Chapter 2

Outgrowth of Human Immunodeficiency Virus Type I Induced by Hairpin Polyamides Targeted to the RCS of the LTR: A Potential Role for the Host Transcription Factors LSF and YY1 in Regulating Latency

The text of this chapter was taken in part from two manuscripts co-authored with Professor David Margolis and co-workers of UT, Southwestern Medical Center

Ylisastigui, L., Coull, J.J., Rucker, V.C., Melander, C.M., Bosch, R.J., Brodie, S.J., Sodora, D.L., Dervan, P.B., Margolis, D.M. (2003) Derepression by DNA-binding Polyamides Allows Outgrowth of Human Immunodeficiency Type 1 from Resting $CD4^+$ T Cells: A Role for Host Factors in HIV Latency. *In preparation*.

Coull, J.J., He, G., Melander, C.M., Rucker, V.C., Sodora, D., Dervan, P.B., and Margolis,D.M. (2002) Derepression of Quiescent Human Immunodeficiency Type 1 by DNA Binding Polyamides: A Potential Role for Host Factors in HIV Latency. *J . Virol.*, **76**, 12349

Abstract

This chapter describes the design and biological effect, *in vitro* and *in vivo*, of hairpin polyamides targeted to the repressor complex sequence (RCS) of HIV-1 stably integrated into the genome of $CD4^+$ T-cells. Previous research in the Dervan lab has demonstrated the efficacy of polyamides towards repressing HIV-1 expression *in vitro*. The work presented in this chapter describes the use of hairpin polyamides as agents that act in a sequence specific manner to *derepress* the stably integrated HIV-1 genome and upregulate LTR expression. HIV-1 derepression is achieved by disrupting the binding of the host cellular repressor LSF_2/YY1 transcription factor complex bound to the RCS. We characterize the interactions of RCS specific polyamides against LSF and YY1 *in vitro* and in *in vivo* Hut78 cell model systems. With this data in hand, we show that RCS specific hairpin polyamides are able to upregulate basal transcription of the LTR in quiescent HIV-1^+ $CD4^+$ lymphocytes. This data from outgrowth in quiescent $CD4^+$ T-cells suggests a role for LTR repression by the host transcription factors LSF and YY1 as the polyamides used are, as shown in our early *in vitro* and *in vivo* experiments, to be quite specific for disrupting the LSF_2/YY1 complex.

This chapter is divided into two portions. Chapter 2A discusses the design of compounds targeted to the RCS, the first time polyamides have been successfully designed to target within the coding region of the LTR to affect a biological outcome. Chapter 2B provides biological results about the ability of polyamides to antagonize the DNA bound LSF_2/YY1 complex and upregulate LTR expression *in vivo* and *in vitro*. We also present micrographs of bodipy labeled polyamides entering the nucleus of resting $CD4^+$ T-cells.

Background

HIV-1 infection is characterized by cycles of virus production and reinfection within activated $CD4^+$ T-lymphocytes. However, HIV establishes a persistent, nonproductive state within a small population of memory/resting/unactivated $CD4^+$ cells.[1-3] Infection is relatively short-lived in activated T-cells, but it is estimated that as much as 60 years of continuous antiretroviral therapy would be required to eliminate the reservoir of HIV infection within memory $CD4^+$ cells. Classic highly active anti-retroviral therapy (HAART) often only diminishes blood plasma HIV-1 RNA concentrations to very low levels without fully eradicating the infection (Figure 2.1).[4,5] Cellular signaling pathways that allow the provirus to establish latency are unknown, but they may be the result of an infected, activated T-cell returning to a resting state, or infection of an activated T-cell that has already begun to return to G_0.

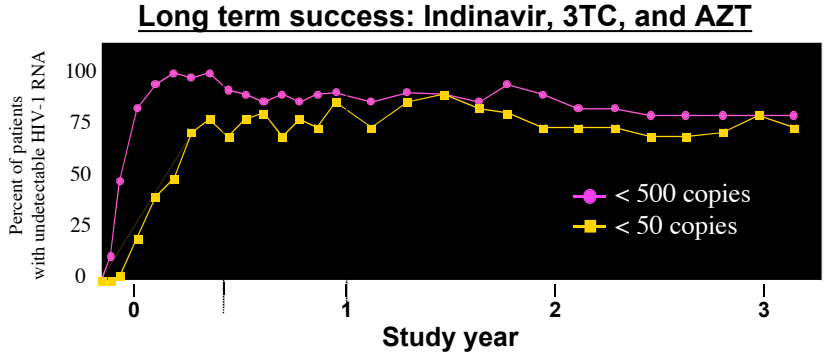


Figure 2.1 Graphical representation of HIV-1 latency phenotype whereby, during intense HAART, HIV-1 RNA becomes virtually undetectable falsely suggesting the infection may be on the decline.

Modulation of histone architecture within the host genome may contribute to HIV latency. The integrated LTR is unresponsive to activation by NF-□B prior to histone acetylation, implying a role for histone deacetylase (HDAC) in establishing or maintaining quiescence.[6] The cellular transcription factors YY1 and LSF form a complex (LSF_2/YY1) that represses transcription from the HIV-1 LTR through recruitment of

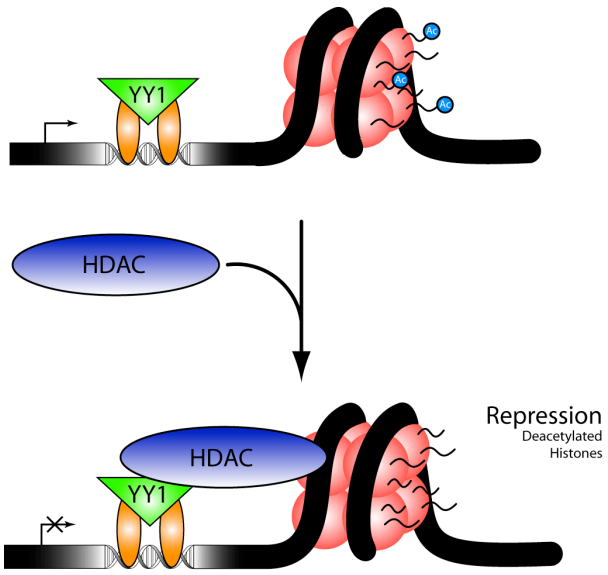


Figure 2.2 Current model for RCS bound LSF (orange ellipses) and YY1 mediated repression of LTR expression through recruitment of HDAC1. (Thanks to EJF for graphic.)

HDAC1 (Figure 2.2).[7-9] Thus chromatin remodeling, as evinced from the active role played by a nucleosome in repressing the LTR, is implicated in generation of latency.[10-12]

The LSF_2/YY1 complex bound to the RCS of the LTR is directly responsible for LTR downregulation through recruitment of HDAC1. Antagonizing LSF_2/YY1 binding would achieve upregulation of the HIV-1 LTR. Specifically targeting LSF binding sites might disrupt the LSF_2/YY1 complex, because YY1 shows no DNA recognition properties in the absence of LSF. Both YY1 and LSF are major groove binding proteins, and YY1 is known to be a zinc finger protein. Previous work in our group has demonstrated the ability of polyamides to allosterically antagonize major groove binding zinc-finger proteins suggesting the potential for success.[13]

Chapter 2A

Design of Polyamides to Target the Repressor Complex Sequence (RCS) of the Stably Integrated HIV-1 Long Terminal Repeat (LTR)

The project of designing polyamides targeted to the RCS of the LTR incorporates several interesting polyamide design issues of potential importance to the small molecule gene regulation community. The design of biologically relevant polyamides targeted to DNA within the coding region of the LTR has never been fully explored. The reason for this may be that polymerase II (PolII), the most obvious enzyme to inhibit within the coding region, is reported to be unrefractory towards polyamides that bind downstream of the transcription start site.[14] Thus, this project is novel because it covers the design of polyamides that successfully target the DNA downstream of the LTR transcription start site to successfully displace inhibitory transcription factors that bind there, i.e., the LSF_2/YY1 complex.

Additionally, this project is novel because of the nature of the cellular phenotype we are attempting to chemically elicit. There is only one report of polyamides that exert *de*repression or *anti*repression activity towards gene expression.[15] In this previous study, the human cytomegalovirus UL122 (CMV) repressor immediate early protein 2 (IE86), that normally represses CMV expression through interaction with binding sites upstream of the transcription start site, was competed from its binding site leading to production of CMV gene product.

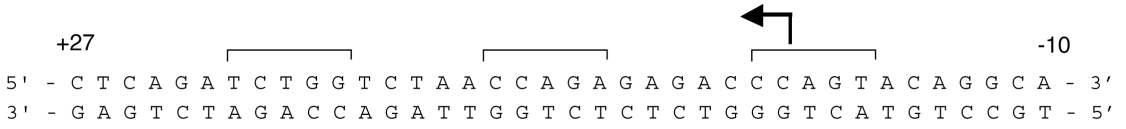


Figure 2.3 Sequence of RCS within the LTR transcription start site. Conserved LSF recognition sites indicated by brackets.

Results and discussion

The sequence of the RCS is shown in Figure 2.3. The sequences in brackets are the conserved LSF_2/YY1 binding sites as ascertained by mutation analysis.[8,9] The transcription start site is indicated with an arrow. Polyamides were designed to target the DNA minor groove of the RCS (Figure 2.4). Shown in Figures 2.4 and 2.5 are the first generation compounds and the binding site where they reside as determined by DNAse I or methidiumpropyl-EDTA (MPE) footprinting techniques. The affinities were determined by quantitative DNAse I footprint titration analysis. Footprinting was performed on a 5' radiolabeled 229 base pair PCR product from pIBI20 provided by Professor David Margolis (see experimental below).

The compounds in Figures 2.4 and 2.5 were designed to target as much of the RCS as possible (in particular the conserved recognition sites of LSF_2/YY1 indicated with brackets), but often the DNA recognition behavior of the polyamide was different from the expectation. The often surprising DNA recognition properties of some of the polyamides discussed in this chapter has led to interesting insight into DNA recognition by polyamides in general, and in particular the new 3-□-3 motif presented here (**3** – **8**, **M3/4**). Details of the DNA recognition properties of these polyamides are discussed below. Because not all polyamides in Figure 2.4 are "well-behaved," some (**5**, **6**, **7**, **8**, **11**,

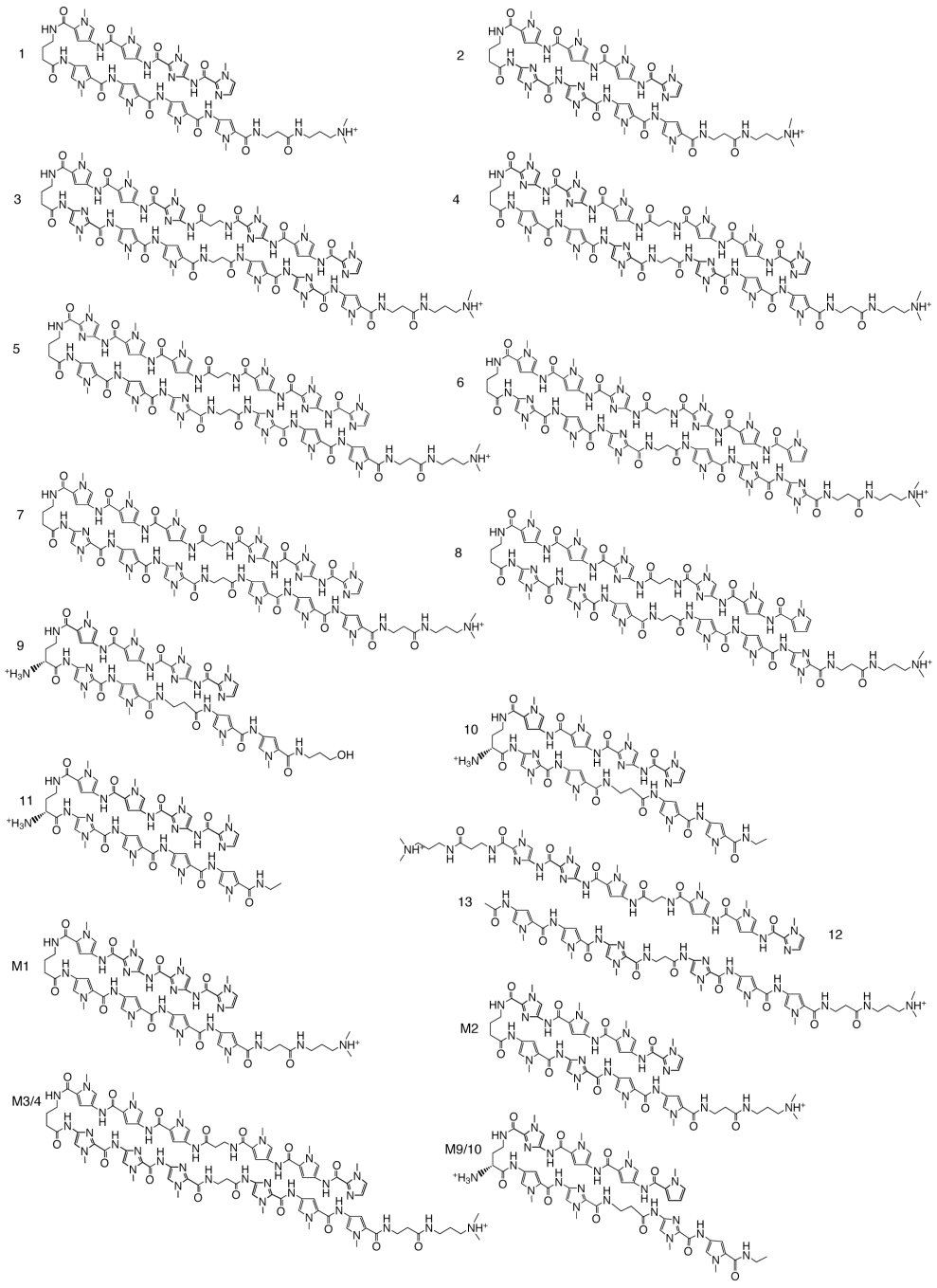


Figure 2.4 Structure of compounds designed to study polyamide interactions with the RCS of the LTR.

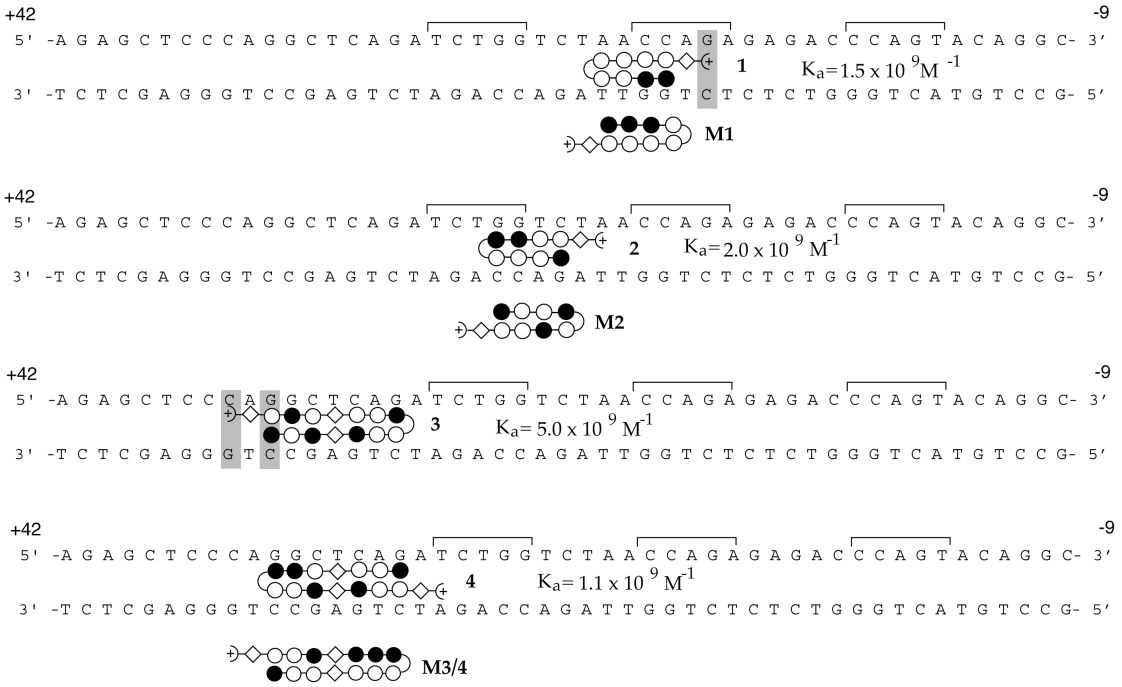


Figure 2.5 Polyamides targeted to the RCS that antagonize the LSF_2/YY1 complex.[18] Shown is the determined binding site for each polyamide and the affinity. The affinity of the controls **M1**, **M2**, and **M3/4** is < 10^8 M^{-1}. Grey indicates unfavorable interaction.

12, **13**) were not used in biological assays and are not discussed in Chapter 2B. Additionally, a few polyamides in Figure 2.4 (**9**, **10**, **M9/10**) have yet to be rigorously characterized in a biological system. Biological results for the other compounds are given in Chapter 2B.

1 and **2** were synthesized and footprinted showing high affinity and specificity towards their match binding site (Figure 2.5, Figure 2.6). The control compounds **M1** and **M2**, constitutional isomers of **1** and **2**, exhibit no affinity for the RCS under experimental conditions (K_a < 10^8 M^{-1}). Polyamides **3** - **8**, and mismatch control **M3/4**, are of a new motif termed the "3-□-3 motif," named so because there are 3 aromatic

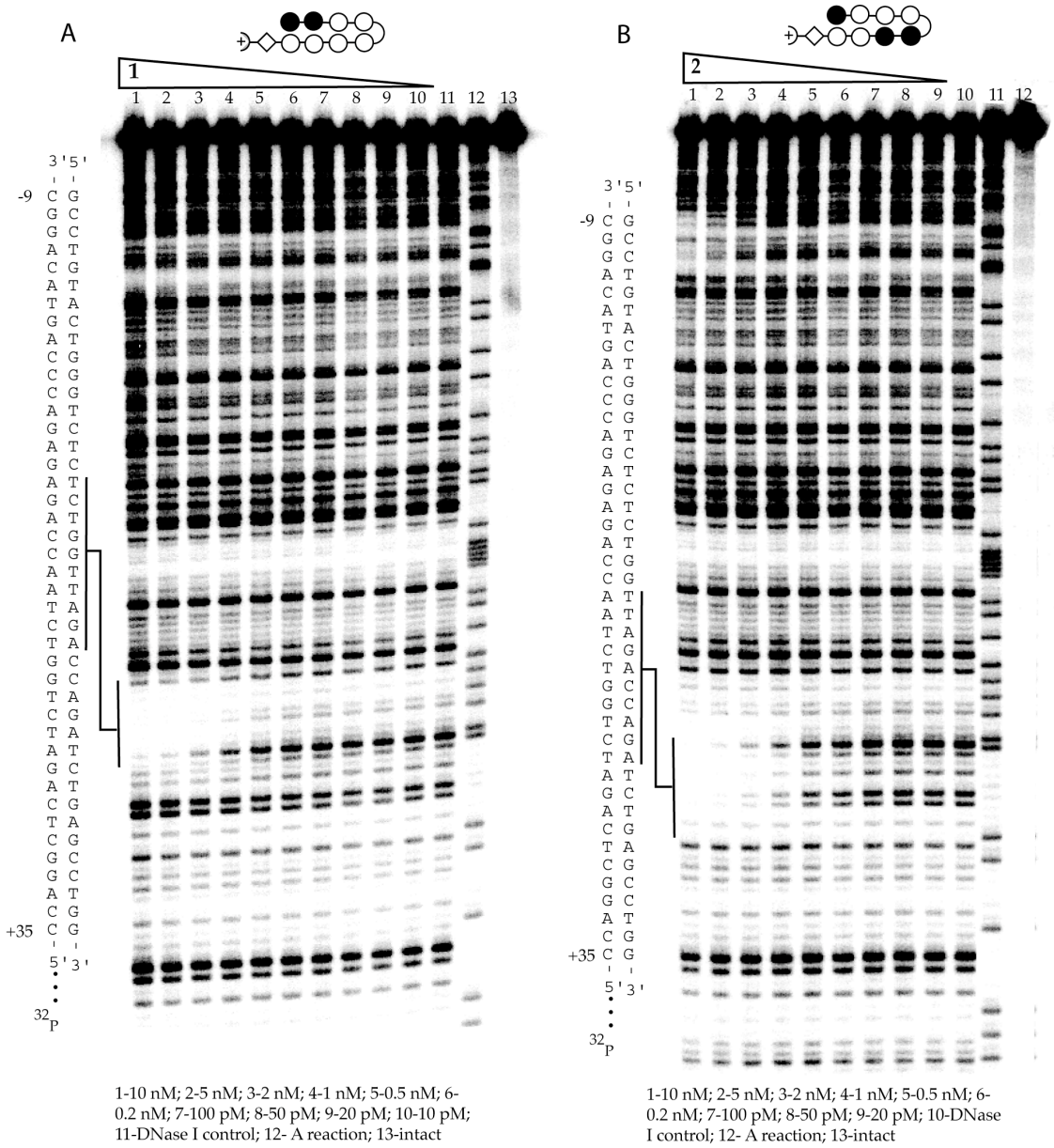


Figure 2.6 DNAse I footprint titration for **1** (A) and **2** (B) binding to the 229 base pair fragment of the 5'-^{32}P labeled PCR product from pIBI20.

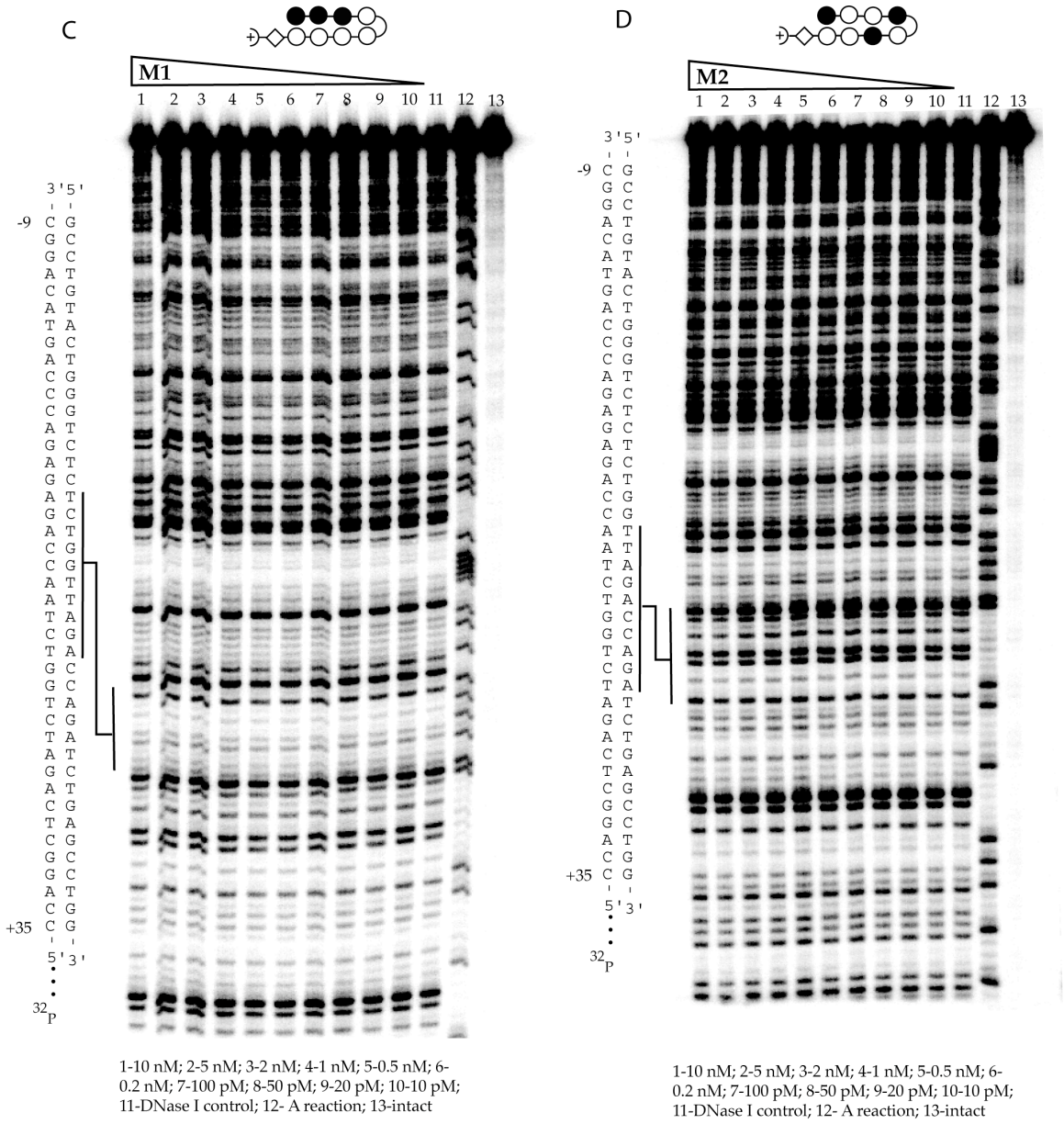


Figure 2.6 (continued) DNAse I footprint titration for controls **M1** and **M2** equilibrated against the 5'-^{32}P labeled PCR product from pIBI20.

Figure 2.7 Designed binding site for **5** targeted to the most upstream conserved LSF binding site.

rings, a β-alanine monomer, and 3 aromatic rings. These compounds are an extension of the 2-β-2 motif that has received much attention in our group because it allows for relaxation of the overcurvature of a polyamide composed of 5 or more aromatic rings.[16,17] The allure of 3-β-3 polyamides is the huge length of DNA protected when these compounds bind in the minor groove (10 base pairs).

The first 3-β-3 designed for this project is shown in Figure 2.7. Polyamide **5** was expected to bind the sequence predicted as a pairing rules match. It did not bind, however, as shown by DNAse I footprint titration PAGE analysis (data not shown). Interestingly, the compound that was to be the control for **5**, called **3** in this thesis and elsewhere,[18] bound in anomalous fashion to the sequence 5'-AGGCTCAGAT-3' with high affinity and specificity (Figure 2.8). Aligning the N→C of polyamide **3** with the 5'→3' of the DNA phosphate backbone presents many mismatches between the aromatic DNA recognition elements of the polyamide and the edges of the DNA bases in the minor groove. The motivation for alignment of the DNA 5'→3' phosphate backbone parallel to the N→C polarity of the polyamide recognition peptide, in spite of the numerous mismatches as shown in Figure 2.8, arises from studies conducted in this group indicating this as the preferred alignment for polyamide recognition of DNA in the minor groove.[19]

Alternatively, one can imagine, as shown in Figure 2.8, that alignment of C→N polarity of polyamide **3** with the 5'→3' polarity of the phosphate backbone results in

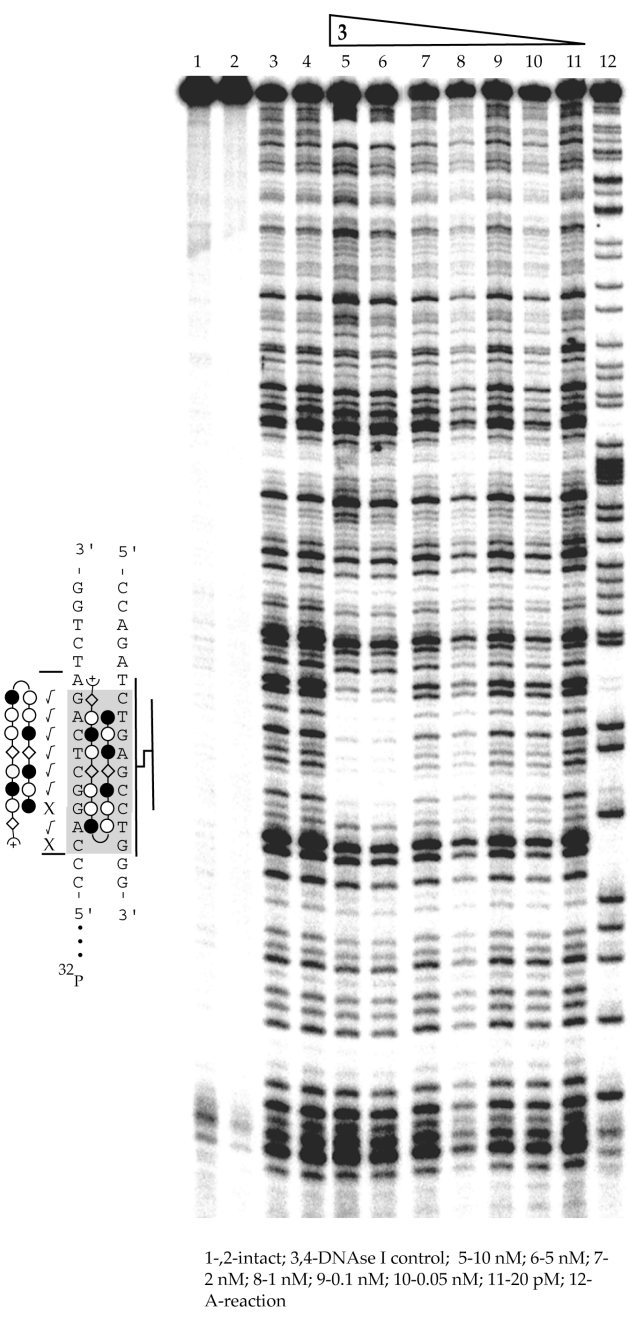


Figure 2.8 DNAse I footprint titration of **3** with the 5'-^{32}P labeled PCR product from pIBI20. Note alignment differences between N→C, 5'→3' (inside sequence) and C→N, 5'→3' (outside sequence).

significantly less mismatches between the aromatic rings of **3** and the edges of the DNA bases in the minor groove. Indeed, redrawing **3** such that it is in an orientation "reversed" from that of the N→C, 5'→3' alignment provides for only 2 mismatches between the polyamide and the DNA bases in the minor groove. To test the hypothesis that **3** may be recognizing the minor groove in a reversed orientation, the EDTA affinity cleavage analog of **3** (**3E**), was synthesized and the orientation preference of **3** was examined (Figure

2.9). We observed that the hydroxy radical cleavage pattern does allow for a reversed binding mode of **3** as postulated, i.e., the polyamide N-terminus of each recognition peptide is aligned with the 3'-end of the DNA strand it recognizes.

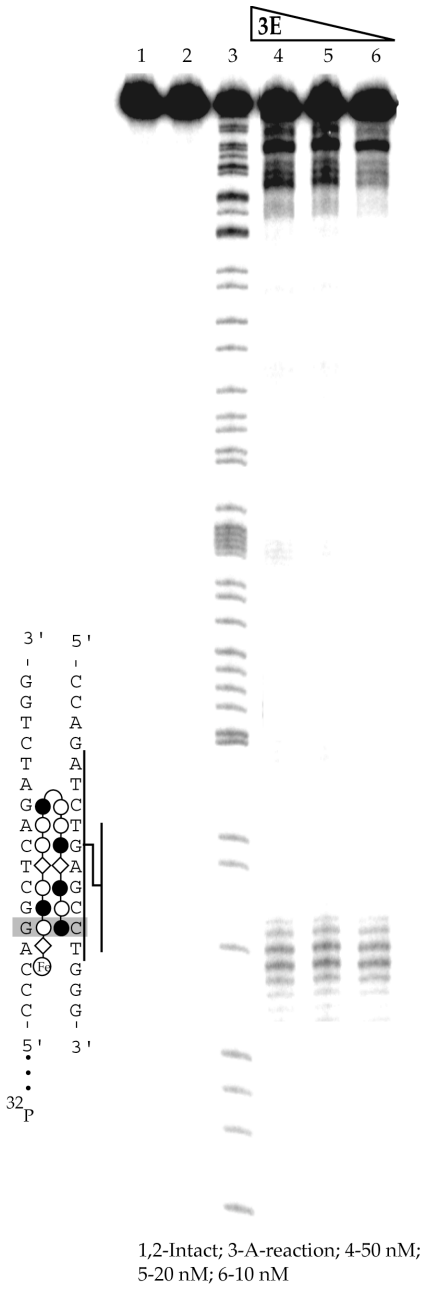


Figure 2.9 Affinity cleavage pattern for **3E** indicating a "reversed" binding mode.

Because of the aromatic ring-DNA mismatch in **3** (Im/Py ring pairing over C:G base pair), we set out to design 4 additional compounds that would not have aromatic ring•base, tail-2, and tail-1[20] mismatches. Polyamides **4**, **6**, **7**, and **8** were synthesized to explore the DNA recognition properties of all 4 constitutional isomers of **3** (Figure 2.10). The Im/Py mismatch over a C:G pair in **3** is corrected in **4**, **6**, **7**, and **8,** and the tail-2 and tail-1 mismatches are corrected in compounds **4** and **8**. Compounds **4** and **6** explore the newly observed orientation, C N, 5' 3', while compounds **7** and **8** provide for the previously published studies that suggest polyamides prefer to align N C, 5' 3'.

4, **6**, **7**, and **8** all bind DNA with high affinity ($K_a \geq 10^9$ M^{-1}), but only **4** shows good specificity towards only the match site (Figure

2.11). The power of screening the DNA recognition properties of these 4 polyamides is that the polyamide is telling us exactly what peptide polarity and connectivity of rings is most suitable for recognizing the sequence 5'-AGGCTCAGAT-3'.

Considering that **4** would bind the sequence 5'-AGGCTCAGAT-3' as a perfect pairing rules match in a reversed orientation, we performed MPE footprinting and affinity cleavage with the EDTA analog of **4** (**4E**) to determine the exact binding site size and orientation of **4**. What we found is that **4** indeed binds in a reversed orientation as shown by the single DNA cleavage pattern corresponding to a reversed orientation (Figure 2.12).

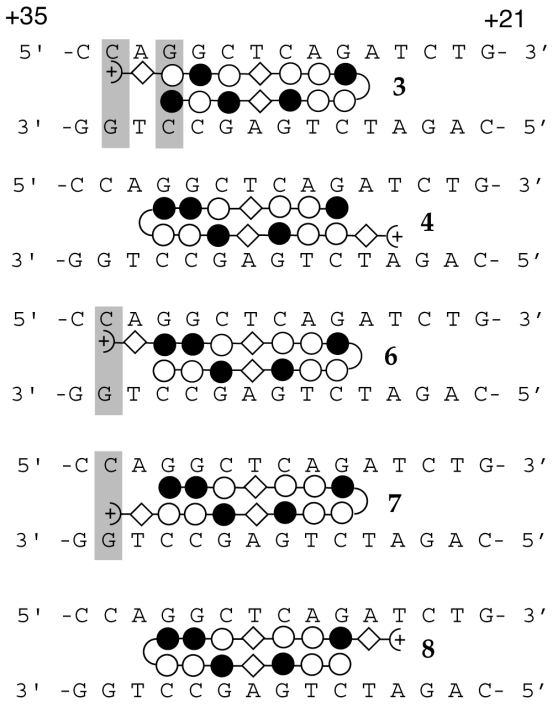


Figure 2.10 4 constitutional isomers of **3**, corrected for Im/Py mismatch in **3**. These "connectamers" probe all 4 tail and turn possible attachments and binding orientation.

Returning to our efforts to design compounds targeted to the RCS, as shown in Figures 2.3 and 2.4, the conserved recogntion sites for the LSF_2/YY1 complex comprise the sequences 5'-TCTGG-3' (+24 to +20), 5'-CCAGA-3' (+14 to +10), and 5'-CCAGT-3' (+4 to 0). With compounds **1** through **4** (and their controls **M1**, **M2**, and **M3/4**) we effectively target the two most downstream LSF_2/YY1 recognition sites. However, none of the compounds in Figure 2.4 target the most upstream LSF_2/YY1 recognition element that

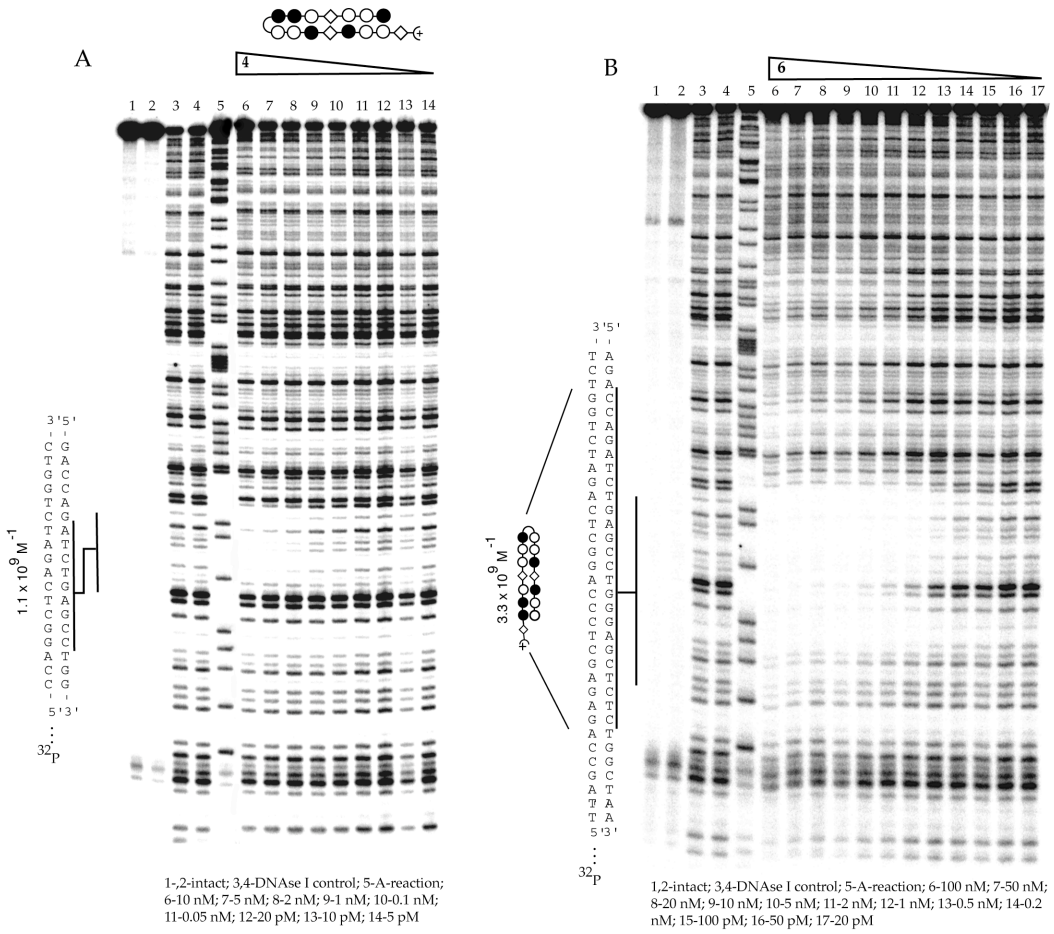


Figure 2.11 DNAse I titration experiments for compounds **4** and **6** designed to bind in a C N, 5’ 3’ orientation.

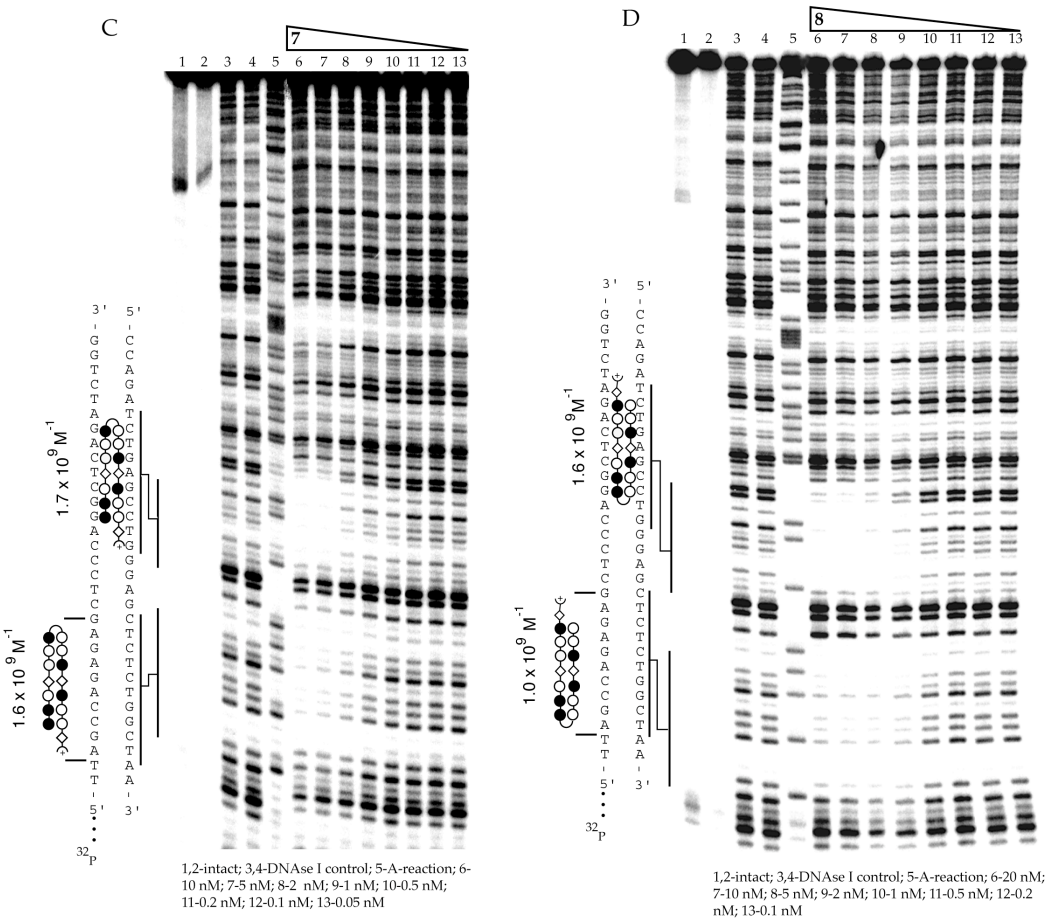
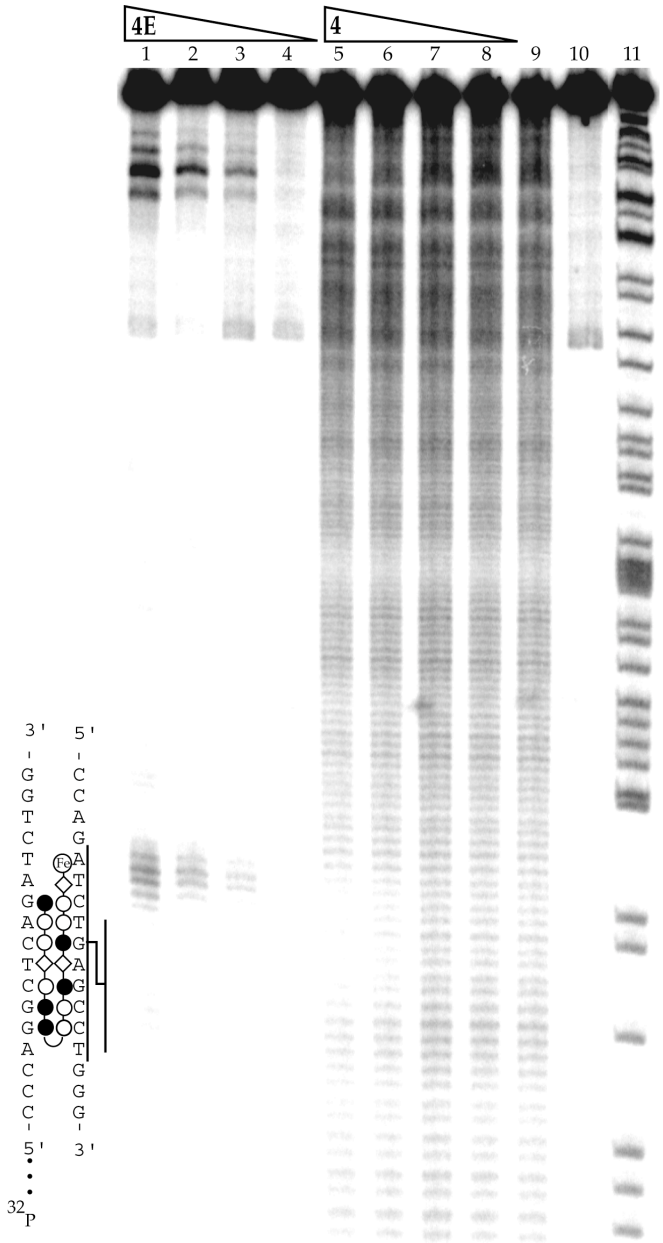


Figure 2.11 (continued) DNAse I titration experiment with compounds **7** and **8** designed to bind in a N C, 5' 3' orientation.



1-20 nM; 2-10 nM; 5 nM; 1-nM; 5- 10 nM; 6-5 nM; 7-1 nM;
8-0.1 nM; 9-MPE control; 10-intact; 11-A-reaction

Figure 2.12 MPE and affinity cleavage assay of **4** and **4E** binding against LTR. Note single cleavage pattern of **4E** consistent with a C N, 5' 3' orientation.

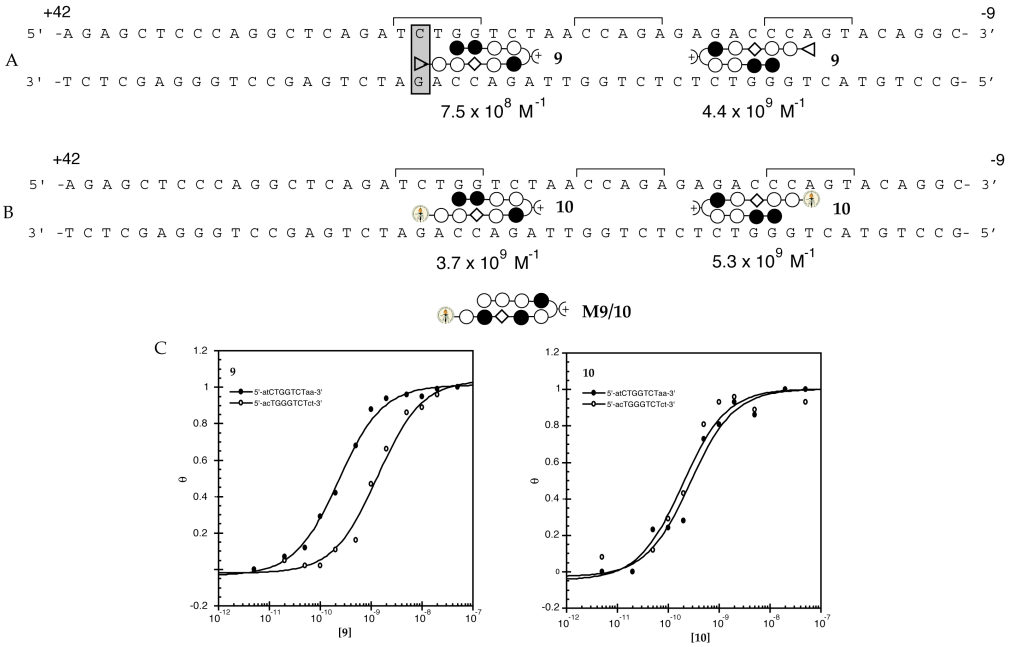


Figure 2.13 A. Binding site and affinity for **9**. B. Binding affinity and site for **10**. C. Isotherms for **9** and **10** Clearly indicated is the energetic cost associated with interaction between the propanol amide tail of **9** with a G:C base pair in the minor groove.

encompasses the transcription start site, 5'-CCAGT-3'. Compound **5** was designed for this site (Figure 2.7) but there was no binding observed under experimental conditions. To achieve the goal of designing polyamides to recognize DNA encompassed by all three LSF recognition domains, polyamide **9** (Figure 2.4) was synthesized using oxime resin techniques[20] developed in this laboratory.

Containing a pendent, unpaired, C-terminal pyrrole residue, **9** was expected to bind two LSF_2/YY1 recognition domains simultaneously with equal affinity, thus providing for single molecule antagonization of two LSF_2/YY1 binding domains (Figure 2.13A). It was observed, however, that **9** bound the sequence 5'-CTGGTCT-3' with ~ 6 fold lower affinity than the sequence 5'-AGACCCA-3' (Figure 2.13C). Closer inspection of **9** revealed that a disfavored interaction existed between the hydroxyl group

of the propanol amide unit at the C-terminus of **9** and the G:C base pair it resides over. This energetic cost is predicted from previous studies.[20] Compound **10** was designed to alleviate this problem. The ethyl amide tail of **10** (Figure 2.4, Figure 2.13B) alleviated the steric clash at the C-terminus of **9** and equal affinity was observed for binding at both sites by **10** (Figure 2.13C).

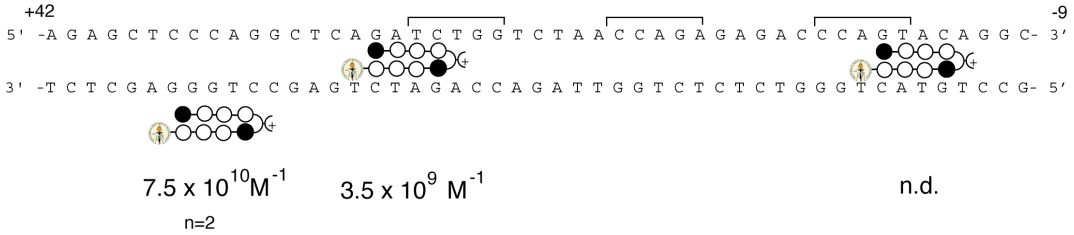


Figure 2.14 Designed binding sites (shown with compound within DNA sequence) and unanticipated binding site (polyamide outside of sequence) for compound **11**. Association constants (K_a values) are shown beneath each polyamide.

Another 8-ring hairpin polyamide we designed to bind upstream and downstream sites of the RCS was the compound ImPyPyPy-β-ImPyPyPy-(ethyl amide) which should recognize the sequence 5'-GWWC-3' within the RCS.[14] This compound was designed to bind the most upstream and downstream of the LSF_2/YY1 conserved recognition elements (Figure 2.14 as shown for compound aligned with putative match sites). What we observed, however, was anomalous binding to the RCS (Figure 2.15) such that weak affinity (< 10^8 M^{-1}) was observed for the most upstream site 5'-GTAC-3', good affinity (3.5 x 10^9 M^{-1}) for the most downstream RCS site 5'-GATC-3', and high affinity (7.5 x 10^{10} M^{-1}) at an unanticipated binding site of the sequence 5'-AGAGCTC-3' (Figure 2.14). The binding isotherm at this unexpected site, 5'-AGAGCTC-3', fit a n = 2 Hill isotherm well indicating a possible 2:1 (opened hairpin) binding motif. This underscores, as with the design of the 3-β-3 compounds, the unexpected surprises we found when designing compounds that target the coding region of the LTR.

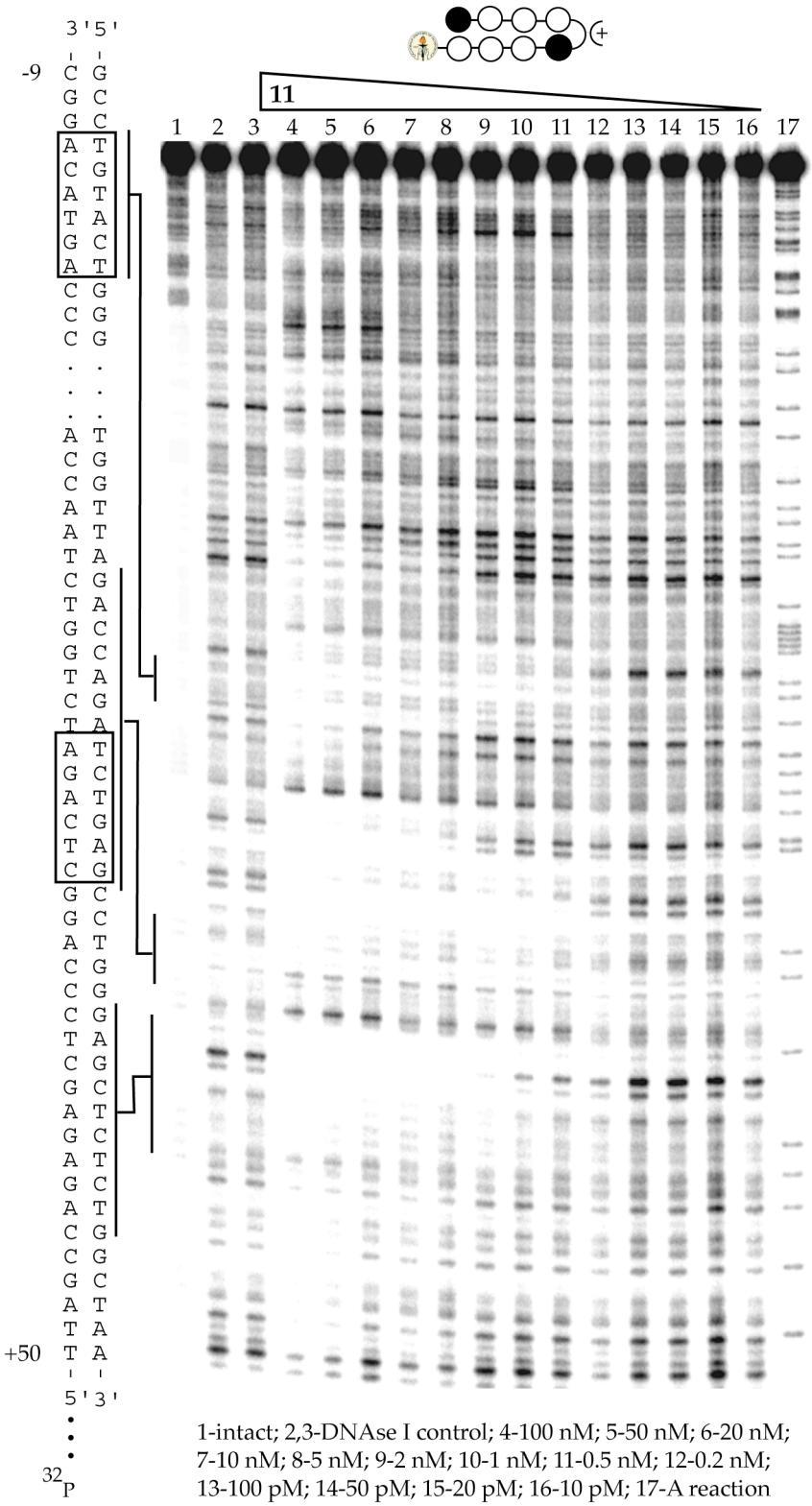


Figure 2.15 Designed (boxes) and observed binding sites for **11** on the LTR as ascertained by DNAse I footprint titration analysis.

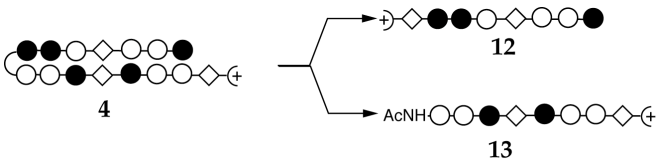
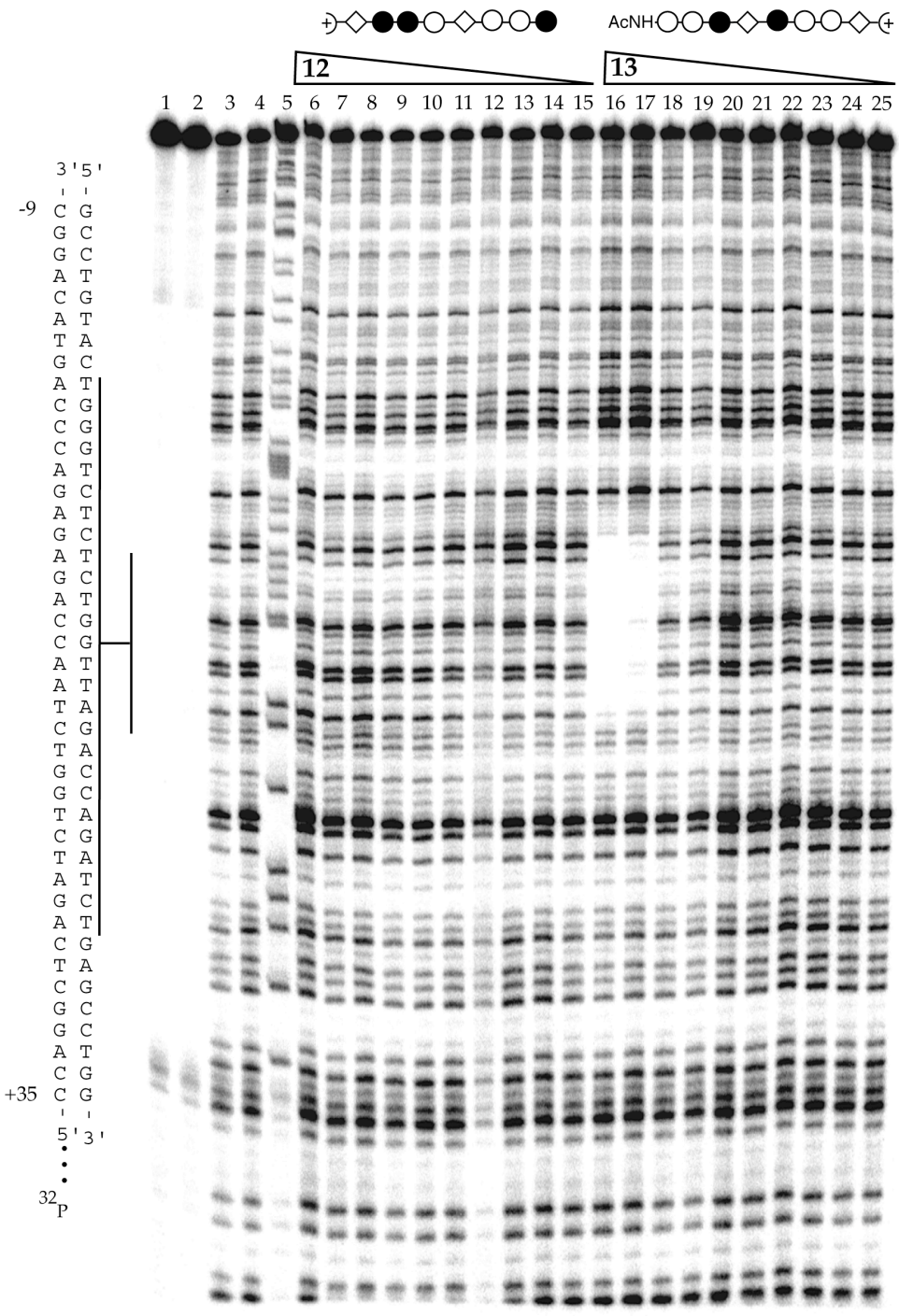


Figure 2.16 Representation of division of **4** into two fragments, **12**, a N-terminal fragment and **13**, a C-terminal fragment.

Along this same vein, two additional compounds **12** and **13**, representing the N-terminal (**12**) and C-terminal (**13**) halves of **4** were synthesized to establish that **4** was not binding in an anomalous 1:1 fashion (Figure 2.16).[21] Because the N→C, 5'→3' polyamide alignment doctrine has been so well supported,[19] we felt thoroughness required full investigation of **4**'s DNA recognition properties by exploring the two antiparallel strands of **4** separately. We observed no binding of **12** to the RCS, but **13** quite specifically bound to nearly the entire RCS with a $K_a = 5 \times 10^8\ M^{-1}$ (Figure 2.17). Affinity cleavage experiments with the EDTA analog of **13** revealed a very complex binding behavior likely consisting of both 2:1 and 1:1 binding to the minor groove. Several models may be proposed (Figure 2.18), but no recognition mode of **13** suggests that **4** binds DNA in any fashion other than as a hairpin.

Conclusions

In closing, the design of compounds for targeting the RCS of the HIV-1 LTR is clearly a complex issue that obscures the direct simplicity of the pairing rules and other guidelines for polyamide design.[22] In the course of designing polyamides targeted to the RCS, we have explored a new motif, the 3-β-3 (**3** - **8**, **M3/4**), examined reverse binding in the context of the 3-β-3, explored a simple set of polyamides with pendent C-terminal unpaired pyrrole residues (**9**, **10**, **M9/10**), observed that some of our best characterized



1,2-intact; 3,4-DNAse I control; 5-A-reaction; 6-10 nM; 7-5 nM; 8-2 nM; 9-1 nM; 10-0.5 nM; 11-0.2 nM; 12-0.1 nM; 13-50 pM; 14-20 pM; 15-10 pM; 16-10 nM; 17-5 nM; 18-2 nM; 19-1 nM; 20-0.5 nM; 21-0.2 nM; 22-0.1 nM; 23-50 pM; 24-20 pM; 25-10 pM

Figure 2.17 DNAse I footprint titration analysis of **12** and **13** equilibrated against the 5'-^{32}P bound RCS. Indicated is the sequence of DNA protected by binding of **13**.

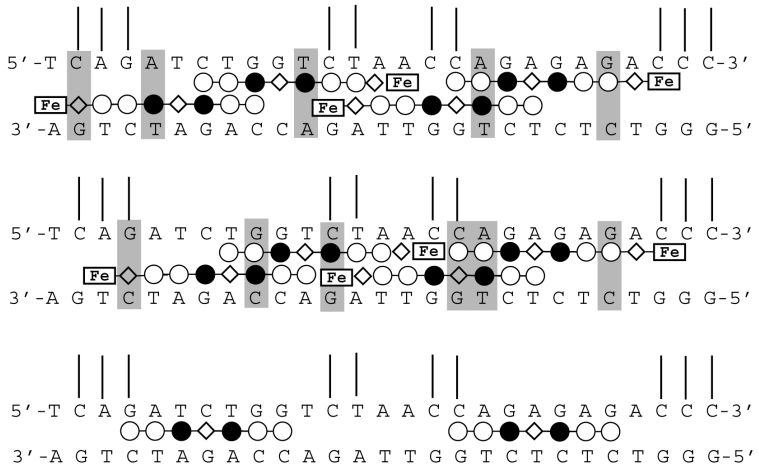


Figure 2.18 Observed cleavage data from affinity cleavage experiment conducted with EDTA derivative of **13**. Indicated are the cleavage sites (gel not shown) and possible orientations that may correspond to cleavage pattern. For simplicity, polyamides are shown without NHAc at N-terminus.

polyamides exhibit anomalous binding behavior (**11**), and touched upon a non-hairpin motif (**13**) that is (inexplicably) specific for a large sequence of DNA within the RCS.

Experimental

Polyamide synthesis

All polyamides were synthesized using standard Boc-β-Ala PAM resin protocols or oxime resin protocols.[20,23] Liberation of polyamides from resin was achieved by aminolysis with the appropriate amine. All compounds were purified by reversed phase HPLC (C_{18} Waters, 0.1% TFA/ACN), and identity was verified by analytical HPLC and MALDI/TOF MS analysis. For **1** (monoisotopic) 1223.47 (1223.58 calc'd for M + H); **2** 1224.33 (1224.57 calc'd for M + H); **3**, 1856.79 (1856.83 calc'd for M + H); **4**, 1856.21 (1856.83 calc'd for M + H); **5**, 1856.34 (1856.83 calc'd for M + H); **6**, 1856.88 (1856.83 calc'd for M + H); **7**, 1856.45 (1856.83 calc'd for M + H); **8**, 1856.11 (1856.83 calc'd for M + H); **9**, 1212.64 (1212.54 calc'd for M + H); **10**, 1182.67 (1182.52 calc'd for M + H); **11**, 1111.53 (1111.49 calc'd for M + H); **12**, 966.98 (966.46 calc'd for M + H); **13**, 1022.24 (1022.49 calc'd for M + H); **M1**, 1223.86 (1223.58 calc'd for M + H); **M2**, 1224.88 (1224.57 calc'd for M + H); **M3/4** 1856.10 (1856.83 calc'd for M + H); **M9/10**, 1182.48 (1182.52 calc'd for M + H). Compounds were stored at –80 ºC and used as necessary.

Radiolabeled PCR product preparation

pIBI20 was received from Professor David Margolis at UT Southwestern Medical Center. Two primers, CM-A: 5'-TTG AGG CTT AAG CAG TGG GTT C-3' and CM-B: 5'-AGC TGC ATC CGG AGT ACT ACA A-3' were used to generate a 229 base pair PCR product labeled with ^{32}P at the 5' end of primer CM-A. Protocols for labeling, subsequent work-up, Iverson sequencing reaction, affinity cleavage, MPE, and DNAse I footprinting have been described elsewhere.[24] Gels were stored using phosphor

radiography techniques, data analyzed by densitometry, and binding isotherms determined by fitting to a modified Hill equation for zero-order polyamide and first-order DNA concentrations.[24]

Chapter 2B

Hairpin Polyamides Antagonize LSF Binding *in vitro*, Occupy the RCS *in vivo*, and Elicit Viral Outgrowth in Resting $CD4^+$ Cells

The experiments described here were conducted at UT Southwestern Medical Center under the direction of Professor David Margolis

Polyamides block binding of LSF to the RCS *in vitro*

Electrophoresis mobility shift assays (EMSA) demonstrate that 25-75 nM polyamides **1** and **2** each specifically inhibit LSF binding to the RCS (-10 to +27) *in vitro*, while mismatched control polyamides **M1** and **M2** do not (Figure 2.19A). Polyamide **3** (Figure 2.19B) and **4** targeted to the 5' end of the RCS potently inhibit LSF binding to RCS-S (+4 to +35). Polyamides **1** or **2** in combination with either polyamides **3** or **4** potently and specifically inhibit LSF binding to RCS-L (-10 to +34) probe (Figure 2.19C). Polyamides **1** and **2** were not tested in combination as they target partially overlapping sequences. 0.5 μM of match polyamides were required to inhibit complex formation at the RCS when EMSA assays were performed using nuclear extract (not shown). Nonspecific multiprotein complexes in the extract may bind the oligo and be less affected by polyamide binding in the minor groove.

Polyamides block binding of LSF to the RCS *in vivo*

The ability of polyamides to specifically alter promoter occupancy *in vivo* was investigated using chromatin immunoprecipitation (ChIP) assays.[25] These were performed using a HeLa cell line containing a single chromosomally integrated copy of the LTR linked to a CAT reporter gene.[26] To validate the ability to measure changes in occupancy in the vicinity of the RCS by ChIP assay, acetylated histone H4 near the LTR

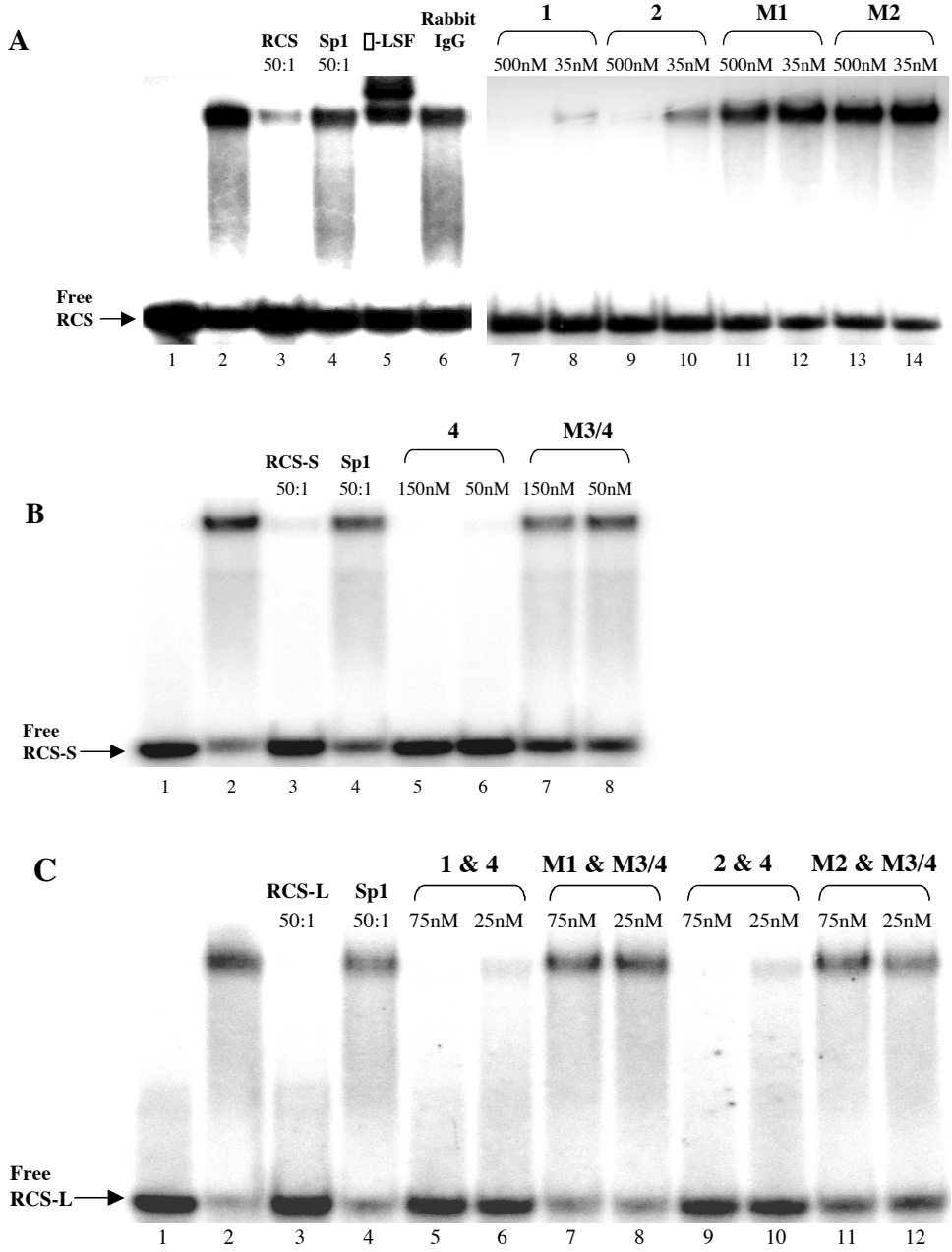


Figure 2.19 A. EMSA of RCS (-10/+27) alone (lane1), and retarded by his-LSF (lane 2). Lanes 3-6 show competition by unlabelled RCS or Sp1 consensus binding sequence oligonucleotides (50:1 ratio of competitor to probe), and supershift induced by the addition of antibody to LSF or non-specific IgG. Lanes 7-14 show competition by polyamides **1**, **2**, **M1**, and **M2** (concentration of polyamide is shown). B. EMSA of RCS-S (+4/+35) alone (lane1), and retarded by his-LSF (lane 2), and competition by RCS or Sp1 (lanes 3 and 4), polyamide **4** (lanes 5-6), or **M3/4** (lanes 7-8). C. EMSA of RCS-L (-10/+34) alone (lane1), and retarded by his-LSF (lane 2), and competition by RCS or Sp1 (lanes 3 and 4), polyamides **1** and **3** (lanes 5-6), **M1** and **M3/4** (lanes 7-8), **2** and **4** (lanes 9-10), or **M2** and **M3/4** (lanes 11-12).

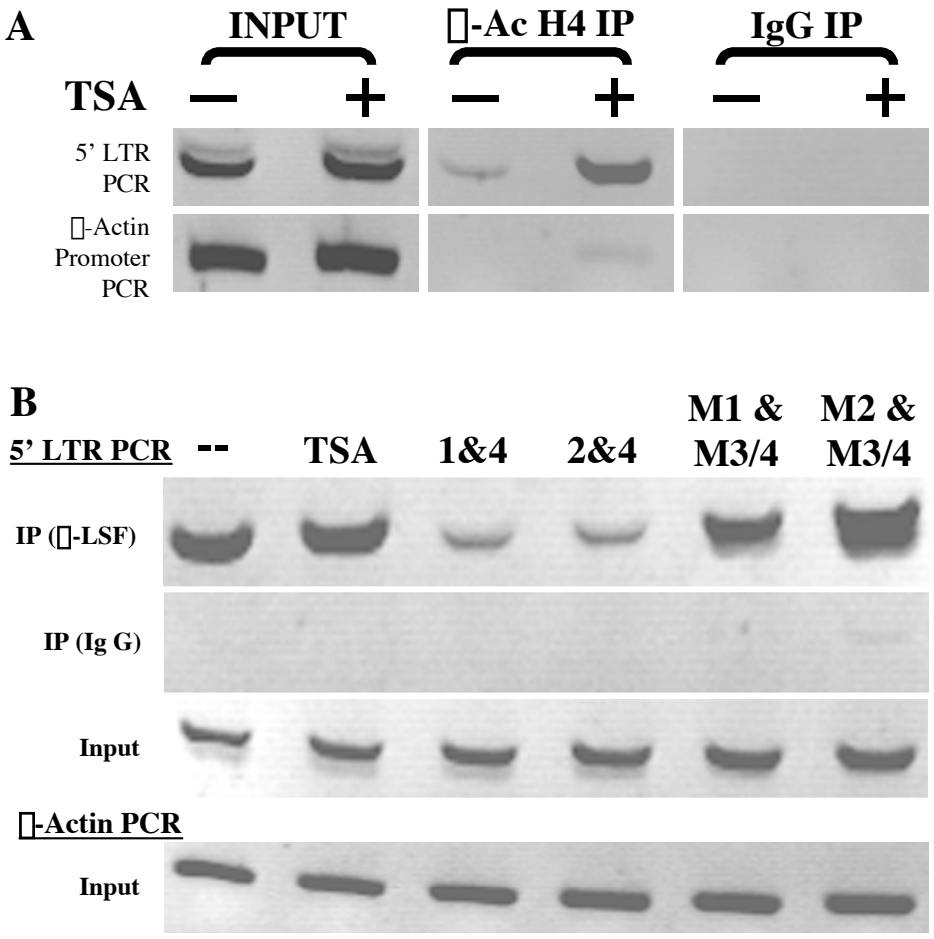


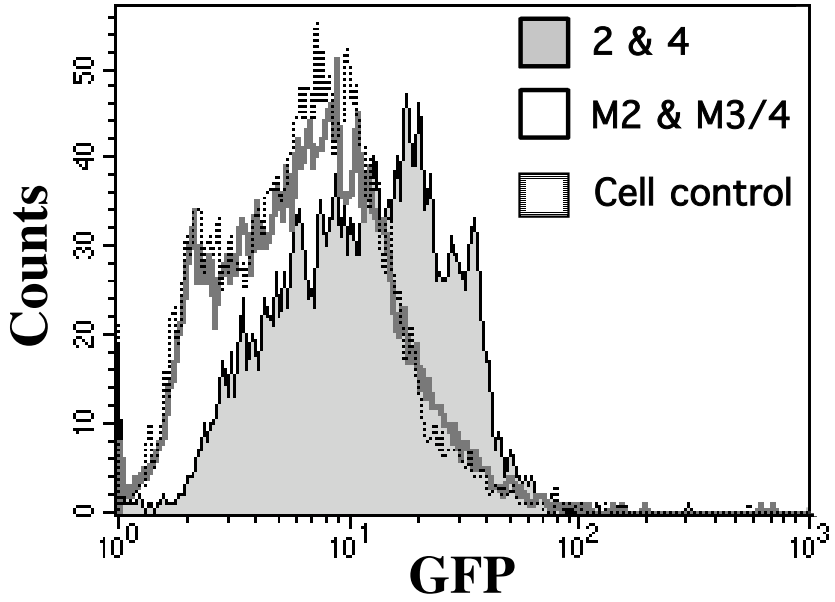
Figure 2.20 Chromatin immunoprecipitation (ChIP) assays show an increase of acetylated histone H4 near Nuc-1 of the LTR after treatment with TSA, and a decrease in LSF at this region of the LTR after exposure to RCS-binding polyamides. A. PCR amplification of the 5' LTR or -actin promoter performed using cell extracts (-) and cells exposed to TSA (+). As indicated, products of ChIP and PCR amplification of the 5' LTR or -actin promoter using anti-acetylated histone H4 or mock IP with rabbit IgG, and control PCR amplification of the 5' LTR or -actin promoter using extract prior to IP. Results are representative of three independent experiments B. ChIP using anti-LSF. PCR products of extracts of untreated cells, cells exposed to TSA, polyamides **1** and **4**, **2** and **4**, **M1** and **M3/4**, or **M2** and **M3/4**. As indicated, products of ChIP and PCR amplification of the 5' LTR using anti-LSF or mock IP with rabbit IgG, and control PCR amplification of the 5' LTR or -actin promoter using extract prior to IP. PCR amplification of the -actin promoter using extract following immunoprecipitation with anti-LSF yielded no product (not shown). These two experiments are representative of four independent experiments.

Nuc-1 DNAse cleavage site was measured before and after exposure of cells to the histone deacetylase inhibitor trichostatin (TSA). Consistent with previous DNAse protection studies,[6, 12] ChIP repeatedly demonstrated that histone acetylation near Nuc1 and the RCS input increased following treatment with TSA (Figure 2.20A, center panel. "Input" is quantitative PCR product for control of gel loading, "IgG IP" is an antibody control). A greater than 4-fold increase in acetylated histone H4 near the LTR Nuc-1 site, as measured by titration of DNA input, was induced by TSA treatment. The β-actin promoter, an input control on amount loaded per well, was less affected by the global effects of TSA, showing changes of 2-fold or less (Figure 2.20A, center panel).

To document that polyamides act directly to antagonize LSF binding to the LTR RCS site in the natural context of host chromatin, we then performed ChIP assays using an α-LSF antibody. Cells were either treated with 400 nM TSA, or exposed to 2 μM polyamides **1** and **4**, **2** and **4**, **M1** and **M3/4**, or **M2** and **M3/4**, and ChIP assays performed (Figure 2.20B). The β-actin PCR product was equivalent before ChIP, but, presumably due to the lack of LSF sites at the β-actin promoter, no product was detected after ChIP with α-LSF.

As expected, TSA treatment resulted in no change in ChIP with α-LSF (as compared with treatment with α-AcH4IP, Figures 2.20A and 2.20B). However polyamides **1** and **4** or **2** and **4** reproducibly displaced LSF as evidenced by significant decrease in ChIP (Figure 2.20B). A modest, nonspecific increase in ChIP was sometimes observed when cells are treated with mismatch control polyamides. This modest, nonspecific increase was also observed using polyamides targeted to the Ets-1 site of the

LTR,[27] a factor that binds in an upstream region of the LTR. Nevertheless, the potent ability of specific RCS-binding polyamides to inhibit LSF binding *in vivo* results in an overall decrease in ChIP.



	Polyamides 1 and 4		Polyamides 2 and 4	
	2.0 μM	**5.0 μM**	**2.0 μM**	**5.0 μM**
% Change log mean fluorescence vs. controls	**14.2 ± 5.8**	**41.9 ± 9.6**	**22.1 ± 7.5**	**76.7 ± 9.4**
	N=10	N=5	N=10	N=5
	P< 0.01	P< 0.01	P< 0.01	P< 0.01

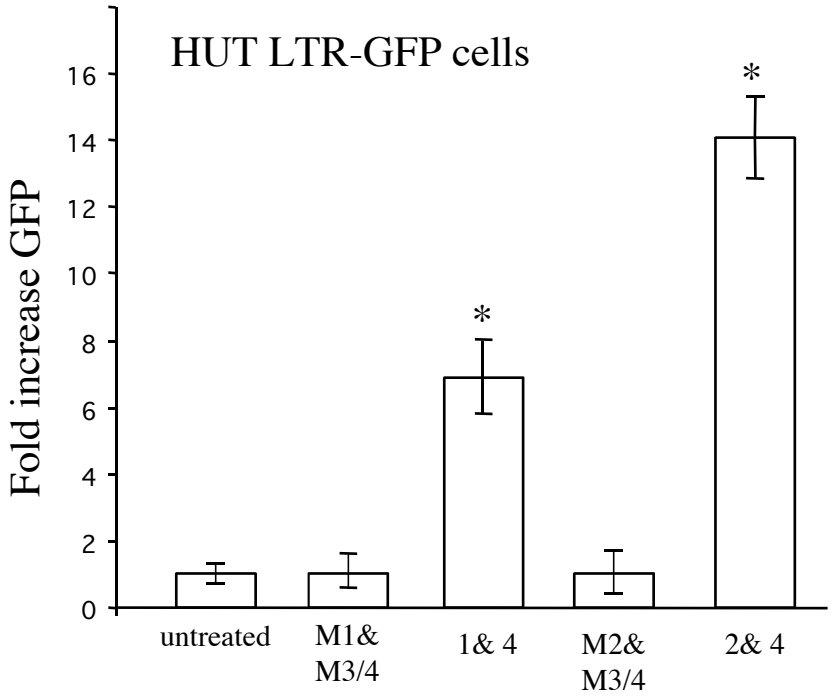
Figure 2.21 RCS-binding polyamides induce an increase in expression of an integrated LTR-GFP reporter gene in HUT78 cells in the absence of Tat, as shown by an increase on a log scale. Log change in mean fluorescence was analyzed using the unpaired two-tailed Student's T-test.

LTR expression is increased by polyamides targeted to the RCS

As T-cells are the primary targets of HIV infection, CEM-LTR-GFP and HUT78-LTR-GFP (provided by J.V. Garcia) cells, which each contain integrated HIV-1 LTR promoters linked to a green fluorescence protein (GFP) reporter gene, were used to

directly measure the effect of RCS-binding polyamides on the LTR within the context of host chromatin. As these cell lines lack Tat, only basal LTR expression occurs and elongation of LTR transcripts is inefficient.[28] Though polyamide induced derepression may not achieve the levels of LTR expression comparable to those following Tat activation, an evaluation of the effect of RCS-binding polyamides on basal LTR expression can still be made

Cells were exposed to RCS-binding polyamides or mismatch control polyamides. GFP expression was measured in a log scale, and mean log fluorescence intensity (MLF) was calculated. MLF of cells was not increased significantly (<5%) by mismatch control polyamides, but was increased by up to 50% in CEM cells (not shown) and 76.6% in HUT78 cells by RCS-binding polyamides (Figure 2.21). Expressed in different fashion, addition of **1** and **4** evinced a 7.1-fold in GFP fluorescence in HUT78-LTR-GFP cells while addition of **2** and



Graph 2.1 Graphical representation of GFP expression increase, inferred from fluorescence increase, in presence of RCS binding polyamides **1** & **4** or **2** & **4** as compared to untreated cells or cells treated with mismatch controls.

4 evinced a larger 14.2-fold increase, both values relative to GFP fluorescence in untreated cells (Graph 2.1).

Replication of latent HIV within resting CD4$^+$ T-cells from HIV$^+$ donors is induced upon exposure to RCS-binding polyamides

The *in vitro* and tissue culture model system findings presented above suggest that polyamides targeted to the RCS may disrupt the LSF_2/YY1 repressor complex, and upregulate LTR expression. However there is no suitable model system in which to study quiescent HIV infection within resting CD4$^+$ T-cells. Cultures of primary CD4$^+$ resting T-cells taken from HIV$^+$ individuals have been used to document and quantitate the reservoir of latent, replication competent HIV.[2,3,5,29-32] HIV replication induced following the activation of these cells with polyhemaglutinin (PHA) or combinations of cytokines has demonstrated that T-cell activation results in the production of virus from this reservoir.[2,3,31,32] These studies do not, however, address what transcription factors are involved in repression of HIV-1 expression prior to stimulation. By using the polyamides developed here and shown by EMSA assays *in vitro*, ChIP analysis of LTR occupation *in vivo*, and observation of derepressed LTR driven GFP production *in vivo* to antagonize the LSF_2/YY1•DNA complex, the observation of LTR expression in quiescent CD4$^+$ cells would support the role of the cellular transcription factor complex LSF_2/YY1 in maintaining latency in HIV-1 infection. We therefore tested the ability of polyamides targeted to the RCS to induce expression of HIV from purified resting CD4$^+$ T-cells obtained from HIV$^+$ donors.

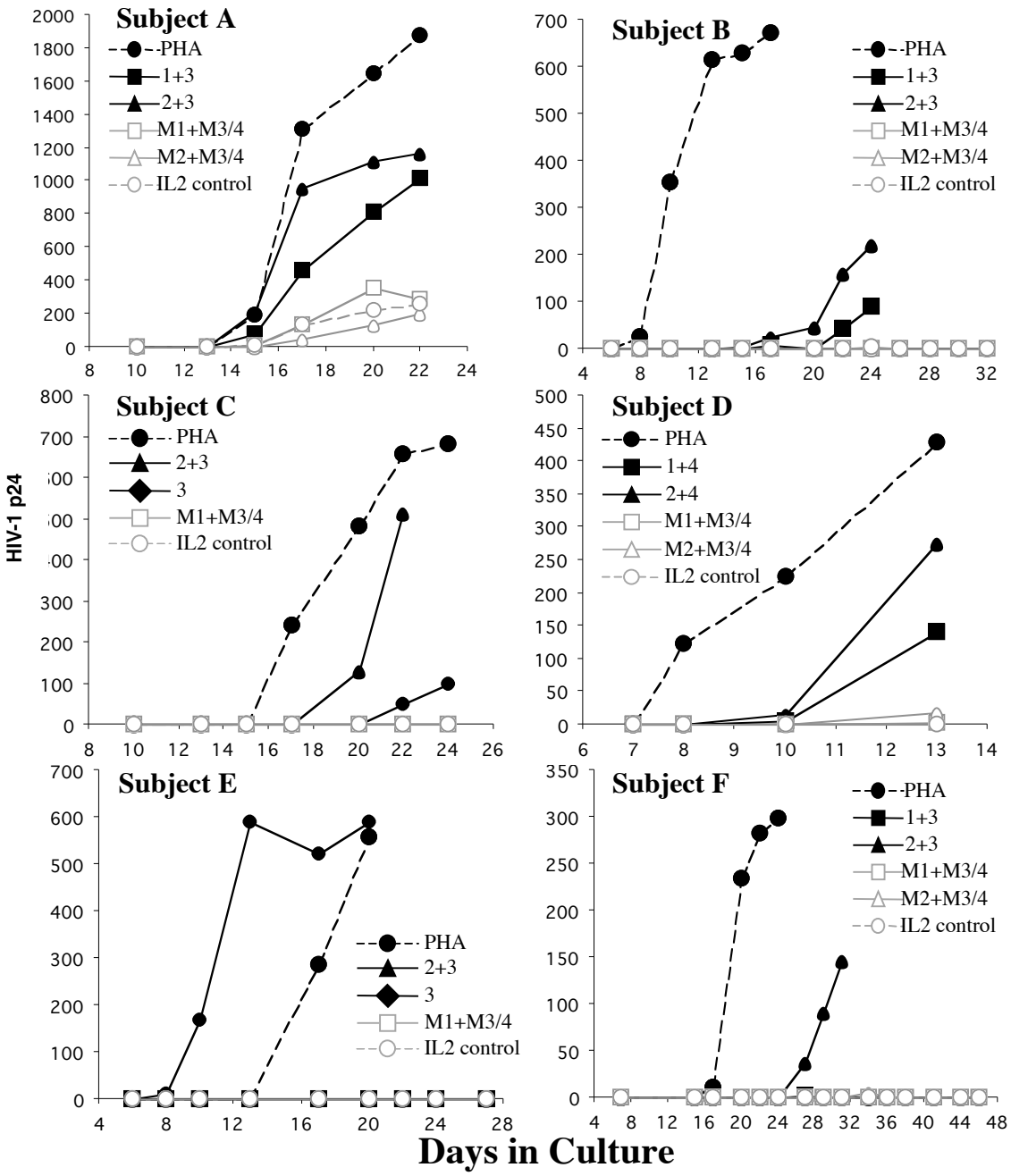


Figure 2.22 Replication-competent HIV is recovered from the resting cells of 22 HIV$^+$ donors with similar frequency when cultures are exposed to PHA or polyamides **1**-**4**, and infrequently when unstimulated or treated with mismatched polyamides. The overall frequency of detection of replication-competent HIV in cultures treated with PHA or polyamides **1**-**4** was significantly greater than the frequency of HIV recovery in control cultures ($p < 0.001$ by two-tailed, unpaired T-test).

Blood was obtained (at UT Southwestern Medical Center) by standard phlebotomy from HIV^+ volunteers, and resting $CD4^+$ T-cells isolated by negative antibody selection. We selected subjects with $CD4^+$ cell counts > 300 cells/μl so that adequate numbers of resting $CD4^+$ cells could be obtained. HIV was grown from resting $CD4^+$ cells as described.[30] Only 4 to 12 million resting T-cells could be obtained by phlebotomy from each patient. This resulted in 1 to 2 million resting T-cells cultured per well. In addition to rare, integrated, replication competent provirus, such cells may also contain an unstable pool of replication-competent proviral DNA in pre-integration complexes.[33,34]

To limit the contribution of this labile pool to viral outgrowth detected in this culture system, cells were incubated for 7 to 10 days without IL2 prior to activation or exposure to polyamides. As latently infected $CD4^+$ cells may be as rare as 0.5 - 10 per million T-cells, and in some cases less, in these studies it is possible that some wells with only 1 to 2 million resting T-cells may not have received infected cells at all.

Results representative of outgrowth assays of HIV from cultures of resting primary $CD4^+$ cells from HIV^+ donors are displayed (Figure 2.22, Table 2.1). We recovered HIV from cells exposed to one or two RCS-binding polyamides at least as often as cells activated with PHA. HIV was infrequently recovered from cells exposed to one or two mismatched control polyamides, at a frequency similar to or less than cells treated with IL2 alone. Overall, HIV was recovered significantly more frequently (one-

Subject	CD4 cells per µl	Months < 50 copies HIV RNA/ml	Anti-retroviral therapy†	Recovery of replication-competent HIV with PHA/IL2	Recovery of replication-competent HIV with binding polyamides/IL2	Recovery of replication-competent HIV with nonbinding polyamides/IL2 or IL2 alone	Integrated HIV copies per million CD4 cells by Alu-PCR (C.I. 95%)
1	858	6	2N, NN	+	+	-	830 ± 290
2	782	8	2N, NN	+	+	-	41 ± 13
3	671	12	2N, bPI	+	+	-	43 ± 12
4	584	>23	2N, bPI	+	+	-	42±17
5	436	6*	2N, bPI	+	+	-	270 ± 66
6	723	15	4 N	-	-	-	50 ± 16
7	758	25	2N, bPI	+	-	-	< 5
8	643	28	2N, NN	+	+	-	130 ± 36

Table 2.1

† Antiretroviral therapy denoted by number of nucleoside/tide RT inhibitors (N), non-nucleoside RT inhibitors (NN), and ritonavir-boosted protease inhibitor (bPI)

*Single viral load of 64 copies/ml 3 months prior to leukopheresis during self-limited viral illness

tailed Fisher's exact test) when cultures were stimulated with PHA or exposed to polyamides **1**-**4**, than when unstimulated or exposed to polyamides **M1**, **M2**, or **M3/4** (Figure 2.22, Table 2.1). Consistent with previous studies,[1,29] HIV outgrowth was less frequent in patients with plasma HIV RNA < 50 copies/ml, but was recovered in cultures activated by PHA or exposed to RCS binding polyamides. Viral replication was significantly detected within 15 days of stimulation or exposure to polyamides, but for 4 cases, it was detected at 17 to 26 days. Cell cultures were usually discarded if HIV was not recovered within 30 days; in some cases cultures were carried for up to 49 days to insure that control cultures remained negative. Due to limited cell numbers, every condition could not be tested in every subject. HIV outgrowth from cultures treated with mismatch polyamides or IL2 alone may have occurred due to non-specific activation of chromosomal provirus by allogeneic feeder cells or tissue culture antigens, survival of residual replication-competent cytoplasmic provirus, or incomplete removal of small numbers of activated, infected cells.

HIV was recovered less frequently from cells exposed to 0.1 or 0.5 μM of RCS-binding polyamides than from parallel cultures exposed to 2.5 μM of RCS-binding polyamides (not shown). Cells exposed to pairs of RCS-binding polyamides produced virus more rapidly than parallel cultures exposed to single RCS-binding polyamides.

Polyamides had no measurable effect on cultured resting T-cell homeostasis. No morphological effect could be observed in cultures following exposure to RCS-binding or mismatched polyamides; no lymphoblasts were observed until the addition of activated feeder cells. In contrast to PHA stimulation, following 48 hrs of exposure to RCS-binding polyamides (the duration of exposure used for outgrowth experiments) no

significant changes in expression of cell surface activation markers CD38, CD25, HLA-DR, CD69 or the nuclear proliferation antigen Ki67 could be detected in resting cells selected from two HIV^+ donors and one seronegative control (not shown).

As a preliminary descriptive analysis, we measured the frequency of integrated HIV genomes in a number of our patient samples by quantitative nested Alu-PCR (Table 2.1). As reported,[1] results ranged from $<10^1$ to 10^3 integrated genomes per 10^6 resting $CD4^+$ cells. Also as reported, the results of this assay correlated with the frequency of viral detection following stimulation with PHA or exposure to RCS-binding polyamides: virus outgrowth was detected infrequently in samples with fewer integrated genomes.

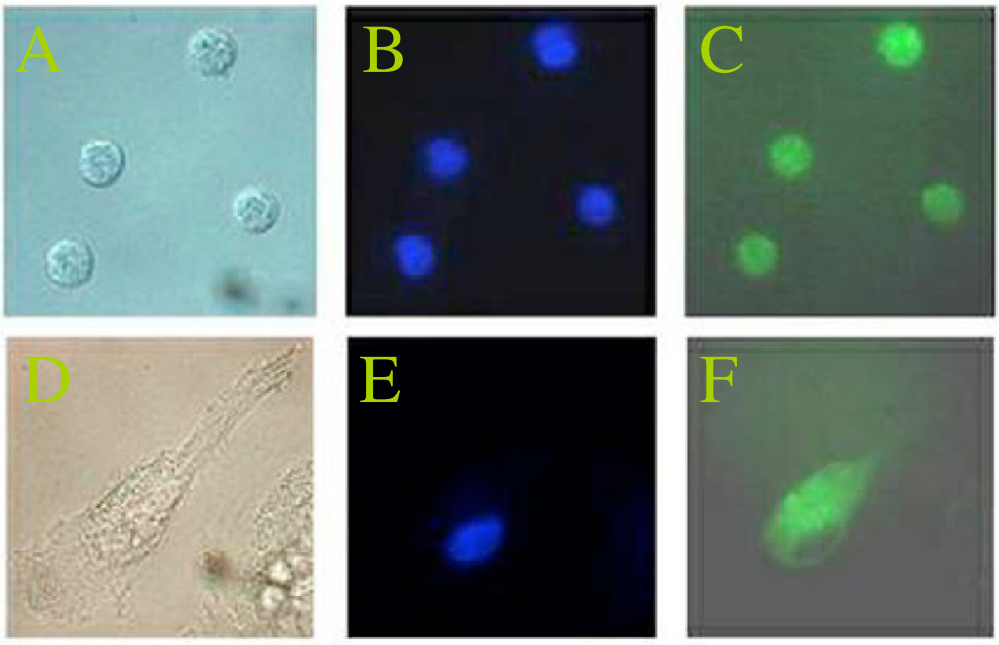


Figure 2.23 Polyamides enter the nucleus of primary resting CD4+ T cells: Resting CD4+ cells seen in A) visible light, B) Hoechst nuclear dye staining, C) bodipy-labeled polyamide fluorescence. Primary macrophages seen in D) visible light, E) Hoechst nuclear dye staining, F) bodipy-labeled polyamide fluorescence. 100-X magnification.

Polyamides enter the nucleus of resting primary CD4$^+$ T-cells

To demonstrate that polyamides access the nucleus of primary resting T lymphocytes, blood was obtained from HIV$^+$ donors and resting CD4$^+$ T-cells isolated by negative selection as described.[35] Cells were incubated in 20 U/ml IL2 or IL2 and PHA for 4 to 18 hrs in the presence of 5 μM bodipy-labeled polyamides.[43] To delineate nuclear morphology, 5 μg/ml Hoechst 33342 (Molecular Probes, OR) was added to cells 30 min before microscopy. Bodipy-labeled polyamides access the nucleus of primary resting CD4$^+$ T-cells (Figure 2.23 A-C). Activation of these cells with PHA did not significantly alter nuclear bodipy staining (not shown). In contrast, polyamides enter the cytoplasm of primary macrophages but enter the nucleus of these cells poorly (Figure 2.23 D-F). This possible exclusion of □-Dp polyamides from the nucleus of non-T-lymphocytes may confer drug selectivity towards T-cell types.[44]

Conclusions

RCS-binding polyamides **1** - **4** targeted to the HIV LTR can inhibit LSF binding to the RCS *in vitro* (Figure 2.19), access nuclear DNA *in vivo* as shown by ChIP (Figure 2.20), and increase the basal level of expression of the LTR as observed by GFP expression *in vivo* (Figure 2.21, Graph 2.1). Direct observation of nuclear localization in micrographs of T-cells treated with bodipy labeled polyamides suggests polyamides traffic to the nucleus where they may sequence specifically bind the RCS of the LTR (Figure 2.23).

RCS-binding polyamides upregulate replication of quiescent HIV within primary resting CD4$^+$ cells harvested from HIV$^+$ donors. As we could not obtain sufficient

numbers of resting cells from individual donors to perform limiting dilution assays, we recruited a large number of donors and regard our results in a binary manner: the presence or absence of viral outgrowth. The differences observed in outgrowth kinetics (Figure 2.22) within individual patient samples could be caused by the outgrowth of different viral species, but are more likely due to differential response to PHA or polyamides. Further experiments are required to address this question.

Our findings suggest that YY1, LSF, and HDAC1 play an ongoing, active role in the maintenance of viral persistence within the resting $CD4^+$ T-cells. In latently infected primary cells capable of producing Tat, viral outgrowth is reproducibly seen at a frequency comparable to that seen following PHA treatment. It is unlikely that HIV outgrowth is the secondary result of upregulation of cellular genes by specific RCS-binding polyamides, as four distinct polyamides targeted to three different sequences within the RCS specifically activate LTR expression, but control mismatched polyamides did not. Further, exposure of cells to polyamides has no effect on cell morphology or surface marker expression.

We are now in a position to modify the model earlier presented in Figure 2.2 (Figure 2.24). Occupancy of the repressor complex sequence of the LTR clearly exists in a state of equilibrium, and polyamides appear capable of entering into that equilibrium to disfavor LSF_2/YY1 complex formation. In the purely biological context, LSF recruits YY1, YY1 then recruits HDAC-1, and transcriptional repression is mediated, at least in part, via deacetylation of local nucleosomes. RCS-binding polyamides antagonize this process, allowing LTR expression in the presence of constitutive host activators, and potentially full activation when Tat is produced. In addition to maintaining LTR

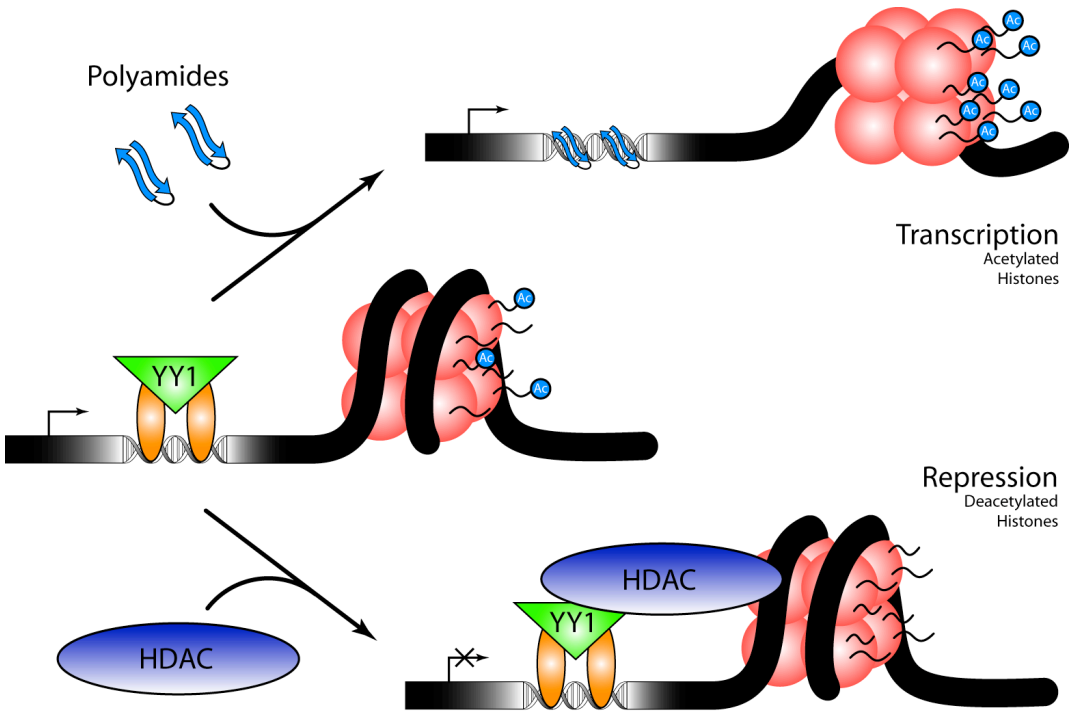


Figure 2.24 Dynamic model for polyamide intervention in LSF_2/YY1 induction of latency. (Thanks to EJF for figure.)

quiescence once it is established, host repressors may act to downregulate the LTR as an infected, activated cell returns to the resting state.[4,6]

Beyond furthering our understanding of how host cellular transcription factors downregulate LTR expression, the most powerful aspect of these findings is that targeting the conserved DNA binding sites of cellular transcription factors may circumvent mutations in HIV-encoded gene products that are often problems for developing anti-HIV therapeutics. As a potential therapeutic, problems caused by DNA binding polyamides perturbing other biological pathways involving LSF[45] and YY1[46-48] may cast the medicinal use of polyamides in a an unattractive light, but recent evidence in our group suggests (and our micrographs confirm) that polyamides containing □-Dp successfully traffic to the nucleus of T-cells but are excluded by non-T-lymphocytes.[43,44]

Thus polyamide action on non-T-lymphocyte health, e.g., through downregulation of the LSF regulated thymidylate synthase gene,[45] may be minimal.

Of note, the RCS is conserved within HIV-1 consensus sequences. This may be due to sequence constraints imposed by the TAR secondary structure (TAR is the early RNA product of LTR transcription, TAT acts upon the stem-loop secondary structure formed by TAR to upregulate LTR transcription, the sequence of TAR overlaps with that of the RCS).[37,38] TAR must be conserved to allow Tat binding, and perhaps, as we suggest here, conservation of the TAR sequence allows the binding of conserved cellular transcription factors to the TAR-encoding DNA allowing for induction of latency.

As activation of resting T-cells can result in the emergence of HIV from the unactivated T-cell reservoir, simultaneous immune activation and intensive antiretroviral therapy have been attempted.[39] Subsequent studies suggest that global T-cell activation may induce viral replication to a level that cannot be contained by antiretroviral therapy.[40] Specific derepression of *only* repressed HIV LTR in infected, resting $CD4^+$ T-cells may offer an alternative approach to global T-cell activation.

Therefore, in combination with immunotherapeutics and/or antiretroviral chemotherapy, interventions that derepress the LTR may improve control of HIV infection through immunity mediated clearance. Future pharmacological or immunomodulatory therapy potent enough to terminate viral replication, e.g., HAART, in combination with interventions that result in derepression of the quiescent LTR, e.g., polyamides **1** – **4**, may allow for clearance of HIV infection.

Experimental

Polyamide synthesis and electrophoretic mobility shift assays

Polyamides were synthesized as described in Chapter 2A. Histidine-tagged LSF (His-LSF) was prepared as previously described.[41] Oligonucleotides probes were labeled and incubated in EMSA buffer as described[9] with polyamides, or unlabeled RCS or Sp1 (Stratagene, San Diego, CA) oligonucleotides at room temperature for 45 min. 60 ng of His-LSF was added and incubated for 15 min. Supershifts were performed by the addition of rabbit α-CP2 antibody[42] or control rabbit IgG (Sigma, St Louis, MO). EMSA was performed as described.[9]

Chromatin immunoprecipitation assay (ChIP)

$5x10^5$ HeLa CD4-LTR-CAT cells[26] were seeded in DMEM, 0.5% FBS for 16 hr. Selected cultures were treated with TSA (400 nM) or polyamides (2 μM) for 4 hr in DMEM/10% FBS. Cells were cross-linked in 1% paraformaldehyde, washed, and lysed in 100μl lysis buffer (Upstate Biotechnology, Lake Placid, NY) supplemented with 5μl protease inhibitor cocktail (Sigma, St Louis, MO). DNA was sheared by sonication, centrifuged, and chromatin fragmentation (300-1000bp) confirmed. Either α-acetyl-histone H4 (Upstate Biotechnology), α-CP2 (LSF antibody[41]), α-HDAC1[7] or rabbit serum (Sigma) was incubated with 50 μg of DNA for 16 hr at 4 ºC. Immunoprecipitates were recovered with salmon sperm DNA/protein-A agarose beads at 4 ºC for 1 hr and washed. Cross-linking of DNA was reversed by PCR grade proteinase K (Boehringer Mannheim) at 56 ºC for 1 hr in 1% SDS, 0.1 M $NaHCO_3$ and DNA precipitated. 28-32 cycles of semiquantitative duplex PCR were performed (94 °C, 30 s; 55 °C, 20 s; 72 °C

15 s), using HIV-1 LTR promoter primers LTR-108F (5'-TACAAGGGACTTTCCGCT GG-3') and LTR+80 (5'-AGCTTTATTGAGGCTTAAGC-3') and internal control primers P-β-Actin-F (5'-TGCACTGTGCGGCGAAGC-3') and P-β-Actin-R (5'-TCGAGCCATAAAAGGCAA-3'). Serial two fold dilutions of input were subjected to PCR to confirm equivalence and linear amplification in each experiment. Gel products quantified by densitometry.

Analysis of GFP expression by PCR and flow cytometry

HUT 78 LTR-GFP (gift of J.V. Garcia) or CEM LTR-GFP[35] cells were held in RPMI/0.5% FBS for 16 hr. Cells were then refed with RPMI/10% FBS with or without 2 μM polyamides. GFP expression in the live cell population was determined by flow cytometry at 48 hr. Mean log fluorescence of 10^4 live cells for each condition was recorded. Data were subjected to analysis by Student's paired T-test. At 4 hr total RNA was isolated from $4x10^5$ cells using TRIZOL™ (Gibco BRL, Grand Island, NY) according to manufacturer's instruction. cDNA was primed from DNase I treated total RNA (5 μg) with 125 ng random hexamers (Promega, Madison, WI) using Omniscript RT (Qiagen, Valencia, CA) as per the manufacturer. $1/20^{th}$ of the cDNA was PCR amplified (PCR High Fidelity supermix, Gibco BRL): 2 min 94 ºC and 20 cycles of 20s 94 ºC, 15s 57 ºC and 20s 72 ºC followed by 7 min 72 ºC using primers sets, GFP-291F (5'-CACCATCTTCTTCAAGGACG-3') and GFP-610R (5'-TGCTCAGGTAGTGGTTGT CG-3'). β-actin was co-amplified to insure equal input of mRNA in PCR reactions. 1 μl was subjected to PCR (25 cycles) under identical conditions with primers GFP-308N (5'-ACGACGGCAACTACAAGACC-3') and GFP-

484N (5'-TGCCGTTCTTCTGCTT GTCG-3'). Products separated by PAGE and visualized by ethidium bromide staining.

Polyamide induction of HIV from resting CD4$^+$ cells

HIV-infected volunteers were enrolled following IRB-approved informed consent. PBMCs were isolated by Ficoll density centrifugation. Highly purified preparations of resting CD4$^+$ cells were obtained by negative selection[30] using a CD4$^+$ selection cocktail (Stem Cell Technologies, Vancouver, Canada) with α-CD25, α-HLA-DR and α-CD41. FACS analysis (α-CD4, α-CD3, α-CD69, α-HLA-DR, α-CD25, Becton Dickinson, San Jose, CA) was used to determine cell surface phenotypic markers of each preparation. 95% or greater purity of CD4$^+$ cells was routinely observed with < 0.5% exhibiting activation markers. Cells were maintained for 7 days in RPMI/10% FBS (Hyclone laboratories, Logan, UT) before stimulation of 1-2x10^6 resting cells with 4 μg/ml PHA-L (Boehringer Mannheim), 2x10^6 allogeneic irradiated PBMCs, and 100units ml^{-1} IL2 or 2μM of indicated polyamide and IL-2, or IL2 alone. 48 hrs later cultures were expanded with the addition of 10^6 CD8$^+$-depleted PBMCs. Cultures were fed (RPMI/10% FBS/20 U/ml IL2) and supernatants harvested every 2-3 days. Feeder cells were added every 7 days; virus was detected by p24 antigen capture ELISA (AIDS Research and Reference Reagent Program, NIH).

References

1. Chun, T. W. et al. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 13193-7.
2. Finzi, D. et al. (1997) *Science* **278**, 1295-1300.
3. Wong, J. K. et al. (1997) *Science* **278**, 1291-5.
4. Pierson, T., McArthur, J. and Siliciano, R. F. (2000) *Annu. Rev. Immunol.* **18**, 665-708.
5. Finzi, D. et al. (1999) *Nat. Med.* **5**, 512-7.
6. El Kharroubi, A., Piras, G., Zensen, R. and Martin, M. A. (1998) *Mol. Cell. Biol.* **18**, 2535-44.
7. Coull, J. J. et al. (2000) *J. Virol.* **76**, 6790-9.
8. Margolis, D. M., Somasundaran, M. and Green, M. R. (1994) *J. Virol.* **68**, 905-10.
9. Romerio, F., Gabriel, M. N. and Margolis, D. M. (1997) *J. Virol.* **71**, 9375-82.
10. El Kharroubi, A. and Verdin, E. (1994) *J. Biol. Chem.* **269**, 19916-24.
11. Sheridan, P. L., Mayall, T. P., Verdin, E. and Jones, K. A. (1997) *Genes Dev.* **11**, 3327-40.
12. Verdin, E., Paras, P., Jr. and Van Lint, C. (1993) *EMBO J.* **12**, 3249-59.
13. Nguyen, D. *Caltech Thesis* (2002).
14. Dickinson, L. A. et al. (1998) *Proc. Natl. Acad. Sc.i USA* **95**, 12890-95.
15. Dickinson, L. A., Trauger, J. W., Baird, E. E., Ghazal, P., Dervan, P. B., Gottesfeld, J.M. (1999) *Biochemistry* **38**, 10801-7.
16. Kelly, J. J., Baird, E. E., Dervan, P. B. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 6981-6985.

17. Turner, J. M., Baird, E. E., Dervan, P. B. (1997) *J. Am. Chem. Soc.* **119**, 7636-7644.

18. Coull, J. J. et al. (2002) *J. Virol.* **76**, 12349-12354.

19. White, S., Baird, E. E., Dervan, P. B. (1997) *J. Am. Chem. Soc.* **119**, 8756-8765.

20. Belitsky, J. M., Nguyen, D. H., Wurtz, N. R., Dervan, P. B. (2002). *Bioorg. & Med. Chem.* **10**, 2767-2774.

21. Urbach, A. R. and Dervan, P. B. (2001) *Proc. Natl. Acad. Sci. USA* **98**, 4343-4348.

22. Dervan, P. B. (2001). *Bioorg. & Med. Chem.* **9**, 2215-2235.

23. Baird, E. E. and Dervan, P. B. (1996) *J. Am. Chem. Soc.* **118**, 6141-6146.

24. Trauger, J. W. and Dervan, P. B. (2001). *Methods Enz.* **340**, 450-466.

25. Orlando, V., Strutt, H. and Paro, R. (1997) *Methods* **11**, 205-14.

26. Ciminale, V., Felber, B. K., Campbell, M. and Pavlakis, G. N. (1990) *AIDS Res. Hum. Retroviruses* **6**, 1281-7.

27. Dickinson, L. A. et al. (1999) *J. Biol. Chem.* **274**, 12765-73.

28. Nahreini, P. and Mathews, M. B. (1995) *J. Virol.* **69**, 1296-301.

29. Chun T. W., et al. (1999) *Nature Med.* **5**, 651-5.

30. Chun, T. W. et al. (1995) *Nature Med.* **1**, 1284-9027.

31. Chun, T. W. et al. (1997) *Nature* **387**, 183-8.

32. Chun, T. W., Engel, D., Mizell, S. B., Ehler, L. A. and Fauci, A. S. (1998) *J. Exp. Med.* **188**, 83-91.

33. Stevenson, M., Stanwick, T. L., Dempsey, M. P., Lamonica, C. A. (1990) *EMBO J.* **9**, 551-60.

34. Zack, J. A., Arrigo, S. J., Weitsman, S. R., Go, A. S., Haislip, A., Chen, I. S. (1990) *Cell* **61**, 213-222.

35. Gervaix, A. et al. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 4653.

36. Ramratnam, B. et al. (2000) *Nat. Med.* **6**, 82-5.

37. Pereira, L. A., Bentley, K., Peeters, A., Churchill, M. J. and Deacon, N. J. (2000) *Nucleic Acids Res.* **28**, 663-8.

38. Garcia, J. A. and Gaynor, R. B. (1994) *Prog. Nucleic Acid Res. Mol. Biol.* **49**, 157-96.

39. Prins, J. M. et al. (1999) *AIDS* **13**, 2405-10.

40. Fraser, C. et al. (2000) *AIDS* **14**, 659-69.

41. Volker, J. L., Rameh, L. E., Zhu, Q., DeCaprio, J. and Hansen, U. (1997) *Genes Dev.* **11**, 1435-46.

42. Swendeman, S. L. et al. (1994) *J. Biol. Chem.* **269**, 11663-71.

43. Belitsky, J. M., Leslie, S. J., Arora, P. S., Beerman, T. A., Dervan, P. B. (2002) *Bioorg. & Med. Chem.* **10**, 3313-18.

44. Best, T. B., Edelson, B. E., Dervan, P. B. *unpublished results*

45. Powell, C. M. H., Rudge, T. L., Zhu, Q., Johnson, L. F., Hansen, U. (2000) *EMBO J.* **19**, 4665-75.

46. Austen, M., Cerni, C., Lüscher-Firzlaff, J. M., Lüscher, B. (1998) *Oncogene* **17**, 511-20.

47. Oei, S. L. and Shi, Y. (2001) *Biochem. Biophys. Res. Comm.* **284**, 450-54.

48. http://caroll.vjf.cnrs.fr/cancergene/CG1010.html