence-specific targeting of duplex DNA by peptide nucleic acids via triplex strand invasion. *Methods.* 23:108–122.
18. Lohse, J., O. Dahl, and P. E. Nielsen. 1999. Double duplex invasion by peptide nucleic acid: a general principle for sequence-specific targeting of double-stranded DNA. *Proc. Natl. Acad. Sci. USA.* 96:11804–11808.
19. Nielsen, P. E. 1999. Peptide nucleic acids as therapeutic agents. *Curr. Opin. Struct. Biol.* 9:353–357.
20. Ganesh, K. N., and P. E. Nielsen. 2000. Peptide nucleic acids. Analogs and derivatives. *Curr. Org. Chem.* 4:931–943.
21. Demidov, V. V., M. V. Yavnilovich, B. P. Belotserkovskii, M. D. Frank-Kamenetskii, and P. E. Nielsen. 1995. Kinetics and mechanism of polyamide (''peptide'') nucleic acid binding to duplex DNA. *Proc. Natl. Acad. Sci. USA.* 92:2637–2641.
22. Peffer, N. J., J. C. Hanvey, J. E. Bisi, S. A. Thomson, C. F. Hassman, S. A. Noble, and L. E. Babiss. 1993. Strand-invasion of duplex DNA by peptide nucleic acid oligomers. *Proc. Natl. Acad. USA.* 90:10648–10652.
23. Cherny, D. Y., B. P. Belotserkovitskii, M. D. Frank-Kamenetskii, M. Egholm, O. Buchardt, R. H. Berg, and P. E. Nielsen. 1993. DNA unwinding upon strand displacement of binding of PNA to double-stranded DNA. *Proc. Natl. Acad. USA.* 90:1667–1670.
24. Wittung, P., P. Nielsen, and B. Nordén. 1996. Direct observation of strand invasion by peptide nucleic acid (PNA) into double-stranded DNA. *J. Am. Chem. Soc.* 118:7049–7054.
25. Griffith, M. C., L. M. Risen, M. J. Greig, E. A. Lesnik, K. G. Sprankle, R. H. Griffey, J. S. Kiely, and S. M. Freier. 1995. Single and bis

peptide nucleic acids as triplexing agents: binding and stoichometry. *J. Am. Chem. Soc.* 117:831–832.

26. Kuhn, H., V. V. Demidov, M. D. Frank-Kamenetskii, and P. E. Nielsen. 1998. Kinetic sequence discrimination of cationic bis-PNAs upon targeting of double-stranded DNA. *Nucleic Acids Res.* 26:582–587.
27. Brown, S. C., S. A. Thomson, J. M. Veal, and D. G. Davis. 1994. NMR solution structure of a peptide nucleic acid complex with RNA. *Science.* 265:777–780.
28. Leijon, M., A. Gräslund, P. E. Nielsen, B. Nordén, S. K. Kristensen, and M. Eriksson. 1994. Structural characterization of PNA-DNA duplexes by NMR. Evidence for DNA in a B-like conformation. *Biochemistry.* 33:9820–9825.
29. Chen, S.-M., V. Mohan, J. S. Kiely, M. C. Griffith, and R. H. Griffey. 1994. Molecular dynamics and NMR studies of single-stranded PNAs. *Tett. Lett.* 35:5105–5108.
30. Sen, S., and L. Nilsson. 2001. MD simulations of homomorphous PNA, DNA, and RNA single strands: characterization and comparison of conformations and dynamics. *J. Am. Chem. Soc.* 123:7414–7422.
31. Kumar, V. A. 2002. Structural preorganization of peptide nucleic acids: chiral cationic analogues with five- or six-membered ring structures. *Eur. J. Org. Chem.* 2021–2032.
32. Govindaraju, T., V. A. Kumar, and K. N. Ganesh. 2004. (1*S*,2*R*/1*R*,2*S*)-*cis*-Cyclopentyl PNAs (*cp*PNAs) as constrained PNA analogues: synthesis and evaluation of *aeg-cp*PNA chimera and stereopreferences in hybridisation with DNA/RNA. *J. Org. Chem.* 69: 5725–5734.
33. Pokorski, J. K., M. A. Witschi, B. L. Purnell, and D. H. Appella. 2004. (S,S)-*trans*-Cyclopentane-constrained peptide nucleic acids. A general backbone modification that improves binding affinity and sequence specificity. *J. Am. Chem. Soc.* 126:15067–15073.
34. Haaima, G., A. Lohse, O. Buchardt, and P. E. Nielsen. 1996. Peptide nucleic acids (PNAs) containing thymine monomers derived from chiral amino acids: hybridisation and solubility properties of D-lysine PNA. *Angew. Chem. Int. Ed. Engl.* 35:1939–1942.
35. Püschl, A., S. Sforza, G. Haaima, G. Dahl, and P. E. Nielsen. 1998. Peptide nucleic acids (PNAs) with a functional backbone. *Tett. Lett.* 39:4707–4710.
36. Sforza, S., G. Haaima, R. Marchelli, and P. E. Nielsen. 1999. Chiral peptide nucleic acids (PNAs): helix handedness and DNA recognition. *Eur. J. Org. Chem.* 197–204.
37. Sforza, S., R. Corradini, S. Ghirardi, A. Dossena, and R. Marchelli. 2000. DNA binding of a D-lysine chiral PNA: direction control and mismatch recognition. *Eur. J. Org. Chem.* 2905–2913.
38. Menchise, V., G. De Simone, T. Tedeschi, R. Corradini, S. Sforza, R. Marchelli, D. Capasso, M. Saviano, and C. Pedone. 2003. Insights into peptide nucleic acid (PNA) structural features: the crystal structure of a D-lysine-based chiral PNA-DNA duplex. *Proc. Natl. Acad. Sci. USA.* 100:12021–12026.
39. Betts, L., J. A. Losey, J. M. Veal, and S. R. Jordan. 1995. A nucleic acid triple helix formed by a peptide nucleic acid-DNA complex. *Science.* 270:1838–1841.
40. Rasmussen, H., J. S. Kastrup, J. N. Nielsen, J. M. Nielsen, and P. E. Nielsen. 1997. Crystal structure of a peptide nucleic acid (PNA) duplex at 1.7 Å. *Nat. Struct. Biol.* 4:98–101.
41. Haaima, G., H. Rasmussen, G. Schmidt, D. K. Jensen, J. S. Kastrup, P. Wittung Stafshede, B. Nordén, O. Buchardt, and P. E. Nielsen. 1999. Peptide nucleic acids (PNA) derived from N-(N-methylaminoethyl)-glycine. Synthesis, hybridization and structural properties. *New J. Chem.* 23:833–840.
42. Eldrup, A. B., C. Christensen, G. Haaima, and P. E. Nielsen. 2002. Substituted 1,8-Napthydrin-2(1*H*)ones are superior to thymine in the recognition of adenine in duplex as well as triplex structures. *J. Am. Chem. Soc.* 124:3254–3262.
43. Rasmussen, H., T. Liljefors, B. Petersson, P. E. Nielsen, and J. S. Kastrup. 2004. The influence of a chiral amino acid on the helical handedness of PNA in solution and in crystals. *J. Biomol. Struct. Dyn.* 21:495–502.
44. Eriksson, M., and P. E. Nielsen. 1996. Solution structure of a peptide nucleic acid-DNA duplex. *Nat. Struct. Biol.* 3:410–413.
45. Topham, C. M., and J. C. Smith. 1999. The influence of helix morphology on co-operative polyamide backbone conformational flexibility in peptide nucleic acid complexes. *J. Mol. Biol.* 292:1017–1038.
46. Zhou, P., M. Wang, L. Du, G. W. Fisher, A. Waggoner, and D. H. Li. 2003. Novel binding and efficient cellular uptake of guanidine-based peptide nucleic acids (GPNA). *J. Am. Chem. Soc.* 125:6878–6879.
47. Almarsson, Ö., and T. C. Bruice. 1993. Peptide nucleic acid (PNA) conformation and polymorphism in PNA-DNA and PNA-RNA hybrids. *Proc. Natl. Acad. Sci. USA.* 90:9542–9546.
48. Torres, R. A., and T. C. Bruice. 1996. Interresidue hydrogen bonding in a peptide nucleic acid-RNA heteroduplex. *Proc. Natl. Acad. Sci. USA.* 93:649–653.
49. Soliva, R., E. Sherer, F. J. Luque, C. A. Laughton, and M. Orozco. 2000. Molecular dynamics simulations of PNA·DNA and PNA·RNA duplexes in aqueous solution. *J. Am. Chem. Soc.* 122:5997–6008.
50. Kabsch, W., and C. Sander. 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers.* 22:2577–2637.
51. Carter, P., C. A. F. Andersen, and B. Rost. 2003. DSSPcont: continuous secondary structure assignments for proteins. *Nucleic Acids Res.* 31:3293–3295.
52. Sen, S., and L. Nilsson. 1998. Molecular dynamics of duplex systems involving PNA: structural consequences of the nucleic acid backbone. *J. Am. Chem. Soc.* 120:619–631.
53. Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne. 2000. The Protein Data Bank. *Nucleic Acids Res.* 28:235–242.
54. Bruenger, A. T., and M. Karplus. 1988. Polar hydrogen positions in proteins: empirical energy and neutron diffraction comparison. *Proteins: Struct. Funct. Genet.* 4:148–156.
55. Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: A program for macromolecular energy, minimisation, and dynamics calculations. *J. Comp. Chem.* 4:187–217.
56. MacKerell, A. D., Jr., J. Wiórkiewicz, and M. Karplus. 1995. An all-atom empirical energy function for the simulation of nucleic acids. *J. Am. Chem. Soc.* 117:11946–11975.
57. MacKerell, A. D. Jr., D. Bashford, M. Bellott, R. L. Dunbrack Jr., J. Evenseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph, L. Kuchnir, et al. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B.* 102:3586–3616.
58. Frisch, M. J., G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. A. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, et al. 1995. Gaussian 94. Revision D.4, Gaussian, Inc., Pittsburgh PA.
59. Frisch, M. J., G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, et al. 1998. Gaussian 98. Revision A.9, Gaussian, Inc., Pittsburgh PA.
60. Nina, M., J. C. Smith, and B. Roux. 1993. Ab initio quantum chemical analysis of retinal Schiff base hydration in bacteriorhodopsin. *J. Mol. Struct.* 286:231–245.
61. Nina, M., B. Roux, and J. C. Smith. 1995. Functional interactions in bacteriorhodopsin: a theoretical analysis of retinal hydrogen bonding with water. *Biophys. J.* 68:25–39.
62. Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.
63. Lavery, R. 1988. Junctions and bends in nucleic acids: a new theoretical modelling approach. *In* Structure and Expression, Vol. 3. DNA Bending and Curvature. W. K. Olson, M. H. Sarma, R. H. Sarma, and M. Sundaralingam, editors. Adenine Press, New York. 191–211.

64. Lavery, R., K. Zakrzewska, and H. Sklenar. 1995. JUMNA (junction minimisation of nucleic acids). *Comp. Phys. Comm.* 91:135–158.

65. Lavery, R., and H. Sklenar. 1988. The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.* 6:63–91.

66. Lavery, R., and H. Sklenar. 1989. Defining the structure of irregular nucleic acids: conventions and principles. *J. Biomol. Struct. Dyn.* 6:655–667.

67. Chatfield, C., and A. J. Collins. 1980. Introduction to Multivariate Analysis. Chapman and Hall, London and New York.

68. Schmidt, M. W., K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery. 1993. General atomic and molecular electronic structure system. *J. Comput. Chem.* 14:1347–1363.

69. Hubbard, S. J., and J. M. Thornton. 1993. NACCESS,Computer Program, Department of Biochemistry and Molecular Biology, University College, London, UK.

70. Eriksson, M., and P. E. Nielsen. 1996. PNA-nucleic acid complexes. Structure, stability and dynamics. *Q. Rev. Biophys.* 29:369–394.

71. Baker, E. N., and R. E. Hubbard. 1984. Hydrogen bonding in globular proteins. *Prog. Biophys. Mol. Biol.* 44:97–179.

72. McDonald, I. K., and J. M. Thornton. 1994. Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* 238:777–793.

73. Nadassy, K., S. Wodak, and J. Janin. 1999. Structural features of protein-nucleic acid recognition sites. *Biochemistry.* 38:1999–2017.

74. Luscombe, N. M., R. A. Laskowski, and J. M. Thornton. 2001. Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. *Nucleic Acids Res.* 29:2860–2874.

75. Prabakaran, P., J. An, M. M. Gromiha, S. Selvaraj, H. Uedaira, H. Kono, and A. Sarai. 2001. Thermodynamic database for protein-nucleic acid interactions (ProNIT). *Bioinformatics.* 17:1060–1061.

76. Plaxco, K. W., and W. A. Goddard III. 1994. Contributions of the thymine methyl group to the specific recognition of poly- and mononucleotides: an analysis of the relative free energies of solvation of thymine and uracil. *Biochemistry.* 33:3050–3054.

77. Durland, R. H., T. S. Rao, V. Bodepudi, D. M. Seth, K. Jayeram, and G. R. Revanker. 1995. Azole substituted oligonucleotides promote antiparallel triplex formation at non-homopurine duplex targets. *Nuc. Acids Res.* 23:647–653.

78. Gee, J. E., G. R. Revankar, S. Rao, and M. E. Hogan. 1995. Triplex formation at the rat *neu* gene utilizing imidazole and 2′-deoxy-6-thioguanosine base substitutions. *Biochemistry.* 34:2042–2048.

79. Jayaraman, K., R. H. Durland, T. S. Rao, G. R. Revanker, V. Bodepudi, N. Chaudhary, and J. Guy-Caffey. 1995. Approaches to enhance the binding affinity and nuclease stability of triplex forming oligonucleotides. *Nucleosides Nucleotides.* 14:951–955.

80. Shields, G. C., C. A. Laughton, and M. Orozco. 1998. Molecular dynamics simulation of a PNA-DNA-PNA triple helix in aqueous solution. *J. Am. Chem. Soc.* 120:5895–5904. Additions and corrections. 1999. *J. Am. Chem. Soc.* 121:1625.

81. Gowers, D. M., and K. R. Fox. 1999. Towards mixed sequence recognition by triple helix formation. *Nucleic Acids Res.* 27:1569–1577.

82. Buchini, S., and C. J. Leumann. 2003. Recent improvements in antigene technology. *Curr. Opin. Chem. Biol.* 7:717–726.

83. Goldbech Olsen, A., O. Dahl, and P. E. Nielsen. 2004. Synthesis and evaluation of a conformationally constrained pyridazinone PNA-monomer for recognition of thymine in triple-helix structures. *Bioorg. Med. Chem.* 14:1551–1554.

84. Evans, S. V. 1993. SETOR: Hardwire-lighted 3-dimensional solid model representations of macromolecules. *J. Mol. Graph.* 11:134–138.

85. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14:33–38.