The EMBO Journal vol.13 no.24 pp.6162-6171, 1994

The λ terminase enzyme measures the point of its endonucleolytic attack 47 ± 2 bp away from its site of specific DNA binding, the *R* site

R.Rachel Higgins and Andrew Becker

Department of Molecular and Medical Genetics, University of Toronto, Toronto, Ontario, Canada M5S 1A8

Communicated by H.Eisen

λ terminase is an ATP-interactive, site-specific endonuclease comprising the products of λ genes *Nu1* and *A*. Terminase binds to *cos*, at the junction of two chromosomes in a concatemer, catalyzes *cos* cleavage and initiates the packaging of λ DNA into proheads. *cos* consists of a nicking domain, *cosN*, where terminase cleaves to regenerate the 12 nucleotide cohesive ends of mature λ chromosomes and a binding domain, *cosB*, where terminase binds to 16 bp repeat sequences called *R3*, *R2* and *R1*. Evidence is presented that terminase is a single-strand endonuclease that can nick DNA by one of two mechanisms, both of which require ATP. (i) When bound to any *R* site, terminase nicks the strand which, within that *R* site, is purine-rich; the position of this nick is 47 ± 2 nucleotides away from the mid-point of that *R* site, measured in the 3′ direction; (ii) enzymes that are not bound to *R* sites nick DNA within certain specific sequences that resemble *cosN* half sites. These two modes of action are nicely combined for the *R3*-bound protomer that nicks the bottom strand at position N1 in *cosN* since the interval between N1 and the *R3* midpoint is 47 nucleotides. Within *cosN*, the bottom and top strand nicks are generated by a rigid protein couple with a 2-fold rotational symmetry. The location of both of these nicks, however, is gauged asymmetrically from *R3*, 47 nucleotides away. Again, *R1* and *R2* are separated by 47 bp and orient bound protomers towards each other but, unless the DNA between these *R* sites is lengthened, the enzymes do not nick, indicating an inhibitory gpA–gpNu1 apposition.

Key words: *cos*/lambda/terminase

Introduction

The *Nu1* and *A* gene products of bacteriophage λ form the hetero-oligomeric enzyme known as terminase which plays a central role in the assembly of the mature virion (for reviews see Feiss and Becker, 1983; Feiss, 1986; Becker and Murialdo, 1990). During late stages of the λ lytic life-cycle, the phage DNA is replicated into tandem concatemers consisting of many genomes joined head-to-tail in a linear array. At the same time empty phage capsids, or proheads, are formed in a separate pathway. Terminase binds to specific sequences in λ DNA, called *cos*, near a potential chromosomal end, and to a prohead. It then fashions the end by nicking and melting the DNA strands (Higgins *et al.*, 1988) and proceeds to either package (pump) the DNA into the prohead or scan the incoming DNA for the next chromosomal end which it will cut (Feiss and Widner, 1982).

cos is the specific DNA terminase site of action where DNA packaging is initiated and where the endonucleolytic cleavage of the DNA takes place, separating the concatemers into single genome lengths. *cos* comprises the *cosB* region, where terminase binds through its gpNu1 subunits, and *cosN* where the DNA is nicked specifically, 12 nucleotides apart, on opposite strands of the double helix to generate the sticky ends of mature λ DNA.

When viewed as a one-dimensional array of base pairs, *cosN* has an imperfect dyadic symmetry comprising the two half elements, *cosNL* on the left of the symmetry axis and *cosNR* on the right [see the accompanying paper by Higgins and Becker (1994) for an illustration of *cos*]. The two nicks, N2 in the top strand within *cosNL* and N1 in the bottom strand in *cosNR*, are placed symmetrically within this dyad by terminase. *cosB* is a region of some 160 bp between *cosN* and the *Nu1* gene of terminase. Within *cosB* there are three 16 bp segments called *R3*, *R2* and *R1* that are known to bind gpNu1 (Feiss *et al.*, 1983; Miwa and Matsubara, 1983; Bear *et al.*, 1984; Shinder and Gold, 1988). Another *R*-like element, *R4*, situated to the left of *cosN*, fails to bind gpNu1 *in vitro* (Shinder and Gold, 1988) but it does bind terminase holoenzyme (Higgins and Becker, 1994).

In the nicking activities at *cosN*, the element *R3* in *cosB* has been shown to play a special *cis* regulatory role (Higgins and Becker, 1994), but the functional relationship between these two sites was not well understood. In the present study we show that terminase, when bound to any *R* site, is positioned to nick the strand which is purine-rich within that *R* site. The nick introduced by this mechanism is placed 47 ± 2 nucleotides away from the mid-point of that *R* site, measured in the 3′ direction. If the point of nicking does not resemble *cosN* by sequence, the reaction is dependent upon ATP addition. Terminase also nicks DNA at sites which, by sequence, resemble one or the other of the two *cosN* half sites; this reaction does not require binding to any of the *R* sites, but is ATP-dependent. These two modes of operation are accurately brought together in the case of the *R3*-bound terminase protomer which is poised to nick the bottom strand at *cosNR*, since the N1 nick site and the reference binding point within *R3* are separated by exactly 47 nucleotides.

Results

The effect of DNA deletions in the cosNR–R3 interval on nicking at cosN

Recent studies have shown that the *R3* site of *cosB* is a dominant regulatory element in the nicking reactions at

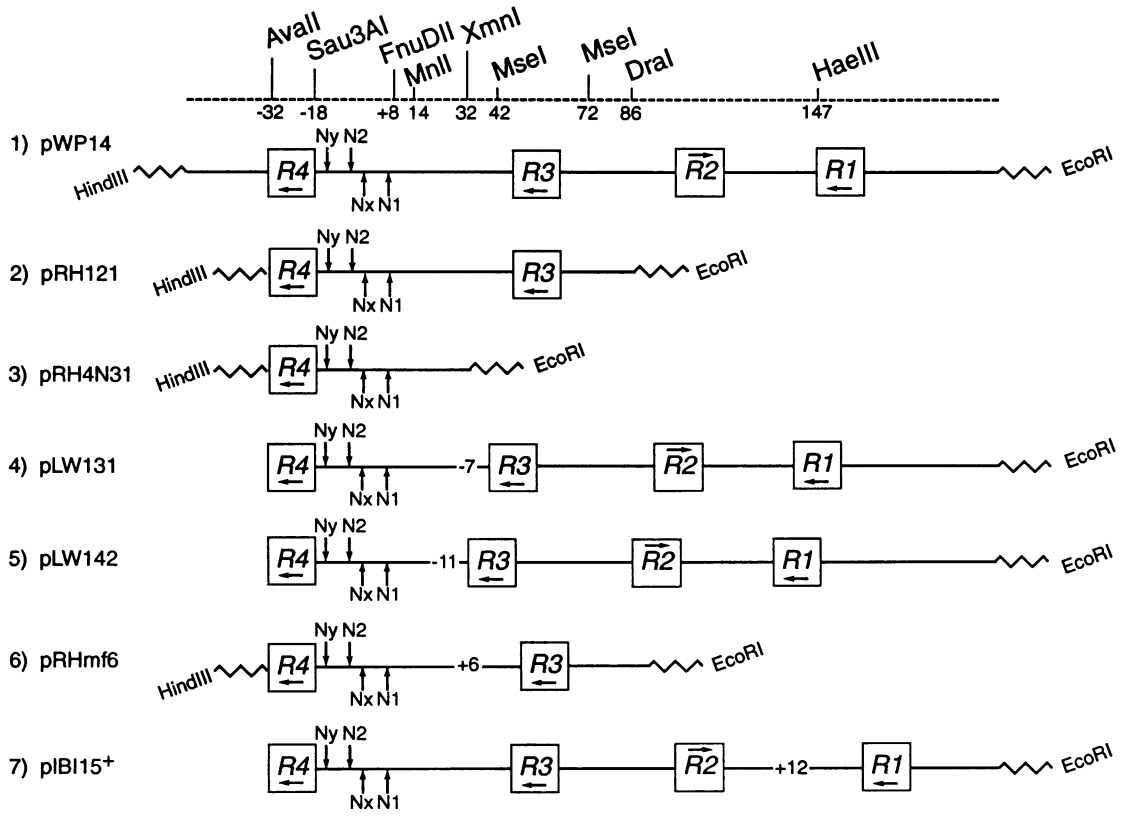

Fig. 1. Schematic diagram of *cos*-containing fragments used as substrates in *cosN* cleavage studies. The plasmid source of each and the corresponding *cos*-containing fragments are shown. The restriction fragments are drawn to approximate scale; the locations of *R* sites are indicated, as are the nicking sites N1, N2, Nx and Ny in *cosN*. λ-derived DNA is drawn as solid lines; polylinker DNA is zigzagged. The fragment in (1) from pWP14, contains the complete λ *cos* site, *cosN* in continuity with *cosB*. The fragment in (2) from pRH121, lacks *R2* and *R1*. The fragment in (3) lacks the *R3*, *R2* and *R1* sites of *cosB*. The fragments in (4) and (5) from pLW131 and pLW142, carry deletions of 7 and 11 bp, respectively, in the *R3*−*cosN* interval (Miller and Feiss, 1988). The fragment in (6) from pRHmf6, carries a 6 bp insertion in the *R3*−*cosN* interval, a duplication of base pairs +9 to +14. The fragment in (7) from pIBI15, carries a 12 bp insertion in the *R2* to *R1* interval (Cue and Feiss, 1992a). Plasmid pRHxm17 and pRHxf23 (not shown) have the same structure as pRHmf6 (6) but carry insertions of 17 and 23 bp, respectively, in the *R3*−*cosN* interval.

cosN catalyzed by the phage λ terminase enzyme, and that terminase, when bound specifically to *R3* through its gpNu1 subunit, has characteristic *cosN*-nicking properties that are different from those of an enzyme which is not bound to *R3* (Higgins and Becker, 1994).

How might *R3* exert its specific controls over the nicking reactions at *cosN*? Two possibilities are that (i) the terminase holoenzyme spans the *R3*−*cosN* interval, with gpNu1 binding to *R3* and gpA to *cosN* or (ii) the terminase holoenzyme first binds to *R3*, is somehow altered to acquire different nicking properties, then diffuses on DNA to *cosN* where it nicks. One guesses that the first mechanism would be sensitive to changes in the distance between *cosN* and *R3*, but that the second might not. Feiss *et al.* (1983) and Miller and Feiss (1988) first demonstrated the importance of correct spacing between *cosN* and *cosB*. Very short deletions in the *cosN*−*R3* interval were viable in IHF$^+$ hosts, but longer deletions of 7 or 11 bp were not.

We tested these Δ7 and Δ11 substrates for *cosN* nicking and obtained the result given in Figure 2. Whereas DNAs bearing short additions or deletions (1 or 2 bp) in the *cosN*−*R3* interval (data not shown) or the normal substrate responded typically with the ATP-induced Nx−N1 adjustment (Figure 2A, lanes 2 and 3) (Higgins and Becker, 1994), the shortening of the *cosNR*−*R3* interval by either the Δ7 or the Δ11 deletion repositioned the points of endonucleolytic attack from the normal position of N1 to that of Nx on the bottom strand (Figure 2A, Δ7 lanes 5 and 6; Δ11 lanes 8 and 9), and from N2 on the top strand in the normal substrate (Figure 2B, lanes 2 and 3), to the site Ny, 12 bp to the left of Nx (Figure 2B, Δ7 lanes 5 and 6; Δ11 lanes 8 and 9).

The results indicate first, that the nicking operations at *cosN* are catalyzed by an enzyme that is bound to *R3*, since terminase can only nick at *cosN* in the absence of ATP when bound to *R3* (Higgins and Becker, 1994). Secondly, that the top and bottom strand nicking centers of terminase have been displaced to the left of the *cosN* symmetry axis due to shortening of the *R3*−*cosN* DNA interval in the face of the dominant binding interaction at *R3*. Note how the addition of ATP fails to elicit the Nx−N1 adjustment in the substrates that bear the deletions.

The above data suggest that the binding of terminase to *R3* controls the location of both the top and bottom strand nicks at *cosN* by a mechanism that gauges the nick sites from *R3*. The binding of terminase to *R3* through its gpNu1 subunit appears to be the dominant site-specific protein−DNA interaction; the binding of the gpA subunits to *cosN* and the resultant nicking activities take place secondarily to *R3* binding, *R3* being the point from which the location of the nicks is measured. The measured distance may not always be exact, apparently being

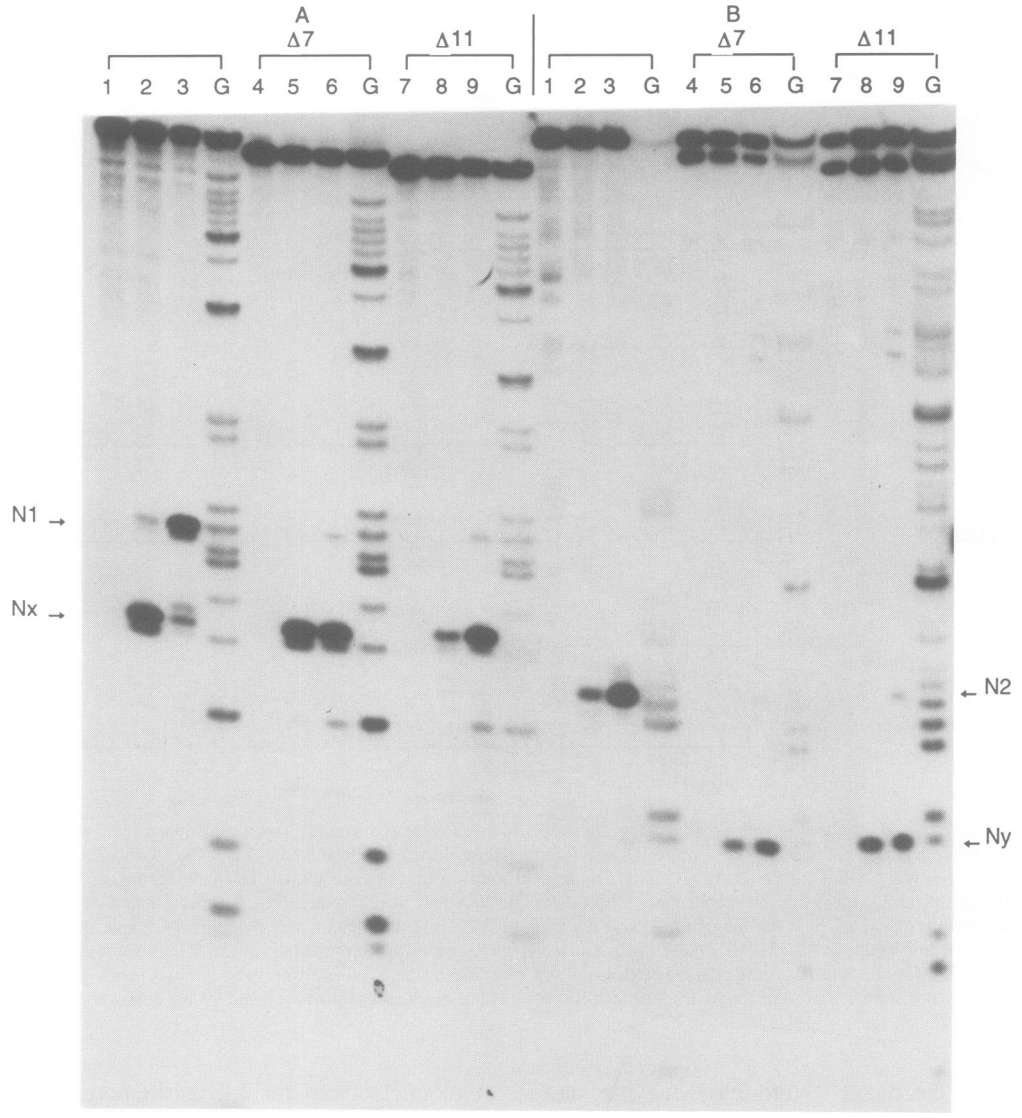

Fig. 2. Shortening of the *R3–cosN* interval displaces the nicking complex at *cosN* to the left, to nick at sites Nx and Ny. The following *cos*-DNA substrates were used: (**A**) the *Ava*II–*Dra*I piece from plasmids pWP14 (normal *cos*, lanes 1–3), the *Ava*II–*Dra*I piece from pLW131 (Δ7, lanes 4–6) and the *Ava*II–*Dra*I piece from pLW142 (Δ11, lanes 7–9), all uniquely labeled at the 3′ end of the bottom strand, at the *Ava*II site; (**B**) the same *Ava*II–*Dra*I DNA pieces as in (A) were used and were labeled on the other strand, at the 5′ end of the *Ava*II site. For lanes 1, 4 and 7, terminase was omitted during incubation; for lanes 2, 5 and 8, incubation was in the presence of terminase, but in the absence of ATP; for lanes 3, 6 and 9, incubation was in the presence of terminase and ATP; lanes marked G are the Maxam and Gilbert (1980) G-specific sequencing ladders. The positions of the terminase nicks, N1, Nx, N2 and Ny are indicated. Analysis was on an 8% denaturing gel as described in Materials and methods.

subject to certain sequence preferences on the part of the nucleolytic recognition domains; Nx and Ny are, for example, preferred nicking sites, even in the absence of *R3* (Higgins and Becker, 1994). The data also indicate that terminase's top and bottom strand nicking centers respond as a rigid couple when displaced from *cosN* by the deletion, suggesting that the dual nicking operation is performed by an enzyme with a constrained geometry between the two nicking active centers as would be the case for a symmetric terminase dimer.

The effect of DNA insertions in the cosNR–R3 interval on nicking at cosN

If the above model of an *R3*-anchored *cosN*-nicking system is correct, then the insertion of extra DNA into the *cosN–R3* interval should have the effect of displacing the nicking centers rightward with respect to *cosN* and, if a suitable substrate is placed in the vicinity of the endonucleolytic center(s), nicks will be catalyzed at new positions. Figure 3 shows the results of nicking experiments with two substrates, one with a normal *cos* configuration (Figure 3A) and the other, called the ins-6 substrate (Figure 3B), comprising a 6 bp insertion lengthening the space between *cosN* and *R3*. Examining the bottom strand with the normal substrate, the results are typical of an $R3^+$ system. In the absence of ATP the nick is placed at Nx, almost exclusively, in the presence of ATP the point of nuclease attack is largely shifted to N1, although in this experiment considerable nicking at Nx still takes place.

With the ins-6 substrate, a different pair of nicks, N1 (+6/+7) and, to a lesser extent, a nick in the interval +4/+5, respond to ATP addition as would the nick at Nx in a normal substrate. Thus, there is considerable nicking at

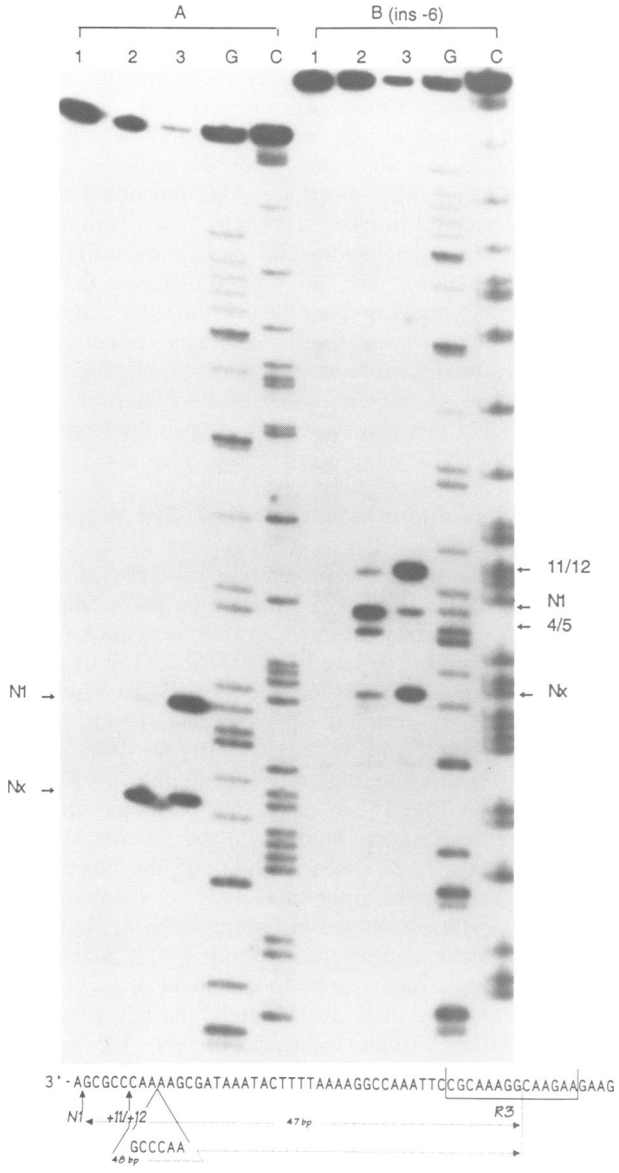

Fig. 3. Lengthening the *R3*−*cosN* interval displaces the nicking complex at *cosN* to the right, to nick at site N1 in the absence of ATP, or +11/+12 in the presence of ATP. The following *cos*−DNA substrates were used: (**A**) the *Hin*dIII−*Eco*RI piece from plasmids pRH121 (normal *cos*), labeled at the 3′ end of the bottom strand at the *Hin*dIII site; (**B**) the *Hin*dIII−*Eco*RI piece from plasmid pRHmf6 (ins-6) also labeled at the 3′ end of the bottom strand at the *Hin*dIII site. Normal *cos* fragments in (A) contain 47 bp in the *R3*−N1 interval, and ins-6 fragments in (B) bear an insertion of 6 bp in this interval. For lanes 1, terminase was omitted during incubation; for lanes 2, incubation was in the presence of terminase, but in the absence of ATP; for lanes 3, incubation was in the presence of terminase and ATP; lanes G and C are the Maxam−Gilbert, G- and C-specific sequencing ladders. The positions of the terminase nicks Nx, N1, 4/5 and 11/12 are indicated. Part of the sequence present in these fragments is given below the gel; the positions of the 6 bp insertion and the *R3*-dependent nicks N1 and 11/12 are shown.

N1 in the absence of ATP which is markedly attenuated upon ATP inclusion; further, the nick at +4/+5 is eliminated upon ATP addition. On the other hand, a novel nick is seen in the +11/+12 interval and this nick is markedly stimulated by ATP addition. Thus, the ATP-dependent system for adjusting bottom strand nick-site specificity is behaving as if displaced rightward by the DNA insertion, away from the *cosN* symmetry axis, the enzyme holding fast at *R3*. The binding to *R3* taking place in this reaction is based on the observation that nicking at *cosN* takes place in the absence of ATP (Higgins and Becker, 1994). In this ins-6 substrate, the top strand nick is also displaced five nucleotides to the right, to position −1/−2 (data not shown), again implying a coupling between the bottom strand- and top strand-directed activities. Note how the nick at Nx is now stimulated, rather than suppressed, by the presence of ATP, an indication of nicking at Nx by an enzyme that is not bound at *R3* (Higgins and Becker, 1994 and see below).

R1- and R2-bound enzymes can also nick DNA

The model being suggested, based partly on previous results using a *cosNR* half-substrate (Higgins *et al.*, 1988), is that terminase binds to the asymmetric *R3* through its gpNu1 subunit and, in so doing, is oriented in the 3′ direction, positioning the nicking domain of the gpA subunit to interact with, and nick, the bottom strand at *cosNR*, 47 bp away (Figure 3A). This rule of directionality and distance implies that the terminase protomers occupying *R1* and *R2* might be complexed in analogous ways. The enzyme bound at *R1* is directed toward *R2*, since *R1* is in the same orientation as *R3*, and the one bound at *R2* is directed in opposite orientation towards *R1*. Curiously, equivalent positions in *R1* and *R2* are separated from each other by exactly 47 bp implying a clash (or interaction) between the gpA and gpNu1 subunits of the diametric enzymes (Figure 6). The model predicts that, if the length of the DNA interval between *R1* and *R2* were to be increased, uncomplexed DNA might become accessible to the nicking domains of these two terminase molecules and that nicking might take place according to the following rule: on the bottom strand, ~47 nucleotides 3′ of *R1* and, on the top strand, ~47 nucleotides 3′ of *R2*. DNA insertions in the *R1*−*R2* interval are available (Cue and Feiss, 1992) and such a test was performed. The results are in general agreement with this model, as next described.

The experiment of Figure 4 shows the nicking by terminase of a substrate that carries an insertion of 12 nucleotide pairs between *R1* and *R2*. In addition to the usual nicks generated within *cosN* (as well as the N117 nick, see below), there is now further nicking of both the bottom and the top strands in the *R1*−*R2* DNA interval. As predicted by the model, the nick on the bottom strand, at position +118/+119, is situated 46 nucleotides 3′ of the mid-point of *R1* (Figure 4A, lane 3).

The situation on the top strand is somewhat more complex with nicking taking place at three positions. One of these nicks, at position +153/+154, is 46 nucleotides 3′ of *R2* as the model might predict; the others are 35 and 54 bp to the right of *R2*, with the former the most prominent (Figure 4B, lane 3). The results with the top strand can be explained by postulating that the nucleolytic active center of the *R2*-bound enzyme is made to confront DNA which by sequence is not a favorable substrate, and that the enzyme inefficiently catalyzes nicks at any of three adjacent sites that reside on roughly the same face of the double helix. Both top and bottom strand nicking in the *R1*−*R2* interval is dependent upon ATP addition, which we take as an indication of *R* site directed, but otherwise sequence non-specific nucleolytic activity. In

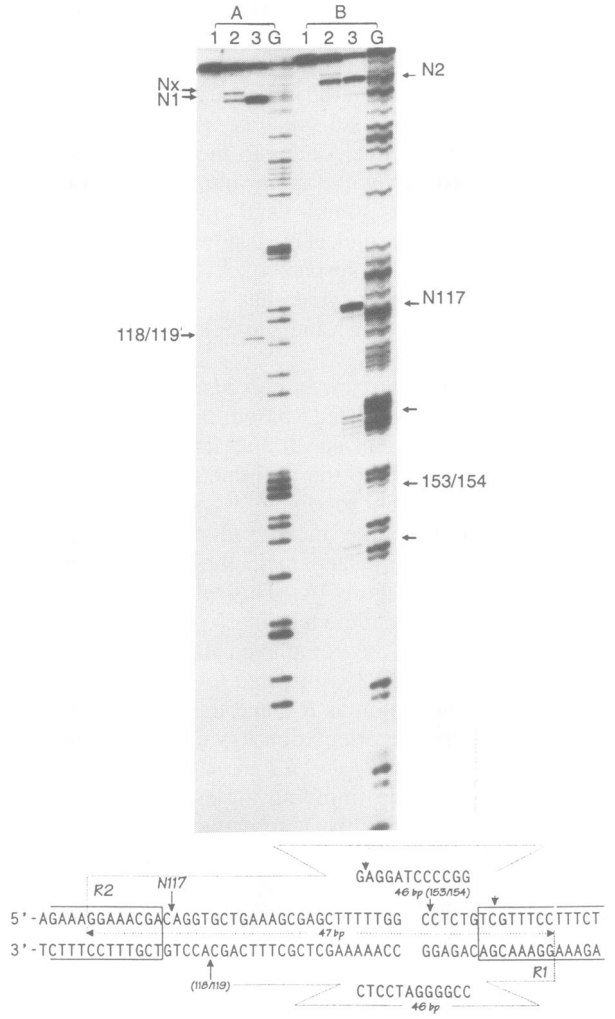

Fig. 4. *R2*- and *R1*-bound protomers nick DNA strands at sites that are 46 bp 3′ of their respective *R* site midpoints. The *cos*−DNA substrates used were: (**A**) the *Ava*II−*Eco*RI piece from plasmid pIBI15$^+$ (+12 insertion in the *R2* to *R1* interval) labeled at the 5′ end of the bottom strand, at the *Ava*II site; in (**B**) the same *Ava*II−*Eco*RI piece was used and was labeled at the 3′ end of the top strand, at the *Eco*RI site. For lanes 1, terminase was omitted during incubation; for lanes 2, incubation was in the presence of terminase, but in the absence of ATP; for lanes 3, incubation was in the presence of terminase and ATP. The positions of the *ter* nicks, Nx, N1, N2 and N117 are shown. The dinucleotide intervals 118/119 and 153/154 show novel nicks by terminase, introduced upon the lengthening of the *R2*−*R1* interval. The sequence of a portion of the *cos* fragment is given below the gel; the positions of the insertion and the *R2*- and *R1*-dependent nicks are indicated. Unmarked arrows indicate putative *ter* nicks that may occur by an *R* site-independent mechanism.

considering these nicking operations in the top strand, one should note the prominent *R* site-independent nick that is always made at N117; this nick would create a swivel in the DNA duplex and the 'measuring' system of an *R2*-bound terminase protomer might have more relaxed properties in this case. Furthermore, if the N117 and +118/+119 nicks were to be present on the same DNA molecule, then the *R2* element would become detached from the potential nick site and the DNA would be lost as a substrate, a possible reason for the apparent inefficiency of nicking the top strand.

Analogous results were obtained using substrates that carried DNA insertions of 17 and 23 bp in the *cosN*−*R3* interval. For a terminase protomer bound at *R3*, its nicking domain was displaced to the right of N1 by 16 and 22 bp for the +17 and +23 substrates, respectively. In each case, bottom strand nicks were observed 48 nucleotides 3′ of the *R3* midpoint (data not shown) and the generation of these nicks was dependent upon the addition of ATP. In the +17 and +23 substrates ATP-dependent nicking was also observed in *cosN*. The latter result is diagnostic of an *R3*-independent nicking taking place at *cosN* (Higgins and Becker, 1994), with *R3* now subtending the novel bottom strand nicking activities within the lengthened *cosN*−*R3* intervals of these substrates. Finally, when specifically searched, one can observe nicking of the bottom strand 46 nucleotides 3′ of *R4* (data not shown) indicating that this less typical *R* site can bind terminase through its gpNu1 subunit *in vitro*.

R site-independent nicking of cosN-like sequences by terminase

Terminase catalyzes the production of nicks at several specific positions in the *cos*−containing DNA fragments that we have studied. In *cosB*, a prominent nick (named N117) is placed in the interval C+116/A+117 of the top strand (Figure 5 *cos*). To the left of *cosN*, two nicks are placed in the bottom strand−in the C−117/A−118 interval (named N−118) and in the G−194/G−195 interval (named N−195) (Figure 5 *cos*). In addition to Nx, N1, N2 and Ny (−14/−15) in *cosN*, two additional nicks are sometimes made nearby. One is in the top strand in the interval G+10/G+11 (named N11) and the other in the bottom strand in the interval G−8/T−9 (named N−9) (Figure 5 *cosN*). As shown in Figure 5, the production of all of these nicks requires the addition of ATP. We consider these nicking activities to be independent of any specific *R* site interactions since, for example, the DNA fragment −199(48310)/−77(48432) bearing the N−195 site, but detached from the *R4*−*cosN*−*cosB* module, is nicked at the N−195 position by terminase in the presence of ATP. Nicking at N117 was similarly examined in truncated *cos*-derived substrates from which the *R4*−*cosN* or the *R4*−*cosN*−*R3* modules had been deleted with identical results, namely, ATP-dependent nicking at N117 (data not shown). There exist other unique positions in the *cos*-containing substrates where terminase can be seen to nick on occasion.

The base sequences surrounding each of these nicks were examined, and most were found to resemble the canonical half sites of *cosN*. For example, the N117 nick is situated within the sequence 5′CGAC↓AGGT that resembles the N1 site, 5′CGCG↓AGGT. The site Ny is identical to the N117 site. The N−118 and N11 nicks also reside within similar sequences, 5′ACGC↓AGGT and 5′CGCG↓GGTT, respectively. The N−195 nick, on the other hand, is placed within a run of eight nucleotides, 5′TACG↓GGGC, which is identical to that of the canonical N2 site. The N−9 segment, 5′CCCG↓TAAC, is somewhat atypical in respect of these base sequence similarities, as is the site Nx.

We do not know if nicking of the *cos*-like half sites takes place *in vivo*. The demonstration of such nicks, however, emphasizes again the intrinsic sequence specificity, however degenerate, of the nicking domains of gpA (Higgins and Becker, 1994).

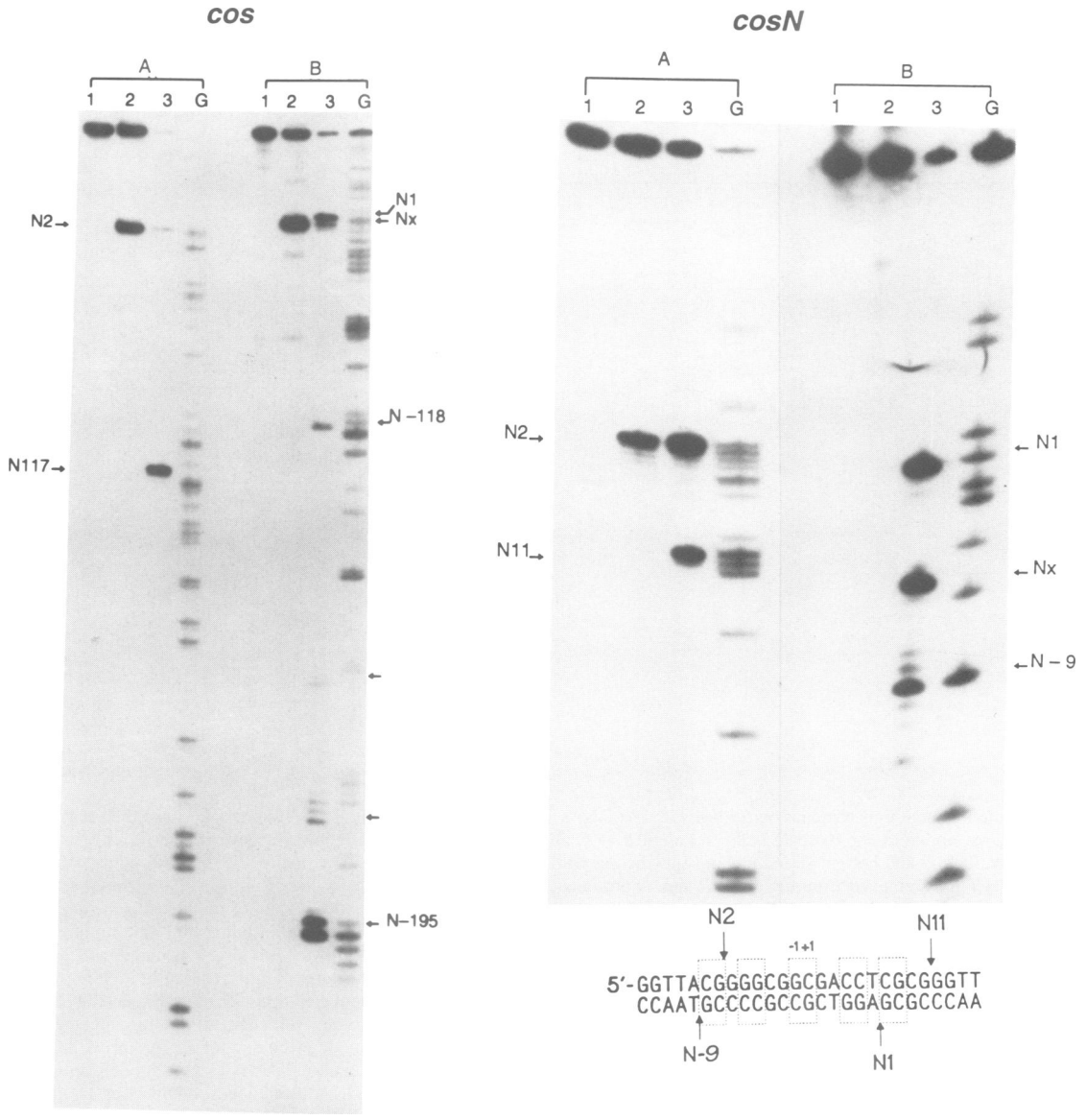

Fig. 5. ATP-dependent nicks in the sequence of *cos* and *cosN*. The *cos*-DNA fragments used as substrates were: in *cos*, (**A**) the *Hin*dIII−*Eco*RI piece from plasmid pWP14 labeled at the 3′ end of the top strand, at the *Eco*RI site; (**B**) the same *Hin*dIII−*Eco*RI piece was used and labeled at the 3′ end of the bottom strand, at the *Hin*dIII site. In *cosN*, (**A**) the *Hin*dIII−*Eco*RI piece from plasmid pRH121 radiolabeled at the 3′ end of the top strand, at the *Eco*RI site; (**B**) the *Hin*dIII−*Eco*RI piece from plasmid pRH4N31 labeled at the 3′ end of the bottom strand, at the *Hin*dIII site. For lanes 1, terminase was omitted during incubation; for lanes 2, incubation was in the presence of terminase but in the absence of ATP; for lanes 3, incubation was in the presence of terminase and ATP; lanes G, are the Maxam and Gilbert (1980) G-specific sequencing ladder. The positions of the terminase nicks at *cosN*, N1, N2 and Nx, are indicated. The ATP- and terminase-dependent nicks, N117, N−118, N−195 (*cos*), N−9 and N11 (*cosN*) nicks occur by an *R* site-independent mechanism. The sequence of *cosN* is given below the gel; depicting the positions of *ter* nicks N1, N2, N−9 and N11.

Discussion

Previously we had shown that *R3* is a dominant *cis* regulatory element in the nicking reactions at *cosN*, catalyzed by terminase (Higgins and Becker, 1994). Terminase holoenzyme, when bound specifically to *R3* through its gpNu1 subunit, has characteristic *cosN*-nicking properties that are different from those of an enzyme which is not bound to *R3*. Based on this notion that terminase has to interact with two different, rather widely separated, sites for proper function, we asked how this nucleolytic system might respond to deletions and/or insertions of DNA placed between these two different sites of protein−DNA interaction.

When Feiss *et al.* (1983) and Miller and Feiss (1988) first demonstrated the importance of correct spacing between *cosN* and *cosB*, they observed that certain longer deletions of 7 or 11 bp, which rendered the DNAs non-packagable, were nevertheless good substrates for cleavage by terminase. The *cos* cleavage seen in these experiments using the Δ11 substrate could be explained by invoking the so-called helical-phasing rule (Hochschild and Ptashne, 1986) in that a *cosB*-bound enzyme might successfully confront *cosN* which, although placed closer to *R3*, resided on approximately the same face of the double helix as in the normal situation. Such had been the case, for example, with another similarly bipartite DNA target, *att*, the site of action of the λ integrase enzyme (Thompson *et al.*,

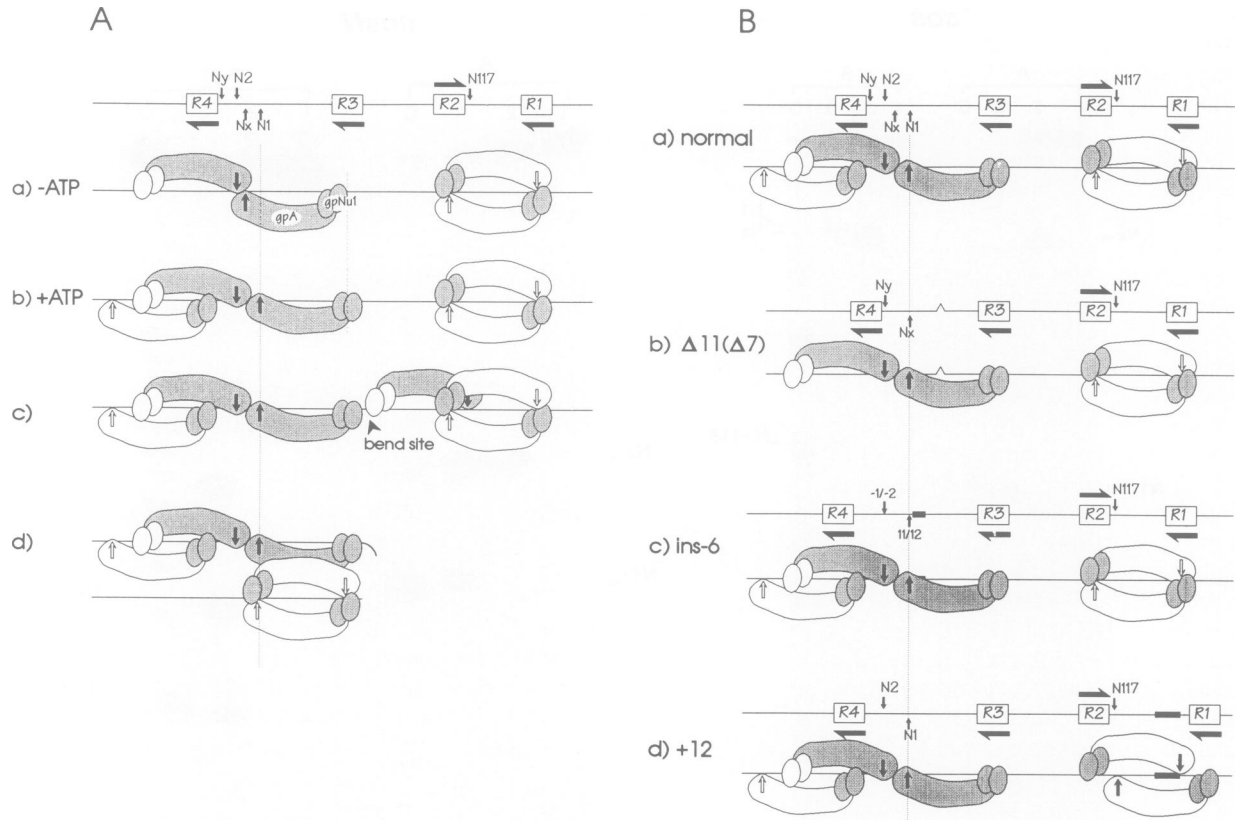

Fig. 6. Models for the interaction of terminase enzymes with *cos* DNAs. In both (A) and (B), the straight horizontal lines represent DNA of the *R4*−*cosN*−*cosB* region, drawn to approximate scale and aligned vertically with respect to *R3*. The horizontal arrows above or below the *R* sites indicate, respectively, the top and bottom strand nicking specificities of the protomer bound; the direction of these arrows point to the site of nicking which is 47 ± 2 nucleotides removed from the given *R* site. A protomer of terminase is portrayed to be comprised of one gpA subunit (the bent cigar-like structure) and two gpNu1 subunits (the smaller ovals), labeled as such in (A). The functional span of the holoenzyme as described in the text is drawn to scale. Shading of a subunit means site-specific DNA binding, while lack of shading implies either no DNA binding or non-specific DNA binding. Vertical arrows indicate potential (unfilled arrows) or actual (filled) sites of nicking of the top (arrow points down) or bottom (arrow point up) strands of DNA. In (**A**) terminase binding and nicking of normal *cos* substrates are illustrated. In the absence of ATP (A, a) nicking at Nx and N2 is catalyzed by a pair of protomers, only one of which is bound to an *R* site—the one bound to *R3*, nicking at Nx. Interaction between the two gpAs at *cosN* is postulated to explain nicking at Nx and N2 in the absence of ATP. In (A, b) the ATP-charged protomer nicking the bottom strand alters its mode of binding so that its nicking domain confronts the N1 site; the alternate mode of binding of *R3* by gpNu1 is also illustrated (see Discussion). In (A, c) the N117 nick is catalyzed by a terminase protomer that specifically binds DNA via its gpA subunit. In (A, d) one imagines how the static bend and/or the binding of an IHF molecule at *I1*, sharply bending the DNA at this locus, might bring the *cosN*-bound pair of protomers into contact with the *R1*−*R2* pair forming the uniquely configured termisome. In (**B**) terminase binding and nicking of mutant *cos* substrates are depicted. In (B, a) normal *cos* DNA is nicked at N1, 47 bp 3′ from the *R3* center. In (B, b) Δ11/Δ7 DNA, shown deleted for 7 bp (notch) in the *R3*−N1 interval, is nicked at *Nx*, 48 bp 3′ from *R3*. In (B, c) ins-6 DNA, which has an insertion (black block) of 6 bp in the *R3*−N1 interval, is nicked at the +11/+12 site, also 48 bp 3′ to *R3*. In B(d) +12 DNA, which has an insertion of 12 bp in the *R2*−*R1* interval, is nicked at sites 46 bp 3′ to each of *R2* and *R1*. The *R2*- and *R1*-dependent nicking can only occur when the gpNu1−gpA clash or interaction shown above is disengaged by the DNA insertion. Only within *cosN* is there indication of the assembly of a symmetric dimer of gpA molecules to concomitantly nick the top strand with the typical 12 bp stagger, at N1 and N2 in (B, a); at Ny, pairing with Nx in (B, b); and at −1/−2, pairing with +11/+12 in (B, c).

1988; Moitoso de Vargas *et al.*, 1989). The *cos* cleavage seen with the Δ7 substrate seemed more problematical in this respect.

We tested the substrates in which the *R3*−*cosN* interval was shortened by 7 or 11 bp and showed that bottom strand nicking took place almost exclusively at Nx. Top strand nicking was also displaced by 8 bp to the left, from N2 to the site Ny. Ny bears the same spatial relation to Nx as N2 does to N1, i.e. 12 bp apart. When the same interval was lengthened by an insertion of 6 bp, top and bottom strand nicks were displaced rightward from *cosN* to new sites. The new points of endonucleolytic attack are also 12 bp apart, but are embedded in DNA sequences that bear no obvious resemblance to *cosN*.

Thus, the binding of terminase holoenzyme to *R3* by the gpNu1 subunit controls the location of both the top and bottom strand nicks at *cosN*, catalyzed by gpA, in a mechanism that gauges the nick sites from *R3*. In this bipartite recognition system, the binding of terminase to *R3* through its gpNu1 subunit is likely the dominant site-specific protein−DNA interaction; the binding of the gpA subunit to *cosN* DNA and the resultant nicking activities likely take place secondarily to *R3* binding, *R3* being the point from which the location of the nicks is measured.

Furthermore, our results show that the *cosN*−*R3* interval does not behave as a simple DNA loop between two different, site-specific protein ligands within one terminase protomer, namely, gpNu1 at *R3* and gpA at *cosN*. One might then ask, what is the difference between a DNA−protein nexus that obeys helical phasing rules and one that does not? One simple answer would be that the former loop would be essentially free, while the latter

'loop' would be constrained by some degree of protein binding along its length. The DNase I footprint of terminase on *R3–cosN* DNA indicates extensive contacts by the enzyme along this interval (R.Higgins and A.Becker, in preparation).

In normal *cos*, the N1 nick in *cosN* is situated 47 nucleotides 3′ of a nucleotide interval (+53/+54) chosen arbitrarily at the mid-point of *R3*, bottom strand. The two other *R* sites, *R1* (oriented toward *R2)* and *R2* (oriented oppositely) also bind terminase protomers but apparently do not nick the DNA. Curiously, equivalent positions in *R1* and *R2* are separated from each other by exactly 47 bp, implying either a steric clash, or a specific interaction between the gpAs and gpNu1s of the apposed holoenzymes. Since the DNA of this region in the termisome appears fully protected in a DNase I footprint (R.Higgins and A.Becker, in preparation), we suggest that the protomers bound to *R1* and *R2* actually touch one another. In this configuration, their respective nicking activities are inhibited.

When the *R2–R1* DNA interval was lengthened by the insertion of DNA of 12 bp, novel nicks were catalyzed within the *R2–R1* interval, in each of the bottom and top strands. The +118/+119 nick in the bottom strand is situated 46 bp to the left of *R1*. The situation on the top strand is more complex, with nicking in three places, but confined to an interval 35–55 nucleotides to the right of *R2*, as though gauged from *R2* in the predictably rightward direction. Notably, these two nicking activities that are subtended by *R2* and *R1* do not result in the generation of nicks on the opposite strand, 12 nucleotides removed, as in the *R3*-subtended nicking dyad at *cosN*. This result indicates that *R* site-subtended nicking activities are strand specific when asked to take place outside of *cosN*. Thus, the dual (top and bottom strand) nicking operation at *cosN* appears to depend on the unique structure of *cosN*, likely its hyphenated dyad symmetry, upon which the gpA subunits of two terminase protomers might assemble to form a symmetric dimer. Interestingly, the regulation of this dual nicking operation at *cosN* is asymmetric in that it emanates from one apparent site, *R3*, on the right-hand side of *cosN*.

Our other studies (Higgins and Becker, 1994) have shown that, whereas *R3* is an important regulatory element in the nicking of *cosN*, *R1* and *R2* are not. Since terminase holoenzyme binds to both *R1* and *R2* (R.Higgins and A.Becker, in preparation), we surmise that the *R1*- and *R2*-bound protomers serve some important function other than nicking. They may, for example, be incorporated into an oligomeric 'termisome' that is needed for the initiation and/or propagation of DNA packaging (Figure 6A; Feiss, 1986; Becker and Murialdo, 1990). Our results show that these protomers were perfectly capable of nicking the DNA made available by the 12 bp insertion. Although the 47 bp spacing between *R1* and *R2* could be regarded as fortuitous, we suggest that the system may have evolved to allow binding, but not nicking, at such positions in *cosB*. Whereas we primarily view the *R3–cosN* system as a substructure for the regulation of *cosN* cleavage, we regard the *R1/R2*-based sub-assembly as contributing to other aspects, e.g. DNA packaging and/or its directionality.

Taking all of the results together, one can formulate a rule that relates the orientation of an *R* element to the point of potential nucleolytic attack by a terminase protomer bound to that *R* element, in the following way: *R* site DNA is characteristic in that one of its strands is purine-rich and a terminase enzyme bound to an *R* site will nick this purine-rich strand. The position of the nick is 47 ± 2 nucleotides away from a point in the middle of that *R* site, measured in the 3′ direction along the same strand. As a consequence, for example, for any putative terminase protomer bound to *R4* (which is in the same orientation as *R3*), the nucleolytic domains would be directed away from *cosN*. Our observations are that purified terminase, indeed, binds to *R4* (R.Higgins and A.Becker, in preparation) and nicks the bottom strand 46 nucleotides 3′ of the *R4* reference point.

The ability of an *R* site-bound enzyme to nick DNA at a measured distance, but without significant regard to sequence at the nick site, is to be contrasted with the ability of terminase to nick the *cos*-like half sites without apparent need to bind to an *R* site. These two reactivities can be viewed, in principle, as the essential components of the coupled, bipartite *R3–cosN* system which can be decomposed by the lengthening of this critical DNA interval. When bound to an *R* site through its gpNu1 subunit, terminase acts as a position-specific, strand-specific but sequence non-specific, nicking enzyme. When not bound to an *R* site, the gpA subunit of terminase seeks out *cosN*-like sites and nicks these in sequence-specific fashion. In a normal *cos* substrate, all these elements correctly come together within the specific geometry that exists between *R3* and *cosN*. Furthermore, the *cosNR*-bound protomer appears to engage a stable partner for interaction/nicking at *cosNL*, through a putative gpA–gpA dimerization. We believe that the dyadic structure of *cosN* may underlie and promote this assembly, since the dual, staggered nicking of complementary strands has not been seen outside of *cosN*.

Sequence non-specific nicking by an *R* site-bound enzyme, as well as the nicking that takes place at half-*cos*-like elements show a strict requirement for ATP. Cue and Feiss (1993) studied the role of ATP in promoting the nicking reactions at *cosN* and concluded that the role of ATP was not to promote the binding of terminase to *cosB* but rather to increase the rate of conversion of a *cosN*–terminase complex into the nicked DNA product. In our studies, only an *R3*-bound system, nicking at *cosN*, can bypass the ATP requirement, but nicking of the bottom strand is incorrect in such cases, taking place predominantly at Nx; ATP addition restores the correct specificity to N1 and stimulates both the rates (up to 10-fold) and yields of the reactions on the top and bottom strands.

What is the mechanism of the ATP-induced Nx–N1 adjustment in the nicking of *cosNR*? We can only speculate on this issue. Previously (Higgins *et al.*, 1988), we suggested that an ATP-dependent change in the tertiary or quaternary structure in the *cosN*-bound gpA complex, or an ATP-induced adjustment in gpA's *cosN* specificity determinants might be responsible. Neither of these two hypotheses has been excluded. In the light of the current results, however, another possibility comes to mind in the following way. For ATP-charged enzymes bound to *R3*, the span from the *R* site reference point to Nx is 45 nucleotides for the Δ11, and 48 nucleotides for the Δ7

substrates, respectively. The analogous span from *R3* to the +11/+12 site, in the ins-6 substrate, is 48 nucleotides. Again, the same interval in the *R1*-subtended nicking operation is 46 nucleotides. Thus, the 47 ± 2 rule appears rather strict and the *R* site-bound systems do not seem very flexible (Figure 6B). We were, therefore, surprised by the spacing of 55 nt, seen in the absence of ATP, both between *R3* and Nx in the normal substrate, and between *R3* and N1 in the ins-6 substrate and we suggest the following alternative idea on this issue.

We suggest that the terminase protomer nicking at *cosN* can be bound at *R3* in two different ways depending on whether or not it is charged with ATP (Figure 6A). We argue that such an enzyme is truly anchored by its gpNu1 subunits since it nicks *cosN* in the absence of ATP and does so even on the half-*cos* substrate that contains only *cosNR* where nicking of the top strand is not observed (Higgins *et al.*, 1988). One can then envisage a model where a pair of gpNu1 subunits in the protomer (Tomka and Catalano, 1993; our unpublished results), arranged more-or-less in tandem along the DNA axis, can alternately bind the holoenzyme at *R3*. When the gpNu1 subunit furthest from *cosN* is bound to *R3*, the nicking active center of gpA is situated at Nx; under normal circumstances, in the presence of ATP, it is the *cosN*-proximal subunit that engages *R3* so that the nicking center of gpA is at N1 (Figure 6A). We do not understand how ATP binding might regulate such enzyme repositioning, but we know from other studies (R.Higgins and A.Becker, in preparation) that it is the gpA subunits, exclusively, that incorporate [α-^{32}P]8−azido-ATP under these conditions of experiment. Thus, our data suggest that conformational changes within terminase, brought about by ATP binding to gpA, are responsible for the correct positioning of the protomer that is to nick the bottom strand at *cosN*.

Our current ideas are summarized in the models of Figure 6. The models emphasize how the binding of a terminase protomer to an *R* site orients the enzyme to nick a specific strand, at a precise distance (47 ± 2 bp) from the site of binding. Each of the protomers bound in *cos* is unique in these respects. The enzyme bound at *R3* is directed to nick the bottom strand in *cosN*. We propose that the gpA subunit of this *R3*-bound protomer recruits a terminase partner within *cosN* to form the enzymatic dyad that will fashion the cohesive end. We also suggest that it is the underlying 2-fold rotational symmetry of *cosN*, as well as the intrinsic specificity of terminase for *cosN*, that is responsible for this dimerization. By contrast, the enzymes bound at *R1* and *R2* are apposed in such a way that their respective nicking domains within their gpA subunits encounter DNA to which gpNu1 is bound, resulting in the inhibition of their nicking activities. One can then imagine how a sharp bend in the *R3*−*R2* interval would bring the *cosN*-bound pair of protomers into contact with the *R1*−*R2* pair, forming the uniquely configured termisome (Figure 6A, d). Sharp bending here is predicted by the presence of a static DNA bending locus and by the binding of IHF to *I1* (Xin and Feiss, 1988, 1993).

Materials and methods

Sequence designations

The numbering convention that we have used previously (Higgins *et al.*, 1988) gives the interval, in bases, extending away from the axis of symmetry of *cosN* to emphasize distances from the nicking domain of *cos*. It uses negative numbers to the left of the symmetry axis, beginning with −1 and proceeding leftward, and positive numbers to the right of the axis beginning with +1 and proceeding rightward. In addition, the standard numbering system for λ DNA of Daniels *et al.* (1983) is often included in parentheses.

cos-containing plasmids and cos fragment resections

Plasmids were constructed from small restriction fragments containing *cos* DNA, or portions thereof, which could be resected with restriction enzymes for use as DNA substrates; some of these have been described in Higgins *et al.* (1988) and in Higgins and Becker (1994). The plasmids that were used in this study are given in Figure 1, where the *cos*-containing DNA substrates derived from them are also sketched. Plasmid pRHmf6 has a *cos* that carries a duplication of base pairs +9(15) to +14(20); it was constructed by restricting plasmid pRH121 with *Eco*RI and *Fnu*DII or with *Hin*dIII and *Mnl*I. The *Hin*dIII−*Mnl*I and *Fnu*DII−*Eco*RI fragments were electroeluted from 15% polyacrylamide gels, precipitated and ligated to pGEM4Z which had been linearized by restriction with *Eco*RI and *Hin*dIII. Plasmid pRHxm17 contains a duplication of base pairs +15(21) to +31(37) and was constructed as above. *Hin*dIII−*Xmn*I and *Mnl*I−*Eco*RI fragments, resected from pRH121, were ligated to the *Eco*RI and *Hin*dIII sites of pGEM4Z. Plasmid pRHxf23 contains a duplication of base pairs +9(15) to +31(37). *Hin*dIII−*Xmn*I and *Fnu*DII−*Eco*RI fragments, resected from pRH121, were ligated to the *Eco*RI and *Hin*dIII sites of pGEM4Z. pIBI15^{+} carries an insertion of 12 bp in the *R2*−*R1* DNA interval (Cue and Feiss, 1992a). pLW131 and pLW142 carry DNA deletions in the *R3*−*cosN* interval of 7 and 11 bp, respectively (Miller and Feiss, 1988). pLW104 has a 1 bp deletion in the *R3*−*cosN* interval and pGM53 has a 2 bp insertion in the *R3*−*cosN* interval (Miller and Feiss, 1988). Plasmids pIBI15^{+}, pLW131, pLW142, pLW104 and pGM53 are the gifts of Drs Feiss and Kosturko.

Enzymes and reagents

Terminase holoenzyme was purified as outlined in Higgins *et al.* (1988). Restriction enzymes were purchased from New England Biolabs, Boehringer Mannheim and Pharmacia. T4 DNA ligase, T4 polynucleotide kinase and calf intestinal alkaline phosphatase were from New England Biolabs. AMV reverse transcriptase was from Boehringer Manheim and DNase I from Promega Biotechnologies. dNTPs and ATP were from Sigma Chemical Co. [α-^{32}P]dATP and [γ-^{32}P]ATP (3000 Ci/mmol) were from Amersham.

Preparation of end-labeled DNA fragments

^{32}P end-labeled restriction fragments were 3′ end-labeled using AMV reverse transcriptase and [α^{32}P]dATP, or were 5′ end-labeled using T4 polynucleotide kinase and [γ-^{32}P]ATP after dephosphorylation of the ends with calf intestinal alkaline phosphatase. Fragments labeled at one end were then prepared by digestion with a second restriction enzyme. The labeled fragments were isolated by electrophoresis on 5–10% polyacrylamide gels, followed by electroelution and were concentrated by ethanol precipitation.

Other methods

The assays for nicking at *cosN* catalyzed by terminase were performed as described previously (Higgins *et al.*, 1988). For electrophoresis on sequencing gels, samples were heated at 90°C for 5 min in a solution of 80% formamide, 10 mM NaOH, 1 mM EDTA, 0.1% xylene cyanol and 0.1% bromophenol blue (gel-loading buffer). Autoradiography of dried gels was done at −80°C using XAR-5 X-ray film with a Du Pont Lightning Plus intensifying screen. When required, band intensities were measured by scanning autoradiographs, using a Bio-Rad 620 video densitometer. The results were integrated using a Bio-Rad 3302A integrator.

Acknowledgements

We wish to thank Paul Sadowski and Clarence Fuerst for valuable discussions and Michael Feiss and Linda Kosturko for the gift of plasmids. The work was supported by a grant from the Medical Research Council of Canada.

References

Bear,S., Court,D. and Friedman,D. (1984) *J. Virol.*, **52**, 966–972.
Becker,A. and Murialdo,H. (1990) *J. Bacteriol.*, **172**, 2819–2823.

Cue,D. and Feiss,M. (1992) *J. Mol. Biol.*, **228**, 58–71.
Cue,D. and Feiss,M. (1993) *J. Mol. Biol.*, **234**, 594–609.
Daniels,D., Schroeder,J., Szybalski,W., Sanger,F., Coulsen,A., Hong,D., Hill,D., Peterson,G. and Blattner,F. (1983) In Hendrix,R.W., Roberts,J., Stahl,F. and Weisberg,R. (eds), *Lambda II*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 519–676.
Feiss,M. (1986) *Trends Genet.*, **2**, 100–104.
Feiss,M. and Becker,A. (1983) In Hendrix,R.W., Roberts,J., Stahl,F. and Weisberg,R. (eds), *Lambda II*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 305–330.
Feiss,M. and Widner,W. (1982) *Proc. Natl Acad. Sci. USA*, **79**, 3498–3502.
Feiss,M., Widner,W., Miller,G., Johnson,G. and Christiansen,S. (1983) *Gene*, **24**, 207–218.
Higgins,R.R. and Becker,A. (1994) *EMBO J.*, **13**, 6152–6161
Higgins,R.R., Lucko,H. and Becker,A. (1988) *Cell*, **54**, 765–775.
Hochschild,A. and Ptashne,M. (1986) *Cell*, **54**, 681–687.
Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499–559.
Miller,G. and Feiss,M. (1988) *Mol. Gen. Genet.*, **212**, 157–165.
Miwa,T. and Matsubara,K. (1983) *Gene*, **24**, 199–206.
Moitoso de Vargas,L., Kim,S. and Landy,A. (1989) *Science*, **244**, 1457–1461.
Shinder,G. and Gold,M. (1988) *J. Virol.*, **62**, 387–392.
Thompson,J.F., Snyder,U.K. and Landy,A. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 6323–6327.
Tomka,M.A. and Catalano,C.E. (1993) *J. Biol. Chem.*, **268**, 3056–3065.
Xin,W. and Feiss,M. (1988) *Nucleic Acids Res.*, **16**, 2015–2030.
Xin,W. and Feiss,M. (1993) *J. Mol. Biol.*, **230**, 492–504.

Received on June 27, 1994; revised on September 22, 1994