1930–1934 *Nucleic Acids Research, 1997, Vol. 25, No. 10*

Effect of the 1-(2′-deoxy-β-D-ribofuranosyl)-3-nitropyrrole residue on the stability of DNA duplexes and triplexes

Olga Amosova, Jay George[1] and Jacques R. Fresco*

Department of Molecular Biology, Princeton University, Princeton, NJ 08544, USA and [1]Codon Pharmaceuticals, Gaithersberg, MD 20884, USA

Received January 30, 1997; Revised and Accepted March 24, 1997

ABSTRACT

3-Nitropyrrole (M) was introduced as a non-discriminating 'universal' base in nucleic acid duplexes by virtue of small size and a presumed tendency to stack but not hydrogen bond with canonical bases. However, the absence of thermally-induced hyperchromic changes by single-stranded deoxyoligomers in which M alternates with A or C residues shows that M does not stack strongly with A or C nearest neighbors. Yet, the insertion of a centrally located M opposite any canonical base in a duplex is sometimes even less destabilizing than that of some mismatches, and the variation in duplex stability is small. In triplexes, on the other hand, an M residue centrally located in the third strand reduces triplex stability drastically even when the X·Y target base pair is A·T or G·C in a homopurine·homopyrimidine segment. But, when the target duplex opposition is M-T and the third strand residue is T, the presence of M in the test triplet has little effect on triplex stability. Therefore, a lack of hydrogen bonding in an otherwise helix-compatible test triplet cannot be responsible for triplex destabilization when M is the third strand residue. Thus, M is non-discriminating and none-too-destabilizing in a duplex, but in a triplex it is extremely destabilizing when in the third strand.

INTRODUCTION

Because of the intrinsic sequence specificity of nucleic acid duplex and triplex formation, the development of oligonucleotides to regulate gene expression is a focus of much contemporary research. Oligonucleotides can interact specifically with messenger RNAs via duplex formation, effectively blocking their translation (see for example 1,2), and they can bind to duplex genomic DNA in a sequence-specific manner via triple helix formation, thereby blocking transcription of particular genes (see for example 3,4). However, a serious limitation to exploiting third strand binding to regulate gene expression or to induce site-specific gene repair (5) lies in the requirement that the binding site be homopurine·homopyrimidine. Much effort is being made to circumvent this barrier. One approach is to develop synthetic base analogs specific for 'inverted' base pairs (C·G and T·A) that interrupt homopurine·homopyrimidine continuity (6–8). But those efforts have so far been unsuccessful in that they do not provide *both* meaningful affinity and the required target base pair specificity. Moreover, little is known regarding the features that would enable base analogs in a third strand to bind to inverted target base pairs.

In the present work, we have investigated the nucleoside 1-(2′-deoxy-β-D-ribofuranosyl)-3-nitropyrrole (Fig. 1; the base is M) as a possible 'non-discriminatory' residue for third strand binding in a pyrimidine triplex motif. This residue was designed by Bergstrom and associates (9,10) to serve in PCR primers because it was presumed to be a good, i.e., strong, 'stacker', small enough to fit readily in a double helix opposite all canonical bases, yet unable to hydrogen bond to them. A residue with such properties provides an opportunity to evaluate the relative importance of hydrogen bonding and stacking interactions in triplex stabilization.

A deoxyoligonucleotide system capable of both duplex and triplex formation was used for these studies: A_{10}-X-A_{10}, T_{10}-Y-T_{10} (where X and Y are A, T, G or C) (11) and T_{10}-M-T_{10}. These 21mers were used to form duplexes with X-M and M-Y base oppositions and triplexes with M on an otherwise all-pyrimidine third strand interacting with an A·T, T·A, G·C or C·G base pair in the target duplex. In duplexes, as has been previously noted (9), M opposite a canonical base is no more destabilizing than mismatches between canonical bases. Yet, in single strands, M residues disrupt base stacking. In third strands, not only do they not enhance binding to inverted target base pairs, but, in fact, they destabilize triplexes more than mismatched canonical bases.

MATERIALS AND METHODS

Deoxyoligonucleotides

Twenty-one residue oligonucleotides A_{10}-X-A_{10}, T_{10}-Y-T_{10} (X and Y stand for A, G, C, T or M) were synthesized by the phosphoramidite method, deprotected and purified by denaturing PAGE. Bands were visualized by UV light, eluted with 10 mM Tris–HCl, 2 mM EDTA, pH 7.0, and desalted by C18 reversed phase chromatography. Purity was ascertained by denaturing PAGE

*To whom correspondence should be addressed. Tel: +1 609 258 3927; Fax: +1 609 258 6730; Email: jrfresco@princeton.edu

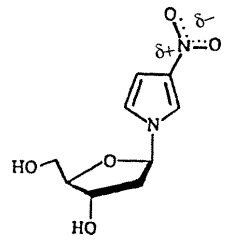

Figure 1. Structure of the residue 1-(2′-deoxy-β-D-ribofuranosyl)-3-nitropyrrole, which contains the base M.

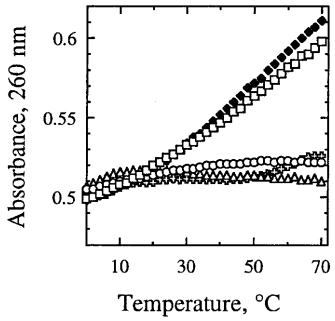

Figure 2. Melting profiles of M-containing single strands in the standard solvent. ◆, A_{21}; □, A_{10}-M-A_{10}; ○, $(A\text{-}M)_{10}$-A; Δ, $(C\text{-}M)_{10}$-C; open plus, $(C\text{-}T)_{10}$.

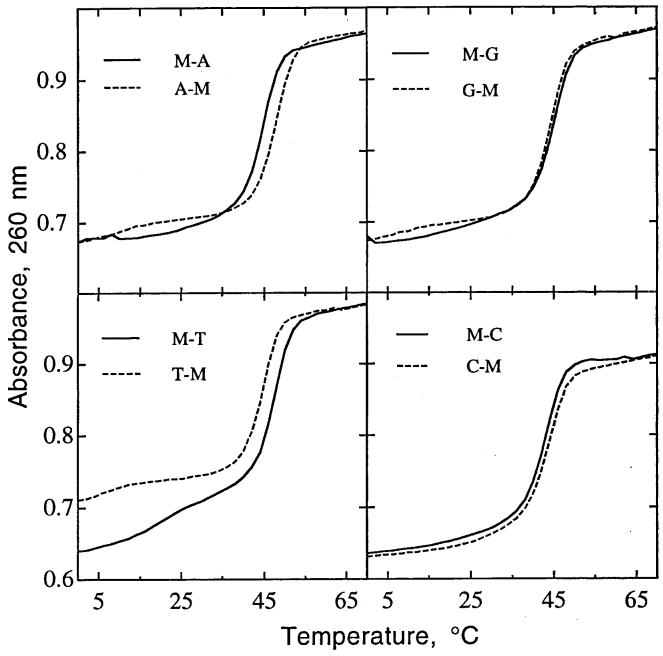

Figure 3. Melting profiles of duplexes with M-containing oppositions in the standard solvent.

of ^{32}P-end-labeled oligomers. The concentrations of A_{10}-X-A_{10} and T_{10}-Y-T_{10} strands were estimated using molar extinction coefficients for poly(dA) at 25°C of ε_{257} = 8600 and for poly(dT) of ε_{265} = 8700. The phosphoramidite of M was from Glen Research Inc.

Solvents

Thermal melting was performed in a standard solvent, 0.15 M NaCl, 0.005 M $MgCl_2$, 0.01 M cacodylate (Na^+), pH 7.0, as in Fossella *et al.* (11).

Duplex and triplex mixtures

Equimolar amounts of strands were mixed in the standard solvent to form each of the four possible target duplexes with X-Y, X-M or M-Y base opposition inserts. For triplex formation, an equimolar amount of the appropriate third strand was added to a duplex at 4°C and incubated for at least 2 h prior to melting.

PAGE analysis

Oligonucleotides were purified by denaturing PAGE (16%, 5% cross- linking) at 1500 V for 4 h in 8 M urea, 90 mM Tris–borate, 2 mM EDTA, pH 8.2, at room temperature. Duplex and triplex formation was monitored by native PAGE (14%, 5% cross-linking) in 20 mM Tris–acetate, pH 6.8, 100 mM NaOAc, 10 mM $Mg(OAc)_2$ at 4°C. Oligonucleotides were 5′-end-labeled with ^{32}P using T4 polynucleotide kinase, and the gel pattern was visualized by autoradiography.

UV melting and 'cooling' experiments

Absorbance–temperature profiles were obtained as described in Fossella *et al.* (11). To obtain equilibrium cooling profiles, temperature was decreased at a rate of 0.1°C/min and the absorbance monitored every 0.1°C; for equilibrium structures, cooling profiles were identical to melting profiles.

RESULTS

M in single strands

To ascertain whether M is truly a strong stacker, as was presumed by Bergstrom and associates (9,10), its ability to stack was evaluated in single-strand sequences where its effect on the thermally-induced hyperchromic change would be readily noticeable. Thus, UV melting profiles were measured on the alternating sequences $d(A\text{-}M)_{10}A$ and $d(C\text{-}M)_{10}C$, as well as on $d(A_{21})$, which, along with $d(C)_n$ oligomers, are known to be well stacked at low temperature (see for example 12,13) and show substantial non-cooperative UV hyperchromic changes and reduction in CD intensity on melting. When similarly examined, the melting of $d(A_{10}\text{-}M\text{-}A_{10})$ (Fig. 2) showed that a single M base in the middle of the A tract, if anything, makes it easier for the oligomer to unstack in comparison with the melting of $d(A)_{21}$. *Alternating* M residues have a much more pronounced effect on the stacking of A or C residues, eliminating their thermally-induced hyperchromic change (Fig. 2). Such behavior would not be expected were the M residues capable of stacking with their A or C nearest neighbors. This effect of M is analogous to the effect of alternating T residues, as in $d(C\text{-}T)_{10}$ (Fig. 2).

M in duplexes

Figure 3 shows melting profiles of M-containing duplexes in the standard solvent. Duplex formation was confirmed by native PAGE under the same ionic conditions (data not shown). Table 1 lists T_m values of M-containing duplexes in cases where M is present in either the homopyrimidine strand, i.e., in A_{10}-X-A_{10}·T_{10}-M-T_{10}

Table 1. Duplexes with X-M or M-Y test oppositions

$A_{10}\cdot T_{10}$ X –M $A_{10}\cdot T_{10}$	T_m values, ° C				$A_{10}\cdot T_{10}$ M –Y $A_{10}\cdot T_{10}$	T_m values, ° C			
Test X–M	Duplex with X–M	Duplex with W–C X–Y	$\Delta T_m = T_{wc}-T_{xm}$	Range for Duplexes with X-Y test mismatches	Test M–Y	Duplex with M–Y	Duplex with W–C X-Y	$\Delta T_m = T_{wc}-T_{my}$	Range for Duplexes with X-Y test mismatches
G–M	44.3	53.0	8.7	46 – 47	M–G	44.6	54.0	9.4	44 – 47
A–M	48.0	51.3	3.3	43 – 45	M–A	44.6	51.3	6.7	43.6 – 46
T–M	44.6	51.3	6.7	43 – 47	M–T	47.6	51.3	3.7	45.3 – 47
C–M	44.1	54.0	9.9	41 – 45	M–C	43.5	53.0	9.5	41 – 43.6
M–M	45.7	–	–	–	M–M	45.7	–	–	–

or in the homopurine strand, A_{10}-M-$A_{10}\cdot T_{10}$-Y-T_{10}, and for comparison the T_m values of duplexes in which the test pair consists of mismatched canonical bases or matched, i.e., Watson–Crick base pairs (taken from ref. 11). It is thereby apparent that M is a truly 'non-discriminating' base, since the stabilities of duplexes with different X-M and M-Y combinations vary very little. All the T_m values are in the range 43–48°C, similar to the values for mismatched oppositions formed with canonical bases, i.e., 41–47°C (11). Comparison of stabilities of duplexes with M on either the homopurine or homopyrimidine strand and the same opposing base on the other strand (e.g., A-M versus M-A or G-M versus M-G) (Fig. 3) provides additional insights. M-G, G-M, M-C and C-M test oppositions all destabilize the duplex by 7–9°C, whether M is on the homopurine or homopyrimidine strand. Moreover, the T_m values of the resulting duplexes are within the range of T_m values for duplexes with mismatched pairs of canonical bases. In contrast, A-M and M-T oppositions destabilize the duplex to a relatively small extent, by only 3.3–3.7°C, and duplexes with both A-M and M-T oppositions are more stable than any with the mismatched canonical X-Y oppositions. This is probably because the small M base does not perturb the geometry of the 'host' d($A_{21}\cdot T_{21}$) duplex. These data support the notion that whatever its stacking tendency, M is well-accommodated within the hydrophobic helical core of base pairs in these duplexes.

M in triplexes

M was also evaluated in a third strand for its ability to bind to target duplex base pairs using the T_{10}-M-T_{10} third strand and A_{10}-X-$A_{10}\cdot T_{10}$-Y-T_{10} target duplexes, with the four possible Watson–Crick X·Y combinations. That such triplexes form was confirmed by the presence of slow migrating triplex bands in PAGE analysis of such mixtures (not shown). Stabilities of the resulting triplexes were determined by UV melting and compared with stabilities of triplexes with all possible matched (Z:X·Y) and mismatched (Z-X·Y) triplets in the test position. Some relevant UV melting profiles are shown in Figure 4. Those profiles display the classical biphasic melting characteristic of triplexes (14). As evident from the T_m values in Table 2, M in the third-strand position of the test triplet drastically weakens third-strand binding to both 'direct' (A·T, G·C) and 'inverted' (T·A, C·G) target base

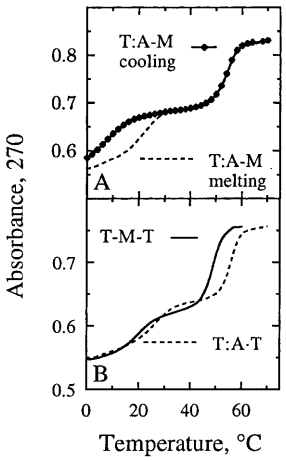

Figure 4. Melting and cooling profiles of triplexes with M-containing test triplets in the standard solvent. (**A**) Note that the melting and cooling profiles differ for the triplex with a T:A-M test triplet. (**B**) Note that the melting and cooling profiles completely coincide for triplexes with a T-M-T and with a T:A·T test triplet.

pairs. The fact that T_m values for all such M-containing triplexes are very similar is in contrast to the wide range of stabilities for triplexes with mismatched test triplets of canonical bases.

To test whether the destabilizing effect of M in the third strand is due to mere absence of hydrogen bonds, the stability of triplexes with M in the test target base pair and T as the third strand residue (T-M-T and T:A-M) were compared with that of the canonical T:A·T triplex. Triplexes with both T-M-T and T:A-M test oppositions were only slightly destabilized and the absence of hydrogen bonds between third strand T and M on the A strand of the duplex hardly matters. The triplex with T:A-M is a metastable structure; to observe its melting, the T_{21} third strand was added to the A-M-containing duplex at 4°C and incubated for just a few hours before melting, to avoid strand exchange. Such exchange does occur, however, after third strand dissociation at 22°C, so that the subsequently observed duplex transition is that of the duplex with A·T rather than A-M in the test position. The fact that the '*cooling*' profile of the T:A-M triplex mixture (Fig. 4) does not coincide with the melting profile of the triplex

Table 2. Triplexes with M-X · Y test triplets

T:A•T TTTTTTTTTTMTTTTTTTTTT
MOTIF AAAAAAAAAAXAAAAAAAAAA
TTTTTTTTTTYTTTTTTTTTT

Test M-X•Y	T_m[1] triplex	T_m[1] range triplexes with Z:X•Y test mismatches	T_m[1] triplexes with canonical Z:X•Y
M-A•T	5.2 ± 0.5	3 – 4	22.6 A:A•T
M-T•A	4.2 ± 0.5	– 5 – 16	–
M-G•C	5.7 ±0.5	9.6 – 18.3	31 C:G •C
M-C•G	5.7 ± 0.5	3.2 –14	–

[1]All T_m values (°C) are for 3→2+1 transitions.

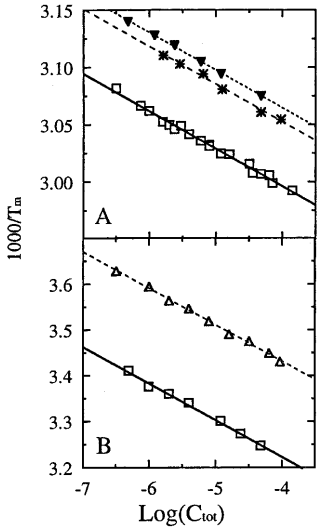

Figure 5. van't Hoff plots for (**A**) the test oppositions A·T (□), M-T (▼) and A-M (✳) in duplexes and (**B**) the test oppositions T:A·T (□) and M-A·T (△) in triplexes.

confirms this exchange and indicates the occurrence of the significantly less stable M-A·T triplex at low temperature.

Thermodynamic analysis

The thermodynamic parameters for helix formation (Table 3) were determined from melting profiles of duplexes with A·T, M-T and A-M test doublets and of a triplex with an M-A·T test opposition in the range of strand concentration from 2×10^{-7} to 10^{-5} M in the standard buffer (15). van't Hoff plots (Fig. 5) for both duplex and triplex association are linear, confirming the all-or-none character of the transitions. ΔH values for the control A·T duplex and T:A·T triplex agree well with those previously reported (16,17). Those for duplexes with M-T and A-M test oppositions are, within experimental error, the same as for the control. The apparent small destabilization of the duplex by M in the test positions can be related to the moderate increase in entropy that must derive from the absence of hydrogen bonding in just one of 21 possible base oppositions. For third strand M test residues, which result in much greater triplex destabilization, ΔH is again only marginally affected, but now ΔS is significantly higher than for the control, with consequent impact on ΔG values.

Table 3. Thermodynamic parameters for helix formation

At test duplex site	At test triplex site	-ΔH, kcal/M	-ΔS, cal/M·°K	-ΔG kcal/M
A•T		140±2	400±8	20.8±3.1
M-T		139±2	408±8	17.4±3.1
A-M		139±2	407±8	17.7±3.1
	T:A•T	58±1	167±3	8.2±1.3
	M-A•T	58±1	179±3	4.6±1.3

DISCUSSION

The finding that 1-(2′-deoxy-β-D-ribofuranosyl)-3-nitropyrrole in the test opposition reduces the stability of DNA *duplexes* only to a small extent that is similar for all its combinations with the four canonical bases suggests that this residue behaves essentially in a sequence-independent manner. Since it is readily spatially accomodated against all canonical bases, it probably does not strain the backbone of the duplex.

In contrast, *triplexes* containing test triplets with M on the *third* strand are highly destabilized relative to those containing test triplets with only canonical bases. As noted, T_m values for all triplexes with M in the third strand are substantially reduced and very similar even when the target pairs are A·T or G·C, i.e., when the target duplex is uninterrupted homopurine·homopyrimidine. Indeed, this destabilization is dramatic (Table 2) even when compared to triplexes with mismatched canonical bases in the third strand. At first sight it seems curious that while M is 'non-discriminating' in both triplexes and duplexes, it is very destabilizing in triplexes but only about as destabilizing in duplexes as canonical base mismatches.

In a single strand, M behaves rather like T in displaying a poor tendency to stack with nearest neighbor canonical bases. This may be so for several reasons. One is that the nitro group is apparently not co-planar with the pyrrole ring, being bent some 7.5° out of plane (9). In addition, the presence of the highly polar nitro substituent must reduce the intrinsic stacking tendency of the pyrrole ring, consistent with many observations (see for example 18) that highly non-polar aromatic structures stack more strongly than more polar ones. Finally, the dipole of M may be counter-productive to stacking with particular nearest neighbor aromatic systems. The effect of a single M will nevertheless be hardly more than the effect of a single T in the center of an A strand, as was observed. A single M-Y or X-M opposition does not lower the stability of a duplex much for the same reason, perhaps only kinking the helix at the locus of the M residue. In a triplex, however, the binding of a third strand with an M residue in the pyrimidine triplex motif may be substantially more sensitive to the additional dipole on the nitro group due to its burden of negative charges, which arises from the crowding of three negatively charged backbones in essentially the same cylindrical volume as a duplex.

Thus, what makes an M-containing *third strand* bind with such reduced affinity is probably not so much that the M base is not a strong stacker (after all, neither is T), nor that it lacks hydrogen bonding capacity, for these same properties have little effect when the M residue is instead buried in one or the other of the target duplex positions (Fig. 4). Rather, it is probably because of the combination of a strong dipole, bulkiness and relative hydrophilicity of the nitro group. The charge density of the test triplet should not be differentially affected by the electronegativity of the M residue either in the target pair or in the third strand. So the differential position effect of M must be due to its relatively greater hydrophobic environment when in a duplex than when on the more water-accessible surface of a triplex. Therefore, small size and aromaticity are probably a necessary but insufficient combination of characteristics to make for a universal 'filler' residue for third strands opposite inverted base pairs. This knowledge should prove instructive in trying to design such a base analog.

ACKNOWLEDGEMENTS

This work was supported by NIH grant GM42936 to J.R.F. and a fellowship from Codon Pharmaceuticals to O.A.

REFERENCES

1 Mizuno,T., Chou,M.Y. and Inouye,M. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 1966–1970.
2 Izant,J.G. and Weintraub,H. (1984) *Cell*, **36**, 1007–1015.
3 Young,S.L., Krawczyk,S.H., Mateucci,M.D. and Toole,J.J. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 10023–10026.
4 Duval-Valentin,G., Thuong,N.T. and Hélène,C. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 504–508.
5 Wang,G., Levy,D.D., Seidman,M.M. and Glazer,P.M. (1995) *Mol. Cell. Biol.*, **15**, 1759–1768.
6 Durland,H.D., Rao,T.S., Bodepudi,V., Seth,D.M., Jayaraman,K. and Revankar,G.R. (1995) *Nucleic Acids Res.*, **23**, 647–653.
7 Kiessling,L.L., Griffin,L. and Dervan,P.B. (1992) *Biochemistry*, **31**, 2829–2834.
8 Berressem,R. and Engels,J.W. (1995) *Nucleic Acids Res.*, **23**, 3465–3472.
9 Bergstrom,D.E., Zhang,P., Toma,P.H., Andrews,P.C. and Nichols,R. (1995) *J. Am. Chem. Soc.*, **117**, 1201–1209.
10 Nichols,R., Andrews,P.C., Zhang,P. and Bergstrom,D.E. (1994) *Nature*, **369**, 492–493.
11 Fossella,J.A., Kim,Y.J., Shih,H., Richards,E.G. and Fresco,J.R. (1993) *Nucleic Acids Res.*, **21**, 4511–4515.
12 Brahms,J., Michelson,A.M. and Van Holde,K.E. (1966) *J. Mol. Biol.*, **15**, 467–488.
13 Brahms,J., Maurizot,J.C. and Michelson,A.M. (1967) *J. Mol. Biol.*, **25**, 465–480.
14 Blake,R.D., Massoulie,J. and Fresco,J.R. (1967) *J. Mol. Biol.*, **30**, 291–308.
15 Marky,L.A. and Breslauer,K.J. (1987) *Biopolymers*, **26**, 1601–1620.
16 Pilch,D.S., Brousseau,R. and Shafer,R.H. (1990) *Nucleic Acids Res.*, **18**, 5743–5750.
17 Dolinnaya,N.G. and Gryaznova,O.I. (1989) *Russian Chem. Rev.*, **58**, 758–777.
18 Guckian,K.M., Schweitzer,B.A., Ren,R.X.-F., Sheils,C.J., Paris,P.L., Tahmassebi,D.C. and Kool,E.T. (1996) *J. Am. Chem. Soc.*, **118**, 8182–8183.

This is paper no. 27 in the series entitled Polynucleotides, of which the last is Dolinnaya,N.G., Ulku,A. and Fresco,J.R. (1997) *Nucleic Acids Res.*, **25**, 1100–1107.