Journal of Cellular Biochemistry 83:1–13 (2001)

ARTICLES

Mapping of Intrinsic Bent DNA Sites in the Upstream Region of DNA Puff *BhC4-1* Amplified Gene

Adriana Fiorini,[1] Luiz Roberto Basso, Jr.,[2] Maria Luisa Paçó-Larson,[2] and Maria Aparecida Fernandez[1]*

[1]Departamento de Biologia Celular e Genética, Universidade Estadual de Maringá, Maringá, Paraná 87020-900, Brazil

[2]Departamento de Biologia Celular, Molecular e Bioagentes Patogênicos, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, São Paulo 14049-900, Brazil

Abstract We have identified bent DNA sites in the distal and proximal DNA puff *BhC4-1* amplified gene promoter region of *Bradysia hygida*. The 2D modeling of the 3D DNA path and the ENDS ratio values calculated in this promoter region resulted in the identification of ten pronounced bent sites named *Bh*C4B −9 to +1. The 1847 bp fragment (−3697 to −1850) in relation to the transcription start site shows multiple bending sites, *Bh*C4B −9 to *Bh*C4B −4, with periodicity ∼300 bp. The analysis of the other identified bent region, starting at position −957, reveals that the *Bh*C4B +1 bent site colocalizes with the putative *BhC4-1* minimal promoter. The sequence analysis of bent site *Bh*C4B −4 shows a distribution of dA•dT at ∼10 bp intervals between the middle of each tract, but intervals with more than one turn, ∼20 bp, two helix turns, were detected in the other bent sites described here. The bent sites *Bh*C4B −6 and *Bh*C4B −4, contain two consensus sequences, with 60 bp each. The apparent molecular weight of fragments in the *BhC4-1* promoter region were estimated in agarose gels and compared with the data obtained in polyacrylamide gels without and with ethidium bromide. The mobility reduction ratios (*R*-values) were determined, and a high *R*-value, 1.80, for a 1215 bp fragment in the distal promoter region and a 1.23 significant *R*-value for a 662 bp fragment in the proximal segment were found. To further analyze the predicted bent DNA sites in these fragments, the 2D trajectories of the 3D DNA path and other parameters, AT percentage, roll angle, ENDS ratio and ΔG, were determined. The role of these bent sites in the *BhC4-1* transcription regulation is discussed. J. Cell. Biochem. 83: 1–13, 2001. © 2001 Wiley-Liss, Inc.

Key words: intrinsic bent site; amplified region; *BhC4-1* gene promoter; C4 puff; *Bradysia hygida*

Intrinsic bent DNA sites promote a curvature in the DNA helix and are established by structural motifs A/T rich in phase with DNA helical ∼10 bp repeats or multiples of them [Crothers et al., 1990]. Bent sites represent important *cis*-acting elements in gene regulatory regions since they permit the interaction between enhancer binding proteins and transcription factors which results in the modulation of the starting and of the levels of transcription [Williams et al., 1988]. Curved DNA was found in replication origins of bacteria and some viruses where the molecule uncoils and opens to initiate DNA synthesis [Calladine and Drew, 1997], and is also present in prokaryotic promoters [Bracco et al., 1989; Hagerman, 1990; Perez-Martin et al., 1994; Perez-Martin and de Lorenzo, 1997; de Souza and Ornstein, 1998]. However, little is known about bent DNA in eukaryotic promoters [Wada-Kiyama and Kiyama, 1994; Delabre et al., 1995; Marilley and Pasero, 1996; Nair, 1998; Bash et al., 2001]. Bent DNA can also be found in recombination events [Milot et al., 1992], fragile sites [Palin et al., 1998], tandemly-repeated DNA sequences [Hibino et al., 1993; Pasero et al., 1993], replication origins [Linial and Shlomai, 1988; Caddle et al., 1990], Matrix Attachment Regions (MARs) [Homberger, 1989; von Kries et al., 1990; Bouliskas, 1993], and MARs associated with replication origins [Anderson, 1986; Snyder et al., 1986;

Grant sponsor: Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP); Grant sponsor: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq); Grant sponsor: The Third World Academy of Sciences (TWAS).

*Correspondence to: Maria Aparecida Fernandez, Departamento de Biologia Celular e Genética, Universidade Estadual de Maringá, Av. Colombo, 5790, 87020-900, Maringá, Paraná, Brazil. E-mail: mafernandez@uem.br

Received 12 December 2000; Accepted 12 March 2001

© 2001 Wiley-Liss, Inc.

DOI 10.1002/jcb.1188

Krajewski and Razin, 1992], indicating a relationship between secondary DNA structure and function.

One possible approach to identify alternative DNA structures is the comparison between the migration rates of a given fragment in polyacrylamide and agarose gel electrophoresis [Marini et al., 1982]. Another approach to identify secondary DNA structures is the use of computer modeling [Pasero et al., 1993; Delabre et al., 1995].

We are interested in the analysis of developmentally amplified DNA puff genes of sciarids, which is a suitable system for investigating replication and transcription. Replication initiation sites active during amplification were mapped near these genes [Liang et al., 1993; Yokosawa et al., 1999]. Furthermore, the analysis of the pattern of expression of these genes has shown that their expression is developmentally regulated, accompanying the formation of the respective DNA puffs [Santelli et al., 1991; Paçó-Larson et al., 1992; Wu et al., 1993; Monesi et al., 1995; Fontes et al., 1999]. Here we report the identification of ten intrinsic bent DNA sites in the promoter region of the DNA puff *BhC4-1* gene of *Bradysia hygida*. Four of the identified bent sites are localized near the putative *BhC4-1* minimal promoter, whereas the other six are localized further upstream from the gene. The possible role of these bent sites in the regulation of *BhC4-1* transcription is discussed.

MATERIALS AND METHODS

Restriction Fragments

The 3990 bp fragment comprises 293 bp of *BhC4-1* gene coding sequences and 3697 bp of upstream flanking sequences. The 3990 bp fragment and the set of deletions employed in this study were cloned in the phagemid pT7T3 18 U (Pharmacia). Plasmid DNA was extracted using the CTAB method [Del Sal et al., 1989] and restriction enzymes were employed according to the manufacturer's instructions.

Agarose and Polyacrylamide Gels Electrophoresis

The 0.7% agarose gels were run in TBE (45 mM Tris–borate, 1 mM EDTA, pH 8.0) without ethidium bromide (EthBr) at 3.5 V/cm at room temperature for 5 hr to determine the apparent molecular weight of the analyzed fragments. The 6% TBE polyacrylamide gels without EthBr were run in TBE at 7 V/cm at 4°C, for 18–44 hr, depending on the molecular weight of the analyzed fragment. TBE polyacrylamide gels containing 1 μg/ml EthBr were run in TBE at 7 V/cm at room temperature. All gels were stained after running with 1 μg/ml of EthBr and photographed under UV light. The 1 kb ladder (GIBCO BRL) was utilized as molecular weight marker. In all three gels systems, an index was calculated by the ratio between the real molecular weight, determined by the number of base pairs of the fragments, and the apparent molecular weight of the fragment in the gel (index = real/apparent molecular weight). To determine the mobility reduction of each fragment, an *R*-value was calculated from the ratio between the agarose index and the polyacrylamide index in each system, with or without ethidium bromide.

Computational Analysis of DNA Curvature

The 2D projection of the 3D DNA path was calculated by using the algorithm of Eckdahl and Anderson [1987] and the helical parameters as described by Bolshoy et al. [1991], Pasero et al. [1993], and Marilley and Pasero [1996]. One of these parameters was the ENDS ratio, which shows the ratio of the contour length of the segment helical axis to the shortest distance between the fragments ends and was analyzed with a 120 bp window width and a 10 bp step. A second parameter was ΔG, which was calculated with a 10 bp window width and a 1 bp step, to predict the stability of the DNA duplex. Variations of roll angle, which estimates the base pairs rolling-open along their long axes, was calculated with a 100 bp window width and a 10 bp step. By convention, if base pairs open toward the minor-groove side, the roll angle is positive. On the other hand, negative roll angles that cause a narrow minor groove suggest a bent region [Calladine and Drew, 1997]. The AT percentage was calculated with a 100 bp window and a 5 bp step.

Sequencing and DNA Analysis

The 3990 bp *Eco*RI fragment was subcloned into the vector pT7T3 18U (Pharmacia). Overlapping deletions were generated by exonuclease III using the double-stranded Nested Deletion Kit (Pharmacia). Both strands were completely sequenced using the T7 sequencing Kit (Pharmacia). Specific primers were

designed in both strands to cover gaps left by the exonuclease generated deletions. The obtained sequences were processed and analyzed using the softwares PCGENE version 6.01 (Intelligenetics, Mountains View), DNASIS version 6.00 (Pharmacia) and BLASTN 2.0 (www.genome.ad.jp). The Genbank accession number for 3990 bp is U13892.

RESULTS

Electrophoretic Behavior of the *BhC4-1* Promoter Region

In order to identify potential bent DNA sites in the *BhC4-1* promoter region, we have analyzed, in agarose and polyacrylamide gels, fragments from a 3990 bp region that contains 293 bp of coding sequence and 3697 bp of upstream regulatory sequences. Polyacrylamide gels, which form orderly connections that hinder the migration of irregular DNA structures, are used to characterize DNA fragments according both to their size and form [Marini et al., 1982]. In contrast, the analysis of the same fragments in agarose gels, which form an irregular mesh, provides information only about the fragment size. The comparison between the electrophoretic behavior of a given fragment in both gel systems enables the identification of fragments containing atypical DNA structures. The use of the intercalating agent ethidium bromide in polyacrylamide gels confirms the presence of these structures by abolishing the intrinsic DNA curvature and providing a straight path in the fragments [Diekmann and Wang, 1985]. The mobility reduction of different fragments in the *BhC4-1* promoter region was measured in all three gel systems, agarose gels (AG), polyacrylamide gels (PA) and polyacrylamide gels with ethidium bromide (PA+EthBr). For each analyzed fragment, an index was calculated by the ratio between the real molecular weight determined by sequencing and the apparent molecular weight of the fragment in the gel (index = real/apparent molecular weight). After that, to determine the fragment mobility reduction, the R-value was calculated by the ratio between the agarose index and the polyacrylamide index, without and with ethidium bromide. R-values between 0.9 and 1.19 indicate no alteration in the fragment mobility and R-values ≥ 1.2 a reduced mobility indicative of bent DNA [Homberger, 1989]. R-values < 0.9 indicate faster mobility, characteristic of fragments with bent regions in both fragment ends [Wu and Crothers, 1984].

The electrophoretic behavior analysis of fragments in the *BhC4-1* promoter region is shown in Figures 1 and 2. The 891 bp fragment (-3697 to -2806) localized at the extremity of the 3990 bp fragment, did not show mobility alterations, R-value 1.01 (Fig. 1a, lanes 1 and 2). The analysis in polyacrylamide gel with ethidium bromide (PA+EthBr) shows that the 891 bp fragment migrates asexpected from its known size, which is confirmed by R-value around 1.0 (data not shown). These results indicate that this fragment either does not contain bent DNA regions or that the bent regions are not centralized in this fragment. On the other hand, the 1418 bp (-2780 to -1362) and the 1215 bp (-2780 to -1565) fragments show a high reduction in electrophoretic mobility with R-values 1.70 and 1.82, respectively, indicative of the presence of bent sites in these fragments (Fig. 1a, lanes 1 and 2). When the same fragments were analyzed in PA+EthBr gels, R-values around 1.0 were obtained, confirming the presence of bent sites. The pT7T3 18U phagemid DNA migrates at 2800 bp molecular weight in AG gels, and up to 3200 bp in PA gels, which results in R-values of 0.95 and 0.90, respectively, compatible with the 2890 bp vector's actual size.

Figure 1b shows the results of the analysis of two other fragments in the *BhC4-1* promoter region, the 1207 bp fragment (-2005 to -798) and the 839 bp fragment (-1637 to -798). As can be observed, these two fragments present distinct migration behaviors (Fig. 1b, lanes 3 and 4). The 1207 bp fragment did not show a significant reduction in mobility (R-value 1.13) but when 368 bp are removed from the 5′ end of this fragment, to generate the 839 bp fragment, an R- value of 0.84 is found. Analysis of this 839 bp fragment in polyacrylamide gels shows that it presents faster mobility, indicating that this fragment should contain a bent region in both of its extremities [Wu and Crothers, 1984]. If this is correct, a bent region must be present near the center of the 1207 bp fragment, which wasn't detected in the analysis of this fragment. In fact, the low R-value 1.13 observed for the 1207 bp fragment suggests that if a bent region is present in this fragment it isn't centralized in it. Hence, the faster mobility observed for the 891 bp fragment can only be due to one bent

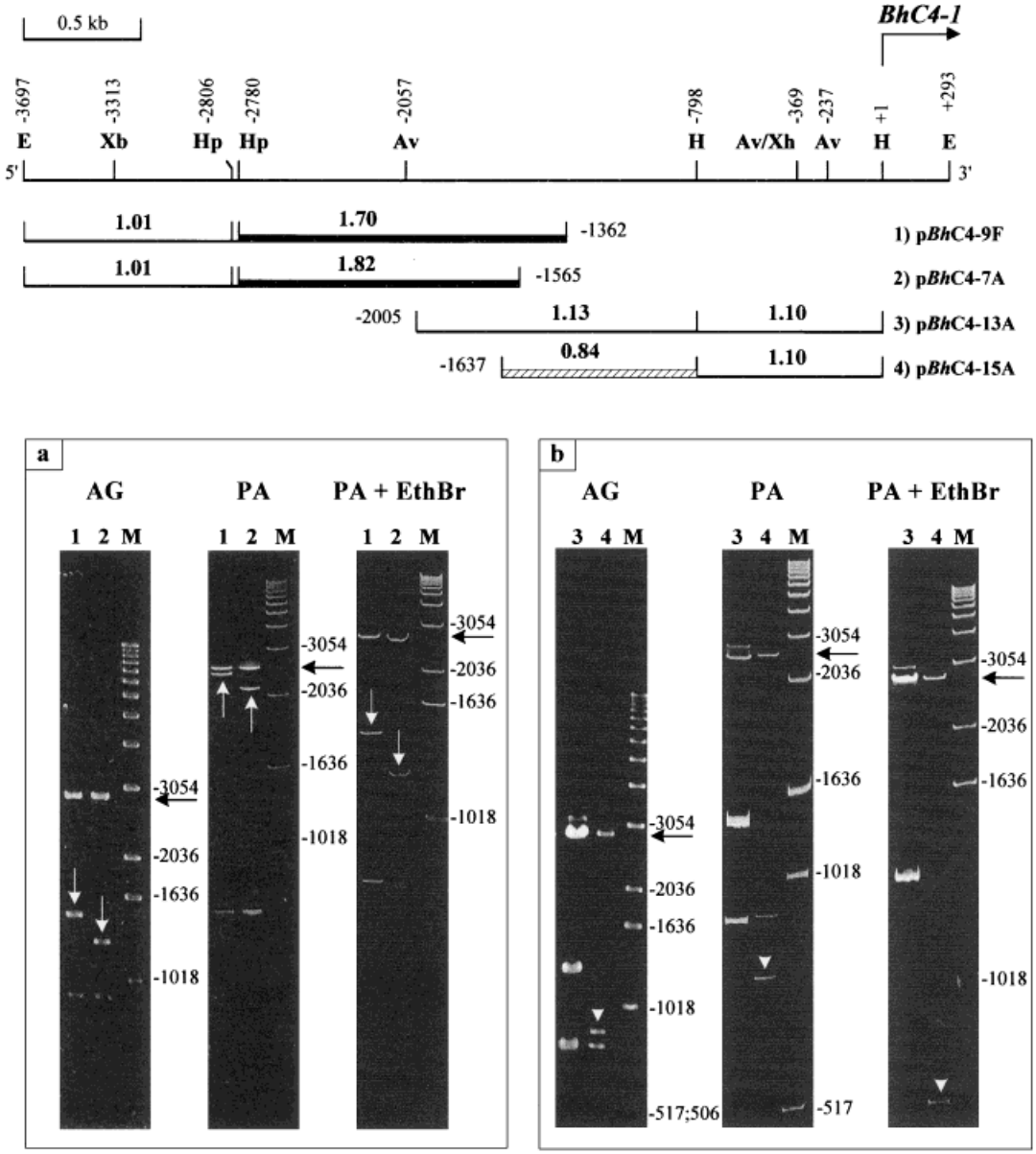

Fig. 1. Electrophoretic behavior analysis of fragments in the *BhC4-1* gene promoter region. The restriction map of the *BhC4-1* promoter region (− 3697 to + 293) is shown on top of the figure. The arrow on top of the diagram indicates the *BhC4-1* start site and the direction of transcription. The four analyzed deletions are represented bellow and the numbers on the ends of the diagrams represent the position of the end of each deletion, in relation to the transcription start site (+ 1). The bold numbers above each diagram are the *R*-values, calculated by the ratio agarose/polyacrylamide without EthBr index (see Material and Methods). The thicker lines represent fragments with reduced mobility, indicated by white arrows in the gels (bellow), which present *R*-value ≥ 1.2. The thinner lines represent fragments with no mobility alteration, which present *R*-values between 0.9 and 1.19. The hatched box represents a fragment with increased mobility, indicated by a white arrowhead in the gels (bellow), which presents *R*-value < 0.9. The black arrow points to the 2890 bp linearized pT7T3 18U phagemid. **A, lane 1:** deletion clone p*Bh*C4-9F after digestion with *Hpa*I, and **lane 2**, deletion clone p*Bh*C4-7A after digestion with *Hpa*I. **b, lane 3:** deletion clone p*Bh*C4-13A after digestion with *Hin*dIII, a fragment of weaker intensity observed at 3054 bp is a partially digested fragment, and **lane 4**, deletion clone p*Bh*C4-15A after digestion with *Hin*dIII. M, 1 kb ladder molecular weight (Gibco BRL). E, *Eco*RI; Xb, *Xba*I; Hp, *Hpa*I; Av, *Ava*I; H, *Hin*dIII; Xh, *Xho*I. (AG) 0.7% agarose gel, (PA) 6% polyacrylamide gel without ethidium bromide and (PA + EthBr) 6% polyacrylamide gel with ethidium bromide.

region that should be localized near position − 798. The analysis of the 799 bp fragment (− 798 to+ 1) indicates small mobility alteration, *R*-value 1.10 (Fig. 1b, lanes 3 and 4).

The existence of no centralized bent regions in the 799 bp fragment was further investigated by the analysis of smaller restriction fragments in this region (Fig. 2). Initially the 3990 bp fragment, contained in clone p*Bh*C4-4, was digested with *Eco*RI and *Xho*I, which resulted in a 3328 bp fragment (− 3697 to − 369) and in a 662 bp fragment (− 369 to + 293). The 3328 bp fragment shows low reduced mobility, *R*-value 1.11, that suggest the existence of bent DNA sites in this segment whereas centralized bent DNA was found in the 662 bp fragment, *R*-value 1.23 (Fig. 2, lane 1). The electrophoretic analysis of the p*Bh*C4-9A deletion clone cleaved by *Xho*I+*Hin*dIII (Fig. 2, lane 2) or by *Xho*I+*Hin*dIII+*Ava*I (Fig. 2, lane 3) shows three fragments, localized upstream position − 798 with reduced mobility. Fragments localized downstream position − 798 did not show significant alteration in mobility, which is probably due to the fact that the bent site observed in the previously investigated 662 bp fragment (− 369 to + 293) is no longer centralized in these smaller fragments. The 1688 bp fragment (− 2486 to − 798) shows a high *R*-value 1.50, which is lower than the one observed for the 1418 bp fragment (− 2780 to − 1362), R-value 1.70 (Fig. 1a, lane 1). This R-values difference could be attributed to the decentralization of the bent sites in 1688 bp fragment since this fragment is 294 bp shorter at 5′ end and 564 bp extended at the 3′end than the 1418 bp fragment. This suggestion is supported by the analysis of the behavior of fragments generated after the digestion of the 1688 bp (− 2486 to − 798) fragment with *Ava*I (Fig. 2, lane 3). The 429 bp fragment (− 2486 to − 2057) presents a well-centralized bent DNA site shown by the high *R*-value 1.37, while the 1259 bp fragment (− 2057 to − 798) shows a low *R*-value, 1.18.

The analysis of the p*Bh*C4-17A deletion clone after being released from the vector provides a 1430 bp fragment (− 1137 to + 293) and extends in 768 bp the previously described 662 bp bent fragment (− 369 to + 293). This fragment shows a reduced mobility, *R*-value 1.20, thus confirming the presence of the bent DNA site or sites in the proximal *BhC4-1* gene promoter region (Fig. 2, lanes 1 and 4). The region spanning from position − 3697 to − 1565 was further investigated after digesting the deletion clone p*Bh*C4-7A with *Hpa*I (Fig. 2, lane 5). Three fragments result from this digestion: an 891 bp, (− 3697 to − 2806), a 723 bp, (− 2780 to − 2057) and a 492 bp, (− 2057 to − 1565). The first one, the 891 bp fragment, *R*-value 0.90, was also shown in Figure 1a (lanes 1 and 2) *R*-values 1.01, confirming the absence of bent DNA sites in this fragment or the existence of decentralized ones. The two other fragments, of 723 bp and 492 bp, *R*-values of 1.37 and 1.45 respectively, centralize and prove that at least two important bent DNA sites occur in the 1215 bp fragment (− 2780 to − 1565), shown in Figure 1a, (lane 2), which display the highest *R*-value observed in this work.

2D Trajectory of the Fragments Analyzed by Electrophoresis

To visualize the 2D trajectory of some fragments and their relationship with the experimentally determined *R*-values, four of the fragments analyzed by agarose and polyacrylamide gels were submitted to theoretical 2D modeling (Fig. 3a–d). In the Figure 3a, the 891 bp fragment (− 3697 to − 2806) shows an interesting 2D model structure. Two symmetric bent regions, in inverted directions, were predicted by the 2D modeling, which results in a non-curved DNA major segment. In agreement with this result, in the electrophoretic behavior analysis of this fragment, no mobility alteration was found, *R*-value 1.01 to 0.90 (Fig. 1a, lanes 1 and 2 and Figure 2, lane 5). Similar structure, with no mobility reduction, was previously reported [Goodsell and Dickerson, 1994]. Figure 3c, shows a representative bent region, near position − 798, found in the 839 bp fragment (− 1637 to − 798). This fragment displayed increased mobility in the electrophoretic behavior analysis, *R*-value 0.84 (Fig. 1b, lane 4). In contrast to the analysis of Wu and Crothers [1984], using circular permutation, in our work the analyzed fragment didn't require structural bent DNA in both of the fragments ends for faster electrophoretic mobility. Other fragments of this region present bent DNA in both fragments ends but didn't show faster mobility, possibly because their bent extremities are curved to the same direction (data not shown). The other two fragments (2D trajectory shown in Fig. 3b and d) showed pronounced reduction in mobility, and the 2D modeling reveals their curved shape. These are

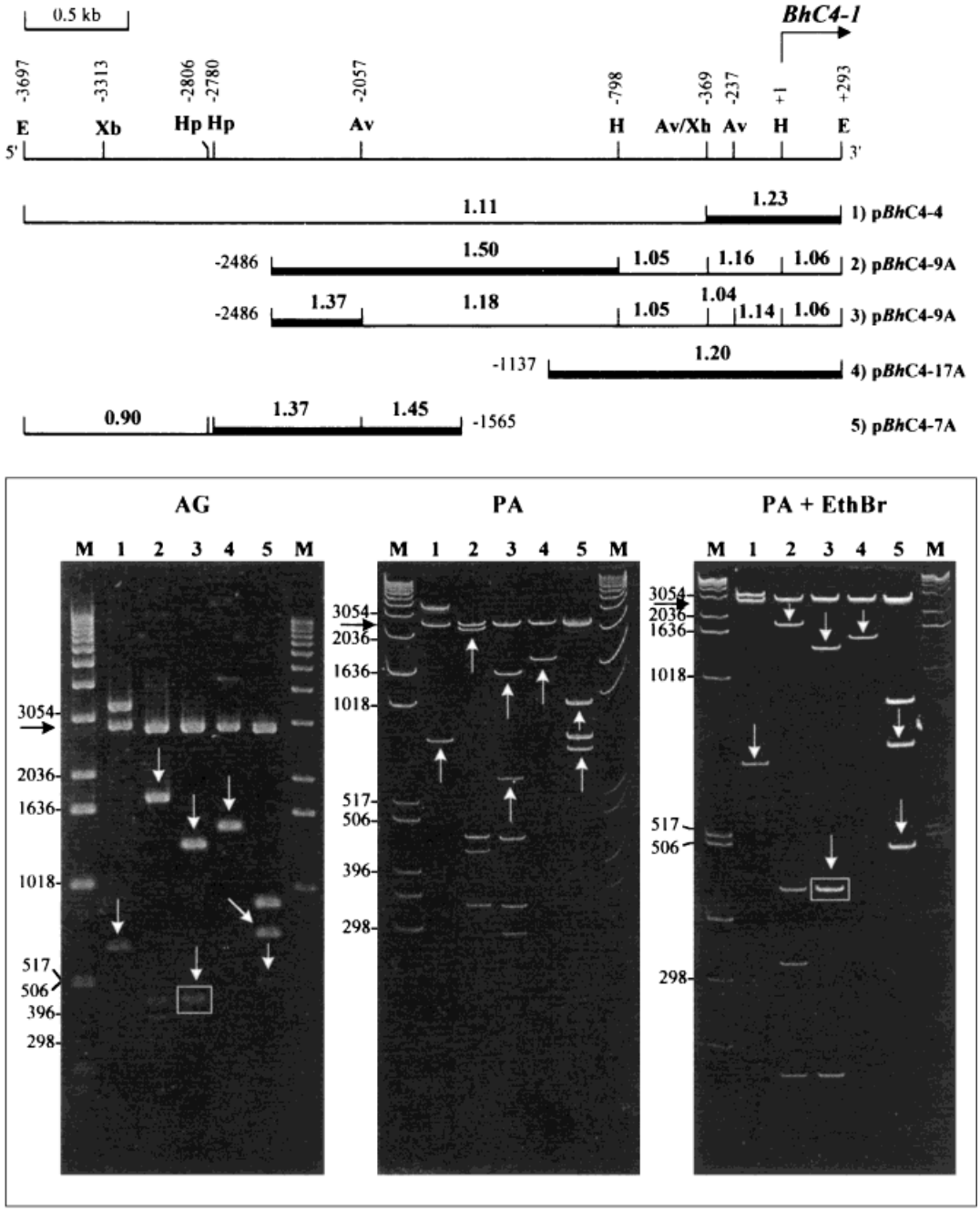

Fig. 2. Electrophoretic behavior analysis of fragments in the *BhC4-1* gene promoter region. The restriction map of the *BhC4-1* promoter region (−3697 to +293) is shown on top of the figure. The arrow on top of the diagram indicates the *BhC4-1* start site and the direction of transcription. The five analyzed fragments are represented bellow and the numbers on the ends of the diagrams represent the position of the end of each deletion in relation to the transcription start site (+1). The thicker lines represent fragments with reduced mobility, *R*-value ≥1.2, indicated by white arrows in the gels, and the thinner lines, fragments with no mobility alteration, values between 0.9 and 1.19. The black arrow at the left hand side of the gels (bellow) indicates the 2890 bp linearized pT7T3 18U phagemid. **Lane 1**, clone p*Bh*C4-4 after digestion with *Xho*I; **lane 2**, deletion p*Bh*C4-9A after digestion with *Hin*dIII + *Xho*I; **lane 3**, deletion p*Bh*C4-9A after digestion with *Hin*dIII + *Xho*I + *Ava*I; **lane 4**, deletion p*Bh*C4-17A after digestion to liberate the whole insert; **lane 5**, deletion p*Bh*C4-7A after digestion with *Hpa*I + *Ava*I. M, molecular weight 1 kb ladder (Gibco BRL). E, *Eco*RI; Xb, *Xba*I; Hp, *Hpa*I; Av, *Ava*I; H, *Hin*dIII; Xh, *Xho*I. (AG) 0.7% agarose gel, (PA) 6% polyacrylamide gel without ethidium bromide and (PA + EthBr) 6% polyacrylamide gel with ethidium bromide.

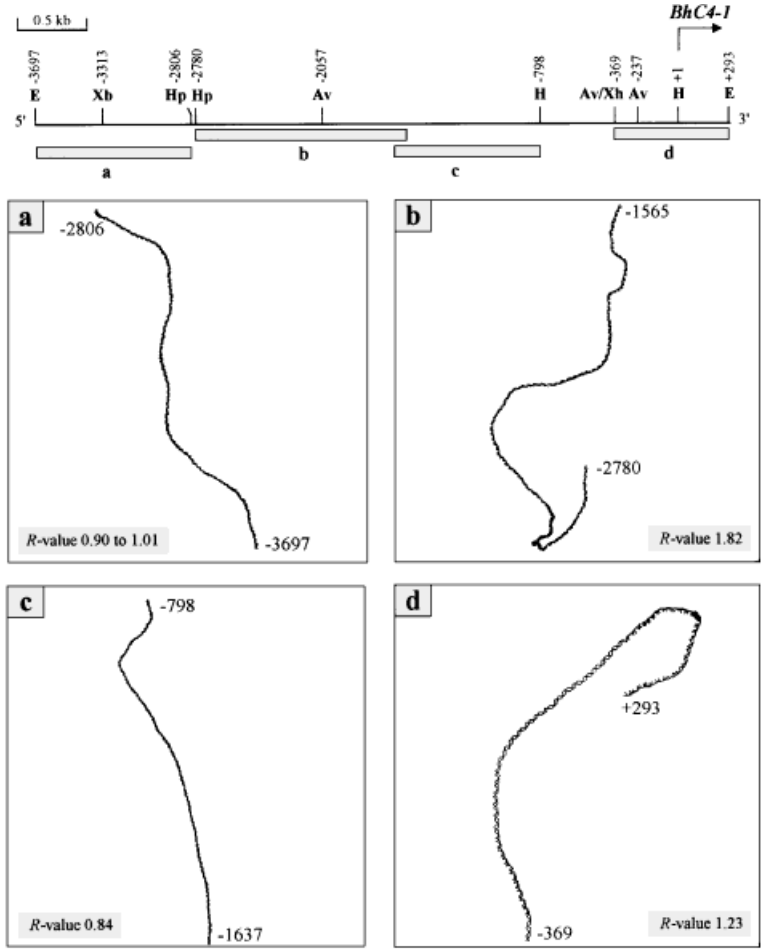

Fig. 3. Relationship between electrophoretic behavior and 2D projection of 3D DNA path of four representative fragments. The restriction map of the *BhC4-1* promoter region (−3697 to +293) is shown on top of the figure. The arrow on top of the diagram indicates the *BhC4-1* start site and the direction of transcription. The bars below the restriction map represent the four analyzed fragments. **a:** The 891 bp fragment, −3697 to −2806 position, displays two curved regions in inverted directions, which suggests no reduction in mobility, *R*-values 0.90 to 1.01. **b, d:** show typical DNA curved segments in computer modeling, *R*-values 1.82 and 1.23, 1215 bp fragment (−2780 to −1565), and 662 bp fragment (−369 to +293) respectively. **c:** The 839 bp fragment, −1637 to −798, displays a bent region in one of the fragment ends, and presents fast electrophoretic mobility, *R*-value 0.84. E, *Eco*RI; Xb, *Xba*I; Hp, *Hpa*I; Av, *Ava*I; H, *Hin*dIII; Xh, *Xho*I.

respectively, the 1215 bp fragment (−2780 to −1565), which was submitted to electrophoretic analysis in Figure 1a, (lane 2, *R*-value 1.82), and the 662 bp fragment (−369 to +293), the analysis of which is shown in Figure 2 (lane 1, *R*-value 1.23).

Computational and Bent DNA Sites Sequence Analyses of the Entire 3990 bp Fragment

To visualize theoretical intrinsic bent sites in the entire studied 3990 bp fragment, the 2D trajectory and the ENDS ratio were calculated as described in Material and Methods (Fig. 4a and b). Ten pronounced theoretical bent sites were found which were named *Bh*C4B −9 to +1. In the distal *BhC4-1* promoter segment (position −3697 to −1850), six bending sites (*Bh*C4B −9 to *Bh*C4B −4) presenting a periodicity of ~300 bp, were found (Fig. 4b). The cumulative effect of these bending sites contributes to a complete curvature in this fragment (Fig. 4a). In the proximal *BhC4-1* promoter segment, four significant bent sites at position −957, −467, −169 and −39 were predicted by the ENDS ratio analysis (Fig. 4b). The sequences of these ten bent sites are shown in Figure 5. A distribution of the dA•dT tracts in 10 bp intervals, between the middle of each tract, can be observed in *Bh*C4B −4 site. However, intervals with more than one turn, ~20 bp, which corresponds to 2 helix turns, are detected in other bent regions described here, and in

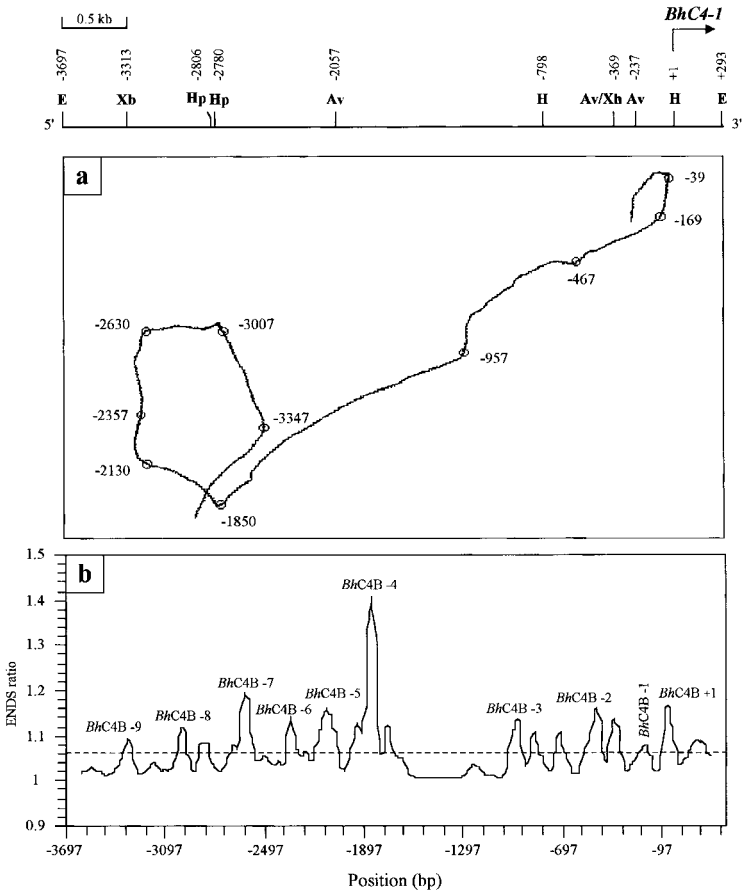

Fig. 4. 2D projection of 3D DNA path and ENDS ratio of the 3990 bp fragment. The restriction map of the *BhC4-1* promoter region (− 3697 to + 293) is shown on top of the figure. The arrow on top of the diagram indicates the *BhC4-1* start site and the direction of transcription. The 2D projection (**a**) and the ENDS ratio (**b**) were calculated as described in Materials and Methods. The ten pronounced identified bent sites were named *Bh*C4B − 9 to + 1. The 1847 bp fragment (− 3697 to − 1850) presents multiple bending sites with periodicity ~300 bp, which contributes for a complete curvature in this region. The 1250 bp fragment (− 957 to + 293) presents four important bent sites at positions − 957, − 467, − 169 and − 39 that result in changes in the trajectory of the fragment. E, *Eco*RI; Xb, *Xba*I; Hp, *Hpa*I; Av, *Ava*I; H, *Hin*dIII; Xh, *Xho*I.

bent sequences reported elsewhere [Dlakic and Harrington, 1998]. All bent DNA sites considered in this work display the ability to form a narrow minor groove that causes a major curved fragment. The most characteristic is the *Bh*C4B − 4, in which two stretches with 4 adenine tracts flanked by C on the 5′ side and T on the 3′ side, (positions − 1866 to − 1850), can lead to a stronger curvature [Crothers et al., 1990]. A less pronounced bent region (*Bh*C4B − 5) reveals the presence of two stretches with A_4T_4, positions − 2173 to − 2157, where the A-tract is found before the T-tract, a motif previously reported as relevant for anomalously electrophoretic migration and considered as bent sequence [Hagerman, 1990]. Two bent regions, the *Bh*C4B − 6 and *Bh*C4B − 4, showed two consensus sequences, with 60 bp each, cons 1 and cons 2, which display 88% identity between them. A region between positions − 1833 to − 1812, part of the *Bh*C4B − 4 bent region, presents 85% of identity with an eukaryote Ori motif and is characterized by an imperfect palindrome sequence AAATTA(N8)T-TAATTT [Dobbs et al., 1994].

Detailed Computer Analysis of the Representative Proximal and Distal Bent DNA Regions

Two representative bent DNA regions, one located in the distal *BhC4-1* gene promoter segment and another in the proximal region, were analyzed by computer 2D modeling of the 3D DNA path and by four distinct parameters, ΔG, AT percentage, roll angle and ENDS ratio, which permit detailed theoretical studies of the bent sites in these curved segments. The 1215 bp fragment (− 2780 to − 1565) displays three bent sites, *Bh*C4B − 6 to − 4 (Fig. 6). This fragment presents the strongest *R*-value, 1.82 and ENDS ratio curve, 1.41, observed in this work. There are three pronounced curvatures in this fragment at positions − 2630, − 2130 and − 1850, which characterize the center of each bent site, causing changes in the fragment trajectory and defining their major bending regions. A straight region in 2D fragment

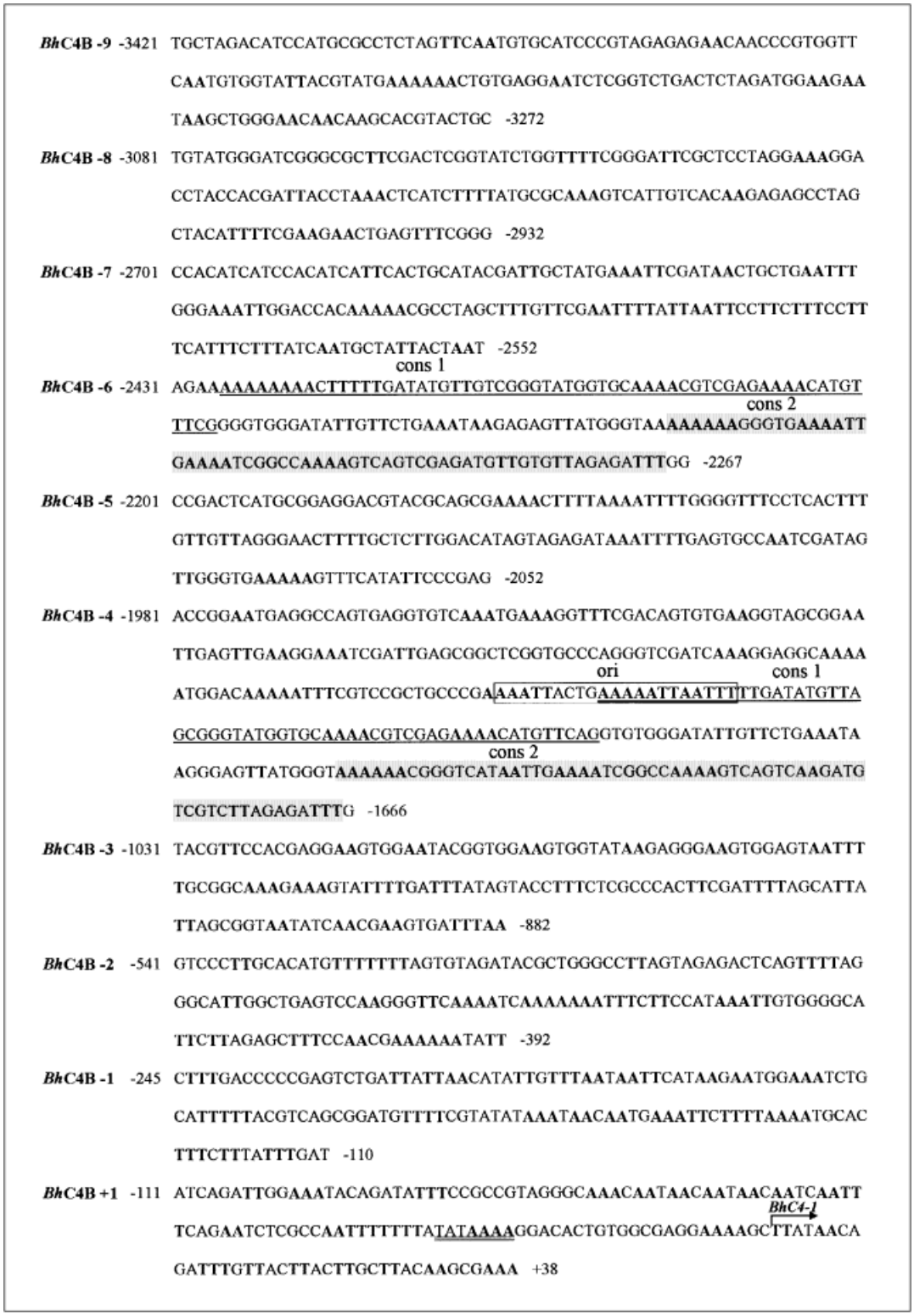

Fig. 5. Bent sites sequence analysis. The sequences of approximately 150 to 300 nucleotides around the center of the ten bent sites, detected in the 3990 bp fragment, named *Bh*C4B − 9 to + 1, are shown and the dA•dT tracts with two or more nucleotides are bolded. Consensus 1 and 2, found both in bent regions *Bh*C4B − 6 and *Bh*C4B − 4 are underlined and shadowed, respectively. The sequence with 85% of identity with a eukaryote Ori motif [Dobbs et al., 1994] found in the *Bh*C4B − 4 bent region is boxed. The TATA box, in the bent DNA *Bh*C4B + 1 site is double underlined, and the arrow indicates the star site (+ 1) and the direction of *BhC4-1* transcription.

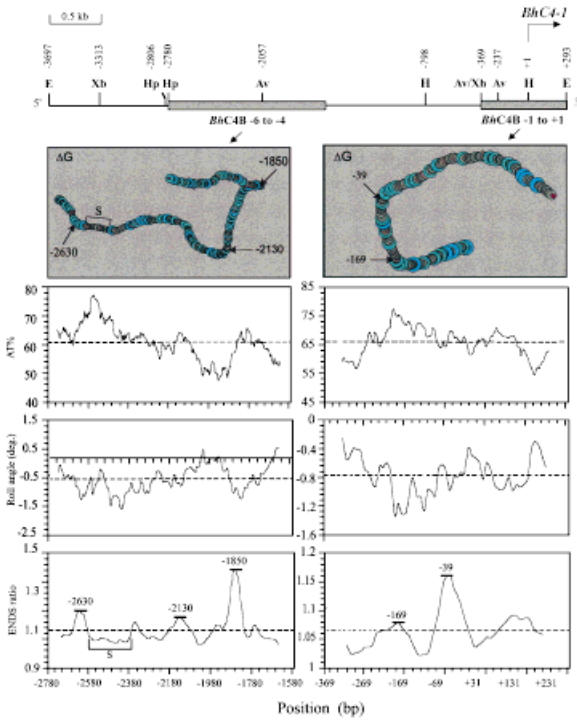

Fig. 6. Theoretical fragments analysis in the 1215 bp fragment (− 2780 to − 1565), and in the 662 bp fragment (− 369 to + 293) of the *BhC4-1* promoter region. The restriction map of the *BhC4-1* promoter region (− 3697 to + 293) is shown on top of the figure. The arrow on top of the diagram indicates the *BhC4-1* start site and the direction of transcription. The 1215 bp, which contains the *Bh*C4B − 6 to − 4 and 662 bp, *Bh*C4B − 1 to + 1 bent sites, are boxed in the diagram. In the projected figures, circles of different sizes and colors graphically represent the ΔG, which predicts the stability of the DNA duplex. The blue larger circles represent high ΔG and gray circles represent low ΔG values. The ΔG, AT percentage, roll angle and ENDS ratio were calculated using the parameters described in Material and Methods. The strait region (S) identified in the 1215 bp fragment is indicated in the projected figure of this fragment. Bent DNA sites identified at positions − 2630, − 2130, − 1850, − 169 and − 39 in the analyzed fragments are indicated. E, *Eco*RI; Xb, *Xba*I; Hp, *Hpa*I; Av, *Ava*I; H, *Hin*dIII; Xh, *Xho*I. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

trajectory (S) (positions − 2580 to − 2380) displays an interesting feature with higher AT percentage, low ΔG, and less pronounced ENDS ratio and roll angle. In this region the dA•dT tracts are not phasing at 10 bp intervals or multiples of them, indicating that the sequence distribution along the analyzed fragments plays an important role in generating a curved segment. The 662 bp fragment (− 369 to + 293) contains two bent sites, *Bh*C4B − 1 to + 1 (Fig. 6). A defined "U" shaped figure, formed by two important deflections, at positions − 169 and − 39, was predicted by the computer modeling. The studied parameters showed that these two bent segments don't have similar features, with different ΔG, AT percentage and ENDS ratio values. The roll angle analysis showed that the values are always negative which results in the major structural curving of the fragment.

With the aim of visualizing the distribution of particular sequences in the 1215 bp and 662 bp fragments, the 2D modeling of the 3D DNA path and the anticlockwise rotation of these fragments were calculated (Fig. 7). The distal *BhC4-1* gene promoter region contains two consensus sequences , cons1 and cons2, twice repeated, that are localized between the *Bh*C4B − 6 and *Bh*C4B − 5 bent sites, and after the *Bh*C4B − 4 bent site (Fig. 7a). The modeling of the 662 bp fragment (Fig. 7b) suggests that the bending in the promoter proximal region causes a narrow minor groove that co-localizes with the TATA box, and which has been reported to promote the association between the TATA box and the

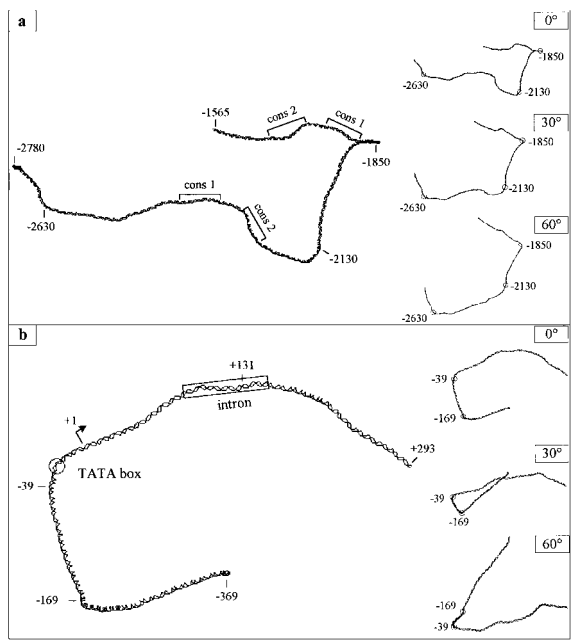

Fig. 7. 2D projection of 3D DNA path of the 1215 bp fragment (− 2780 to − 1565) and the 662 bp fragment (− 369 to + 293) of the *BhC4-1* gene promoter region. The anticlockwise rotation of these fragments in 30 and 60 degrees, shown at the right hand side of the figures, confirms that the bent sites localized at positions − 2630, − 2130 and − 1850 in the 1215 bp fragment (**a**) and at positions − 169 and − 39 in the 662 bp fragment (**b**), are important for the fragment shape maintenance. In a, the squared bracket indicates the localization of the cons1 and cons2 consensus, which are twice repeated. In b, the red circle indicates the TATA box, and the arrow indicates the *BhC4-1* start site and the direction of transcription. The region containing the 62 bp intron is indicated by a rectangle.

TATA box binding protein [Kim et al., 1993]. Analyses in prokaryotic and eukaryotic cells have shown that the promoter strength is dependent, in part, on the intensity of the curvature [Plaskon and Wartell, 1987; Delic et al., 1991]. Moreover, several studies have pointed out a correlation between the rotational orientation of the curved DNA relative to the promoter and its ability to enhance transcription [McAllister and Achberger, 1989]. The rotation of the distal and proximal *BhC4-1* gene promoter region fragments in 30 and 60 degrees (Fig. 7a and b) confirms that the bent sites are important for the maintenance of the shape of these fragments and should be essential for the maintenance of the function of these regions.

DISCUSSION

In this work we have identified both experimentally and through modeling predictions bent DNA sites localized in the *BhC4-1* gene promoter region. The results of the electrophoretic behavior, supported by 2D modeling, suggest the existence of at least two strong intrinsic curved DNA regions, one located in a more distal 1847 bp fragment (− 3697 to − 1850) and another 1250 bp fragment in the promoter proximal region (− 957 to + 293). The 1847 bp fragment presents six bent DNA sites, which promote a cumulative effect and result in a major curvature of this fragment. Two bent DNA sites in this 1847 bp fragment, *Bh*C4B − 6 and *Bh*C4B − 4, contain two DNA motifs, the consensus segments, cons1 and cons2. The functional relevance of these consensus sequences remains an open question. An eucaryotic consensus sequence, characterized by an imperfect palindrome sequence AAATTA(N8)TTAATTT overlaps with the 5′ end of cons 1, identified in the *Bh*C4B − 4 bent site, which could be indicative of a functional role for this bent site. In fact, the analysis of replication intermediates using 2D neutral/alkaline electrophoresis in the 3990 bp fragment has shown that this fragment is replicated by unidirectional forks entering the fragment and has suggested that an origin of replication exists in the vicinity or is part of this fragment [Paçó-Larson, unpublished data]. The existence of bending sites in a DNA origin segment was reported for the amplified chromosomal DHFR replicon in Chinese hamster cell strain CHOC 400, where tracts with three or four dA were found [Caddle et al., 1990]. For the yeast *Saccharomyces cerevisae*, a bent segment with 40–55 bp in the autonomously replicating sequence ARS1, was reported [Snyder et al., 1986]. However, an exception was recently described for the *Yarrowia lipolytica* replication origin [Vernis et al., 1999], where no bent DNA is associated with the characterized origin.

The characterization of *Drosophila* transformed with a 3606 bp *BhC4-1* fragment (− 3313 to + 293) has demonstrated that the *BhC4-1* expression occurs specifically in the salivary gland of early prepupae [Monesi et al., 1998]. A series of progressive deletions of the 3606 fragment have been fused to the reporter gene *lac Z* and employed to transform *Drosophila*. The analysis of these transgenic lines has revealed that the *BhC4-1* gene is also expressed in the ring gland. A 2020 bp fragment (− 3313 to − 1293) has been identified as participating in the repression of *BhC4-1* in the ring gland in developmental stages prior to 0h prepupae. Since this 2020 bp fragment overlaps with one of the strong intrinsic curved DNA regions located in the 1847 bp fragment (− 3697 to − 1850), it could be suggested that this 1847 bp fragment participates in the temporal regulation of *BhC4-1* in the ring gland. The 1250 bp fragment (− 957 to + 293) in the proximal bent DNA region presents four bent sites, *Bh*C4B − 3, *Bh*C4B − 2, *Bh*C4B − 1 and *Bh*C4B + 1. Transgenic *Drosophila* transformed with a 410 bp fragment (− 370 to + 40), which does not include the *Bh*C4B − 2 bent site, display ectopic reporter activity in late embryos, which suggests that the *Bh*C4B − 2 bent site might be involved in repressing *BhC4-1* expression in late embryos. The bent DNA *Bh*C4B − 1 site partially overlaps with a 66 bp fragment (− 253 to − 187), which seems to be involved in promoting *BhC4-1* expression in the ring gland. Moreover, bent DNA *Bh*C4B − 1 and *Bh*C4B + 1 sites contain a 226 bp fragment (− 186 to + 40), which has been characterized as sufficient to regulate *BhC4-1* expression in the prepupal salivary gland and a 97 bp (− 57 to + 40) fragment which is the putative *BhC4-1* minimal promoter.

The 2D modeling of the 662 bp fragment displays a canonical figure described for eukaryotic promoters and establishes the structure for the proximal *BhC4-1* gene promoter in this region (Figs. 6 and 7b). This structure is similar to the one described for scorpion α-Toxin gene

promoter [Delabre et al., 1995] which codifies a venom gland neurotoxin. The production of the secretory proteins is also a common feature between these genes. The *BhC4-1* protein is a putative cocoon protein, which is present both in the salivary gland and in the gland secretion. [Monesi et al., 1995; Paçó-Larson, unpublished results]. In conclusion, the distal and proximal bent regions described here might play an important role in regulating *BhC4-1* transcription. The transformation of *Drosophila* with constructs containing mutated versions of the bent sites identified in this work might contribute to the understanding of the participation of this structural feature in transcription regulation.

ACKNOWLEDGMENTS

We thank Valmir Peron and Jimes R. da S. Santos for their dedicated technical assistance. The authors are indebted to Dr. Philippe Pasero (Institut de Génétique Moléculaire, CNRS, Montpellier, France) for the help in the theoretical bent DNA analysis and Dr. Nadia Monesi for critical review of the manuscript, the encouragement and for sharing unpublished results. We also thank Vanessa Pinatto Gaspar for revising the manuscript. Fiorini, A. received a master fellowship from Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq and Basso Jr., L.R. from Conselho de Aperfeiçoamento do Pessoal do Ensino Superior, CAPES.

REFERENCES

Anderson JN. 1986. Detection, sequence patterns, and function of unusual DNA structures. Nucleic Acids Res 14:8513–8533.

Bash RC, Vargason JM, Cornejo S, Ho S, Lohr D. 2001. Intrinsically bent DNA in the promoter of the Yeast *GAL1-10* and *GAL80* genes. J Biol Chem 276:861–866.

Bolshoy A, McNamara P, Harrington RE, Trifonov EN. 1991. Curved DNA without A-A: experimental estimation of all 16 DNA wedge angles. Proc Natl Acad Sci USA 88:2312–2316.

Bouliskas T. 1993. Nature of DNA sequences at the attachment regions of genes to the nuclear matrix. J Cell Biochem 52:14–22.

Bracco L, Kotlarz D, Kolb A, Diekmann S, Buc H. 1989. Synthetic curved DNA sequences can act as transcriptional activators in *Escherichia coli*. EMBO J 8:4289–4296.

Caddle MS, Dailey L, Heintz NH. 1990. RIP60, a mammalian origin-binding protein, enhances DNA bending near the dihydrofolate reductase origin of replication. Mol Cell Biol 10:6236–6243.

Calladine CR, Drew HR. 1997. Understanding DNA: the molecule and how it works. London: Academic Press.

Crothers DM, Haran TE, Nadeau JG. 1990. Intrinsically bent DNA. J Biol Chem 265:7093–7096.

Del Sal G, Manfioletti G, Schneider C. 1989. The CTAB-DNA precipitation method: a common mini-scale preparation of template DNA from phagemids, phages or plasmids suitable for sequencing. Biotechniques 7:514–520.

Delabre ML, Pasero P, Marilley M, Bougis PE. 1995. Promoter structure and intron-exon organization of a scorpion α-Toxin gene. Biochemistry 34:6729–6739.

Delic J, Onclercq R, Moisan-Coppey M. 1991. Inhibition and enhancement of eukaryotic gene expression by potential non-B DNA sequences. Biochem Biophys Res Commun 180:1273–1283.

de Souza ON, Ornstein RL. 1998. Inherent DNA curvature and flexibility correlate with TATA box functionality. Biopolymers 46:403–415.

Diekmann S, Wang JC. 1985. On the sequence determinants and flexibility of the kinetoplast DNA fragment with abnormal gel electrophoretic mobilities. J Mol Biol 186:1–11.

Dlakic M, Harrington RE. 1998. Unconventional helical phasing of repetitive DNA motifs reveals their relative bending contributions. Nucleic Acids Res 26:4274–4279.

Dobbs DL, Shaiu WL, Benbow RM. 1994. Modular sequence elements associated with origin regions in eukaryotic chromosomal DNA. Nucleic Acids Res 22:2479–2489.

Eckdahl TT, Anderson JN. 1987. Computer modelling of DNA structures involved in chromosome maintenance. Nucleic Acids Res 15:8531–8545.

Fontes AM, Conacci ME, Monesi N, de Almeida JC, Paçó-Larson ML. 1999. The DNA puff *BhB10-1* gene encodes a glycine-rich protein secreted by the late stage larval salivary glands of *Bradysia hygida*. Gene 29:67–75.

Goodsell DS, Dickerson RE. 1994. Bending and curvature calculations in B-DNA. Nucleic Acids Res 22:5497–5503.

Hagerman PJ. 1990. Sequence-directed curvature of DNA. Ann Rev Biochem 59:755–781.

Hibino Y, Tsukada S, Sugano N. 1993. Properties of a DNA-binding from rat nuclear scaffold fraction. Biochem Biophis Res Commun 197:336–342.

Homberger HP. 1989. Bent DNA is a structural feature of scaffold-attachment regions in *Drosophila melanogaster*-interphase nuclei. Chromosoma (Berl) 98:99–104.

Kim Y, Geiger JH, Hahn S, Sigler PB. 1993. Crystal structure of a yeast TBP/TATA-box complex. Nature 365:512–520.

Krajewski WA, Razin SV. 1992. Organization of specific DNA sequence elements in the region of the replication origin and matrix attachment site in the chicken alpha-globin gene domain. Mol Gen Genet 235:381–388.

Liang C, Spitzer JD, Smith HS, Gerbi SA. 1993. Replication initiates at a confined region during DNA amplification in *Siara* DNA puff II/9A. Genes Dev 7:1072–1084.

Linial M, Shlomai J. 1988. Bent DNA structures associated with several origins of replication are recognized by a unique enzyme from trypanosomatids. Nucleic Acids Res 16:6477–6492.

Marilley M, Pasero P. 1996. Common DNA structural features exhibited by eukaryotic ribosomal gene promoters. Nucleic Acids Res 24:2204–2211.

Marini JC, Levene SD, Crothers DM, Englund PT. 1982. Bent helical structure in knetoplast DNA. Proc Natl Acad Sci USA 79:7664–7668.

McAllister CF, Achberger EC. 1989. Rotational orientation of upstream curved DNA affects promoter function in *Bacillus subtilis*. J Biol Chem 264:10451–10456.

Milot E, Belmaaza A, Wallenburg JC, Gusew N, Bradley WE, Chartrand P. 1992. Chromosomal illegitimate recombination in mammalian cells is associated with intrinsically bent DNA elements. EMBO J 11:5063–5070.

Monesi N, Fernandez MA, Fontes AM, Basso LR Jr, Nakanishi Y, Baron B, Buttin G, Paçó-Larson ML. 1995. Molecular characterization of an 18kb segment of DNA puff C4 of *Bradysia hygida* (Diptera, Sciaridae). Cromosoma 103:715–724.

Monesi N, Jacobs-Lorena M, Paçó-Larson ML. 1998. The DNA puff gene *BhC4-1* of *Bradysia hygida* is specifically transcribed in early prepupal salivary glands of *Drosophila melanogaster*. Chromosoma 107:59–569.

Nair TM. 1998. Evidence for intrinsic DNA bends within the human cdc2 promoter. FEBS Lett 422:94–98.

Paçó-Larson ML, de Almeida JC, Edström JE, Sauaia H. 1992. Cloning of a developmentally amplified gene sequence in the DNA puff C4 of *Bradysia hygida* salivary gland. Insect Biochem Mol Biol 22:439–446.

Palin AH, Critcher R, Fitzgerald DJ, Anderson JN, Farr CJ. 1998. Direct cloning and analysis of DNA sequences from a region of the Chinese hamster genome associated with aphidicolin-sensitive fragility. J Cell Sci 11:1623–1624.

Pasero P, Sjakste N, Blettry C, Got C, Marilley M. 1993. Long-range organization and sequence-directed curvature of *Xenopus laevis* satellite 1 DNA. Nucleic Acids Res 21:4703–4710.

Perez-Martin J, de Lorenzo V. 1997. Clues and consequences of DNA bending in transcription. Ann Rev Microbiol 51:593–628.

Perez-Martin J, Rojo F, de Lorenzo V. 1994. Promoters responsive to DNA bending: a common theme in prokaryotic gene expression. Microbiol Rev 58:268–290.

Plaskon RR, Wartell RM. 1987. Sequence distributions associated with DNA curvature are found upstream of strong *E. coli* promoters. Nucleic Acids Res 15:785–796.

Santelli RV, Machado-Santelli GM, Pueyo MT, Navarro-Cattapan LD, Lara FJ. 1991. Replication and transcription in the course of DNA amplification of the C3 and C8 DNA puffs of*Rhynchosciara americana*. Mech Dev 36:59–66.

Snyder M, Buchman AR, Davis RW. 1986. Bent DNA at a yeast autonomously replicating sequence. Nature 324: 87–89.

Vernis L, Chasles M, Pasero P, Lepingle A, Gaillardin C, Fournier P. 1999. Short DNA fragments without sequence similarity are initiation sites for replication in the chromosome of the yeast *Yarrowia lipolytica*. Mol Biol Cell 10:757–769.

von Kries JP, Phi-Van L, Diekmann S, Stratling WH. 1990. A non-curved chicken lysozyme 5′ matrix attachment site is 3′ followed by a strongly curved DNA sequence. Nucleic Acids Res 18:3881–3885.

Wada-Kiyama Y, Kiyama R. 1994. Periodicity of DNA bend sites in human epsilon-globin gene region. Possibility of sequence-directed nucleosome phasing. J Biol Chem 269:22238–22244.

Williams JS, Eckdahl TT, Anderson JN. 1988. Bent DNA functions as a replication enhancer in *Saccharomyces cerevisiae*. Mol Cell Biol 8:2763–2769.

Wu HM, Crothers DM. 1984. The locus of sequence-directed and protein-induced DNA bending. Nature 308:509–513.

Wu N, Liang C, DiBartolomeis SM, Smith HS, Gerbi SA. 1993. Developmental progression of DNA puffs in *Sciara coprophila*amplification and transcription. Dev Biol 160:73–84.

Yokosawa J, Soares MAM, Dijkwel PA, Stocker AJ, Hamlin FJSL. 1999. DNA replication during amplification of the C3 puff of *Rhynchosciara americana* initiates at multiple sites in a 6kb region. Chromosoma 108:291–301.