ARTICLE

Why Is the Temperature of the Universe 2.726 Kelvin?

Michael S. Turner

The Cosmic Background Explorer satellite has recently made the most accurate measurement of the temperature of the universe, determining it to be 2.726 ± 0.01 kelvin. In trying to understand why the temperature has this value, one is led to discover the most fundamental features of the universe—an early, radiation-dominated epoch, enormous entropy per nucleon, synthesis of the light elements around 3 minutes after the bang, and a small excess of matter over antimatter—as well as some of the most pressing issues in cosmology today—the development of structure in the universe and the identification of the nature of the ubiquitous dark matter.

The existence of the cosmic background radiation (CBR) is one of the cornerstones of the standard cosmology, or hot big bang model (*1*). Indeed, its very existence provides the evidence that the universe began from a hot state (*2*). The temperature of the CBR has recently been measured to unprecedented precision by the Far-Infrared Absolute Spectrophotometer (FIRAS) instrument on the National Aeronautics and Space Administration's (NASA's) Cosmic Background Explorer (COBE) satellite (*3*)

$$T_0 = 2.726 \pm 0.01 \text{ K} \qquad (1)$$

Measurements of the CBR temperature, made over the period of almost 30 years since its discovery by Penzias and Wilson (*4*), now span almost three and a half decades in wavelength, from about 0.04 to 70 cm, and are all consistent with the COBE temperature. Deviations from a perfect blackbody spectrum are less than 0.03% over the wavelengths probed by COBE, 0.05 to 0.5 cm (Fig. 1) (*3*). The CBR is probably the most well studied and best blackbody known; indeed, the COBE collaboration plans to use their data to test the form of the Planck law itself (*5*).

With a number density of 411 cm^{-3}, the photons in the CBR by a wide margin account for most of the (known) particles in the universe, outnumbering atoms by a factor of around a billion. The surface of last scattering for the CBR is the universe itself at an age of a few 100,000 years (Fig. 2), and thus, the CBR provides a fossil record of the infant universe. As such, its every property has been studied—spectrum, polarization, and spatial isotropy—revealing important information about the evolution of the universe (*6*). Just trying to answer the simple question, "Why is the temperature of the CBR 2.726 K?" reveals the most fundamental features of the universe, as well as the most pressing problems in cosmology.

To begin, it is imprecise to say that the universe has a temperature, as it is not in thermal equilibrium today. At an age of less than a few 100,000 years, the matter was ionized and a state of thermal equilibrium existed; at about this time, the temperature was about 3000 K, and the equilibrium ionization fraction of matter became very small. The universe is said to have "recombined"; because neutral matter is transparent to the radiation, the CBR photons we detect today last scattered a few 100,000 years after the bang. After last scattering, the expansion simply redshifted the energy of CBR photons and diluted their number density; because of a remarkable feature of the expansion, a Planck distribution was maintained with a temperature that decreased in proportion to the size of the universe. For this reason, the universe today is filled with thermal radiation of temperature 2.726 K despite the fact that the universe is no longer in thermal equilibrium.

Because the temperature of the universe is decreasing—and has been for some 15 billion years or so—the original question must be rephrased: Why did the temperature of the universe reach about 3 K at an age of about 15 billion years old? [Several independent measures of the age, based on the evolution of stars in the oldest globular clusters, the cooling of the oldest white dwarfs in the galaxy, and the dating of certain radioactive isotopes, indicate that the universe is between 12 and 18 billion years old (*7*).]

According to Einstein's equations, the present age of the universe t_0, that is, time since the bang, is related to the present energy density ρ_0

$$t_0 = \frac{c}{\sqrt{6\pi G\rho_0}} \qquad (2)$$

where $G = 6.67 \times 10^{-8}$ $\text{cm}^3\ \text{s}^{-2}\ \text{g}^{-1}$ is Newton's gravitational constant, $c = 3.00 \times 10^{10}$ cm s^{-1} is the speed of light, and for simplicity, I have assumed that the universe is spatially flat ($\Omega_0 = 1$). The quantity $\Omega_0 = \rho_0/\rho_{\text{crit}}$ is the ratio of the mean density to the critical, or closure, energy density; $\rho_{\text{crit}} = 3H_0^2 c^2/8\pi G$ corresponds to a mass density of 1.88×10^{-29} $(H_0/100 \text{ km s}^{-1} \text{ Mpc}^{-1})^2$ g cm^{-3}, and $H_0 = 40$ to 100 km s^{-1} Mpc^{-1} is Hubble's constant, whose value is still only known to within a factor of 2. "Low-density" universes, $\Omega_0 < 1$, are negatively curved and expand forever, whereas "high-density" universes, $\Omega_0 > 1$, are positively curved and eventually recollapse. The "critical" universe, $\Omega_0 = 1$, is spatially flat and also expands forever. In the general case, $t_0 = \sqrt{3\Omega_0 f^2(\Omega_0)c^2/8\pi G\rho_0}$, where the function $f(\Omega_0)$ varies between 1 and 2/3 for Ω_0 between 0 and 1.

We know at least one component of the energy density today: the CBR blackbody radiation itself, which contributes an energy density

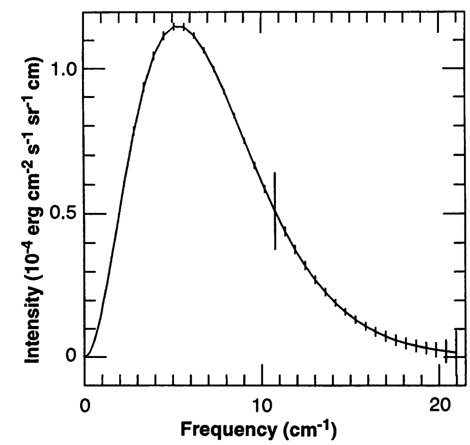

Fig. 1. The COBE FIRAS measurements of the CBR spectrum (*3*) and the spectrum of a 2.726 K blackbody (curve). The frequency values have been divided by the speed of light. Note the COBE 1σ error flags have been enlarged by a factor of 100.

The author is in the Departments of Physics and of Astronomy and Astrophysics, Enrico Fermi Institute, The University of Chicago, Chicago, IL 60637–1433, and Theoretical Astrophysics, Fermi National Accelerator Laboratory, Batavia, IL 60510–0500.

$$\rho_{CBR} = \frac{\pi^2 k_B^4 T_0^4}{15\hbar^3 c^3} \simeq 4.18 \times 10^{-13} \text{ erg cm}^{-3} \quad (3)$$

where $\hbar = 1.05 \times 10^{-27}$ erg s is Planck's constant divided by 2π, $k_B = 1.38 \times 10^{-16}$ erg K^{-1} is Boltzmann's constant, and $\pi^2 k_B^4/15\hbar^3 c^3 = 7.56 \times 10^{-15}$ erg cm^{-2} s^{-1} K^{-4} is four times the Stefan-Boltzmann radiation constant divided by the speed of light. If the CBR were the only contribution to the energy density, Eq. 2 would imply an age of about 1300 billion years, a factor of about 100 too large. Put another way, for the age to be consistent with the energy density in the CBR alone, the temperature would have to be closer to 30 K.

Matter in the Universe

By asking a simple question, we have learned that the CBR blackbody radiation must make a minor contribution to the energy density today, $\rho_{CBR} \sim 10^{-4}\ \rho_0$. What then accounts for the bulk of the present energy density? It could exist in other thermal backgrounds of relativistic particles; however, that would require the existence of several thousand additional massless particle species, and we know of at most three, the electron, muon, and tau neutrinos, which together contribute an energy density comparable to that of the CBR (provided all three neutrino species are massless or nearly massless).

It is almost certain that the bulk of the energy density exists in the form of nonrelativistic matter (8). Taking the age of the universe to be 15 billion years, Eq. 2 implies a matter density of about 3×10^{-30} g cm^{-3} (energy density of about 3×10^{-9} erg cm^{-3}). Today, the energy density in matter is more than 10,000 times greater than that in the CBR, but that was not always the case. As the universe expands, the matter density decreases as R^{-3}, the factor by which the volume increases; $R(t)$ is the cosmic scale factor, which describes the linear expansion of the universe. The energy density in radiation decreases faster, as R^{-4}, because the energy of each photon is also redshifted by the expansion, accounting for the additional factor of R^{-1}. Owing to the different scalings of the matter and radiation energy densities, when the universe was about 10^{-4} of its present size and a few thousand years old, the two energy densities were equal. Earlier than this, the energy density in radiation exceeded that in matter, and the universe is said to have been "radiation dominated."

Early on, matter was a trace constituent in a universe dominated by a hot plasma of thermal particles; at the earliest times, $t << 10^{-5}$ s, the hot plasma was a soup of the fundamental particles—quarks, leptons, and gauge bosons (the photon, $W^{\pm}$ and Z^0, and gluons, the carriers of the forces). This is an extremely important feature of the universe that has profound implication for the study of its earliest history. Among other things, it means that the formation of structure in the universe—galaxies, clusters of galaxies, voids, superclusters, and so on—through the gravitational amplification of small inhomogeneities in the matter density began only a few thousand years after the bang (9). This is because during the radiation-dominated phase, the self-gravitational attraction of the matter was no match for the rapid expansion driven by the enormous energy density in radiation, and density perturbations could not grow (Fig. 3).

A year ago, another instrument on the COBE satellite, the Differential Microwave Radiometer (DMR), detected tiny differences in the CBR temperature measured in different directions, on average about a part in 10^5 (or 30 μK) between directions separated by 10° (Fig. 4) (10). Inhomogeneities in the matter density give rise to temperature variations of a similar size, and this COBE discovery provided evidence for the existence of the primeval density fluctuations that seeded all the structure in the universe. Moreover, because density fluctuations grow in proportion to the cosmic scale factor and the level of inhomogeneity exceeds unity today (that is, $\delta\rho/\rho > 1$), the amplitude of the primeval fluctuations needed to seed the observed structure is set roughly by the size of the scale factor when the matter and radiation energy densities were equal, about 10^{-5} or so, a number that is determined by the present ratio of the energy density in radiation to that in matter. In a very real sense, the CBR temperature set the amplitude of temperature fluctuations that were expected.

The extreme uniformity of the temperature of the CBR across the sky, to better than a part in 10^4 on angular scales from arcminutes to 180° (Fig. 4) (11), reveals an important property of the universe—its smoothness, or isotropy and homogeneity—and raises another question—why is it so smooth? Though the universe was very small at early times, its rapid expansion limited the distance over which even photons could travel. At the epoch of last scattering, this distance, known as the horizon distance, corresponded to an angle of only about 1° on the sky (Fig. 2); this fact precludes any causal physical process from explaining the temperature uniformity, and hence, the smoothness of the universe, on angular scales greater than 1°. Further, it raises the same question about the origin of the primeval density inhomogeneities; they too could not have been created on such large distances by causal processes operating at early times.

The smoothness and the primeval inhomogeneity needed to seed structure could have existed since the beginning. However, Guth showed that both can be explained by a very rapid period of expansion, called

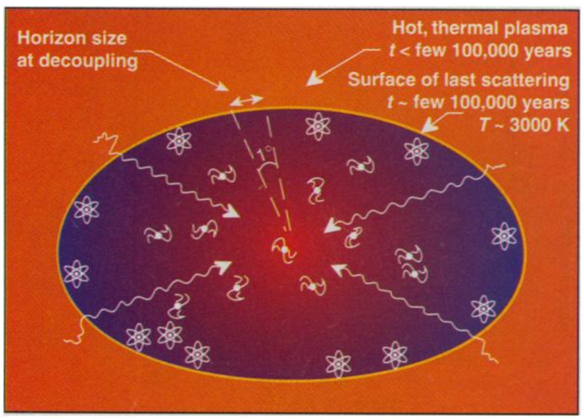

Fig. 2. A schematic diagram illustrating the last-scattering surface. Also shown is the size of the horizon at that epoch, which subtends an angle of about 1°.

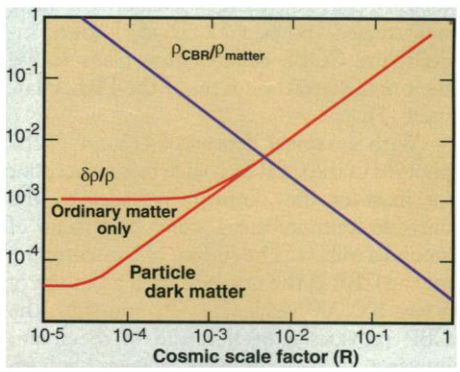

Fig. 3. Primeval density perturbations grow in proportion to the cosmic scale factor R (whose value today is taken to be one). With ordinary matter only, perturbations begin growing when matter and radiation decouple ($R \simeq 10^{-3}$); with particle dark matter, perturbations begin growing much earlier, as soon as the universe becomes matter dominated ($R \simeq 3 \times 10^{-5}$), and thus, smaller primeval density inhomogeneities are required. Also shown is the ratio of energy density in the CBR to that in matter, which decreases as R^{-1}.

cosmic inflation, that may have taken place about 10^{-34} s after the bang (*12*). This rapid expansion is driven by the false-vacuum energy (particle physics analog of latent heat) associated with a first-order phase transition. The basic idea is that a tiny patch of the universe, which could have been made smooth early on, grew exponentially to a size that encompassed all that we see today and well beyond. The enormous growth of the scale factor also allows quantum mechanical fluctuations arising during inflation on very small length scales to become density perturbations on length scales large enough to account for the primeval density inhomogeneities needed to seed the structure seen today (*13*) (the COBE DMR results are consistent with the temperature variations predicted in inflationary models, as are two other models for the origin of the density fluctuations). In addition, the tremendous growth in the size of the universe, by a factor greater than that by which the universe has grown since, also leads to a universe that, regardless of its initial curvature, today still appears flat, making $\Omega_0 = 1$ a "prediction" of inflation.

The Nucleon-to-Photon Ratio

Assuming that the present mass density exists in the form of ordinary matter, atoms made of nucleons (neutrons and protons) and electrons, a present nucleon density of about 2×10^{-6} cm^{-3} is implied. From this, we can form the dimensionless ratio of the nucleon number density to the photon number density

$$\eta \equiv \frac{n_N}{n_\gamma} \sim 5 \times 10^{-9} \qquad (4)$$

This ratio indicates that CBR photons outnumber nucleons by a factor of around a billion. The inverse of η, the ratio of photons to nucleons, measures the entropy in radiation per nucleon (in units of k_B). The radiative entropy per nucleon in a star like our sun is only about 10^{-2}; even in the highest entropy environment known, the center of a newly born neutron star, the entropy per nucleon is only a few. The universe has such an extremely high entropy that it is very difficult to imagine that any physical process could have produced the CBR or added significantly to it. Further, because the CBR spectrum is so accurately Planckian, there are severe restrictions on any process that produces photons, for example, radiation from an early generation of stars or the decay of relic neutrinos (if they are massive and unstable). The entropy per nucleon seems to be an initial condition, rather than a quantity that can be readily explained by astrophysical processes.

The nucleon-to-photon ratio η also quantifies the net excess of nucleons over antinucleons, or net baryon number, per photon. The net baryon number per photon is equal to η because there is no significant amount of antimatter in the universe today ($n_{\bar{N}} << n_N$)

$$\frac{n_B}{n_\gamma} \equiv \frac{n_N - n_{\bar{N}}}{n_\gamma} \approx \frac{n_N}{n_\gamma} \qquad (5)$$

Baryon number, like charge, is known empirically to be conserved to a high degree of precision (the longevity of the proton, lifetime greater than 10^{32} years, attests to this; were baryon number not conserved, the proton would be expected to decay in a fraction of a second). Conservation (or even approximate conservation) of baryon number and the value of η imply that earlier than about 10^{-5} s, when it was hot enough for matter and antimatter to be freely created, there was approximately one more baryon than antibaryon for every billion or so of both. Looking at it the other way around, in the absence of this tiny excess, all the baryons and antibaryons would have annihilated as the universe cooled, leaving essentially no matter or antimatter today.

Though the details have not been worked out, many believe that this excess of matter over antimatter, so crucial to the existence of matter today, evolved as a result of particle interactions in the very early universe ($\lesssim 10^{-12}$ s) that respected neither the symmetry between matter and antimatter nor the conservation of baryon number (*14*) (violation of the conservation of baryon number is an almost universal prediction of theories that attempt to unify the forces of nature, and also arises in the standard model of particle physics because of subtle quantum mechanical effects; the symmetry between matter and antimatter is violated by a small amount in the decays of the K^0, $\bar{K}^0$ mesons). Explaining the small net baryon number, quantified by η, appears to be much more promising than trying to explain the large entropy, quantified by η^{-1}.

The high entropy plays a crucial role in the determination of the chemical composition of the universe. Were the entropy per nucleon even a thousand times smaller, nuclear reactions taking place when the universe was only a fraction of a second old and the energy equivalent of the temperature k_BT was a few megaelectron volts would have quickly processed all the nucleons into tightly bound nuclei such as carbon, oxygen, and on up to iron. Instead, most of the nucleons remain in the form of protons with only the lightest isotopes, D, ^{3}He, ^{4}He, and ^{7}Li, being produced (it is generally believed that the other elements were produced in stars or spallation reactions in the interstellar medium). The lack of significant nucleosynthesis beyond the light elements traces directly to the high entropy: The enormous number of high-energy photons per nucleon delayed the onset of nucleosynthesis until a temperature of order $k_BT \sim 0.1$ MeV because earlier photons rapidly dissociated nuclei as they formed; when nucleosynthesis did begin, coulomb repulsion between light nuclei prevented their fusion into the heavier, more tightly bound nuclei [this fact was appreciated before the discovery of the CBR and led Gamow and others to predict the existence of a relic radiation with about the correct temperature (*15*)].

The predictions of primordial nucleosynthesis agree with the inferred primordial abundances of the light elements provided that the nucleon-to-photon ratio lies in the interval

$$3 \times 10^{-10} \lesssim \eta \lesssim 4 \times 10^{-10} \qquad (6)$$

The very existence of a "concordance interval" is an important test of the standard

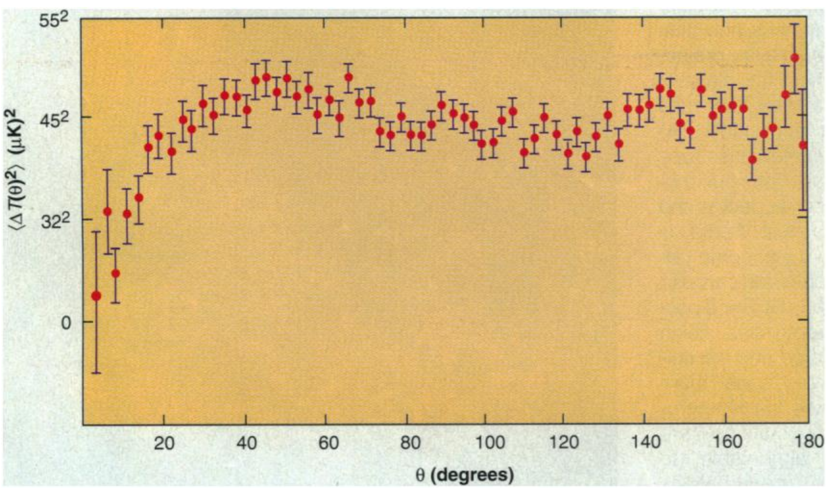

Fig. 4. The COBE DMR measurements of $\langle \Delta T(\theta)^2 \rangle$, the temperature difference squared between two points on the sky separated by angle θ and averaged over the entire sky (*10*). The much larger temperature anisotropy of about 3400 μK attributable to our motion with respect to the cosmic rest frame has been removed.

cosmology, and as a bonus, it provides the most accurate determination of the nucleon-to-photon ratio (16) (Fig. 5).

The success of the theory of primordial nucleosynthesis not only provides the earliest test of the big bang model, but it also leads to a startling suggestion: Most of the matter in the universe is something other than nucleons. From primordial nucleosynthesis and the temperature of the CBR, the mass density contributed by nucleons can be computed

$$\rho_N = m_N \eta n_\gamma \simeq 2.7 \times 10^{-31} \text{g cm}^{-3} \quad (7)$$

where $m_N \simeq 1.7 \times 10^{-24}$ g is the mass of a nucleon, $n_\gamma = 2\zeta(3)k_B^3T_0^3/\pi^2\hbar^3c^3 = 411$ cm^{-3} is the number density of photons, and $\zeta(3) = 1.20206 \cdots$. This mass density is significantly lower than the earlier estimate of the total mass density derived from the age of the universe, though to be sure, we made certain assumptions at the time. In any case, the small mass density in nucleons leads one to ask whether the mass density of the universe is greater than that contributed by ordinary matter alone.

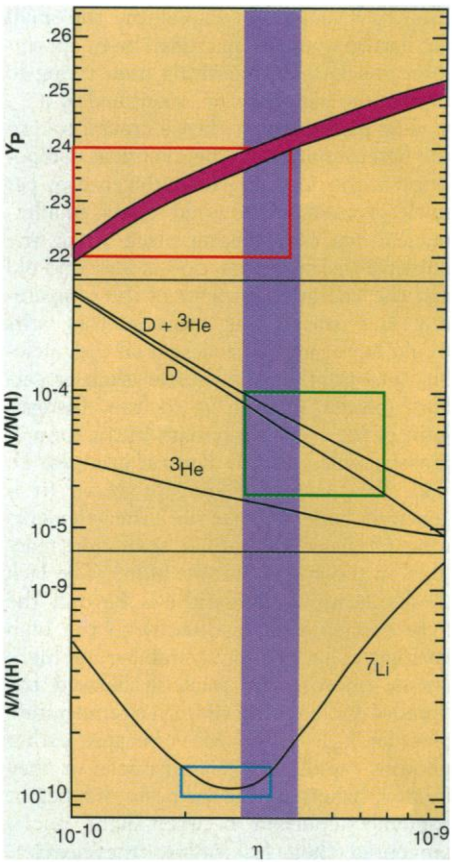

Fig. 5. The predictions of primordial nucleosynthesis and the inferred primordial abundances of D, ^{3}He, ^{4}He, and ^{7}Li (*16*). The ^{4}He abundance is the mass fraction of nucleons in ^{4}He, Y_P, and is shown on a linear scale; the thickened line indicates the theoretical uncertainty in Y_P, which is all attributable to the uncertainty in the neutron lifetime. Abundances for the other elements are given as the number of atoms per hydrogen atom and are shown on logarithmic scales. The boxes indicate the observational uncertainties in the inferred primordial abundances and the concordance intervals; the overall concordance interval (95% CL) is shaded.

Dark Matter in the Universe

Let me very briefly review what we know about the mass density of the universe (*17*). On the basis of the above determination of the density of ordinary matter and our imperfect knowledge of the Hubble constant, it follows that ordinary matter contributes between 1% and 10% of the critical density (the larger value for the lower value of the Hubble constant). From astronomical observations, we know (i) luminous matter, in the form of stars, contributes less than 1% of critical density; (ii) the gravitational effects of mass on the motion of stars in spiral galaxies (*18*), the motions of galaxies in clusters, and so on, indicate that the total amount of mass is at least 10 to 20% of the critical density (*19*); (iii) our motion with respect to the CBR suggests that the density is near critical; and (iv) no definitive measurement of the total amount of matter has yet been made.

The third point deserves further discussion; the CBR is hotter in the general direction of the constellations Hydra and Centaurus, by about 3.4 mK, and cooler in the opposite direction by the same amount (*20*) (Fig. 6). The simplest, and now standard, interpretation is that our galaxy is moving with respect to the "cosmic rest frame" at a speed of about 620 km s^{-1} (COBE detected a much smaller yearly modulation of the same kind arising from Earth's motion around the sun at 30 km s^{-1}; this should convince any remaining "geocentrists" that the Earth does indeed move!). The motion of the Milky Way arises because of the gravitational tugs exerted on it by the thousands of galaxies within a hundred megaparsecs or so. Because the distribution of galaxies is not precisely homogeneous, the sum of these tugs does not cancel but results in a net force in the direction of Hydra-Centaurus. Because the gravitational force on the Milky Way attributable to another galaxy is proportional to that galaxy's mass, an estimate for the mass in this volume, and for the average mass density, can be made by relating our velocity to the observed distribution of galaxies in this volume. This technique samples the largest volume of space of any method yet and indicates a value for Ω_0 that is close to unity (*21*).

Though our knowledge of the mass density of the universe is still incomplete, we can already conclude that (i) most of the matter in the universe is dark; that is, it does not emit or absorb radiation of any wavelength; (ii) because nucleosynthesis indicates that ordinary matter contributes more than is accounted for by stars, some of the dark matter is baryonic; (iii) if the mass density of the universe is at the lower limit of current estimates and if the density of ordinary matter is at its upper limit, then ordinary matter can account for all the mass with Ω_0 being around 0.1; (iv) on the other hand, if the mass density is significantly greater than 10% of the critical density,

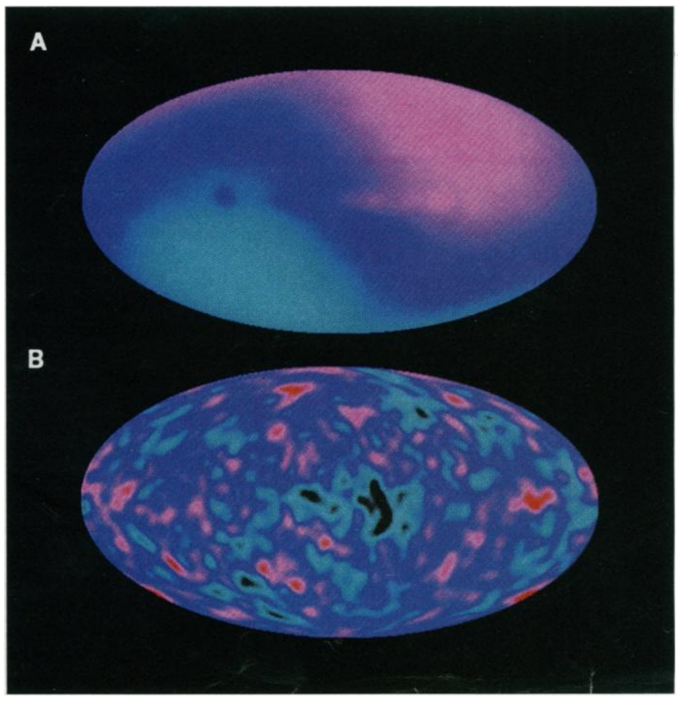

Fig. 6. COBE DMR temperature maps of the sky. The variation in the CBR temperature is represented on a color scale (pink is hot, blue is cold) on a sky projection where the plane of our the Milky Way runs across the middle. (**A**) The dipole anisotropy resulting from our motion with respect to the cosmic rest frame is clearly seen; some galactic emission can also be seen. (**B**) The dipole anisotropy has been subtracted and the color scale made more sensitive; the temperature fluctuations are partly attributable to density perturbations on the last-scattering surface and partly attributable to instrumental noise in the DMR.

then the bulk of the dark matter must be something other than ordinary matter. This possibility is favored by many cosmologists, mainly the theorists, as theoretical considerations, including cosmic inflation and theories of structure formation, argue strongly for the critical universe ($\Omega_0 = 1$). I hasten to add that the observational situation is far from settled, and many, if not most, astronomers would say that the case for $\Omega_0 = 0.1$ is the more compelling one at present.

Note the crucial role played by the CBR temperature in reaching these conclusions. The outcome of primordial nucleosynthesis depends only on the nucleon-to-photon ratio. Therefore, the primordial abundances of the light elements serve to determine η rather than the nucleon mass density itself. To determine nucleon mass density, the photon number density, and hence CBR temperature, must be known. Were the CBR temperature a factor of 3 or so higher, the mass density contributed by ordinary matter would be close to the critical density.

If most of the mass in the universe is not ordinary matter, what is it? The most promising idea is that it exists in the form of elementary particles left over from the early, fiery moments of the universe (*22*). In this case, another dimensionless ratio can be formed, the ratio of the number density of "exotic particles" to CBR photons

$$\eta_X \equiv \frac{n_X}{n_\gamma} \simeq 7 \times 10^{-9} \left(\frac{m_N}{m_X}\right) \qquad (8)$$

where m_X is the mass of the exotic particle, and for simplicity, I have assumed that exotic particles contribute critical density and a Hubble constant of 50 km s^{-1} Mpc^{-1}.

As it turns out, there are a handful of interesting candidates for the dark matter. They include a massive neutrino, the neutralino, and the axion. All three possibilities are motivated by particle physics considerations first, with their important cosmological consequence as a bonus, perhaps a hint that the particle dark matter hypothesis is on the right track.

How do these particles arise as relics of the big bang? In the early universe, thermodynamics dictated a kind of particle democracy, with all species being roughly equally abundant. As the universe cooled, pair creation of massive particles became energetically forbidden, and massive-particle species disappeared through particle-antiparticle annihilations. If a particle species is stable, it can have a significant relic abundance because, in the expanding universe, annihilations eventually cease as particles and antiparticles become too sparse to encounter one another and annihilate. The relic abundance depends on the potency of annihilations, quantified by the annihilation cross section, σ_{ann}, which has units of area.

In the case of neutrinos, annihilations became ineffective before they could start significantly reducing the neutrino abundance relative to photons, and so η_X is expected to be around one (more precisely 3/11). Thus, the contribution of neutrinos to the mass density is dictated by their mass: They contribute critical density for a mass of about $2.5 \times 10^{-8}\, m_N$, or a mass energy of about 20 eV. Such a mass is in the ballpark predicted for neutrino masses by many unified theories of particle interactions (*23*). Although experimental evidence rules out a mass this large for the electron neutrino, it is still possible that either the muon or tau neutrino has such a mass.

The neutralino is a particle that is predicted to exist in supersymmetric extensions of the standard model of particle physics (*24*); predictions for its mass are rather uncertain, ranging from 10 to 1000 times that of the nucleon (supersymmetry dictates a spin-½ partner for every integer-spin particle, and vice versa; in the simplest supersymmetric models, the neutralino is the spin-½ partner of the photon). In the case of the neutralino, annihilations significantly decreased the number of neutralinos from their early abundance of one per photon. Their relic abundance is inversely proportional to their annihilation cross section, very roughly

$$\eta_X \sim \frac{(\hbar/c)^2}{m_X m_{Pl} \sigma_{ann}} \qquad (9)$$

where $m_{Pl} = \sqrt{\hbar c/G} \simeq 2.2 \times 10^{-5}$ g is the Planck mass. Note that the relic abundance depends inversely on the neutralino mass, so it cancels out in the computation of the relic mass density of neutralinos. Remarkably, the condition that the neutralino contribute critical density becomes a condition on its annihilation cross section alone

$$\sigma_{ann} \sim \frac{10^{-2}\hbar^2}{k_B T_0 m_{Pl}} \sim 10^{-36}\ \text{cm}^2 \qquad (10)$$

The cross section required is of the order of magnitude of a weak-interaction cross section, which is the general size expected for the neutralino annihilation cross section.

The axion is a particle whose existence can be traced to an attempt to solve a nagging problem of the standard model of particle physics, the strong-CP problem. (CP is the symmetry that relates particles and antiparticles.) Subtle quantum mechanical effects associated with quantum chromodynamics (QCD), the theory of the strong interactions that bind quarks together, result in a predicted value for the electric-dipole moment of the neutron that is nine orders of magnitude larger than the current experimental upper limit. In 1977, Peccei and Quinn proposed an elegant solution: the introduction of a new symmetry (now referred to as PQ symmetry) that solves the problem and leads to the prediction of a new particle, the axion (*25*). The axion interacts more feebly than neutrinos do, which explains why its existence has yet to be verified or falsified, and for the same reason, it would not have been produced in the thermal plasma during the earliest moments.

Relic axions arose in a different and rather unusual way. Because the axion interacts so weakly, the value of the axion field is left undetermined at early times, taking on whatever random value it had at the beginning; eventually, at about 10^{-5} s, because of QCD effects, the axion field began to relax to its equilibrium value. In so doing, it overshot that value and was left oscillating. These cosmic harmonic oscillations correspond to an extremely high density of very low momentum axions that should still be with us today. If the rest mass energy of the axion is around 10^{-5} eV, relic axions provide closure density (*26*). Theoretical considerations do little to pin down the mass of the axion; however, a host of laboratory experiments and astrophysical-cosmological arguments have narrowed the allowed window for its mass to 10^{-6} to 10^{-3} eV, roughly the range where it would contribute close to the critical density (*27*).

All three particle candidates for the dark matter are sufficiently attractive that experimental efforts are underway to test their candidacies (*28*); in the case of the axion and neutralino, the experiments involve actual detection of the particles that make up the dark halo of our own galaxy (*29*). For the neutrino, direct laboratory measurements restrict the electron neutrino mass to be less than about 7 eV, too small to account for the critical density. Direct measurements of the muon and tau neutrino masses are far more difficult and cannot come close to probing a mass as small as 20 eV; indirect experiments, such as neutrino-oscillation experiments and solar-neutrino observations, can provide some information, but thus far, there is no conclusive positive evidence (*30*).

Development of Structure in the Universe

One of the most pressing questions in cosmology concerns the details of how the abundance of structure seen in the universe today came to be. If the bulk of the matter in the universe exists in the form of particle relics from the big bang, there are profound implications for how structure formed. First, the process could have begun earlier, as soon as the universe became matter dominated, a few thousand years after the bang; if there was only ordinary matter, the growth of the primeval density perturbations could not have begun until matter and radiation decoupled, a few 100,000 years after the bang, when matter was freed from the drag of the radiation. Because density inhomogeneities could have started growing sooner, their initial amplitude could have been smaller, leading to smaller predicted variations in the CBR temperature.

The COBE DMR result is consistent with this smaller prediction but by no means confirms the existence of exotic dark matter. One of the three viable scenarios of structure formation involves ordinary matter only. This minimalist picture, proposed by Peebles (*31*), postulates a universe with baryonic matter only, the dark matter existing in the form of "dark" stars (low-mass stars or the remnants of high-mass stars—neutron stars or black holes). The density fluctuations arise from local fluctuations in the number of baryons (of unknown origin), and the spectrum is adjusted to both explain the observed structure and to be consistent with the level of CBR anisotropy. The weak point of this model is that Ω_0 must be about 0.2 if the observed structure is to form, which violates the nucleosynthesis bound because all the matter is baryonic.

There are two broad classes of models for structure formation with particle dark matter: hot dark matter models, where the dark matter exists in the form of neutrinos, and cold dark matter models, where it exists in the form of neutralinos or axions. In the case of hot dark matter, the primeval density fluctuations on small length scales are erased by the streaming of fast-moving neutrinos from regions of higher density into regions of lower density, and the structures that form first are very large (superclusters), and smaller structures (galaxies and so on) must be formed by fragmentation. This so-called "top-down" scenario is disfavored as structures as large as superclusters are just forming today, making it difficult to explain the existence of distant galaxies that must have formed long ago (*32*).

The erasure of fluctuations on small length scales does not occur with cold dark matter because the dark matter particles move very slowly—neutralinos because they are so heavy, and axions because they were born with very low momentum. With cold dark matter, structure develops "bottom-up," from galaxies to clusters of galaxies to superclusters. Cold dark matter seems to work much better, though not perfectly (*33*). It has been suggested that the cold dark matter scenario could be improved by "mixing" in a small amount of hot dark matter, in the form of neutrinos of mass 7 to 10 eV, a model referred to as mixed dark matter (*34*).

To complete the description of a scenario for structure formation, the origin of the primeval fluctuations must be specified. One possibility involves quantum fluctuations arising during inflation. This leads to the fairly successful and very well studied "cold dark matter" scenario. Another possibility is that the primeval fluctuations involve topological defects—monopoles, string, or texture—that act as gravitational seeds and were produced in a cosmological phase transition that occurred about 10^{-36} s after the bang. These scenarios are less well developed but look promising (*35*). The different models for structure formation make very different predictions for the variation in the CBR temperature on angular scales of order 1°. Measurements of CBR anistropy on this scale should soon whittle down the list of scenarios and perhaps further strengthen the case for nonbaryonic dark matter, which is required in all but one of the models for structure formation.

Conclusion

The cosmic background radiation is arguably the most important cosmological relic yet discovered, and much has and will be learned from its study. The CBR is so fundamental to the standard cosmology that just trying to understand why its temperature is 2.726 K today leads one to discover the most fundamental features of the Universe as well as some of the most pressing cosmological problems—the origin of structure and the nature of the dark matter. In the end, we have no firm explanation as to why the Universe even has a temperature; that is, where the fiery radiation came from. According to the inflationary scenario its existence traces to the decay of the false-vacuum energy. However, its explanation, like that of the expansion itself, may well involve physics yet to be understood.

Note added in proof: Two groups have just reported evidence for the micro-lensing of stars in the Large Magellanic Cloud by dark stars in the halo of our galaxy (*36*). If this interpretation of their data proves correct, then the form of the baryonic dark matter will have been identified, leaving only the task of identifying the nonbaryonic dark matter (or falsifying that hypothesis). Because the case for two kinds of dark matter always existed, the discovery of baryonic dark matter in no way lessens the motivation for nonbaryonic dark matter.

REFERENCES AND NOTES

1. For a textbook treatment of the standard cosmology, see S. Weinberg, *Gravitation and Cosmology* (Wiley, New York, 1972); P. J. E. Peebles, *Principles of Physical Cosmology* (Princeton Univ. Press, Princeton, NJ, 1993); E. W. Kolb and M. S. Turner, *The Early Universe* (Addison-Wesley, Redwood City, CA, 1990). At a more readable level, though slightly out of date, see S. Weinberg, *The First Three Minutes* (Basic Books, New York, 1979).
2. R. H. Dicke *et al.*, *Astrophys. J.* **142**, 414 (1965).
3. J. Mather *et al.*, *ibid.* **354**, L37 (1990); *ibid.*, in press. Another very accurate measurement of the CBR temperature was made from a rocket-borne instrument [H. Gush *et al.*, *Phys. Rev. Lett.* **65**, 537 (1990)]. Reviews of earlier measurements of the CBR spectrum include R. Weiss, *Annu. Rev. Astron. Astrophys.* **18**, 489 (1980); G. Smoot *et al.*, *Astrophys. J.* **331**, 653 (1988).
4. A. A. Penzias and R. W. Wilson, *Astrophys. J.* **142**, 419 (1965).
5. J. Mather, personal communication.
6. See, for example, R. A. Sunyaev and Ya. B. Zel'dovich, *Annu. Rev. Astron. Astrophys.* **18**, 537 (1980).
7. See, for example, J. J. Cowan, F. K. Thielemann, J. W. Truran, *ibid.* **29**, 447 (1991); W. A. Fowler, *Q. J. R. Astron. Soc.* **28**, 87 (1987).
8. The possibility that the universe today is radiation dominated is discussed by M. S. Turner, G. Steigman, and L. Krauss [*Phys. Rev. Lett.* **52**, 2090 (1984)].
9. See, for example, P. J. E. Peebles, *The Large-Scale Structure of the Universe* (Princeton Univ. Press, Princeton, NJ, 1980); G. Efstathiou, in *The Physics of the Early Universe*, J. A. Peacock, A. F. Heavens, A. T. Davies, Eds. (Adam-Higler, Bristol, 1990), p. 361.
10. G. F. Smoot *et al.*, *Astrophys. J.* **396**, L1 (1992).
11. See, for example, A. C. S. Readhead and C. R. Lawrence, *Annu. Rev. Astron. Astrophys.* **30**, 653 (1992).
12. A. H. Guth, *Phys. Rev. D* **23**, 347 (1981); A. D. Linde, *Phys. Lett. B* **108**, 389 (1982); A. Albrecht and P. J. Steinhardt, *Phys. Rev. Lett.* **48**, 1220 (1982).
13. J. M. Bardeen, P. J. Steinhardt, M. S. Turner, *Phys. Rev. D* **28**, 679 (1983); A. H. Guth and S.-Y. Pi, *Phys. Rev. Lett.* **49**, 1110 (1982); A. A. Starobinskii, *Phys. Lett. B* **117**, 175 (1982); S. W. Hawking, *ibid.* **115**, 295 (1982).
14. See, for example, E. W. Kolb and M. S. Turner, *Annu. Rev. Nucl. Part. Sci.* **33**, 645 (1983); A. Cohen, D. Kaplan, A. Nelson, *ibid.* **43**, 27 (1993).
15. G. Gamow, *Phys. Rev.* **70**, 527 (1946); R. Alpher, H. Bethe, G. Gamow, *ibid.* **73**, 803 (1948); Ya. B. Zel'dovich, *Sov. Phys. Usp.* **6**, 475 (1963). Zel'dovich's paper is interesting because not only did he predict the existence of the CBR on the basis of primordial nucleosynthesis, but misinterpreting observational upper limits to a background temperature, he rejected the hot big bang model.
16. T. P. Walker *et al.*, *Astrophys. J.* **376**, 51 (1991).
17. See, for example, V. Trimble, *Annu. Rev. Astron. Astrophys.* **25**, 425 (1987); K. Ashman, *Proc. Astron. Soc. Pac.* **104**, 1109 (1992).
18. See, for example, V. Rubin, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 4814 (1993).
19. See, for example, *Dark Matter in the Universe*, J. Kormendy and G. R. Knapp, Eds. (Reidel, Dordrecht, Netherlands, 1985).
20. A. Kogut *et al.*, *Astrophys. J.*, in press. Earlier measurements of the dipole anisotropy include G. Smoot, M. Gorenstein, R. Muller, *Phys. Rev. Lett.* **39**, 14 (1977); D. J. Fixsen *et al.*, *ibid.* **50**, 620 (1983); R. Fabbri *et al.*, *ibid.* **44**, 1563 (1980); A. A. Klypin *et al.*, *Sov. Astron. Lett.* **13**, 104 (1987).
21. N. Kaiser *et al.*, *Mon. Not. R. Astron. Soc.* **252**, 1 (1991); M. Strauss *et al.*, *Astrophys. J.* **385**, 421 (1992).
22. See, for example, M. S. Turner, *Phys. Scr.* **T36**, 167 (1991).
23. M. Gell-Mann, P. Ramond, R. Slansky, in *Supergravity*, D. Freedman and P. van Nieuwenhuizen, Eds. (North-Holland, Amsterdam, 1979), p. 315; T. Yanagida, in *Proceedings of the Workshop on Unified Theories and Baryon Number in the Universe*, O. Sawada and A. Sugamoto, Eds. (National Laboratory for High-Energy Physics, Tsukuba, Japan, 1979), p. 95.
24. H. E. Haber and G. L. Kane, *Phys. Rep.* **117**, 75 (1985); J. Ellis *et al.*, *Nucl. Phys. B* **238**, 453 (1984).
25. R. D. Peccei and H. R. Quinn, *Phys. Rev. Lett.* **38**,

1440 (1977); S. Weinberg, *ibid.* **40**, 223 (1978); F. Wilczek, *ibid.*, p. 279; J.-E. Kim, *ibid.* **43**, 103 (1979).
26. J. Preskill, M. Wise, F. Wilczek, *Phys. Lett. B* **120**, 127 (1983); L. Abbott and P. Sikivie, *ibid.*, p. 133; M. Dine and W. Fischler, *ibid.*, p. 137.
27. M. S. Turner, *Phys. Rep.* **197**, 67 (1990).
28. J. R. Primack, D. Seckel, B. Sadoulet, *Annu. Rev. Nucl. Part. Sci.* **38**, 751 (1988); P. F. Smith and J. D. Lewin, *Phys. Rep.* **187**, 203 (1990).
29. D. O. Caldwell, *Mod. Phys. Lett. A* **5**, 1543 (1990); K. van Bibber *et al.*, in *Trends in Astroparticle Physics*, D. Cline and R. D. Peccei, Eds. (World Scientific, Singapore, 1992), p. 154.
30. See, for example, F. Boehm and P. Vogel, *Annu. Rev. Nucl. Part. Sci.* **34**, 125 (1984).
31. P. J. E. Peebles, *Nature* **327**, 210 (1987); *Astrophys. J.* **315**, L73 (1987); R. Cen, J. P. Ostriker, P. J. E. Peebles, *ibid.* **415**, 423 (1993).
32. S. D. M. White, C. Frenk, M. Davis, *Astrophys. J.* **274**, L1 (1983); *ibid.* **287**, 1 (1983); J. Centrella and A. Melott, *Nature* **305**, 196 (1982).
33. See, for example, G. R. Blumenthal *et al.*, *Nature* **311**, 517 (1984); J. P. Ostriker, *Annu. Rev. Astron. Astrophys.* **31**, 689 (1993).
34. Q. Shafi and F. Stecker, *Phys. Rev. Lett.* **53**, 1292 (1984); A. van Dalen and R. K. Schaefer, *Astrophys. J.* **398**, 33 (1992); M. Davis, F. Summers, D. Schlegel, *Nature* **359**, 393 (1992); A. Klypin *et al.*, *Astrophys. J.*, in press.
35. A. Albrecht and A. Stebbins, *Phys. Rev. Lett.* **69**, 2615 (1992); D. Bennett *et al.*, *Astrophys. J.* **399**, L5 (1992); N. Turok, *Phys. Rev. Lett.* **63**, 2652 (1989); A. Gooding *et al.*, *Astrophys. J.* **372**, L5 (1991).
36. E. Aubourg *et al.*, *Nature* **365**, 623 (1993); C. Alcock *et al.*, *ibid.*, p. 621.
37. Supported in part by the Department of Energy (at Chicago) and by NASA (NAGW-2381) (at Fermilab).

RESEARCH ARTICLES

Multiple RNA Polymerase Conformations and GreA: Control of the Fidelity of Transcription

Dorothy A. Erie, Omid Hajiseyedjavadi, Mark C. Young, Peter H. von Hippel

Pre–steady state kinetics of misincorporation were used to investigate the addition of single nucleotides to nascent RNA by *Escherichia coli* RNA polymerase during transcription elongation. The results were fit with a branched kinetic mechanism that permits conformational switching, at each template position, between an activated and an unactivated enzyme complex, both of which can bind nucleotide triphosphates (NTPs) from solution. The complex exists most often in the long-lived activated state, and only becomes unactivated when transcription is slowed. This model permits multiple levels of nucleotide discrimination in transcription, since the complex can be "kinetically trapped" in the unactivated state in the absence of the correct NTP or if the 3′ terminal residue is incorrectly matched. The transcription cleavage factor GreA (or an activity enhanced by GreA) increased the fidelity of transcription by preferential cleavage of transcripts containing misincorporated residues in the unactivated state of the elongation complex. This cleavage mechanism by GreA may prevent the formation of "dead-end" transcription complexes in vivo.

Transcription of RNA is the first step in the chain of events leading to expression of the genetic information encoded in double-stranded DNA. The role of RNA polymerase in transcription is to synthesize, under the direction of the DNA template, the nascent RNA chain with high fidelity and at reasonable rates. The passage of the polymerase along the DNA template is uneven, with the "dwell time" at a given template position ranging from 10 ms to seconds or even minutes. These sequence-dependent differences in the transit rate of the polymerase along the template are "programmed" into the sequence in ways that are not yet understood. It is becoming apparent that these sequence-dependent dwell times comprise the central components of a regulatory network that controls gene expression at the level of transcription (*1, 2*). Accordingly, it is important to identify all of the steps in the transcription pathway and to determine which might be rate limiting (and thus subject to regulation) for both correct and incorrect nucleotide incorporation.

The authors are at the Institute of Molecular Biology and Department of Chemistry, University of Oregon, Eugene, OR 97403.

Pre–steady state kinetic studies, which permit the characterization of the individual steps of an enzyme mechanism, should reveal the rate-limiting steps of the single nucleotide addition process of RNA synthesis, just as such studies have been used to elucidate DNA synthesis mechanisms. Kinetic studies of the single-nucleotide addition cycle of several DNA polymerases (*3–5*) have shown that the processive addition of a nucleotide residue to an elongating polynucleotide chain occurs stepwise (*2*) by (i) NTP binding, (ii) phosphotransfer, (iii) pyrophosphate product release, and (iv) translocation of the catalytic active site of the enzyme relative to the 3′ terminus of the growing chain. In addition, these studies have identified conformational changes that regulate the overall rate, fidelity, and processivity of nucleotide incorporation in DNA synthesis.

Application of quench-flow methods to studies of the pre–steady state kinetics (*6*) of RNA synthesis may not yet be practical, since some properties of stalled transcription elongation complexes differ even for complexes that have been purified by slightly different methods (*2*). Nevertheless, if selected steps of the single-nucleotide addition cycle are slowed, this process can be studied without a quench flow apparatus. Accordingly, we have used specific nucleotide misincorporation to probe the detailed mechanisms of nucleotide incorporation and discrimination in RNA synthesis catalyzed by *E. coli* RNA polymerase.

Misincorporation of a nucleotide residue increases the duration of certain steps of the nucleotide addition cycle from milliseconds to minutes. The slow rates of these processes permit the observation and study of conformational states of elongation complexes that may not be significantly populated during rapid synthesis, but which may be physiologically important in regulation, particularly for processes involving transcriptional pausing. We describe two experiments that provide sufficient information to define the mechanism of nucleotide discrimination at specific loci, as well as to identify several elementary steps in the overall nucleotide incorporation cycle. Experiments at other template positions suggest that our findings are likely to be general. The mechanism shares several features of the DNA polymerase single-nucleotide addition cycle, but the overall pathway differs in critical respects. Unlike the basically sequential ($A \rightarrow B \rightarrow C$) DNA polymerase mechanisms (*3–5*), our experiments indicate a branched kinetic pathway for RNA polymerase characterized by several alternative conformational states at each template position. The processivity and great stability of the transcription elongation complex allow us to observe kinetically the different conformational states that comprise the components of this branched pathway. These qualities ensure that the rates of change between conformational states are faster than the rate of