



Conserved Residues Make Similar Contacts in Two Repressor-Operator Complexes

Carl O. Pabo, Aneel K. Aggarwal, Steven R. Jordan, Lesa J. Beamer, Upul R. Obeysekare, Stephen C. Harrison

Science, New Series, Volume 247, Issue 4947 (Mar. 9, 1990), 1210-1213.

Stable URL:

<http://links.jstor.org/sici?sici=0036-8075%2819900309%293%3A247%3A4947%3C1210%3ACRMSCI%3E2.0.CO%3B2-R>

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

Science is published by American Association for the Advancement of Science. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/aaas.html>.

Science

©1990 American Association for the Advancement of Science

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©2002 JSTOR

obligatorily moved from high to low chemical potential with each turn of the conformational cycle. Indeed, this model shows a strong resemblance to a conventional solute transport system, such as the glucose carrier, in the presence of a large gradient for the transported solute, if it were possible to distinguish the various conformational states of single transporter molecules. Of course, with a single-pore channel operating in this way, we would not actually be able to observe a time-asymmetric single-channel record; it is the channel "substates" arising from the double-barreled nature of the Cl⁻ channel that reveal the time asymmetry.

Although useful in rationalizing our results qualitatively, this simple picture is unrealistic. The model does not predict the vastly different asymmetry ratios observed with different Cl⁻ concentrations (Fig. 2). We have elaborated the basic model in several ways to introduce more realism. First, we have considered a double-barreled model (Fig. 3B), in which the two protochannels always inactivate together, as in the Cl⁻ channel. Second, we have relaxed the constraint that the "upper" and "lower" gates can simultaneously close only when Cl⁻ is bound. With these two twists on the simple model, Monte Carlo simulations show that the asymmetry depends on the many rate constants in the model, and that efficiency of coupling between Cl⁻ movement and gating asymmetry varies with the absolute concentration of Cl⁻. We are currently investigating whether these more complex models can be used to understand quantitatively the variations in coupling efficiency we have observed under different Cl⁻ gradient conditions.

The *Torpedo* Cl⁻ channel provides an opportunity to explore nonequilibrium channel gating driven by electrochemical gradients. Previous investigations on nonequilibrium gating, based on current flow through open channels, are either wholly theoretical (6) or have been described only anecdotally (7). The Cl⁻ channel studied here provides a crisp example of a gating process with a clear, manipulable dependence on electrochemical potential of the permeating ion. It is worth noting that a well-documented and puzzling property of a different channel, the inward rectifier K⁺ channel of excitable membranes, is the ability of its gating machinery to "follow" the electrochemical gradient for K⁺, opening only when net K⁺ movement is inward (8). The mechanism presented here, which provides a concrete example of the coupling of ion-channel gating to the transmembrane gradient of the conducting ion, may therefore have a wider relevance than to the Cl⁻ channel alone.

REFERENCES AND NOTES

1. Preliminary discussions of the data and model are found in E. A. Richard and C. Miller, *Biophys. J.* **53**, 266a (1988); C. Miller and E. A. Richard, in *Cl⁻ Channels and Carriers in Nerve, Muscle, and Glial Cells*, F. J. Alvarez-Leefmans and J. Russell, Eds. (Plenum, New York, in press).
2. C. Miller, *Philos. Trans. R. Soc. London Ser. B* **299**, 401 (1982); _____ and M. M. White, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 2772 (1984).
3. T. L. Hill, *Free Energy Transduction in Biology* (Academic Press, New York, 1977).
4. C. Miller and M. M. White, *Ann. N.Y. Acad. Sci.* **341**, 534 (1980); M. M. White and C. Miller, *J. Gen. Physiol.* **78**, 1 (1981).
5. A similar treatment can be performed by specifying that both gates can only close on an empty site.
6. P. Luger, in *Single-Channel Recording*, B. Sakmann and E. Neher, Eds. (Plenum, New York, 1983), pp. 177–198; A. Finkelstein and C. S. Peskin, *Biophys. J.* **46**, 549 (1984).
7. A. Hamill and B. Sakmann, *Nature* **294**, 462 (1981); A. Trautmann, *ibid.* **298**, 272 (1982); S. G. Cull-Candy and M. M. Usowicz, *ibid.* **325**, 525 (1987).
8. R. H. Adrian *et al.*, *J. Physiol. (London)* **208**, 607 (1970); S. Hagiwara and K. Takahashi, *J. Membr. Biol.* **18**, 61 (1974); but see C. A. Leech and P. R. Stanfield, *J. Physiol. (London)* **319**, 295 (1981).
9. O. S. Andersen, *Biophys. J.* **41**, 119 (1983).
10. We thank R. Horn for advice on the statistical analysis of our data and I. Levitan and R. MacKinnon for critical readings of the manuscript. Supported by NIH research grant GM-31768 and training grant NS-07292.

10 October 1989; accepted 28 December 1989

Conserved Residues Make Similar Contacts in Two Repressor-Operator Complexes

CARL O. PABO, ANEEL K. AGGARWAL,* STEVEN R. JORDAN, LESA J. BEAMER, UPUL R. OBEYSEKARE, STEPHEN C. HARRISON

Comparison of a λ repressor-operator complex and a 434 repressor-operator complex reveals that three conserved residues in the helix-turn-helix (HTH) region make similar contacts in each of the crystallographically determined structures. These conserved residues and their interactions with phosphodiester oxygens help establish a frame of reference within which other HTH residues make contacts that are critical for site-specific recognition. Such "positioning contacts" may be important conserved features within families of HTH proteins. In contrast, the structural comparisons appear to rule out any simple "recognition code" at the level of detailed side chain–base pair interactions.

THE λ AND 434 REPRESSORS, WHICH are produced by related *Escherichia coli* phages (1), recognize very distinct operator sites, but both use the HTH unit for recognition (2–4). Alignment of the amino acid sequences reveals significant similarities (5), and there also are clear structural similarities between the proteins: Helices 1 to 4 of λ repressor correspond to helices 1 to 4 of 434 repressor (4). In each case, helices 2 and 3 form the conserved HTH unit that participates in operator recognition and 7 of the 20 residues in the HTH unit are identical (Fig. 1). The NH₂-terminal domain of λ repressor (92 residues) is somewhat larger than the NH₂-terminal do-

main of 434 repressor (69 residues), and 434 repressor lacks an NH₂-terminal arm or any region that corresponds to helix 5 of λ repressor.

Comparing the HTH units from the two cocrystal complexes (residues 33 to 52 of λ repressor and residues 17 to 36 of 434 repressor) shows that these regions are extremely similar (Fig. 2). The α -carbons of these units can be superimposed with root-mean-square (rms) deviations of 0.48 to 0.59 Å (6) depending on which subunits are used. These values are similar to the deviations observed when comparing the two halves of the same repressor (0.31 Å rms for λ and 0.59 Å for 434). In this region, the polypeptide backbones of the two proteins are essentially indistinguishable, consistent with previous observations that the HTH units of the λ Cro protein, the *E. coli* CAP protein, and the λ repressor are almost identical (7, 8).

Although the structures of the λ and 434 repressors are similar in the HTH regions, there are significant differences if one compares the other helices or compares the relative orientation of the protein dimers. When superimposing helices 1 to 4 of the two repressors, it seems most appropriate to

C. O. Pabo and U. R. Obeyesekere, Howard Hughes Medical Institute, Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, MD 21205.

A. K. Aggarwal and S. C. Harrison, Howard Hughes Medical Institute, Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, MA 02138.

S. R. Jordan, Agouron Pharmaceuticals, Inc., 11025 North Torrey Pines Road, La Jolla, CA 92037.

L. J. Beamer, Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, MD 21205.

*Present address: College of Physicians and Surgeons, Columbia University, New York, NY 10032.

		33										41								44										52
Lambda		Gln	Glu	Ser	Val	Ala	Asp	Lys	Met	Gly	Met	Gly	Ser	Gly	Val	Gly	Ala	Leu	Phe	Asn										
434		Gln	Ala	Glu	Leu	Ala	Gln	Lys	Val	Gly	Thr	Thr	Gln	Gln	Ser	Ile	Glu	Gln	Leu	Glu	Asn									
Trp		Gln	Arg	Glu	Leu	Lys	Asn	Glu	Leu	Gly	Ala	Gly	Ile	Ala	Thr	Leu	Thr	Arg	Gly	Ser	Asn									
		Helix				Turn				Helix																				

Fig. 1. Alignment of the HTH units from the λ , 434, and Trp repressors. Conserved residues are shown in boldface, and boxes enclose the conserved residues that contact the DNA.

treat residues 26 to 29 of λ repressor, which occur in the loop between helix 1 and helix 2, as an "insertion." Structural comparison (9) also confirms the original suggestion (5) that λ repressor has an additional residue in the loop between helix 3 and helix 4. These alignments for helices 1 to 4 allow 41 α -carbons to be superimposed with an rms error of 2.39 Å. There are even larger differences when comparing the protein dimers. Such differences are not surprising, however, since the spacing of the operator half-sites is one base pair different in the two complexes, and the HTH units must therefore be about 3.4 Å

further apart in the λ repressor dimer than in 434. Thus, superposition of the individual HTH units gives the best frame of reference for comparing detailed interactions with DNA. This superposition allows us to align the half-sites accurately and to compare the contacts made by individual residues.

The aligned sequences of the HTH units (Fig. 1) and the structures of the complexes show that almost all of the conserved residues in the HTH region make similar contacts with the operator or play a similar role in the packing of the protein interior. The conserved Ala and Gly (residues 37 and 41

in λ repressor) are characteristic of the HTH unit (10) and seem to stabilize the folding of this bihelical unit. The conserved Leu (residue 50 in λ repressor) stabilizes the packing of the HTH unit against the rest of the protein. The conserved Lys (residue 39 in λ repressor) does not contact the DNA and may be favored because a positively charged residue tends to stabilize the COOH-terminal end of an α -helix (11). The conserved residues involved in folding of the HTH unit have already been discussed in detail (10), and the cocrystal structures now allow us to compare conserved residues that contact the DNA.

The first residues in the HTH unit, Gln³³ and Gln⁴³⁴, each make two hydrogen bonds with the DNA backbone. The peptide -NH hydrogen bonds to one phosphodiester oxygen near the outer edge of the operator site, and the side chain -NH₂ hydrogen bonds to an oxygen of a neighboring phosphate group (Fig. 2). In each case, the contact made by the side chain is on the 3' side of the contact made by the peptide -NH. We presume that this is a particularly stable steric and electrostatic arrangement, and it is clear that the partial positive charge on the NH₂-terminal end of the α -helix will give a favorable electrostatic interaction with the negative charge on the phosphate oxygen (12).

The first residues in the second helix of the HTH unit are Gln⁴⁴ and Gln²⁸. In each complex these glutamines contact an adenine near the ends of the operator site, and these contacts provide a useful way of aligning the sequences of the λ and 434 half-sites. The glutamine and adenine make a pair of hydrogen bonds and allow for specific recognition of the A·T base pair (13). The pair of hydrogen bonds can only be formed if adenine is present at this position. A similar interaction has also been observed in the 434 Cro complex (14) and has been predicted to occur in the λ Cro complex (15). Studies of mutant repressors confirm that Gln⁴⁴ makes a major contribution to the stability of the λ repressor-operator complex (16). Gln⁴⁴ and Gln²⁸ also hydrogen bond to the glutamines at the start of the HTH unit. This hydrogen-bonding network directly correlates phosphate and base pair contacts, and appears to be critical for site-specific recognition.

The last residues in the HTH unit are Asn⁵² and Asn³⁶. In each complex, the asparagine forms a hydrogen bond with the same phosphodiester oxygen that is contacted by the side chain -NH₂ group of the first glutamine in the HTH unit. This asparagine also forms a hydrogen bond to a carbonyl oxygen from a preceding turn of the α -helix.

These critical residues at the beginning

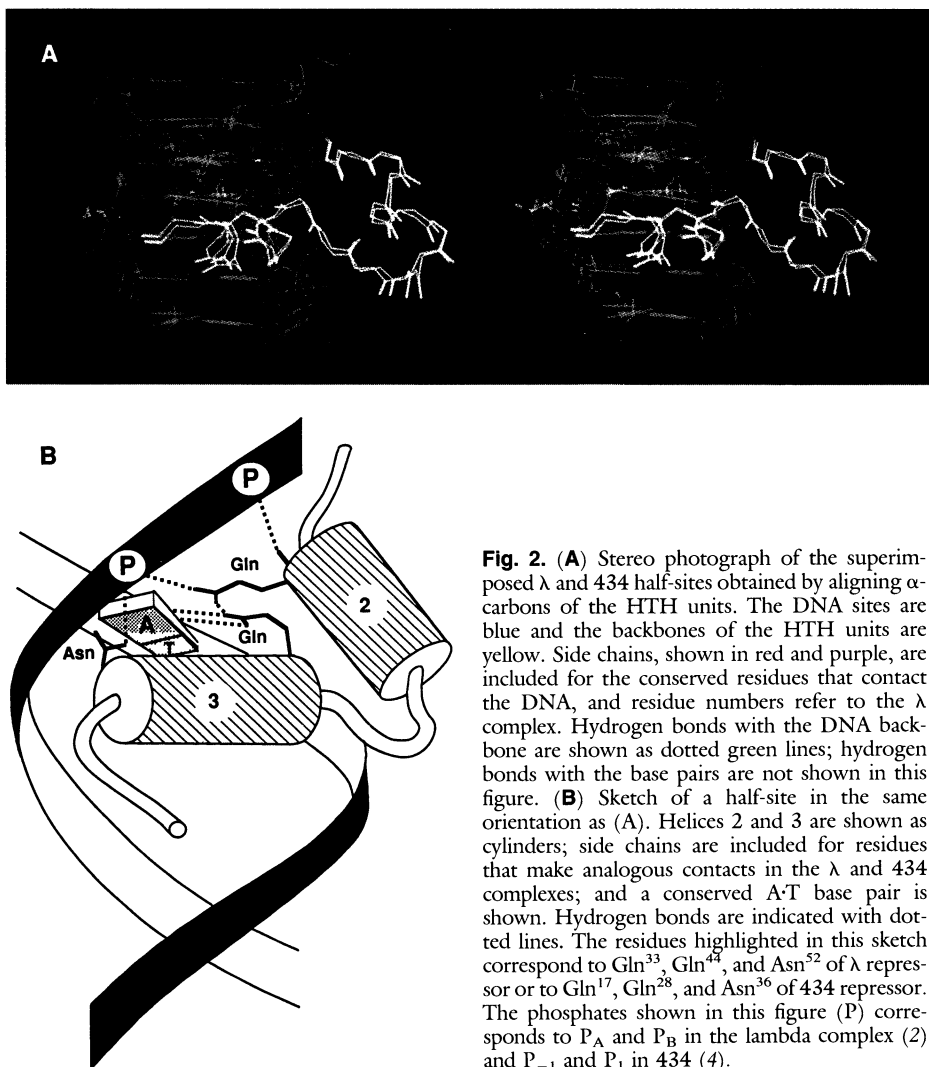


Fig. 2. (A) Stereo photograph of the superimposed λ and 434 half-sites obtained by aligning α -carbons of the HTH units. The DNA sites are blue and the backbones of the HTH units are yellow. Side chains, shown in red and purple, are included for the conserved residues that contact the DNA, and residue numbers refer to the λ complex. Hydrogen bonds with the DNA backbone are shown as dotted green lines; hydrogen bonds with the base pairs are not shown in this figure. (B) Sketch of a half-site in the same orientation as (A). Helices 2 and 3 are shown as cylinders; side chains are included for residues that make analogous contacts in the λ and 434 complexes; and a conserved A·T base pair is shown. Hydrogen bonds are indicated with dotted lines. The residues highlighted in this sketch correspond to Gln³³, Gln⁴⁴, and Asn⁵² of λ repressor or to Gln¹⁷, Gln²⁸, and Asn³⁶ of 434 repressor. The phosphates shown in this figure (P) corresponds to P_A and P_B in the lambda complex (2) and P₋₁ and P₁ in 434 (4).

and ends of the helices provide a conserved framework for recognition and help position the HTH unit in the major groove of the DNA. The conserved phosphate contacts seen in the cocrystal structures and also detected by ethylation interference experiments (17, 18) establish an alignment of the two operators (Fig. 3). Within this framework, there is an approximate correlation between the position of a residue in the HTH unit and the particular base that it contacts in its operator, but the precise pattern of interaction depends on the size of the amino acid side chain and the chemical identity of the nearby base pairs. We take the adenine contacted by the first residue of helix 3 (Gln_λ⁴⁴ or Gln_λ²⁸) as a reference. The second residue of helix 3 (Ser_λ⁴⁵ or Gln_λ²⁹) interacts with the next nearest base on the opposite strand, but Gln_λ²⁹ also contacts two other bases on this strand. More specifically, Ser_λ⁴⁵ forms a hydrogen bond with a guanine, while Gln_λ²⁹ makes van der Waals contacts with two thymine methyl groups and two hydrogen bonds with a guanine base. There are even fewer analogies among the contacts made by other residues. In general, there is a tendency for the proteins to contact the 5' strand near the outside of the operator site and to contact the 3' strand closer to the center of the operator—a simple consequence of fitting a cylindrical object (an α -helix) against a “helical saddle” (the major groove).

In each complex, the first, second, third, fifth, and sixth residues of helix 3 contact base pairs in the operator site. The fourth residue is the only one that is completely buried on the back side of the α -helix, and this residue (Val_λ⁴⁷ or Ile_λ³¹) makes critical hydrophobic interactions with helix 2 (10). It seems that this azimuthal orientation of helix 3 may be optimal for recognition, and this may help explain why the HTH unit is so precisely conserved. Almost any other orientation of helix 3 would bury two residues, leaving only four residues available to make site-specific contacts.

The HTH unit in the Trp repressor-operator complex has a somewhat different arrangement (19). The second helix of the HTH unit has a nearly end-on orientation with respect to the DNA, and several water molecules that hydrogen bond to peptide –NH groups at the end of the α -helix also hydrogen bond to the base pairs. These differences in the base contacts have led to extensive discussion about the differences between Trp repressor and the phage repressors. However, there are interesting analogies between the Trp repressor-operator complex and the phage repressor-operator complexes. Sequence alignments of the HTH units of λ , 434, and Trp (Fig. 1) show

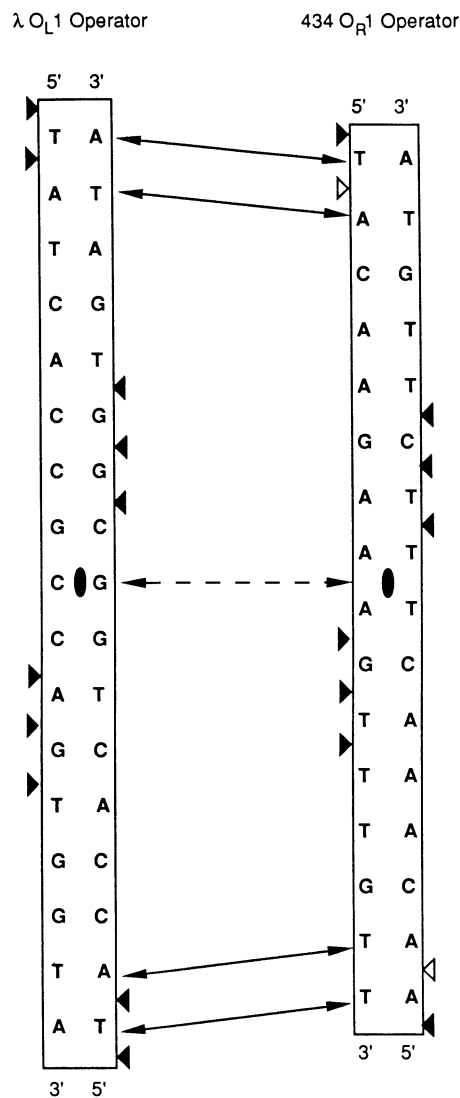
that Gln_{Trp}⁶⁸ corresponds to Gln_λ³³ or Gln_λ⁴⁷ and Asn_{Trp}⁷⁶ corresponds to Asn_λ⁵² or Asn_λ³⁶. Gln_{Trp}⁶⁸ hydrogen bonds to two phosphodiester oxygens near the outside edge of the operator site, as do the Gln residues of the phage repressors. That is, its peptide –NH hydrogen bonds to an oxygen of one phosphate group and its side chain –NH₂ to an oxygen of the 3' neighbor. Asn_{Trp}⁸⁷, like the corresponding Asn residues of the phage repressors, also hydrogen bonds to a phosphodiester oxygen. Instead of sharing the oxygen bonded to the –NH₂ of Gln_{Trp}⁶⁸, however, Asn_{Trp}⁸⁷ contacts an oxygen of the neighboring phosphate on the 3' side. This difference is a result of differences in the orientation of the second helix of the HTH motif.

Another similarity between the HTH units of the Trp and λ repressors is a hydrogen bond from the last peptide –NH in the turn (residue 43 of λ and residue 78 of Trp) to a phosphodiester oxygen. Although the homologous hydrogen bond is absent in the 434 repressor complex, because of a local shift in the DNA backbone at this location, it is present in complexes of 434 Cro with DNA (14, 20). This bond may be another conserved feature of the way many HTH units interact with the DNA backbone.

Comparison of the closely related λ and 434 repressor-operator complexes, the 434 Cro operator complex, and the distinctly less similar Trp repressor complex, demonstrates that there is no simple “code” describing the contacts made with base pairs. The diversity of these interactions is shown by the range of hydrogen-bonding and nonpolar interactions, and it is clear that there is no one-to-one correspondence between amino acid side chains and base pair contacts. However, the comparison does suggest that certain interactions may be particularly important in positioning the HTH unit against the DNA backbone and hence in establishing a frame of reference for base pair recognition. The residues at the NH₂-terminus of the first helix and the NH₂- and COOH-termini of the second helix (positions 1, 12, and 20

Fig. 3. Sketch showing alignment of the λ and 434 operator sites used in the cocrystal studies. Solid triangles mark the phosphate contacts that are detected by ethylation interference experiments and seen in the cocrystal structures; open triangles mark phosphate contacts that are seen in the structures but not detected by ethylation interference experiments. The center of approximate twofold symmetry in each site is marked (●) and indicated with the dashed arrow. Arrows relate corresponding base pairs and contacts in the two half-sites. This figure includes an additional base pair on each end of the conventional (14-base pair) 434 operator site, and shows that the λ operator site is effectively just one base pair longer than the 434 site.

of the conventional HTH alignment) play key roles in the known structures. A number of other HTH proteins have related residues at some of these positions. For example, of the 21 sequences compiled by Pabo and Sauer (10), 8 have Gln at the start of the HTH unit and 7 have Asn at its end. These conserved interactions, which help position the HTH unit in the major groove, may provide a useful guide for thinking about contacts made by homeodomains. Recent nuclear magnetic resonance studies (21) confirm the suggestion that homeodomains also contain a HTH motif (22, 23). In 73 of the 86 homeodomain sequences listed by Scott and co-workers (24), there is an Arg at the position corresponding to Gln_λ³³ or Gln_λ⁴⁷ and there is a Glu corresponding to Gln_λ⁴⁴ or Gln_λ²⁸. These two residues could hydrogen bond to each other as well as to the DNA in a pattern analogous to that seen for the two glutamines in the phage repressors. Our observations may also have important implications for studies of other DNA-binding motifs such as zinc fingers. In gen-



eral, it seems reasonable to expect that the members of such large families of DNA-binding proteins may have several conserved contacts with the DNA backbone, and that closely related members may also have conserved contacts with one or more bases.

REFERENCES AND NOTES

1. M. Ptashne, *A Genetic Switch* (Cell Press, Cambridge, MA, 1986).
2. S. R. Jordan and C. O. Pabo, *Science* **242**, 893 (1988).
3. J. E. Anderson, M. Ptashne, S. C. Harrison, *Nature* **326**, 846 (1987).
4. A. K. Aggarwal, D. W. Rodgers, M. Drottar, M. Ptashne, S. C. Harrison, *Science* **242**, 899 (1988).
5. R. T. Sauer, R. R. Yocum, R. F. Doolittle, M. Lewis, C. O. Pabo, *Nature* **298**, 447 (1982).
6. For the data from 8.0 to 2.5 Å, the *R* factor is 0.204 for the λ complex and 0.179 for the 434 complex.
7. T. A. Steitz, D. H. Ohlendorf, D. B. McKay, W. F. Anderson, B. W. Matthews, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 3097 (1982).
8. D. H. Ohlendorf, W. F. Anderson, M. Lewis, C. O. Pabo, B. W. Matthews, *J. Mol. Biol.* **169**, 757 (1983).
9. A. Mondragon, S. Subbiah, S. C. Almo, M. Drottar, S. C. Harrison, *ibid.* **205**, 189 (1989).
10. C. O. Pabo and R. T. Sauer, *Annu. Rev. Biochem.* **53**, 293 (1984).
11. J. S. Richardson and D. C. Richardson, *Science* **240**, 1648 (1988).
12. W. G. J. Hol, P. T. van Duijnen, H. J. C. Berendsen, *Nature* **273**, 443 (1978).
13. N. C. Seeman, J. M. Rosenberg, A. Rich, *Proc. Natl. Acad. Sci. U.S.A.* **73**, 804 (1976).
14. A. Mondragon and S. C. Harrison, unpublished results.
15. D. H. Ohlendorf, W. F. Anderson, R. G. Fisher, Y. Takeda, B. W. Matthews, *Nature* **298**, 718 (1982).
16. H. C. M. Nelson and R. T. Sauer, *J. Mol. Biol.* **192**, 27 (1986).
17. A. Johnson, thesis, Harvard University, Cambridge, MA (1980).
18. R. Bushman, J. E. Anderson, S. C. Harrison, M. Ptashne, *Nature* **316**, 651 (1985).
19. Z. Orwinowski *et al.*, *ibid.* **335**, 321 (1988).
20. C. Wolberger, Y. Dong, M. Ptashne, S. C. Harrison, *ibid.*, p. 789.
21. Y. Q. Qian *et al.*, *Cell* **59**, 573 (1989).
22. A. Laughon and M. P. Scott, *Nature* **310**, 25 (1984).
23. J. C. W. Shepherd, W. McGinnis, A. E. Carrasco, E. M. DeRobertis, W. J. Gehring, *ibid.*, p. 70.
24. M. P. Scott, J. W. Tamkun, G. W. Hartzell, *BBA Rev. Cancer* **989**, 25 (1989).
25. Supported by the Howard Hughes Medical Institute and by grants from the NIH (GM 31471 to C.O.P.; GM 29109 to S.C.H. and M. Ptashne). We thank R. T. Sauer for helpful discussions and thank K. Kolish for help preparing this manuscript.

22 September 1989; accepted 24 January 1990

An Intron in the Genes for U3 Small Nucleolar RNAs of the Yeast *Saccharomyces cerevisiae*

EVELYNE MYSLINSKI, VÉRONIQUE SÉGAULT, CHRISTIANE BRANLANT*

The origin of the intervening sequences (introns), which are removed during RNA maturation, is currently unknown. They are found in most genes encoding messenger RNAs, but are lacking in almost all small nuclear (sn)RNAs. One exceptional snRNA (U6) is part of the spliceosomal machinery that is involved in messenger RNA maturation. It has been suggested that its intron arose as a result of incorrect splicing of a messenger RNA precursor. This study revealed the presence of an intron, with the characteristic features of nuclear introns from precursors to messenger RNA, in the two genes coding for *Saccharomyces cerevisiae* U3 snRNA. The branch point was GACTAAC instead of the TACTAAC sequence found in all yeast introns examined so far. As U3 is a nucleolar snRNA required for maturation of ribosomal RNA, its intron could not have been acquired from aberrant messenger RNA processing in a spliceosome.

THE NUCLEI OF VERTEBRATE CELLS contain metabolically stable small nuclear RNAs (snRNAs). Five of the major ones, denoted U1, U2, U4, U5, and U6 snRNAs, are involved in the splicing of pre-mRNAs (1). The small nuclear ribonucleoproteins (snRNPs) corresponding to these five RNAs react with antisera of the Sm serotype from patients with the autoimmune rheumatic disease lupus ery-

thematosus (2). Another major snRNA species, denoted U3, is located in the nucleolus (3). Domain A (4) containing the binding site for Sm-antigen proteins is absent in U3 snRNA, and so U3 snRNP does not react with Sm antibodies. Observation of hydrogen bonding between U3 snRNA and nucleolar 28S pre-ribosomal RNA (pre-rRNA) suggested an involvement of U3 snRNA in pre-rRNA maturation (3, 5), and several models of base-pairing have been proposed (1).

Yeast nuclear genes with intervening sequences (introns) interrupting the coding regions are rare and most of them encode ribosomal proteins. Although the pre-

mRNA splicing pathway in yeast is similar to that in higher eukaryotes, some features of introns are significantly different in yeast. First, all known yeast pre-mRNA introns contain the sequence TACTAAC in which lariat formation occurs. Second, the sequence at the 5' end of introns is significantly more variable in mammalian than in yeast mRNA, the 5' sequence GTATGT being highly conserved in yeast introns.

The yeast *Saccharomyces cerevisiae* contains a large variety of snRNAs (6). Two of them, snR17A and snR17B, correspond to vertebrate U3 snRNA (7). Unlike other yeast snRNAs, which are encoded by single-copy genes, snR17 is encoded by two genes having similar nucleotide sequences (7). Before starting a study of snR17 function by site-directed mutagenesis, we sequenced *Saccharomyces cerevisiae* snR17 at the RNA level and discovered that both snR17A and snR17B genes contain an intron.

The nucleotide sequence of *S. cerevisiae* snR17 from the strain named by the American Type Culture Collection ATCC 28383 (8) was determined by both the chemical method for RNA sequencing (9) and the enzymatic method based on primer extension with reverse transcriptase (10). The RNA extracted from mechanically disrupted cells was separated by ultracentrifugation on a sucrose gradient. The reverse transcriptase sequencing was performed with the 4S–8S RNA mixture recovered from the gradient, by means of synthetic deoxyoligonucleotides (Fig. 1a). The chemical method was used on 3' end-labeled, purified snR17 RNA (11). SnR17B RNA was not detected by either method. Nevertheless, the presence of two distinct snR17 genes in our strain was confirmed by Southern blot analysis of total DNA (Fig. 2). The size of the two hybridizing bands observed with Eco RI restriction endonuclease digestion fits with the physical map established for snR17A and snR17B genes (7). Therefore, as previously observed (7), snR17B gene is only poorly expressed in this strain.

There was good correlation between our RNA sequence and that for the snR17A gene, except for the 5' end of the molecule, where the sequence obtained by reverse transcriptase was completely different from that deduced for the gene (Fig. 1). Thus, either snR17A had a different sequence in our strain or snR17A gene had an intron. According to this second hypothesis, the 5' terminal sequence CGACGUACUUCA found by reverse transcriptase analysis (Fig. 1b) should belong to a first exon.

We isolated an snR17 gene from an *S. cerevisiae* ATCC 28383 genomic library, using 3' end-labeled snR17 as a probe, and determined the nucleotide sequence of the

Laboratoire d'Enzymologie et de Génie Génétique, Université de Nancy I, URA CNRS 457, BP 239, 54506 Vandoeuvre Les Nancy, Cedex, France.

*To whom correspondence should be sent.