# Conclusions

The 1:1 motif for DNA recognition by β-alanine-linked polyamides has been investigated by footprinting, affinity cleavage, and multidimensional NMR techniques in order to further understand the properties of this system and to exploit this knowledge in the design of next-generation DNA-binding ligands. A set of rules for 1:1 recognition were developed, which are based on the sequence specificity of individual ligand residues, the sequence-dependent orientation of the ligand, ligand size, placement and number of β residues, and the DNA structure itself. The ambiguity of sequence targeting based on stoichiometry, i.e., 1:1 versus 2:1 binding, established here, has been resolved by employing linker-dependent control over the ligand conformation and, hence, its mode of binding.

The most striking structural feature of the 1:1 binding polyamides discovered by Laemmli and coworkers is the extensive number of aliphatic β linkages. Specificity studies demonstrated β to be A,T specific, similar to Py. The NMR structure reveals that this specificity derives from intimate interaction between the α-methylene group and the C2-H atom of adenine, an interaction that would be sterically excluded at G•C base pairs. Additionally, we found that β in the context of Im-β-Im, as opposed to Im-β-Py, is necessary for high-affinity DNA binding. The NMR structure reveals two distinct conformations for Im-β-Im and Im-β-Py subunits. Im-β-Im requires greater flexibility in order to accommodate the larger, 50° dihedral between the planes of the flanking Im rings, presumably to allow the Im residues to better orient for hydrogen bonding to guanine. Therefore, it is believed that β confers the flexibility required for Im residues to align properly with the DNA helix.

Attempts at developing a recognition code for the 1:1 motif were less successful than those that led to the pairing rules for 2:1 recognition. It is believed that polyamides composed of five-membered heterocyclic amino acids contain insufficient structural information, when bound 1:1, to provide a general recognition code. This is evidenced by lack of single base specificity observed for imidazole, pyrrole, furan, thiophene, thiazole, and hydroxythiophene amino acids. The single site specificity observed for hydroxypyrrole (Hp) in polyamide **4** would suggest otherwise. For this anomalous result, I offer the following explanation: Hp projects the bulky C3-OH group to the minor groove floor, which explains the observed A,T > G,C specificity; the A > T specificity observed for the DNA sequence AAGAGA**A**GAG is probably a result of disrupting a stable, repeating polypurine base stack with a pyrimidine (T). One might argue that, based on this explanation, we should have observed a similar effect for the Hp-containing polyamide **6**. I would say that the Hp residue in **6** has greater conformational flexibility, conferred by the two flanking β residues, and therefore it can tolerate variations in DNA structure to a larger extent, thus resulting in reduced sequence specificity.

The NMR structure presented here offers the first high-resolution look at a polyamide containing imidazole and β residues bound in a 1:1 complex. The hydrogen bonds observed between Im-N3 and G-$NH_2$ provide the first direct evidence of the lexitropsin model as originally envisioned in the 1:1 motif (Kopka et al., 1985). The complex reveals B-form DNA with a narrow minor groove and a large negative propeller twist, which is shown to be stabilized by bifurcated hydrogen bonds donated from each amide NH group to its proximal purine N3

and pyrimidine O2 atoms. Stabilization of the negative propeller twist by these interactions, in addition to the inherently rigid and narrow minor groove, is thought to be the reason that polyamides would bind 1:1 in A-tract-like sequences, but would have difficulty binding as 2:1.

The observed homogeneous register of amide NH groups with respect to the DNA is though to be the driving force for optimal ligand-DNA alignment. If this is so, the G/C-dependent orientation preference of the polyamide could be explained by an inherent asymmetry in the projected angle of the Im-N3 lone pair $sp^2$ orbital with respect to the amide NH groups. Therefore, overlap of this orbital with the propeller-twisted guanine's $NH_2$ group is optimal when the polyamide is oriented N – C with respect to the 3' – 5' direction of the guanine-containing strand.

1:1 and 2:1 binding modes clearly have different rules for recognition. This ambiguity of sequence targeting depending on stoichiometry was addressed in order to eliminate alternative binding modes and therefore improve our accuracy of DNA sequence predetermination. The results presented here indicate that hairpin and 1:1 binding modes, which are dependent on ligand conformation, can be controlled by changing the linkage between subunits. The 1:1 binding mode is favored by using a β linker. Alternatively, the hairpin binding mode is favored by incorporating an α-(R)-amino substituted γ linker, i.e., the Herman turn, to link the flexible polyamide subunits. In addition to disfavoring the 1:1 mode, the acetylated Herman turn substantially improves hairpin binding affinity and sequence specificity.

Although the 1:1 binding mode may be less specific overall, it allows us to target certain DNA sequences that are not accessible to 2:1 binders. Polyamides that integrate 1:1 and 2:1 binding modes, such as the extended hairpins (Trauger et al., 1996c), will be an important future direction for targeting unique sequences in a genomic context. One can imagine that integrated motifs could exploit the preference of 1:1 versus 2:1 binding relevant to the sequence-dependent microstructure of DNA, i.e., certain sequence contexts with a normal B-like minor groove that may prefer pairs of rings (2:1 binding), whereas narrow minor groove tracts would allow the steric fit of only single rings (1:1 binding). The accurate prediction of optimal ligand-DNA complementarity requires an understanding of the geometric and electronic structural parameters for both molecules, knowledge that is based largely on high-resolution structural studies. The Thesis presented here provides a biophysical as well as structural foundation for future applications of the 1:1 motif.