MALDI-MS should, therefore, be very useful for the mass determination of defined RNA analytes such as ribosomal RNA.

All results reported above were obtained with only limited efforts in sample purification. In contrast, more extensive sample purification, particularly removal of alkali cations, is normally required in UV-MALDI-MS (*17*) and ESI-MS (*6*, *8*) of nucleic acids, and similar requirements have been reported for IR-MALDI-MS with succinic acid as matrix (*11*). We therefore speculate that the tolerance against impurities may be a particular feature of the glycerol matrix.

Mass resolution is an important parameter in mass spectrometry. In all spectra, except that of the small 21-nt sample of Fig. 2A, the mass resolution was limited by a sloping high-mass edge of the peaks. It is believed that this peak broadening is caused by adduct ions of as yet unknown origin. It is not likely that glycerol is the major contributor to these adducts, because not even a trace of glycerol adducts peaks was identified in the high-resolution spectrum of a small oligonucleotide (Fig. 1A). More sophisticated purification procedures might result in a substantial improvement in mass resolution.

References and Notes

1. E. Nordhoff, F. Kirpekar, P. Roepstorff, *Mass Spectrom. Rev.* **15**, 67 (1997).
2. P. L. Ross and P. Belgrader, *Anal. Chem.* **69**, 3966 (1997).
3. F. Kirpekar *et al.*, *Nucleic Acids Res.* **22**, 3866 (1994).
4. S. D. Fuerstenau and W. H. Benner, *Rapid Commun. Mass Spectrom.* **9**, 1528 (1995).
5. R. Chen *et al.*, *Anal. Chem.* **67**, 1159 (1995).
6. D. C. Muddiman *et al.*, *ibid.* **68**, 3705 (1996).
7. P. A. Limbach, P. F. Crain, J. A. McCloskey, *J. Am. Soc. Mass Spectrom.* **6**, 27 (1995).
8. Y. Naito *et al.*, *ibid.* **8**, 737 (1997).
9. S. Berkenkamp, C. Menzel, M. Karas, F. Hillenkamp, *Rapid Commun. Mass Spectrom.* **11**, 1399 (1997). The refTOF mass spectrometer has a equivalent flight length of 3.5 m and can also be used in linear TOF mode with 25-kV acceleration potential and for UV-MALDI.
10. S. Nui, Z. Wenzhu, B. T. Chait, *J. Am. Soc. Mass Spectrom.* **9**, 1 (1998).
11. E. Nordhoff *et al.*, *Nucleic Acids Res.* **21**, 3347 (1993).
12. E. Nordhoff *et al.*, *ibid.* **22**, 2460 (1994).
13. Glycerol (0.5 to 1 μl) was mixed with an equal amount of an aqueous analyte solution on the target. The analyte-to-glycerol molar ratio in the sample was 10^{-7}. The mixture was distributed evenly over an area of 1 to 2 mm^2 to form a homogeneous, transparent thin layer on the stainless steel substrate. The water was evaporated at a pressure of 10^{-2} Pa, before sample introduction into the spectrometer.
14. Lasers with wavelengths at 2.79, 2.94, and 10.6 μm were used. Various small organic acids such as succinic, fumaric, lactic, nicotinic, and dihydroxybenzoic acids, urea, thiourea, and triethanolamine were tested as matrices.
15. The mass of these fragments was calculated from the known sequence of the digested plasmid DNA. The average mass of the two complementary strands was taken. Small deviations, caused by base modifications, cannot be excluded but should be well within the resolution of the detected signals.
16. D. A. Melton *et al.*, *Nucleic Acids Res.* **12**, 7035 (1984).
17. D. P. Little *et al.*, *Anal. Chem.* **69**, 4540 (1997).
18. M. Kussman *et al.*, *J. Mass Spectrom.* **32**, 593 (1997).
19. A. Hopert, C. C. Uphoff, M. Wirth, H. Hauser, H. G. Drexler, *J. Immunol. Methods* **164**, 91 (1993).

20. This work was done in partial fulfillment of the requirements for the degree of Ph.D. of S.B. at the University of Münster. Financial support from grant Hi 285/8-2 of the German National Research Council and grant IVA1-800 985 96 of the Department of Education and Research of the State of Northrhine Westfalia is greatly appreciated. We thank J. P. Siebrasse for help with the PCR, T. Becker for assistance with the nanoliter dispensing, and K. Dreisewerd for helpful discussions.

23 February 1998; accepted 4 June 1998

Structure of the *Escherichia coli* RNA Polymerase α Subunit Amino-Terminal Domain

Gongyi Zhang and Seth A. Darst*

The 2.5 angstrom resolution x-ray crystal structure of the *Escherichia coli* RNA polymerase (RNAP) α subunit amino-terminal domain (αNTD), which is necessary and sufficient to dimerize and assemble the other RNAP subunits into a transcriptionally active enzyme and contains all of the sequence elements conserved among eukaryotic α homologs, has been determined. The αNTD monomer comprises two distinct, flexibly linked domains, only one of which participates in the dimer interface. In the αNTD dimer, a pair of helices from one monomer interact with the cognate helices of the other to form an extensive hydrophobic core. All of the determinants for interactions with the other RNAP subunits lie on one face of the αNTD dimer. Sequence alignments, combined with secondary-structure predictions, support proposals that a heterodimer of the eukaryotic RNAP subunits related to *Saccharomyces cerevisiae* Rpb3 and Rpb11 plays the role of the αNTD dimer in prokaryotic RNAP.

Escherichia coli RNAP comprises an essential catalytic core of two α subunits (each 36.5 kD), one β subunit (150.6 kD), and one β′ subunit (155.2 kD), which are conserved in sequence from bacteria to human. In addition to playing key roles in transcription initiation, the α subunit initiates RNAP assembly (*1*) by dimerizing into a platform with which the large β and β′ subunits interact. Deletion mutagenesis and limited proteolysis indicate that the α subunit comprises two independently folded domains, the NH_2-terminal domain (NTD; residues 8 to 235) and COOH-terminal domain (CTD; residues 249 to 329), connected by a flexible, 14-residue linker (Fig. 1A) (*2*). The αCTD is dispensable for RNAP assembly and basal transcription but is required for the interaction with an upstream promoter element (*3*) and is the target for a wide array of transcription activators (*4*). The solution structure of αCTD consists of a compact fold of four short α helices (*5*). The αNTD is essential in vivo and in vitro for RNAP assembly and basal transcription (*2*, *6*). The regions of conserved sequence between α homologs of prokaryotic, archaebacterial, chloroplast, and eukaryotic RNAPs (α motifs 1 and 2) (Fig. 1A) are contained within the NTD (*7*), as are the determinants for α interaction with the RNAP β and β′ subunits (*2*, *6*, *8*–*12*). We crystallized a mutant αNTD with an Arg to Ala substitution at position 45 (αNTD_{R45A}) because we were unable to obtain crystals of wild-type αNTD suitable for structure determination (*13*). The structure was determined by multiple isomorphous replacement (MIR) and refined to a resolution of 2.5 Å (Table 1).

The 26-kD αNTD monomer comprises two domains, each containing a distinct hydrophobic core (Fig. 1B). Domain 1 contains NH_2- and COOH-terminal sequences (residues 1 to 52 and 180 to 235), and domain 2 contains the intervening sequence (residues 53 to 179) (Fig. 2). Each domain has an α/β fold. Domain 1 contains a four-stranded antiparallel β sheet (S1, S2, S10, and S11) and two nearly orthogonal α helices (H1 and H3), whereas domain 2 contains seven β strands in an antiparallel arrangement (S3 to S9) and one α helix (H2) (Fig. 1B).

The αNTD dimer forms an elongated, flat structure with dimensions of about 120 Å by 60 Å by 25 Å (Fig. 1B). Almost all of the monomer-monomer interactions that form the dimer interface arise from H1 and H3 in domain 1 (Fig. 3). The unusual dimer interface can be described as two pairs of nearly orthogonal α helices (one pair from each monomer) that interlock like two V's intersecting through their open ends, resulting in two pairs of antiparallel α helices abutting each other with orthogonal orientations (Figs.

Rockefeller University, 1230 York Avenue, New York, NY 10021, USA.

*To whom correspondence should be addressed. E-mail: darst@rockvax.rockefeller.edu

1B and 3). The primary monomer-monomer interaction occurs through an antiparallel coiled-coil–like interaction between H3 of monomer 1 ($H3_1$) and H3 of monomer 2 ($H3_2$). The other pair of antiparallel α helices ($H1_1$ and $H1_2$) are shifted slightly apart and do not make extensive interactions with each other, but they contribute to the dimer interface by making hydrophobic interactions with H3 of the opposite monomer. The extensive hydrophobic interface sandwiched between the orthogonal pairs of antiparallel α helices accounts for the high stability of the α dimer. The hydrophobic core of the dimer interface is made up of one residue from S2, six residues within or adjacent to H1, and seven residues from H3 (Figs. 2 and 3). Over these 14 positions, hydrophobic residues are nearly absolutely conserved among α homologs from eukaryotes, such as proteins related to *Saccharomyces cerevisiae* Rpb3 and Rpb11 (Fig. 2).

The asymmetric unit of the crystal contains four crystallographically independent αNTD monomers. The overall fold of each monomer is the same, but comparison of different monomers reveals that the linkage between domain 1 and 2 within a monomer is highly flexible. The flexibility can be de-

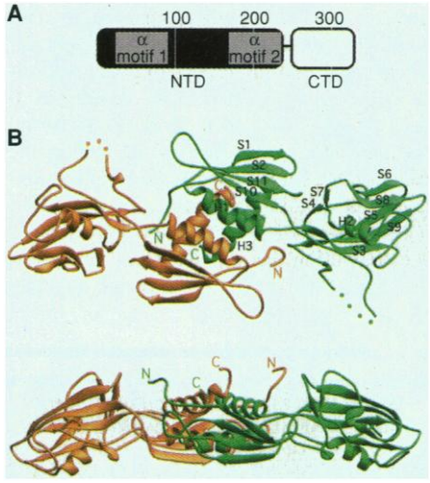

Fig. 1. Structure of the αNTD dimer. (**A**) Schematic diagram showing the domain structure of *E. coli* RNAP α (*2*). The black box indicates the NTD crystallized in this study (α residues 1 to 235). The gray boxes denote regions conserved in sequence between α homologs of prokaryotic, archaebacterial, chloroplast, and eukaryotic RNAPs. (**B**) RIBBONS (*28*) diagram of the three-dimensional structure of the αNTD dimer. One αNTD monomer is colored green and the other is yellow. Unmodeled, disordered regions are indicated as dotted lines. (**Top**) View along the dimer twofold axis; (**bottom**) view perpendicular to the dimer twofold axis.

Table 1. Summary of crystallographic analysis. αNTD_{R45A} (*19*) (15 mg/ml) was crystallized by vapor diffusion against 5 mMβ-mercapotethanol, 0.2 M $MgCl_2$, 100 mM tris-HCl, 100 mM NaCl (pH 8.6 to 8.9), and 18 to 22% polyethylene glycol 400 (PEG400) at 4°C. Heavy-atom derivatives were prepared by soaking crystals for 4 hours in 1 mM $HgCl_2$, 12 hours in 1 mM MetHgCl and $MetHgPO_4$, 24 hours in 5 mM K_2PtCl_4, and 1 week in 10 mM UO_2Ac_2, all dissolved in the crystallization solution. For cryocrystallography, crystals were soaked in steps of increasing PEG400 concentration (2% each step every 30 min) into 40% PEG400 before flash-freezing. Data were collected in the laboratory or at National Synchrotron Light Source beamline X4a (native II, $HgCl_2$ derivative, and Se derivative only) on an *R*-axis IV area detector, and processed with DENZO and SCALEPACK (Z. Otwinowski and W. Minor). Mercury positions were located manually by Patterson methods with PHASES (*20*) and confirmed with HEAVY (*21*) and SHELX-90 (*22*) Additional heavy-atom sites were located in cross-phased difference Fouriers. Heavy-atom parameter refinement and solvent flattening were carried out with PHASES. By use of this initial electron density map, all four copies of domain 1 within the asymmetric unit were nearly completely modeled, whereas only partial polyalanine models of the four copies of domain 2 could be constructed. Three noncrystallographic symmetry (NCS) operators were determined mapping copies 2, 3, and 4 of domain 1 onto copy 1, and the partial models of domain 2 were similarly determined, for a total of six NCS operators. The map was improved by NCS averaging between the four copies of domain 1 and separately between the four copies of domain 2 in the asymmetric unit with DM (*23*). Map interpretation and model building were done with the program O (*24*). The map was improved by cycles of refinement with X-PLOR (*25*) with NCS constraints (six total constraints, as above), and phase combination with SIGMAA (*26*). A final refinement was performed with relaxed NCS restraints. The final model contains residues 1 to 161 and 165 to 232 for molecule one, residues 1 to 159 and 165 to 235 for molecule two, residues 1 to 159 and 168 to 235 for molecule 3, and residues 1 to 160 and 165 to 235 for molecule four. Water molecules were added if they had at least one hydrogen bond with a protein atom or with other waters, and they were kept after refinement if the *B* factor remained below 40 $Å^2$. A total of 250 water molecules were added in the final refinement. Stereochemical values are all within or better than the expected range for 2.5 Å structure, as determined with PROCHECK (*27*). The coordinates have been submitted to the Brookhaven Protein Data Bank.

Diffraction data and MIR statistics

Crystal*	Resolution (Å)	R_{merge}† (%)	Reflections (measured/unique)	Completeness (%)	R_{Cullis}‡	Phasing power§	Number of sites
Native I	3.0	9.0	63816/17832	94.1			
Native II	2.5	6.0	145903/27157	86.0			
$HgCl_2$	3.0	6.9	55489/16625	86.5	0.63	1.67	4
MetHgCl	3.3	9.0	56107/13385	94.0	0.56	1.63	4
$MetHgPO_4$	3.3	11.0	38815/12104	85.7	0.59	1.59	4
UO_2Ac_2	3.5	12.0	39545/10708	91.1	0.73	1.40	9
K_2PtCl_4	3.5	10.0	45383/9839	83.4	0.64	1.39	4
Se Met	4.0	9.7	52982/10640	87.1	0.73	1.21	12
$HgCl_2$ ano	3.0	5.1	56062/28901	79.0	0.35	1.86	4

Overall MIR figure of merit = 0.669 (for 13831 phased reflections)

Refinement (native II)

Resolution (Å)	6–4.36	3.72	3.34	3.08	2.88	2.73	2.60	2.50	Total
No. of reflections	3256	3539	3412	3268	3023	2438	1846	1314	23574
R_{factor}‖	20.00	19.00	21.80	23.45	24.63	27.14	27.88	28.52	22.23
Free *R* factor¶	28.03	27.34	28.57	28.12	36.01	35.97	33.96	36.12	30.20

rms deviations: bonds, 0.009 Å; angles, 1.7°

*Crystal spacegroup *R*32; a = 117.2 Å, c = 350.4 Å; 4 αNTD monomers/asymmetric unit. †$R_{merge} = 100 \times \Sigma_j |Ij - \langle I \rangle| / \Sigma_j Ij$ with Bijvoet pairs treated as equivalent. ‡$R_{Cullis} = \Sigma |F_{deriv} - F_{nat}| - |F_{Hcalc}| / \Sigma |F_{deriv} - F_{nat}|$, summed over centric reflections. §Phasing power = $\langle F_H \rangle / E_{rms}$ or $\langle F_A \rangle / E_{rms}$ for isomorphous and anomalous differences. ‖*R* factor $= \Sigma |F_{obs} - F_{calc}| / \Sigma F_{obs}$ for all amplitudes with $F/\sigma(F) \geq 2$ measured in the indicated resolution bin. ¶Free *R* factor was calculated with 5% of the data in each bin.

scribed as a hinge motion with a range in all directions of at least 15° (*14*), with the hinge centered in the vicinity of Pro^{52}/Gly^{53} (in the strand connecting domain 1 to domain 2) and Pro^{179} (in the connection from domain 2 back to domain 1). The high conservation of Pro and Gly residues at these positions (Fig. 2) suggests that the flexible linkage between domains 1 and 2 is an evolutionarily conserved feature, implying some functional consequence.

Mutagenesis and hydroxyl-radical protein footprinting studies have localized determinants in α that are important for interactions with β and β′, all within α motifs 1 and 2 of the NTD (Fig. 2) (*2*, *6*, *8*–*12*). These are mapped onto the αNTD structure in Fig. 4. It is not known at present how these interactions with β and β′ are distributed between the αNTD monomers. For clarity, the data are presented such that the determinants involved in β interactions are shown on one monomer and the determinants involved in β′ interactions are shown on the other. Of particular interest are two point substitutions, at positions 45 and 48, that cause defects in β binding (*10*). Both of these positions lie on the solvent-exposed face of H1 within the region most strongly protected from hydroxyl-radical cleavage by binding of β (Fig. 4). These observations together strongly support the conclusion that this exposed face of H1, and Arg^{45} and Leu^{48} in particular, directly interact with β. A two–amino acid insertion at position 80 also results in defective β binding (*9*), and this site lies immediately adjacent to a second region of the β footprint on α (Fig. 4).

Small insertions at positions 108 and 200 cause defects in β′ binding without affecting assembly of $\alpha_2\beta$ (*11*). These two sites fall within the regions of α protected from hydroxyl-radical cleavage by the binding of β′. Two point substitutions at positions 86 and 173 also interfere wtih β′ binding (*10*). Although far apart in the sequence, these two sites fall close to each other in the αNTD structure (Fig. 4). However, these sites are far from the regions footprinted by β′ binding,

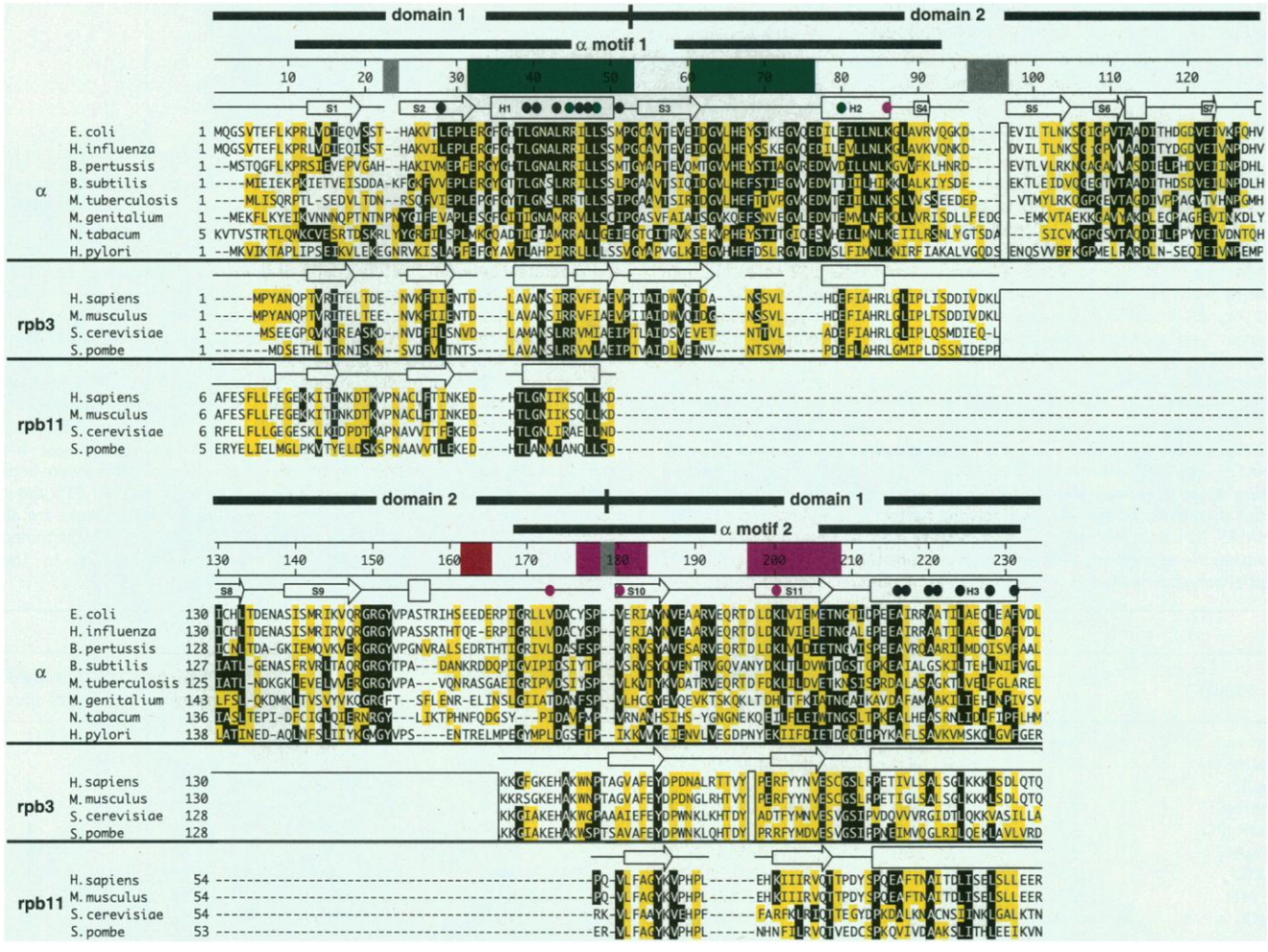

Fig. 2. Partial results of a sequence alignment of α homologs from bacteria and chloroplasts, and eukaryotic Rpb3 and Rpb11 proteins. Numbers at the beginning of each line indicate amino acid positions relative to the start of each protein sequence. Numbers along the bar on top indicate the amino acid position in *E. coli* α. Amino acid identity >50% in the full alignment is indicated by a black background, similarity >50% is indicated by a yellow background. Gaps are indicated by dashed lines, insertions by boxed out regions. The secondary structure of *E. coli* αNTD is indicated schematically above the α sequences, helices H1 to H3 are indicated by rectangles (α helices are labeled, some 3/10 helices are shown but not labeled), β strands S1 to S11 are indicated by arrows, loops are indicated by a black line. The colored areas in the amino acid numbering bar above the α sequences denote regions of the *E. coli* α sequence protected from hydroxyl-radical cleavage by β (green) or β′ (magenta) (*12*), and a region that interacts with CAP at class II CAP-dependent promoters (red) (*15*). Green and magenta dots in the diagram of α secondary structure denote mutations that cause defects in β or β′ binding, respectively (*10*, *11*). The black dots indicate residues participating in the hydrophobic core of the α dimer interface. Shown schematically above the Rpb3 and Rpb11 sequences is the predicted secondary structure (*18*). Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

and one of the substitutions, V173A, replaces a highly conserved buried hydrophobic residue with a less bulky residue, which might be expected to cause a structural perturbation or destabilization, thereby making interpretation of this pair of substitutions in terms of specific effects on β′ binding less certain.

Despite any caveats from the above considerations, it is clear that all of the regions of the α peptide backbone that are protected from hydroxyl-radical cleavage by the presence of β and β′, and all of the mutants that deleteriously affect the binding of β or β′, are exposed on one face of the αNTD dimer (Fig. 4). On the opposite face, sites known to interact with the other RNAP subunits are not found, and located on this face are the COOH-termini of the two αNTD monomers (Fig. 4). Thus, the αCTDs and the β and β′ subunits are located on opposite faces of the αNTD structure.

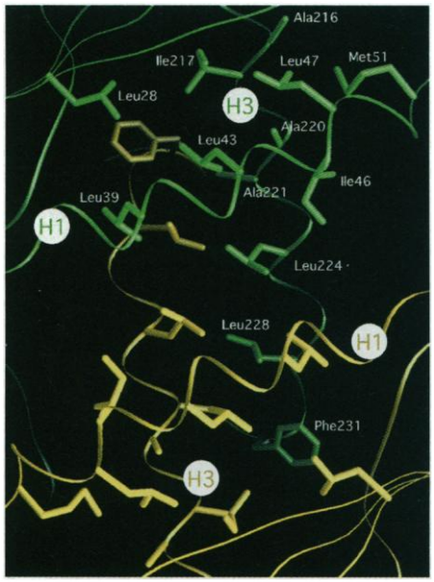

Fig. 3. Dimer interface of αNTD. Ribbon model, viewed along the dimer twofold axis, showing the conserved residues that form the hydrophobic core of the dimer interface (labeled on one monomer only). One αNTD monomer is colored green, and the other is yellow. Helices H1 and H3 of each monomer are labeled. The figure was made with the program GRASP (*29*).

Although the αCTD has been identified as the target for a wide array of transcription activators (*4*), at least one interaction between an activator (catabolite activator protein, or CAP) and αNTD, which is essential for activation at class II CAP-dependent promoters, has been identified. The protein-protein interactions between CAP and αNTD occur between the basic activating region 2 of CAP and a stretch of four acidic residues, Glu^{162}-Glu^{163}-Asp^{164}-Glu^{165}, of αNTD (*15*). This region of the αNTD structure comprises a highly exposed loop (Fig. 4, shown on only one αNTD monomer). A short stretch of residues (160 to 163) in this region is disordered in the crystal structure.

The largest subunits of prokaryotic RNAPs (β′ and β) exhibit strong sequence conservation with homologs in eukaryotic RNAPs (*16*). Less obvious evolutionary relationships have been proposed between α and two families of eukaryotic RNAP subunits related to *S. cerevisiae* Rpb3 and Rpb11, and recent studies suggest that an Rpb3-Rpb11 heterodimer serves as the eukaryotic analog of the prokaryotic α_2 homodimer (*17*). A sequence alignment of 5 Rpb3 homologs, 5 Rpb11 homologs, 17 prokaryotic α sequences, and 6 chloroplast α sequences was performed, taking into consideration the predicted secondary structure for the eukaryotic proteins (*18*) and the structure of *E. coli* αNTD. The sequence and predicted secondary structure of the Rpb3 and Rpb11 homologs align very well with sequences of α corresponding to domain 1. With one exception, gaps or insertions occur only in exposed loops between secondary structural elements and are expected to be compatible with the αNTD fold. The exception is one large gap in the Rpb11 homologs that corresponds to domain 2, which is completely lacking. This also seems to be compatible with the αNTD fold because the COOH-terminal end of the first domain 1 fragment (the COOH-terminal end of H1) is less than 6 Å from the NH_2-terminal end of the second domain 1 fragment (near the NH_2-terminus of S10). These considerations, combined with the observation already noted that all of the hydrophobic residues that comprise the hydrophobic core of the α dimer interface are nearly absolutely conserved in the Rpb3 and Rpb11 homologs, support the suggestion that an Rpb3-Rpb11 heterodimer plays the role of the α_2 homodimer in prokaryotes. The Rpb3-Rpb11 dimer interface is predicted to be structurally very similar to the α_2 dimer interface. Rpb3 is predicted to have a two-domain architecture like that of α, with domain 1 structurally related to α domain 1, whereas the structure of Rpb3 domain 2 diverges from that of α. Rpb11 is predicted to be a single domain with a fold closely related to that of α domain 1.

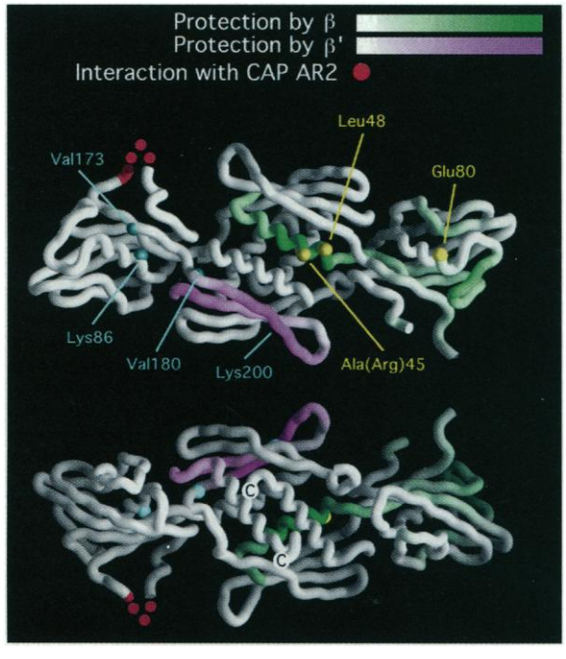

Fig. 4. Protein-protein interactions with αNTD. **(Top)** Backbone representation of the αNTD dimer viewed along the dimer twofold axis as in Fig. 1B. Backbone residues are color-coded according to the hydroxyl-radical footprinting data of (*12*), so that regions protected from hydroxyl-radical cleavage by β or β′ are colored green or magenta, respectively. Shown in yellow or light blue are the α-carbon positions of mutations that cause defects in β or β′ binding, respectively (*10*, *11*). The region of αNTD found to interact with CAP AR2 at class II CAP-sites (αNTD residues 162 to 165) are shown in red (*15*). **(Bottom)** View along the dimer twofold axis from the opposite direction as the top view. The COOH-termini of the two α-NTD monomers are indicated. The figure was made with the program GRASP (*29*).

References and Notes

1. W. Zillig, P. Palm, A. Heil, in *RNA Polymerase*, R. Losick and M. Chamberlin, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1976), p. 101; A. Ishihama, *Adv. Biophys.* **14**, 1 (1981).
2. K. Igarashi and A. Ishihama, *Cell* **65**, 1015 (1991); E. Blatter, W. Ross, H. Tang, R. Gourse, R. Ebright, *ibid.* **78**, 889 (1994).
3. W. Ross *et al.*, *Science* **262**, 1407 (1993).
4. Reviewed in A. Ishihama, *J. Bacteriol.* **175**, 2483 (1993).
5. Y. H. Jeon *et al.*, *Science* **270**, 1495 (1995).
6. R. S. Hayward, K. Igarashi, A. Ishihama, *J. Mol. Biol.* **221**, 23 (1991).
7. R. H. Ebright and S. Busby, *Curr. Opin. Genet. Dev.* **5**, 197 (1995).
8. K. Igarashi, N. Fujita, A. Ishihama, *Nucleic Acids Res.* **18**, 5945 (1990); *J. Mol. Biol.* **218**, 1 (1991).
9. M. Kimura, N. Fujita, A. Ishihama, *J. Mol. Biol.* **242**, 107 (1994).
10. M. Kimura and A. Ishihama, *ibid.* **254**, 342 (1995).
11. ———, *ibid.* **248**, 756 (1995).
12. T. Heyduk, E. Heyduk, K. Severinov, H. Tang, R. H. Ebright, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10162 (1996).
13. The R45A substitution increases the stability of the α dimer (which may have contributed toward obtaining high-quality crystals) but has a decreased affinity for the RNAP β subunit (*10*). However, αNTD_{R45A} is fully competent to assemble active RNAP under appropriate conditions, indicating that it is folded correctly [Y. Wang *et al.*, *J. Mol. Biol.* **270**, 648 (1997)].

14. The range of flexibility observed for domain 2 with respect to domain 1 could be even greater than 15° because it may be restricted by crystal packing.
15. W. Niu, Y. Kim, G. Tau, T. Heyduk, R. H. Ebright, *Cell* **87**, 1123 (1996).
16. J. Archambault and J. D. Friesen, *Microbiol. Rev.* **57**, 703 (1993).
17. Y. Azuma, M. Yamagishi, A. Ishihama, *Nucleic Acids Res.* **21**, 3749 (1993); D. Lalo, C. Carles, A. Sentenac, P. Thuriaux, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 5524 (1993); U. K. Pati, *Gene* **145**, 289 (1994); T. Ulmasov, R. M. Larkin, T. J. Guilfoyle, *J. Biol. Chem.* **271**, 5085 (1996); V. Svetlov, K. Nolan, R. R. Burgess, *ibid.* **273**, 10827 (1998).
18. B. Rost and C. Sander, *J. Mol. Biol.* **232**, 584 (1993); _____ and R. Schneider, *Comput. Appl. Biosci.* **10**, 53 (1994); B. Rost and C. Sander, *Proteins* **19**, 55 (1994).
19. The αNTD (residues 1 to 235) with a point mutation Arg^{45}→Ala (αNTD_{R45A}) was subcloned by polymerase chain reaction into the T7 RNAP-based PET-15b expression vector (Novagen) for overexpression of NH_2 terminally expressed hexahistidine-tagged (His_6) protein. Using *E.coli* BL-21(DE3) host cells, we achieved excellent expression (>20 mg of protein per liter of cell culture after purification to homogeneity from the soluble fraction). After cell lysis through a continuous-flow French press and a low-speed spin, the soluble fraction was loaded onto a Ni^{2+}-chelating affinity column, and His_6-αNTD was eluted with 100 mM imidazole. The protein was then treated overnight at 4°C with thrombin to quantitatively remove the His_6-tag, then was loaded onto a MonoQ (Pharmacia) ion-exchange column. After elution with a NaCl gradient, the protein was homogeneous as judged by Coomassie-stained SDS-polyacrylamide gels.
20. W. Furey and S. Swaminathan, *Methods Enzymol.* **277**, 590 (1997).
21. T. C. Terwilliger, S. H. Kim, D. E. Eisenberg, *Acta Crystallogr.* **A43**, 34 (1987).
22. G. M. Sheldrick, in *Isomorphous Replacement and Anomalous Scattering, Proc. CCP4 St. W.*, W. Wolf, P. R. Evans, A. G. W. Leslie, Eds. (Daresbury Laboratory, Warrington, UK, 1991), p. 23.
23. K. Cowtan, *ESF/CCP4 Newslett.* **31**, 34 (1994).
24. T. A. Jones, J.-Y. Zou, S. Cowan, M. Kjeldgaard, *Acta Crystallogr.* **A47**, 110 (1991).
25. A. Brünger, *X-PLOR (Version 3.1) Manual* (Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT, 1992).
26. R. Read, *Acta Crystallogr.* **A42**, 140 (1986).
27. R. A. Laskowski, M. W. MacArthur, D. S. Moss, J. M. Thornton, *J. Appl. Crystallogr.* **26**, 283 (1993).
28. M. Carson, *ibid.* **24**, 958 (1991).
29. A. Nicholls, K. A. Sharp, B. Honig, *Proteins Struct. Funct. Genet.* **11**, 281 (1991).
30. We are indebted to Y. Wang for cloning αNTD_{R45A} and for preliminary crystallization trials. We thank R. Ebright for plasmid DNA encoding *E. coli* RNAP α, T. Muir for assistance with mass spectroscopy, J. Bonanno for assistance with x-ray measurements, and W. A. Hendrickson and C. Ogata for access to and support at beamline X4A, which is funded by the Howard Hughes Medical Institute at the Brookhaven National Laboratory. We thank L. Niu, J. Goldberg, and D. Jeruzalmi for suggestions on refinement and data analysis. G.Z. was supported in part by a National Research Service Award (NIH GM19441-01). S.A.D. is a Pew Scholar in the Biomedical Sciences. This work was supported in part by grants to S.A.D. from the Irma T. Hirschl Trust, the Pew Foundation, and NIH (GM53759).

27 March 1998; accepted 7 May 1998

Cell Cycle Arrest by Vpr in HIV-1 Virions and Insensitivity to Antiretroviral Agents

Betty Poon, Kathie Grovit-Ferbas, Sheila A. Stewart, Irvin S. Y. Chen*

Expression of human immunodeficiency virus–type 1 (HIV-1) Vpr after productive infection of T cells induces cell cycle arrest in the G_2 phase of the cell cycle. In the absence of de novo expression, HIV-1 Vpr packaged into virions still induced cell cycle arrest. Naturally noninfectious virus or virus rendered defective for infection by reverse transcriptase or protease inhibitors were capable of inducing Vpr-mediated cell cycle arrest. These results suggest a model whereby both infectious and noninfectious virions in vivo, such as those surrounding follicular dendritic cells, participate in immune suppression.

The HIV-1 *vpr* gene encodes a 14-kD nuclear protein (Vpr) that is expressed within infected cells and is packaged into virions (*1*). Vpr is nonessential for viral replication in T cell lines and activated peripheral blood lymphocytes (PBLs) in vitro but it is necessary for efficient infection of nondividing cells such as macrophages (*2, 3*). Simian immunodeficiency virus–induced disease progression in macaques was attenuated when either *vpr* alone or both *vpr* and the related gene *vpx* were mutated, indicating that Vpr plays an important role in viral pathogenesis (*4*). Functions ascribed to Vpr include transport of the viral core into the nucleus of nondividing cells (*3*) and up-regulation of viral gene expression (*5*). One particularly intriguing function of Vpr is the ability to induce cell cycle arrest at the G_2 checkpoint in a variety of mammalian cells, including human PBLs (*6–8*). This cell cycle arrest is characterized by alterations in the activation and phosphorylation state of Cdc2 kinase (*7, 8*) and resembles the G_2 checkpoint induced by genotoxic agents (*9*). Over time, virally infected cells arrested in the G_2 phase by Vpr die by apoptosis (*10*).

In previous studies of Vpr-mediated cell cycle arrest, we used HIV-1 bearing the murine Thy 1.2 reporter gene to allow concurrent visualization of infected cells by cell surface staining for Thy 1.2 and cell cycle status by staining with propidium iodide (*9, 10*). In some experiments, there was significant cell cycle arrest even in cells that were not productively infected (Thy 1.2^-), particularly at early time points before detectable Thy 1.2 expression (*11*). We tested the possibility that virion-associated Vpr can mediate cell cycle arrest by generating a vesicular stomatitis virus G protein (VSV-G) pseudo-typed HIV-1 that carried Vpr but was incapable of synthesizing it because of a frameshift mutation within *vpr* [$HIV\text{-}1_{NL4\text{-}3\text{-}thy}$ env(-)-vprX]. Epitope-tagged Vpr was supplied to the virion by cotransfection, resulting in trans-complementation [$HIV\text{-}1_{NL4\text{-}3\text{-}thy}$ env(-)vprX + BSVpr] (Fig. 1A) (*12*). Virion particles formed in this fashion contained Vpr in amounts comparable to wild-type virus [$HIV\text{-}1_{NL4\text{-}3\text{-}thy}$ env(-)]. Infection with virions complemented in trans with Vpr (VprX mutant + Vpr in trans) resulted in cell cycle arrest in the Thy 1.2^+ population (G_1/G_2 = 0.26) compared with the Thy 1.2^- population (G_1/G_2 = 1.8) and with infection by the *vprX* mutant (G_1/G_2 = 1.9) (Fig. 1B). The level of arrest induced by virion-associated Vpr alone was typically less than that observed with wild-type virus capable of de novo Vpr production (Fig. 1B), probably because of higher levels of Vpr expressed in cells de novo. Similar results were observed after infection of primary human PBLs and SupT1 T cells with an HIV-1 vector that packages Vpr but whose genome is defective for de novo expression of all HIV-1 genes (*13*). Virions pseudotyped with a CXCR-4 tropic HIV-1 envelope also induced cell cycle arrest (*14*) (Fig. 2B). The extent of G_2 arrest induced by virion-associated Vpr varied depending on the virion preparation and cell type. However, HIV-1 particles prepared without envelope, but still containing Vpr within the viral core, did not induce G_2 arrest (*15*). Thus, cell cycle arrest required both Vpr and viral entry into cells and was not the result of contaminating soluble Vpr in our virion preparations. Significantly, the cell cycle arrest observed with virion-associated Vpr was observed at low multiplicities of infection (MOI = 0.15; Fig. 1B) (*16*).

B. Poon, S. A. Stewart, I. S. Y. Chen, Department of Microbiology and Immunology and Medicine, UCLA AIDS Institute, and Jonsson Comprehensive Cancer Center, UCLA School of Medicine, Los Angeles, CA 90095, USA. K. Grovit-Ferbas, Department of Medicine, UCLA AIDS Institute, UCLA School of Medicine, and West Los Angeles VAMC, Los Angeles, CA 90095, USA.

*To whom correspondence should be addressed.