Cell, Vol. 103, 895–898, November 22, 2000, Copyright ©2000 by Cell Press

Bugs on Drugs Go GAGAA

Minireview

Steven Henikoff* and Danielle Vermaak
Howard Hughes Medical Institute
Fred Hutchinson Cancer Research Center
1100 Fairview Avenue North
Seattle, Washington 98109

Sometimes science fact resembles science fantasy, more suitable for the supermarket checkout display than the research library. A case in point is an article in the November issue of *Molecular Cell* that could pass for the subject of a B-grade sci-fi movie: Some unfortunate creature ingests a compound cooked up in the laboratory and undergoes a profound transformation. Implausible as such a scenario may seem, Laemmli's group reports that when certain *Drosophila* larvae are fed a sequence-specific DNA binding compound they undergo homeotic transformations (Janssen et al., 2000a). These transformations involve targeting of the compound to the minor groove of a GAGAA-rich repetitive element, and evidently mimic a mutation in the gene encoding the GAGA factor (GAF) DNA binding regulatory protein. This astonishing result not only has implications for understanding repetitive sequences, heterochromatin, and DNA binding interactions, but also provides a new class of cytological tools. What's more, the strange fate of bugs on drugs represents another milestone on the way to therapeutic drugs that bind DNA in the minor groove.

Minor-Groove Binding Compounds

This breakthrough by Laemmli's group builds upon extensive studies of synthetic compounds that are modeled on the naturally occurring drugs distamycin and netropsin. These are polyamides containing aromatic pyrrole rings, which fit snugly into the narrow minor groove of A/T-rich DNA (Figure 1, top). Over the past decade, a series of reports, principally from the Dervan laboratory, have succeeded in increasing the sequence specificity and range of minor-groove binding compounds, while still maintaining binding affinities in the nanomolar range (reviewed by Wemmer, 2000). Because polyamide subunits are connected by peptide-like amide bonds, they can be synthesized using modified protocols for conventional solid state peptide synthesis. The success of a polyamide derivative in targeting a given sequence can be assessed by standard footprinting assays, and structures can be determined from NMR spectra of polyamide compounds bound to double-stranded oligonucleotides.

A DNA duplex that consists of only A/T base pairs has a narrow minor groove that can accommodate a single molecule of distamycin (Figure 1). However, distamycin also binds as a side-by-side dimer that widens the groove of A/T DNA. Sequences containing G/C base pairs have a wider minor groove and can therefore more easily accommodate a side-by-side dimer, but the protruding NH_2 group of guanidine represents a major steric obstacle for pyrrole rings bound in the minor groove. Replacement of C-H in pyrrole with N: in imidazole allows for the formation of a hydrogen bond with this guanidine NH_2. In the context of a side-by-side dimer, a G/C base pair is contacted by an imidazole (Im) touching guanine and a pyrrole (Py) touching cytosine. For example, GTA/CAT can be recognized by ImPyPy/PyPyPy. Extending beyond a few base pairs requires spacers, because the polyamides become strained when following the DNA trajectory. A β-alanine (β) is generally used as a spacer, because a β can take the place of a pyrrole, allowing recognition of longer sequences. To obtain dimeric compounds with increased specificity and affinity, a linker connects the two side-by-side monomers to form a hairpin.

The subnanomolar affinities for specific short sequences achieved in these studies encouraged in vivo experiments. Transcription factor binding sites provide attractive targets in cases where minor-groove binding would interfere with function. For example, TFIIIA has 9 zinc fingers, some of which contact the minor groove.

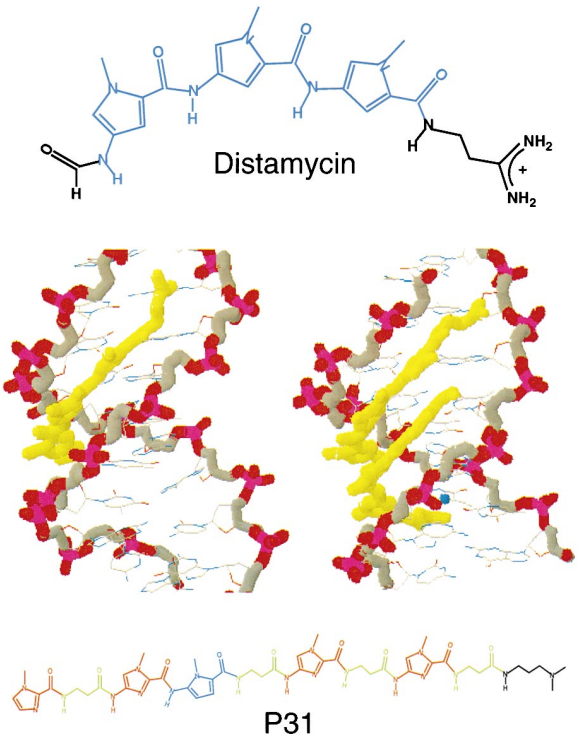

Figure 1. Minor Groove Binding Dyes

Structure of distamycin (top), which targets AT-rich DNA, and P31 (bottom), which targets GAGAAGAGAA. 3D structural representations show binding of a single molecule of distamycin (yellow) in the narrow minor groove (left), and binding of two side-by-side molecules (right), which widen the groove. Subunits of distamycin and P31 are color-coded: pyrroles are blue, imidazoles are red, and β-alanines are green.

*To whom correspondence should be addressed (e-mail: steveh@fhcrc.org).

A polyamide compound targeted to a minor-groove contact site, AGTACT, interfered with targets of TFIIIA transcription in *Xenopus* kidney cells (Gottesfeld et al., 1997), and in subsequent studies, other compounds successfully targeted different transcription binding sites and interfered with in vivo binding. These experiments showed that polyamides are taken up by cells, and can gain access to their cognate sites at concentrations sufficient for interference with transcription to be detected. No general toxic effects were found, raising the hope that these compounds can eventually be used as drugs.

AT Hooks

Some natural peptides also bind in the minor groove. For instance, certain variant histones have tails that contain "SPKK" or similar motifs, which protect the minor groove in footprinting studies (Churchill and Suzuki, 1989). Perhaps the most widespread of these minor-groove binding peptides was termed an "AT hook" by Reeves, who determined that HMG-I(Y) binding to DNA is mediated by a set of three similar 11-12mers that footprinted A/T-rich binding sites and could be competed using a minor-groove binding compound (Reeves and Nissen, 1990). Since then AT hooks have been found in numerous chromatin-associated proteins (Aravind and Landsman, 1998). AT hooks consist of an invariant GRP core motif surrounded by mostly basic residues. In cases of proteins with single AT hooks, minor-groove peptide binding is thought to play an auxiliary role in increasing the specificity of a DNA binding protein.

A few proteins have a large number of AT hooks, allowing for binding to AT-rich repeats, such as AT-rich satellites, which are abundant in *Drosophila* heterochromatin. In fact, the first example of an AT-rich DNA binding protein was described 20 years ago: the *Drosophila* D1 chromosomal protein specifically binds to AT-rich satellite repeats (Rodriguez Alfageme et al., 1980). Upon sequencing, D1 was found to contain 11 AT hooks. An artificial protein that contains 20 HMG-I AT hooks, MATH-20, resembles D1 in also binding to AT-rich *Drosophila* DNA, including scaffold-associated regions (SARs) and satellites (Strick and Laemmli, 1995). The binding of D1 and MATH-20 to naturally occurring satellites emphasizes the potential of satellites as targets for designer peptide compounds.

Minor-Groove Binding Dyes as Cytological Tools

Laemmli's team realized that by conjugating minor-groove binding compounds with dyes, *Drosophila* chromosomes can reveal sequence specificity in situ on a global scale. Compounds targeted specifically to SARs and to two pentameric satellite repeats were synthesized, tested by footprinting, and converted to dyes by addition of a fluorescent group (Janssen et al., 2000b [November issue of *Molecular Cell*]). Addition of the fluor only slightly affected the affinity and specificity of pyrrole-based AT-rich binders.

To target the GAGAA satellite, a polyamide compound was designed in which a monomeric minor-groove binding code was applied to the 10 base sequence GAGAA GAGAA, where Im binds G/C and Py or β binds A/T: Im-β-Im-Py-β-Im-β-Im-β-(dimethylaminopropylamide) (Figure 1, bottom). The authors provided evidence that this compound (P31) binds in a 1:1 ratio in the minor groove, rather than in the conventional side-by-side conformation that others have used to target mixed AT and GC duplexes. It will be interesting to determine the structure of the P31-DNA complex, and whether this Watson-strand-only code applies generally to the design of minor-groove binders of homopurine/homopyrimidine duplexes.

These dyes were used to stain squashed preparations of chromosomes (Janssen et al., 2000b). Two oligopyrroles, P9F and Lex9F, showed similar targeting of known AT-rich regions in both cell line and polytene nuclei. In *Drosophila*, AT-rich satellites can either be complex, with several AT-rich motifs, or simple, consisting primarily of the pentamer AATAT, and both types were stained by the dyes. In addition to these satellites, which are located in the pericentric heterochromatin, numerous euchromatic sites were stained. The banded pattern seen in polytene chromosomes does not correspond exclusively to classical band or interband regions, and perhaps reflects the distribution of SARs in the *Drosophila* genome. Banding patterns seen on polytene chromosomes are extraordinarily detailed, and should be useful in general for assessing the genome-wide sequence specificity of minor-groove binding compounds.

The most impressive evidence for specificity came from utilizing a fly mutation, $brown^{Dominant}$ (bw^D), which is an insertion of ~1–2 Mb of GAGAA into the distally located *brown* gene. In polytene chromosomes, satellite repeats undergo little if any polytenization, and the GAGAA pericentric repeats are undetectable. However, bw^D is sandwiched between distal euchromatic sequences, and undergoes substantial polytenization (Platero et al., 1998). Thus, essentially all of the GAGAA repeats detectable in polytene chromosomes from bw^D mutants are at this single locus, providing an ideal whole-genome assay for the targeting specificity. Indeed, the GAGAA-specific drug P31 targeted bw^D exclusively.

Thus, two completely different minor-groove binding dye conjugates, P9F and P31T, target their substrates in cytological preparations with high specificity, providing a compelling case for the general utility of minor-groove binding dye conjugates. Compounds with different specificities can reveal different landmarks, which together with multiple fluorochromes that emit at different wavelengths (e.g., Sachs et al., 2000), will provide researchers with a new class of versatile cytological tools.

What Happens to Flies on Drugs?

At first glance, the specific targeting of minor-groove binding compounds to satellites would seem unlikely to have functional consequence in vivo, as satellites comprise the most inert subset of gene-poor constitutive heterochromatin. However, inertness can itself provide a functional assay, because reporter genes juxtaposed to heterochromatin are silenced, a phenomenon known as position-effect variegation (PEV; reviewed in Weiler and Wakimoto, 1995). PEV is an extremely sensitive assay for detecting changes in reporter gene activity, providing a readout for alterations in either the state of heterochromatin or the accessibility of the reporter. Therefore, PEV can be used to assess the effect of minor-groove binding drugs in vivo. Evidently, minor-groove binding can affect PEV, because expression of MATH-20 in flies causes PEV suppression (Girard et al.,

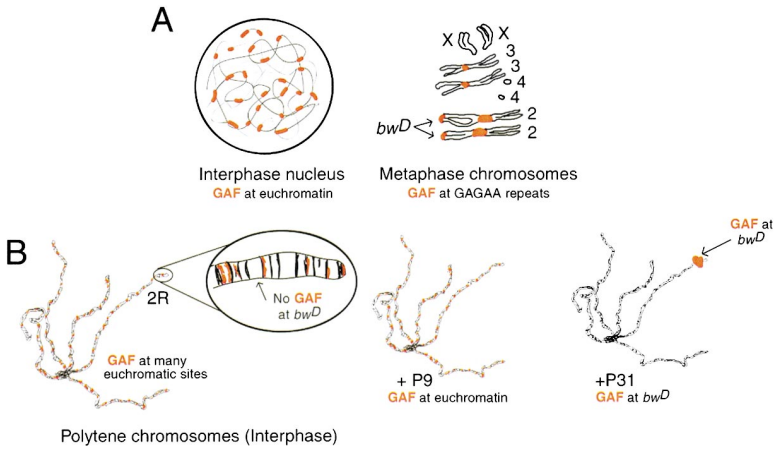

Figure 2. Relocalization of GAF from Euchromatin to *bw*^{D}

(A) Cycling of GAF from dispersed sites in euchromatin to GA-rich satellites in pericentric heterochromatin and to *bw*^{D}, which is located near the distal tip of chromosome arm 2R. (B) GAF is found at numerous sites throughout euchromatin, but is undetectable at *bw*^{D} during interphase. After treatment with P31, but not P9, GAF relocalizes to *bw*^{D} in polytene nuclei.

1998). Given that AT binding compounds show similar specificity to MATH-20, they would be expected to similarly suppress PEV. However, this prediction presupposes that enough of a compound can be ingested, can diffuse into the hemolymph, can be transported into target cells and can reach their DNA targets to have a detectable effect. Impressively, this prediction was fulfilled for larvae grown on food containing 100 μM of the AT binder P9. Emerging adults displayed PEV suppression, i.e., derepression of the classical *white-mottled-4* (*w*^{m4}) PEV mutation, showing no other ill effects (Janssen et al., 2000a).

The effect on PEV is almost certainly attributable to targeted minor-groove binding, because the effect was no different from what was seen for MATH-20. However, both the mechanism of PEV suppression and the identity of in vivo targets of the drugs (and of MATH-20) are uncertain. It is possible that AT-rich sequences in the local vicinity of the *white* reporter gene are targeted. Alternatively, an AT-rich satellite located within the heterochromatin that induces PEV might be altered, although the closest AT-rich satellite to *white* is located on the other side of the entire rDNA locus. As detailed below, we favor a third interpretation, in which satellites behave as sinks for heterochromatin binding proteins. All of these models have in common the idea that derepression results from opening chromatin by minor-groove binding. We stress that uncertainties with respect to mechanism should not detract from the importance of this result, that a developmental change can be caused by ingestion of a DNA binding compound that is known to target specific DNA sequences.

The GAGAA binding compound, P31, had no effect on *w*^{m4}, and some AT-rich binders also failed to show an effect. However, this could be a trivial consequence of reduced uptake, transport, or in vivo target binding affinity. A somewhat different PEV assay for testing drugs was possible using *bw*^{D}, which causes PEV-like *trans*-inactivation of a somatically paired wild-type copy of the *brown* gene. Although no effects on *bw*^{D} were seen with P9, ingestion of P31 may have led to modest derepression of the *brown* gene. However, interpretation of this result was complicated by the surprising response of *bw*^{D}-bearing flies to P31: Wild-type flies showed no ill effects from eating P31, but flies carrying a single copy of *bw*^{D} (*bw*^{D}/+) were feeble, and most flies carrying two copies (*bw*^{D}/*bw*^{D}) died. Considering that the *brown* gene is entirely dispensible, and *bw*^{D}/*bw*^{D} flies are healthy, this is a very odd result.

What Does P31 Do to bw*^{D}*?

Close examination of the P31-fed *bw*^{D} survivors provided intriguing evidence for multiple homeotic transformations (Janssen et al., 2000a). Extra bristles revealed a transformation of the sixth abdominal segment into the fifth, fewer sex combs revealed a *Sex-combs-reduced*-like transformation, and larger halteres in combination with an *Ultrabithorax* gain-of-function mutation revealed enhanced transformation of the third thoracic segment into the second. This unusual combination of phenotypes had been seen before: partial loss-of-function mutations in the gene encoding GAF cause the same syndrome (Farkas et al., 1994). So the question became, does P31 interact with *bw*^{D} in such a way that GAF levels are reduced?

An unexpected connection between *bw*^{D} and GAF had previously been reported (Platero et al., 1998). GAF appears to bind exclusively to GA-rich satellites, including *bw*^{D}, during mitosis (Figure 2A), but does not bind at all during interphase. The absence of GAF from interphase heterochromatin is seen most clearly in polytene chromosomes, where thousands of euchromatic sites have GAF, yet the heterochromatic chromocenter and *bw*^{D} are completely devoid of GAF (Figure 2B). Therefore, the GA-rich satellites provide a potential binding reservoir for GAF that fills up only at the onset of mitosis, during which time the euchromatin condenses and loses all detectable GAF. These observations provide the connection between *bw*^{D} and GAF that can potentially explain the P31-*bw*^{D} syndrome. Perhaps the binding of P31 to *bw*^{D} opens it up, resulting in the unscheduled transfer of GAF from euchromatic sites to *bw*^{D} at interphase, a process that normally occurs only during mitosis.

A critical test of this scenario is to treat with P31 and look for transfer of GAF from euchromatic sites to *bw*^{D} by staining of polytene chromosomes with anti-GAF antibody. Laemmli's group performed this test, and the result was stunning (Figure 2B). *bw*^{D} normally shows no anti-GAF staining, whereas nearly all detectable GAF in P31-treated nuclei was found at *bw*^{D}. Concomitantly, anti-GAF staining of euchromatic bands was nearly abolished (Janssen et al., 2000a). Thus, it appears that

P31-induced opening of bw^D provides a sink for GAF, which leaves its normal sites. Consistent with this interpretation, the authors note that P31 binding and GAF binding to GAGAA repeats are not mutually exclusive. Furthermore, normal GAF binding sites, which are typically GAGAG, are not targeted by P31, so that P31 is expected to open up bw^D preferentially to euchromatic GAF binding sites. Unscheduled GAF binding to bw^D after P31 treatment provides compelling evidence for Laemmli's model that minor-groove binding exerts its effect on chromatin by opening it up.

GAF is not the only mitosis-specific satellite binding protein. *Drosophila* Prod protein, which binds to the AT-rich AATAAGATAC decameric satellite, shows similar cycling behavior (Platero et al., 1998), and reservoirs for other proteins might exist in heterochromatin. We wonder if titration by a drug-opened satellite sink had occurred for the AT binders, resulting in PEV suppression. AT-rich satellites are especially abundant in the *Drosophila* genome, and so it seems reasonable to suspect that if these satellites open up, they would become massive sinks for proteins needed for heterochromatic silencing. Thus, titration of silencing factors by satellites could have led to derepression of the white reporter gene in w^{m4}. This explanation may have precedent in the well-known PEV suppressing effect of extra Y chromosomes, which are thought to derepress w^{m4} and other PEV mutations by titrating out heterochromatin-specific factors, reducing their availability at sites subject to PEV (Dimitri and Pisano, 1989).

The human genome also has satellites in abundance. Indeed, the Ikaros regulatory protein binds to human gamma satellites, which have been proposed to function as Ikaros storage sites in lymphocytes (Cobb et al., 2000). It might not be too far-fetched to imagine that human satellite sinks will someday provide therapeutic drug targets. When converted to dye-conjugated compounds, minor-groove binders have an added advantage, as demonstrated by Laemmli's group, in that they illuminate the DNA target that they open up. Potentially, a therapeutic drug can reveal its DNA target, analogous to illumination of proteins linked to GFP in living organisms. Perhaps future decongestants will not only clear your nose, but also light it up!

Selected Reading

Aravind, L., and Landsman, D. (1998). Nucleic Acids Res. *26*, 4413–4421.

Churchill, M.E., and Suzuki, M. (1989). EMBO J. *8*, 4189–4195.

Cobb, B.S., Morales-Alcelay, S., Kleiger, G., Brown, K.E., Fisher, A.G., and Smale, S.T. (2000). Genes Dev. *14*, 2146–2160.

Dimitri, P., and Pisano, C. (1989). Genetics *122*, 793–800.

Farkas, G., Gausz, J., Galloni, M., Reuter, G., Gyurkovics, H., and Karch, F. (1994). Nature *371*, 806–808.

Girard, F., Bello, B., Laemmli, U.K., and Gehring, W.J. (1998). EMBO J. *17*, 2079–2085.

Gottesfeld, J.M., Neely, L., Trauger, J.W., Baird, E.E., and Dervan, P.B. (1997). Nature *387*, 202–205.

Janssen, S., Cuvier, O., Muller, M., and Laemmli, U.K. (2000a). Mol. Cell *6*, 1013–1024.

Janssen, S., Durussel, T., and Laemmli, U.K. (2000b). Mol. Cell *6*, 999–1011.

Platero, J.S., Csink, A.K., Quintanilla, A., and Henikoff, S. (1998). J. Cell Biol. *140*, 1297–1306.

Reeves, R., and Nissen, M.S. (1990). J. Biol. Chem. *265*, 8573–8582.

Rodriguez Alfageme, C., Rudkin, G.T., and Cohen, L.H. (1980). Chromosoma *78*, 1–31.

Sachs, R.K., Hlatky, L.R., and Trask, B.J. (2000). Trends Genet. *16*, 143–146.

Strick, R., and Laemmli, U.K. (1995). Cell *83*, 1137–1148.

Weiler, K.S., and Wakimoto, B.T. (1995). Annu. Rev. Genet. *29*, 577–605.

Wemmer, D.E. (2000). Annu. Rev. Biophys. Biomol. Struct. *29*, 439–461.