Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 9033–9036, November 1990
Biophysics

Acidic domains around nucleic acids

(DNA/chemical carcinogens/benzo[*a*]pyrene/aflatoxin/pH)

GENE LAMM AND GEORGE R. PACK

Department of Biomedical Sciences, University of Illinois College of Medicine at Rockford, 1601 Parkview Avenue, Rockford, IL 61107

Communicated by Irving M. Klotz, August 27, 1990 (received for review May 1, 1990)

ABSTRACT **The hydrogen ion concentration in the vicinity of DNA was mapped out within the Poisson–Boltzmann approximation. Experimental conditions were modeled by assuming Na-DNA to be solvated in a buffer solution containing 45 mM Tris and 3 mM Mg cations at pH 7.5. Three regions of high H^+ concentration (>10 μM) are predicted: one throughout the minor groove of DNA and two localized in the major groove near N^7 of guanine and C^5 of cytosine for a G·C base pair. These acidic domains correlate well with the observed covalent binding sites of benzo[*a*]pyrene epoxide (N^2 of guanine) and of aflatoxin B_1 epoxide (N^7 of guanine), chemical carcinogens that presumably undergo acid catalysis to form highly reactive carbocations that ultimately bind to DNA. It is suggested that these regions of high H^+ concentration may also be of concern in understanding interactions involving proteins and noncarcinogenic molecules with or near nucleic acids.**

Counterion condensation around nucleic acids has proven to be an extremely useful concept in understanding many aspects of nucleic acid interactions (1). Little mention, however, has been made concerning the concomitant increase in the hydrogen ion concentration at the surface of these macromolecules. To illustrate the importance of this, we present results of calculations suggesting how high H^+ concentrations near the surface of nucleic acids may account for much of the data regarding the interaction of mutagenic epoxides, such as those derived from benzo[*a*]pyrene and from aflatoxin B_1, with DNA.

Considerable progress has been made concerning the metabolic activation and subsequent binding of carcinogenic molecules to single- and double-stranded nucleic acids. The mechanism for benzo[*a*]pyrene-DNA adduct formation, for example, has been shown to involve multiple enzymatic epoxidations leading to the proximate carcinogenic species benzo[*a*]pyrene-7,8-diol-9,10-epoxide (BPDE) that presumably undergoes protonation to yield a highly reactive carbocation as the ultimate carcinogenic form (2–7). It is the covalent binding of this latter species to N^2 of guanine that is thought to be mutagenic. The strong carcinogen aflatoxin B_1 produced by certain fungi has also been shown to undergo metabolic activation to yield an epoxide [aflatoxin B_1 epoxide (AFBE)] that ultimately binds covalently to DNA, in this case at N^7 of guanine (8, 9). Other aliphatic epoxides such as ethylene oxide and propylene oxide are also known to be mutagenic and bind to N^7 of guanine as well as to N^3 of adenine (9, 10).

Although the exact mechanism of binding of these carcinogens to nucleic acids has yet to be elucidated, the N^2 and N^7 sites of guanine appear to be particularly vulnerable to attack by the epoxide species of these molecules. If these epoxides do undergo acid catalysis to form carbocations, then the attraction of hydrogen ions toward the negatively charged macromolecular surface would be expected to play a role. Epoxide-to-carbocation conversion would predominate in regions of higher H^+ concentration followed by possible covalent attachment to a nearby nucleotide base. DNA with its ion atmosphere would then serve as the (auto)catalyst for this presumably mutagenic reaction in a manner similar to that for the hydrolysis of BPDE (4, 11), implying that nucleic acids in solution do indeed act as acids.

To investigate this hypothesis for the interactions of BPDE and AFBE with DNA, the H^+ concentration near a poly(dG)·poly(dC) nucleotide sequence was mapped out within the Poisson–Boltzmann approximation. The numerical solution to the Poisson–Boltzmann equation utilized a modification of a previously described method (12, 13) to divide the space occupied by the solvent into finite grid elements that contour the van der Waals surface of the DNA. These grid elements extend 33.8 Å along the helical axis (corresponding to the helical repeat distance for one base pair of B-DNA) and 100 Å out from the surface. Grid-element centers were placed in layers with distances to the surface matching values chosen for the radii of the ions in solution. This procedure allows the surface Helmholtz layers, from which specific ion types are excluded, to be closely approximated. Grid elements above and below this central plane were obtained by translating these central elements $3.38n$ Å along the helical axis and by rotating them $36n$ degrees, where $-25 \leq n \leq 25$ for a 51-base-pair segment of DNA.

The solution of the Poisson–Boltzmann equation in this cell of 112-Å radius and 172-Å height begins with an initial guess for the electrostatic potential at each grid element. The analytic solution for the Poisson–Boltzmann potential in the presence of a charged cylinder (14) is used for this guess. The iteration sequence then comprises two steps: (*i*) the Boltzmann equation provides the charge distribution from a potential distribution and (*ii*) the solution to Poisson's equation for an array of point charges gives updated potential values at the grid points. The Boltzmann equation giving the amount of charge within a grid element can be identified as an expression of constancy of the activity of each ion type throughout the cell. The charge for ion type i in element j is obtained from

$$q_j^i = z^i \rho_j^i v_j, \qquad [1]$$

where v_j is the volume of the element, z^i is the ionic charge, and the concentration ρ is given by the Boltzmann relation ($\beta = 1/k_B T$)

$$\rho_j^i \gamma_j^i \exp(\beta z^i \phi_j) = \rho_\infty^i \gamma_\infty^i \qquad [2]$$

Here γ_j^i denotes the activity coefficient of the ion type (excluding the electrostatic potential), ϕ_j is the value of the potential, and the subscript ∞ indicates an element at the cell boundary (where we take $\phi_\infty = 0$). The concentration ρ_∞^i at the

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: BPDE, benzo[*a*]pyrene-7,8-diol-9,10-epoxide; AFBE, aflatoxin B_1 epoxide.

boundary is obtained by normalizing distribution ρ_j^i over all grid elements to give the average concentration $\bar{\rho}^i$,

$$\sum_j \rho_j^i v_j = \sum_j \rho_\infty^i v_j (\gamma_\infty^i/\gamma_j^i)\exp(-\beta z^i \phi_j) = \bar{p}^i \sum_j v_j. \quad [3]$$

Thus

$$\rho_\infty^i = \bar{p}^i V / \sum_j v_j (\gamma_\infty^i/\gamma_j^i)\exp(-\beta z^i \phi_j), \quad [4]$$

where $V = \Sigma_j v_j$ is the total volume in the cell available to the ions (and which is slightly dependent upon ion radius). Eq. 4 shows that $v/[\Sigma_j (v_j/\gamma_j^i)\exp(-\beta z^i \phi_j)]$ is the bulk activity coefficient $\bar{\gamma}^i$. The common neglect of the Boltzmann normalization integral is equivalent to assuming $\bar{\gamma}^i = 1$ and is justified only if the potential is everywhere small. Combining Eqs. 2 and 4 gives for the concentration of ion type i in element j

$$\rho_j = \frac{\bar{p}^i V (\gamma_\infty^i/\gamma_j^i)\exp(-\beta z^i \phi_j)}{\sum_j v_j (\gamma_\infty^i/\gamma_j^i)\exp(-\beta z^i \phi_j)}. \quad [5]$$

The activity coefficient for hard sphere ions of radius r_i (15) is generalized to

$$\ln \gamma_j^i = \frac{4\pi}{3} \sum_k (r_i + r_k)^3 \rho_j^k, \quad [6]$$

and is determined in conjunction with Eq. 5. This term is usually neglected but was found in this study to be significant in regions of high cation accumulation.

The total charge q_k ($= \Sigma_i q_k^i$) in each element obtained from Eqs. 1 and 5 provides the value of the potential at each grid element j, according to the solution of Poisson's equation for an array of discrete charges

$$\phi_j = \sum_k \frac{q_k}{\varepsilon r_{jk}} + \sum_n \frac{Q_n}{\varepsilon r_{jn}}, \quad [7]$$

where ε is the dielectric coefficient of the solvent, r_{jk} is the distance between elements j and k, and the sum over n includes all discrete DNA atomic charges Q_n. Iterating between charges and potentials until self-consistency is obtained provides the desired Poisson–Boltzmann solution.

Eq. 2 can be rewritten in terms of the pH at the cell boundary as

$$p[H]_j = pH + 2.3\beta\phi_j + \log(\gamma_j^H), \quad [8]$$

where we have conveniently defined $p[H]_j = -\log(\rho_j^H)$. This result relates the local H^+ concentration to the Poisson–Boltzmann electrostatic potential. If the activity coefficient correction is neglected, this potential can be obtained from readily available programs such as DELPHI (16) where dielectric boundary effects can be included. Eq. 8 is similar to one used in protein titration analysis (17, 18). Here, however, the source of hydrogen ions is bulk water rather than ionizable side groups on the surface of the macromolecule.

The above algorithm was applied to the B-form of a 51-base-pair segment of DNA. Atomic position and charges were those used in our previous Poisson–Boltzmann calculations (12, 13) and were taken from x-ray diffraction results on the B-form of DNA and from quantum chemical descriptions of the electron distribution. Experimental conditions were modeled by assuming Na-DNA to be solvated in water (ε = 78.4, T = 310 K) along with the univalent salts of Tris (45 mM) and Mg (3 mM) (19). The bulk hydrogen ion concentration $\bar{p}^H$ was chosen to yield a pH value of 7.5 at the cell boundary (100 Å from the surface). Finite ion sizes were taken into account by choosing boundaries defining cell layers contouring the DNA molecule to be placed 1 Å (H), 1.5 Å (Mg), 3 Å (Na), and 4.5 Å (Tris) out from the polyelectrolyte surface and excluding ions from inaccessible layers. For simplicity, anions were chosen to be point charges, as they will not be found in predominance at the surface. While these Helmholtz layers, in conjunction with the inclusion of hard sphere activity coefficients, are at best a crude attempt to account for finite ion sizes within Poisson–Boltzmann theory, the results were relatively insensitive to the specific values chosen. Base sequence is also not expected to significantly affect the data since most of the macromolecular charge resides with the phosphate groups along the phosphate-sugar backbone.

In Fig. 1 the average p[H] within the major and minor grooves as well as elsewhere is displayed as a function of the radial distance from the surface of DNA. Because of the helical twist of the molecule, identification of the grooves begins to lose significance >15 Å from the surface. The rapid decrease in p[H] in the grooves starting near this point is apparent. Since it is believed that protons catalyze the irreversible conversion of epoxide to carbocation and the concentration of hydrogen ions is much less than that of epoxide, the conversion rate is proportional to the proton concentration (i.e., to the p[H]). The predicted decrease of 2–3 p[H] units near the surface of DNA could translate into a 100- to 1000-fold increase in the rate of production of carbocation compared to the bulk rate (2). Local p[H] values 3 units below those of the bulk solvent may also have consequences when considerations of the degree of protonation of base sites themselves become important.

To see the variation in p[H] more clearly, a spatial map of the p[H] in the plane of the central G·C base pair is shown in Fig. 2. Regions of lower p[H] are designated by larger dots; stars denote cells with an average p[H] of <4.5. These latter cells are located in the minor groove at the DNA surface. In addition, two regions in the major groove have low p[H] values: one near N^7 of guanine and another near C^5 of cytosine. The approximate symmetry in their locations derives from the attraction of cations toward the phosphate-sugar backbone and demonstrates the relatively small influence the bases have in determining ionic concentrations at

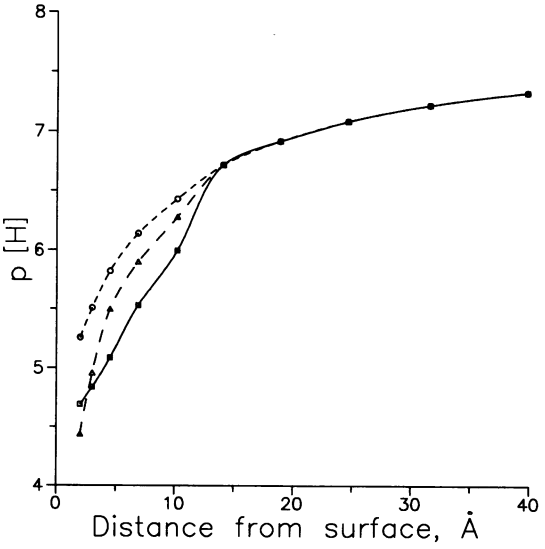

FIG. 1. Average p[H] as a function of the radial distance from the surface of the central base pair of a 51-base-pair segment of poly(dG)·poly(dC) in the B form, calculated according to the Poisson–Boltzmann theory. A solution containing Na-DNA with the univalent salts of 45 mM Tris and 3 mM Mg was assumed. The major (□) and minor (△) grooves as well as elsewhere (○) are shown.

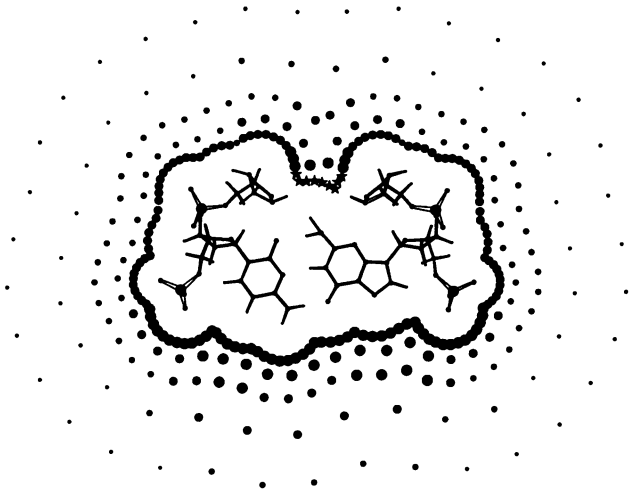

FIG. 2. p[H] in the plane of the central G·C base pair corresponding to the average curves of Fig. 1. The symbols denote the following range of p[H] values for each grid point: stars, p[H] <4.5; progressively smaller circles indicate p[H] ranges from 4.5 to 5; from 5 to 5.5; from 5.5 to 6 and from 6 to 6.5. Grid-element centers are placed in layers 1.25 Å, 2.25 Å, 3.75 Å, 5.63 Å, and 8.45 Å from the DNA van der Waals surface.

the surface. While the approximations involved in assuming the Poisson–Boltzmann algorithm may be questioned, the qualitative nature of the results should hold. Preliminary data from related Monte Carlo calculations support this.

These results are relevant to an understanding of the interaction of carcinogenic molecules that are converted to epoxide intermediates before covalently binding to DNA. Protonation of these intermediates followed by covalent attachment will be more likely in the low p[H] region of the minor groove, where BPDE and propylene oxide bind, and in the major groove near N^7 of guanine, where AFBE binds. Although this correlates well with preferred base sites of these molecules, it does not account for the observed base preference of, for example, aflatoxin B_1 for guanine rather than for adenine (14). Since for BPDE it appears that intercalation into the base-pair stack may be unrelated to covalent binding (4–7, 11), this base preference may derive either from a reaction specificity of the carbocation intermediate with the base site in question or from changes in accessibility to the nucleotide base due to local alterations in the DNA conformation. In any event, failure of the carbocation to locate a site suitable for covalent bond formation would result in hydrolysis of the diol intermediate to a tetrol (2, 3). Reduced p[H] at the surface thus explains the enhanced hydrolysis of BPDE in the presence of DNA (4–7, 11, 20). Furthermore, carbocation production leading to hydrolysis products would occur primarily in regions distinct from those in which covalent bond formation occurs and hence at a different rate. The tetrol-to-adduct product ratio should then be proportional to the relative amounts of hydrogen ions in these domains. If one assumes that carbocation formation leads to covalent bond formation only at the surface of the minor groove (this surface region being represented, say, by the stars in the minor groove in Fig. 2), then the calculated tetrol domain-to-adduct domain hydrogen ion ratio is about 16, clearly of the same order as the observed product ratio of 20 (ref. 6; under slightly different conditions). Calculations performed in the absence of Mg ions show that ionic strength affects p[H] values at the surface of DNA only slightly compared to values farther out into the solvent. This is of course due to the finite ion sizes assumed in the interaction of cations with DNA. Since the covalent binding site for BPDE lies in the minor groove from which cations (except hydronium ions) are essentially excluded, this agrees with the observation that ionic strength affects adduct formation less than it affects tetrol production (6).

Although these conclusions have been made without regard to data on the intercalation of BPDE into DNA, they are consistent with intercalation of unreacted epoxide molecules and possibly of tetrol hydrolysis products. Furthermore, since intercalated epoxides would be subject to protonation depending upon the accessibility of the protonation site to low p[H] solvent regions, covalent bond formation of intercalated epoxides could occur either during intercalation or after reversible ejection back into the solvent. Site specificity of this covalent adduct formation would then naturally correlate with intercalation specificity data.

Although most of the available data concerning epoxide-DNA binding relates to BPDE interactions, the situation regarding aflatoxin derivatives should be similar. An important distinction, however, concerns the preferred covalent binding site at N^7 of guanine in the major groove. This binding is expected to be more sensitive to ionic strength than that for BPDE since the low p[H] domain for this site is more accessible to solvated cations.

BPDE-DNA and AFBE-DNA complexes are important systems in which the influence of increased hydrogen ion concentrations at the surface of DNA can be demonstrated. Recognition of these regions of low p[H] surrounding nucleic acids may prove valuable not only in understanding interactions involving carcinogenic epoxide derivatives but also for describing adduct formation with non-epoxide-forming molecules or with proteins whose conformations are affected by the ionization of side groups. For example, calculations have shown that protonation of proflavin greatly stabilizes its non-covalent association with DNA when either intercalated or externally bound (21). Finally, an appreciation of these acidic domains near double- as well as single-stranded nucleic acids may provide insight into other metabolic mechanisms in which these macromolecules have heretofore been disregarded.

This work was supported by Grant GM29079 from the National Institutes of Health.

1. Anderson, C. F. & Record, M. T., Jr. (1982) *Annu. Rev. Phys. Chem.* **33**, 191–222.
2. Yang, S. K., McCourt, D. W. & Gelboin, H. V. (1977) *J. Am. Chem. Soc.* **99**, 5130–5134.
3. Gelboin, H. V. (1980) *Physiol. Rev.* **60**, 1107–1166.
4. Geacintov, N. E., Ibanez, V., Gagliano, A. G., Yoshida, H. & Harvey, R. G. (1980) *Biochem. Biophys. Res. Commun.* **92**, 1335–1342.
5. Geacintov, N. E., Yoshida, H., Ibanez, V. & Harvey, R. G. (1981) *Biochem. Biophys. Res. Commun.* **100**, 1569–1577.
6. Meehan, T. & Bond, D. M. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 2635–2639.
7. Geacintov, N. E. (1985) *ACS Symp. Ser.* **253**, 107–124.
8. Essigman, J. M., Croy, R. G., Nadzan, A. M., Busby, W. F., Jr., Reinhold, V. N., Buchi, G. & Wogan, G. N. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 1870–1874.
9. Singer, B. & Grunberger, D. (1983) *Molecular Biology of Mutagens and Carcinogens* (Plenum, New York).
10. Lawley, P. D. & Jaman, M. (1972) *Biochem. J.* **126**, 893–900.
11. Koostra, A., Haas, B. L. & Slaga, T. J. (1980) *Biochem. Biophys. Res. Commun.* **94**, 1432–1438.
12. Klein, B. J. & Pack, G. R. (1983) *Biopolymers* **22**, 2331–2352.
13. Pack, G. R., Wong, L. & Lamm, G. (1989) *Int. J. Quantum Chem.* **QBS16**, 1–15.
14. Fuoss, R. M., Katchalsky, A. & Lifson, S. (1951) *Proc. Natl. Acad. Sci. USA* **37**, 579–589.
15. Hill, T. L. (1960) *Introduction to Statistical Mechanics* (Cambridge Univ. Press, Cambridge, U.K.), Chap. 18.
16. Gilson, M. K., Sharp, K. & Honig, B. (1988) *J. Comp. Chem.* **9**, 327–335.
17. Tanford, C. (1961) *Physical Chemistry of Macromolecules* (Wiley, New York).

18. Brenner, S. L. & McQuarrie, D. A. (1973) *J. Theor. Biol.* **39,** 343–361.
19. Yu, F.-L., Bender, W. & Geromino, I. H. (1990) *Carcinogenesis* **11,** 475–478.
20. MacLeod, M. C. & Selkirk, J. K. (1982) *Carcinogenesis* **3,** 287–292.
21. Pack, G. R., Hashimoto, G. M. & Loew, G. H. (1981) *Ann. N.Y. Acad. Sci.* **367,** 240–249.