We should like to thank Dr. Joseph S. Fruton for the use of his automatic amino acid analyzer, and Dr. Karen K. Nilsson and Mr. Warren A. Carlson for performance of the analyses. We are much indebted also to Dr. E. Margoliash for information on the separation of a carbamylated hexadecapeptide and a generous gift of crystalline cytochrome *c;* and to Dr. Lucile Smith and Miss Norma Newton for help with cytochrome oxidase assay procedures.

* This investigation was supported by research grants from the United States Public Health Service and the National Science Foundation.

† Predoctoral fellow of the National Science Foundation.

[1] Margoliash, E., in *Haematin Enzymes*, ed. J. E. Falk *et al.* (London: Pergamon Press, 1961), p. 383.

[2] Smith, L., and H. Conrad, *Arch. Biochem. Biophys.*, **63**, 403 (1956).

[3] Smith, L., and H. Conrad, in *Haematin Enzymes*, ed. J. E. Falk *et al.* (London: Pergamon Press, 1961), p. 260.

[4] Davies, H. C., L. Smith, and A. R. Wasserman, *Biochim. Biophys. Acta*, **85**, 238 (1963).

[5] Minakami, S., K. Titani, and H. Ishikura, *J. Biochem.*, **45**, 341 (1958).

[6] Margoliash, E., in *Enzyme Models and Enzyme Structure*, Brookhaven Symposia in Biology, No. 15 (1962), p. 266.

[7] Takemori, S., K. Wada, K. Ando, M. Hosokawa, I. Sekuzu, and K. Okunuki, *J. Biochem.*, **52**, 28 (1962).

[8] Takahashi, K., K. Titani, K. Furino, H. Ishikura, and S. Minakami, *J. Biochem.*, **45**, 375 (1958).

[9] Margoliash, E., N. Frohwirt, and E. Wiener, *Biochem. J.*, **71**, 559 (1959).

[10] Kurzer, F., and A. Lawson, *Org. Syn.*, **34**, 67 (1954).

[11] Hummel, B. C. W., *Can. J. Biochem. Physiol.*, **37**, 1393 (1959).

[12] Moore, S., D. H. Spackman, and W. H. Stein, *Anal. Chem.*, **30**, 1185 (1958).

[13] Keilin, D., and E. F. Hartree, *Biochem. J.*, **41**, 500 (1947).

[14] Camerino, P. W., and L. Smith, *J. Biol. Chem.*, **239**, 2345 (1964).

[15] Schneider, W. C., and V. R. Potter, *J. Biol. Chem.*, **149**, 217 (1943).

[16] Harbury, H. A., *J. Biol. Chem.*, **225**, 1009 (1957).

[17] Harbury, H. A., and P. A. Loach, these PROCEEDINGS, **45**, 1344 (1959).

[18] Margoliash, E., E. L. Smith, G. Kreil, and H. Tuppy, *Nature*, **192**, 1125 (1961).

[19] Rodkey, F. L., and E. G. Ball, *J. Biol. Chem.*, **182**, 17 (1950).

[20] Harbury, H. A., and P. A. Loach, *J. Biol. Chem.*, **235**, 3640 (1960).

[21] Theorell, H., and Å. Åkeson, *J. Am. Chem. Soc.*, **63**, 1812 (1941).

[22] Harbury, H. A., and P. A. Loach, *J. Biol. Chem.*, **235**, 3646 (1960).

[23] Margoliash, E., J. R. Kimmel, R. L. Hill, and W. R. Schmidt, *J. Biol. Chem.*, **237**, 2148 (1962).

[24] Margoliash, E., and E. L. Smith, *J. Biol. Chem.*, **237**, 2151 (1962).

---

## *THE INFLUENCE OF NONCOMPLEMENTARY BASES ON THE STABILITY OF ORDERED POLYNUCLEOTIDES**

BY E. K. F. BAUTZ AND F. A. BAUTZ†

INSTITUTE OF MICROBIOLOGY, RUTGERS UNIVERSITY

*Communicated by E. L. Tatum, October 22, 1964*

It has been shown that the nearest neighbor frequencies in DNA from several species are distinctly different from a random distribution.[1] Subsequently, a certain relationship was recognized between the G + C content in DNA and some of the dinucleotide frequencies.[2] When it was observed in this laboratory that, in phage T4, the dinucleotide frequencies in DNA are faithfully transcribed into mRNA,[3] we became interested to learn whether this species specific distribution of

nucleotides might be in some way related to structural or functional requirements of mRNA, or whether it reflects the result of a nonrandom mutation or selection of nucleotides in DNA. For either one of these possibilities to apply, the stability of a given helical region must depend not only upon the composition, but also upon the sequence of nucleotides.

Taking into consideration the forces responsible for the helical structure of DNA, it appears that the stacking of bases along the helix is greatly stabilized by Van der Waals forces, hydrophobic interactions, and others, their energy exceeding that of the hydrogen bonds between the base pairs.[4, 5] It is very likely that this stacking energy is not identical for any two of the 16 possible combinations of dinucleotides. Indirect evidence obtained by DeVoe and Tinoco[4] from a calculation of molecular orbitals in derivatives of purines and pyrimidines support such a conclusion.

In order to obtain more direct information on base interactions at the polynucleotide level, we have investigated what influence those bases, which are unable to form hydrogen bonds with U in the Watson-Crick fashion, exert upon the pairing of neighboring adenylic acid residues with poly U.

*Materials and Methods.*—The four nucleoside 5′ diphosphates, ADP, CDP, GDP, and UDP were purchased from Schwarz BioResearch, Inc. IDP was obtained from Calbiochem. Poly U and poly A were purchased from Miles Chemical Corp. Pancreatic ribonuclease A, lyophilized, phosphate-free, bacterial alkaline phosphatase, chromatographically purified, and Micrococcal nuclease were all obtained from Worthington Biochemicals.

T1 ribonuclease was isolated from "Sanzyme R" powder of Sankyo Ltd., Tokyo, according to Takahashi[6] modified such that the heat-treated extract was neutralized, filtered, and directly applied to a DEAE-cellulose column, repeating the chromatography four times.

Polynucleotide phosphorylase was prepared from spray-dried *M. lysodeikticus* cells (from Miles Corp.) according to Steiner and Beers[6a] through the acetone step. The extract was frozen in aliquots and kept at −20°C.

The oligonucleotides $(Ap)_nCp$ and $(Ap)_nUp$ were prepared as follows: ADP + CDP and ADP + UDP at a molar ratio of 4 A's to one C or U were polymerized with polynucleotide phosphorylase, and the resulting copolymers were degraded with pancreatic ribonuclease, followed by column chromatography on DEAE cellulose with a linear gradient from 0.01 to 0.8 *M* ammonium bicarbonate pH 8.4. The peak fractions were desalted, and the base compositions of the oligonucleotides were determined as described elsewhere.[3]

Oligonucleotides of the series $(Ap)_nGp$ were prepared by digestion of an AG copolymer with T1 ribonuclease followed by separation of the oligonucleotides. The oligonucleotides terminating in a free 3′ hydroxyl end were prepared by treatment of the ribonuclease digests with alkaline phosphatase at pH 8.0, followed by DEAE column chromatography as described above. In addition, the purity of each oligonucleotide was checked by electrophoresis at 40 V/cm for 90 min in 0.05 *N* ammonium formate pH 3.5.

ApApA and ApApApA were prepared by a limited digest of poly A with micrococcal nuclease according to Lipsett *et al.*;[7] the digest was terminated by shaking the incubation mixture with an equal volume of buffered phenol. The aqueous layer was extracted with ether to remove residual phenol, followed by treatment with alkaline phosphatase and subsequent fractionation on DEAE cellulose.

The four sets of copolymers of the formula $A_nX_m$ (X being either C, G, I, or U) were prepared as follows: each set was synthesized with the same aliquot of polynucleotide phosphorylase under identical conditions, except for varying the input ratio of the nucleoside diphosphates. On all polymers base ratio analyses were performed in duplicate, and on two copolymers of each set the random distribution of nucleotides was ascertained by oligonucleotide frequency analysis of digests with pancreatic ribonuclease or with T1 ribonuclease.[8] End group analysis showed for all polymers a minimal chain length of 40 nucleotides, setting a lower limit for the average molecular weight of 12,000.[9]

The extinction coefficients of the oligo and polynucleotides were determined as described by Stanley[9] and will be given elsewhere.[8]

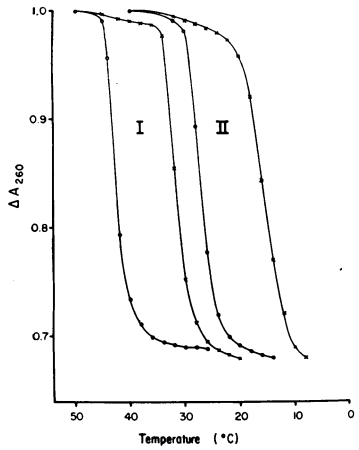


FIG. 1.—Cooling profiles of ApApA (—x—) and of ApApApA (—O—) in the presence of equimolar amounts of poly U; I: in 0.1 *M* $MgCl_2$, II: in 0.5 *M* NaCl, both containing 0.01 *M* Tris, pH 7.5. The relative absorbance at 260 mμ ($A_{260}$) was plotted versus temperature.

The equipment used for the recording of temperature profiles was as described previously.[10]

*Results.—Oligomer-poly U interactions:* The interaction between poly A and poly U to form double- or triple-stranded structures can be observed, if the poly A has a chain length of only a few nucleotides.[7] The oligo A-poly U complexes dissociate over a narrow temperature range showing a characteristic $T_m$ for an oligo A of a given chain length at a given salt concentration. Whereas this interaction has been investigated by others[7] under the assumption that it takes several hours to go to completion, we have observed that under the conditions used here, the oligo A-poly U interaction occurs almost instantly, and that it can be followed most conveniently by measuring the decrease in absorption at 260 mμ during cooling of the sample. Figure 1 shows the cooling profiles of two oligo A + poly U samples at two different salt concentrations. Similar cooling profiles are observed with the oligomers $(Ap)_3C$, $(Ap)_3G$, and $(Ap)_3U$ plus poly U in 0.1 *M* $MgCl_2$ (Fig. 2). The temperature midpoint of the helix-coil transition ($T_m$) is several degrees higher for $(Ap)_3C$ compared to the $T_m$ of the other two. The mixing curves show a minimum in absorption of 260 mμ at a molar concentration of A to poly U of 1:2 for $(Ap)_3C$, whereas $(Ap)_3G$ and $(Ap)_3U$ show a minimum at a ratio of oligomer to poly U closer to 1:1. Therefore, $(Ap)_3C$ seems to establish preferentially a triple-stranded struc-

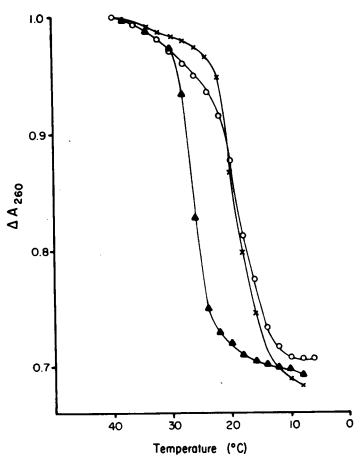


FIG. 2.—Cooling profiles of ApApApC (—▲—), ApApApG (—O—), and ApApApU (—x—) in the presence of equimolar amounts of poly U in 0.1 *M* $MgCl_2$.

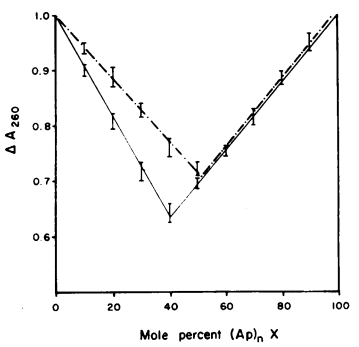


FIG. 3.—Mixing curves for $(Ap)_nX$ with poly U. Samples were mixed in the proportions given on the abscissa, transferred to cuvettes equilibrated for 10 min in the heating block precooled to 10°C, and read at 260 mμ. ——: $(Ap)_4C$ in 0.5 *M* $Na^+$, $(Ap)_3C$, $(Ap)_4C$, $(Ap)_4G$, $(Ap)_4U$ in 0.1 *M* $Mg^{++}$. - - - - -: $(Ap)_3C$, $(Ap)_3G$, $(Ap)_3U$, $(Ap)_4G$, $(Ap)_4U$ in 0.5 *M* $Na^+$; $(Ap)_3G$, $(Ap)_3U$ in 0.1 *M* $Mg^{++}$.

ture together with two strands of poly U, whereas $(Ap)_3G$ and $(Ap)_3U$ appear to favor a double-stranded configuration involving only one strand of poly U (Fig. 3). The three oligomers of the structure $(Ap)_4X$, containing one more adenylic acid residue, all assume a triple-stranded configuration with poly U. In all three cases, the minimum in $A_{260}$ is at approximately 40 mole per cent of oligomer, indicating that the base which is not complementary to U does not contribute to hypochromicity.[11] Table 1 summarizes the $T_m$ values obtained with the oligomers tested: in 0.1 *M* $Mg^{++}$ the oligomers with three adenylic acid residues show the biggest differences; increasing chain length diminishes the differences between C, G, and U. Addition of a phosphate resulting in oligomers of the series $(Ap)_nXp$ also reduces the differential influence of the noncomplementary bases.

TABLE 1

$T_m$ (IN °C) OF EQUIMOLAR MIXTURES OF $(Ap)_nX$ AND POLY U, AND OF $(Ap)_nXp$ AND POLY U

| | In 0.5 *M* $Na^+$ | In 0.1 *M* $Mg^{++}$ | | In 0.5 *M* $Na^+$ | In 0.1 *M* $Mg^{++}$ |
|---|---|---|---|---|---|
| $(Ap)_3C$ | 12 | 26 | $(Ap)_3Cp$ | 5 | 17 |
| $(Ap)_3G$ | 8 | 19 | $(Ap)_3Gp$ | 4 | 13 |
| $(Ap)_3U$ | | 19 | | | |
| $(Ap)_4C$ | 22 | 36 | $(Ap)_4Cp$ | 15 | 30 |
| $(Ap)_4G$ | 18 | 34 | $(Ap)_4Gp$ | 14 | 28 |
| $(Ap)_4U$ | 16 | 32 | $(Ap)_4Up$ | 13 | |

*Copolymer-poly U interactions:* Using different input ratios of ADP versus either CDP, GDP, IDP, or UDP, four series of copolymers have been synthesized with polynucleotide phosphorylase. Each polymer was allowed to complex with poly U at a salt concentration of 0.5 *M* NaCl, and the $T_m$ for each complex was determined by recording its heating profile. The $T_m$ values thus obtained were plotted against mole per cent A of the polymers. As shown in Figure 4, poly AG and poly AI show

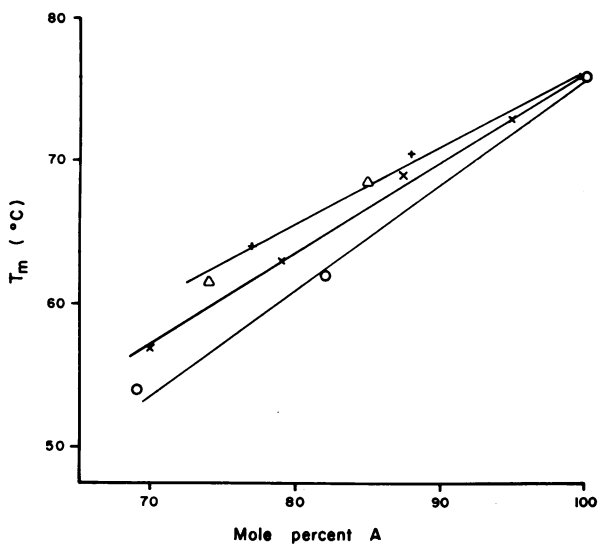


FIG. 4.—Dependence of $T_m$ upon A content of poly AX + U + U. For each copolymer the heating profile in 0.5 *M* NaCl + 0.01 *M* Tris, pH 7.5, was determined at a molar ratio of copolymer to poly U of 1:2. $T_m$ values were plotted versus the percentage of adenylic acid residues in the copolymers. + : poly AG; Δ: poly AI; x: poly AU; O: poly AC.

slightly higher melting temperatures with poly U than the corresponding poly AU polymers, which in turn melt out higher than poly AC. Thus the presence of cytidylic acid residues in poly A seems to weaken the stability of the poly A–poly U complex most strongly, a result which seems to contradict the data obtained from the oligonucleotides, where $(Ap)_nC$ showed the highest $T_m$. The mixing curves in 0.5 *M* NaCl shown in Figure 5 are similar to those obtained with $(Ap)_4X$ (X = C, G, or U) in 0.1 *M* $MgCl_2$ (Fig. 3), suggesting that again three-stranded complexes are preferred, and that the noncomplementary base does not titrate any notable quantity of poly U.

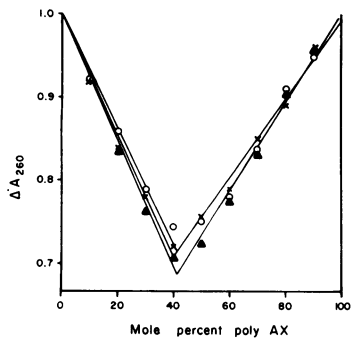


FIG. 5.—Mixing curves of poly U with: poly AC (82% A), —x—; poly AG (77%), —O—; poly AU (79% A), —▲—. The mixtures had been allowed to equilibrate for 1 hr before reading.

*Discussion.*—From the mixing experiments one can deduce that the noncomplementary base does not contribute to the hypochromicity of the complex. Thus, apart from the rather unlikely possibility that the noncomplementary base interacts with poly U under exclusion of the neighboring adenylic acid residue from the complex, there is no indication that the differences in meltings are due to specific interactions between C, G, or U and poly U.

A possible explanation for the data obtained with the oligonucleotides is the difference in the pK values for C, G, and U. Thus the net charge difference of the oligonucleotides might have been responsible for the observed differences in $T_m$. However, all pK values possibly involved are too far from the pH of 7.5 employed to warrant the strong differences observed for $(Ap)_3C$ and $(Ap)_3G,U$; in addition, according to this interpretation, the differences are not expected to diminish as rapidly with increasing chain length of the oligomers; and finally, this interpretation is compatible neither with the $T_m$ values observed for the AG and AI copolymers, nor with the low $T_m$ values obtained for poly AC.

The only interpretation we could conceive to fit both the oligomer and the polymer data involves the assumption that the differences observed are due to variations in interactions between the complementary bases and their neighboring A residues. In that case, on the oligomer level, the melting temperature of an $(Ap)_3X$-poly U complex should be little or not affected, if the noncomplementary base X does not interact with the neighboring A, but swings out freely, and it should be more affected, if the extra base interacts strongly with the neighboring A, pulling it out of the helix and thereby weakening its hydrogen bonding with the opposite uridylic acid residue. In this latter case, the $T_m$ of $(Ap)_3X$ should be reduced markedly. According to this argument we propose that, in a sequence of nucleotides, U and G interact more strongly with a preceding A than does C. This assumption would then account for the low $T_m$ values observed for $(Ap)_3G$ and $(Ap)_3U$ compared to $(Ap)_3C$.

On the polymer level, an extra base interacting strongly with the preceding as well as the following A is likely to participate in stabilizing the helix and should therefore lower the $T_m$ not as much as a base which swings out freely and which thereby in a way interrupts the continuity of the poly A strand.

These assumptions imply that an occasional solitary guanylic acid residue would be held in its place in the helix by the neighboring adenylic acid residues in spite of

its inability to undergo hydrogen bonding with U. The failure of a single guanylic acid residue to loop out of the helix is expected to affect the structure of a poly A-poly U complex in one of the two following ways [Fig. 6 (1)]: the U residue opposite a single G is either forced to loop out of the poly U strand, leading to structure *a*, or it is pushed one place further in the helix to pair with the next A residue, leading to structure *b*. The two neighboring G residues in the bottom half of the complexes 1*a* and 1*b* are depicted as looping out according to the "helix with loop" model by Fresco and Alberts.[11] In the poly AC-poly U complex [Fig. 6 (2)], all C residues are expected to loop out in accord with Fresco's model. While a model of the type 1*b* in Figure 6 is favored through the data obtained from the mixing experiments (Fig. 5), an obstacle to such a model is the question whether the helix is flexible enough to counterbalance the distortion of the angles brought about by the stretching of one strand but not the other. Although this question does not appear to lead to too serious complications as far as the building of molecular models is concerned, it cannot be entirely ruled out.

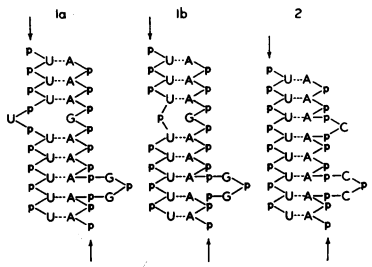


FIG. 6.—Hypothetical structure of a: (1) poly AG-poly U, (2) poly AC-poly U complex.

In case the interpretations offered here will prove to be correct, and if it turned out to be valid to infer from three-stranded complexes of the ribose series to the double-stranded complexes of DNA, then we could expect helix stability to play a predictable role in mutational alterations of DNA as well as in the function of messenger RNA. These more hypothetical assumptions are discussed in more detail elsewhere.[12]

*Summary.*—Model experiments with synthetic polynucleotides have shown the helix-coil transition of an $(Ap)_nX$-poly U or a poly AX-poly U complex to be different for X = C, G, and U. This finding was interpreted as due to differences in stacking energy for different dinucleotides.

Abbreviations used throughout are as follows: A, adenylic acid; C, cytidylic acid; G, guanylic acid, I, inosinic acid, U, uridylic acid; and mRNA, messenger RNA.

* This work was supported by grants USPHS GM 10395 and NSF GB 1882.

† Predoctoral fellow under training grant USPHS 2GM 507.

[1] Josse, J., A. D. Kaiser, and A. Kornberg, *J. Biol. Chem.*, **236**, 864 (1961).

[2] Kaiser, A. D., and R. L. Baldwin, *J. Mol. Biol.*, **4**, 418 (1962).

[3] Bautz, E. K. F., and L. Heding, *Biochemistry*, **3**, 1010 (1964).

[4] DeVoe, H., and I. Tinoco, Jr., *J. Mol. Biol.*, **4**, 500 (1962).

[5] Crothers, D. M., and B. H. Zimm, *J. Mol. Biol.*, **9**, 1 (1964).

[6] Takahashi, K., *J. Biochem.*, **49**, 1 (1961).

[6a] Steiner, R. F., and R. F. Beers, *Polynucleotides* (Elsevier, 1961), p. 374.

[7] Lipsett, M. N., L. A. Heppel, and D. F. Bradley, *J. Biol. Chem.*, **236**, 857 (1961).

[8] Bautz, F. A., in preparation.

[9] Stanley, W. M., Jr., Ph.D. dissertation, University of Wisconsin, 1963.

[10] Bautz, E. K. F., these PROCEEDINGS, **49**, 68 (1963).

[11] Fresco, J. R., and B. M. Alberts, these PROCEEDINGS, **46**, 311 (1960).

[12] Bautz, E. K. F., in *Evolving Genes and Proteins*, ed. H. J. Vogel and V. Bryson (Academic Press), in press.