varied by changing the concentration of bubbles. Overtones are possible, resulting in the often-observed narrow, evenly spaced spectral peaks whose frequencies may change systematically or 'glide'. Two-dimensional cracks with different lengths and widths[2] give rise to sets of frequencies that interact with each other, producing complex spectra and a complicated radiation pattern of the transmitted energy.

Each of these models has a fluid (the magma or water and gas) in contact with rock. A percentage of the energy in the fluid is transmitted to the rock, through which seismic waves propagate to the seismic stations. In one alternative approach[3], the walls are themselves pushed apart by the magma and push back, acting as dampened springs. The resulting self-sustained oscillations are an efficient way to both generate tremor and modulate magma flow, the velocity of which changes as the walls move apart or together (the Bernoulli effect). Yet other formulations consider the effects of source and propagation factors separately; both affect the resulting tremor signal[4].

Jellinek and Bercovici's model[1] offers alternative explanations for many of the features used to formulate these other models. An especially welcome contribution is that the main frequencies produced by wagging are in the 1–5-Hz range, exactly the dominant frequencies observed for most tremor. Importantly, these frequencies are caused by the apparent stiffness of the gas annulus (the spring) and are not related to the dimensions of the conduit. This marks a fundamental distinction between this and previous efforts.

The model also returns similar frequencies for reasonable choices of input parameters such as conduit length, shape or diameter, and is not sensitive to magma composition (andesite, dacite, rhyolite and so on). It also demonstrates that higher frequencies of tremor, up to 7 Hz or more, are produced during explosive eruptions. During eruptions, fragmentation and flow of gases occur in the annulus, causing it to be thinner and stiffer, and hence producing higher wagging frequencies. Such an increase in the frequency of tremor is observed for many eruptions.

There are several limitations to Jellinek and Bercovici's formulation[1]. It may explain only one type of tremor — that during eruptions — and is unlikely to be applicable to deep tremor emanating from around 40 km depth, or tremor caused by hydrothermal boiling. And it does not explicitly address how the wagging system is coupled to the surroundings. Furthermore, the model is simplified to include mainly linear effects: nonlinear effects such as feedback may be relevant in some cases.

Nonetheless, this work[1] provides a fresh perspective on an important and long-standing problem. The basic elements of the model may also provide testable elements to provoke the next generation of field observations. ■

**Stephen R. McNutt** *is at the Alaska Volcano Observatory, Geophysical Institute, University of Alaska Fairbanks, 903 Koyukuk Drive, Fairbanks, Alaska 99775, USA.*
*e-mail: steve@giseis.alaska.edu*

1. Jellinek, A. M. & Bercovici, D. *Nature* **470,** 522–525 (2011).
2. Chouet, B. *Nature* **380,** 309–316 (1996).
3. Julian, B. R. *J. Geophys. Res.* **99,** 11859–11877 (1994).
4. Neuberg, J. *Phil. Trans. R. Soc. A* **358,** 1533–1546 (2000).

BIOPHYSICS

# Flipping Watson and Crick

**Watson–Crick base pairs underpin the DNA double helix. Evidence of transient changes in base-pairing geometry highlights the fact that the information held in DNA's linear sequence is stored in three dimensions.** SEE ARTICLE P.498

BARRY HONIG & REMO ROHS

Linear sequences of the DNA bases adenine, cytosine, guanine and thymine (A, C, G and T) define the amino-acid sequences of proteins through the genetic code. Additional codes have been sought to account for the fact that certain proteins bind to particular base sequences in DNA. It has become clear, however, that protein–DNA interactions involve subtle molecular-recognition phenomena. Further striking evidence of this is provided by Nikolova *et al.*[1] in this issue (page 498). They have used nuclear magnetic resonance (NMR) spectroscopy to reveal intrinsic variations in the hydrogen-bonding patterns of base pairs within DNA double helices. Taken together with recent evidence[2,3] of sequence-dependent variations in DNA shape, it now seems extremely unlikely that a simple linear code exists to determine the binding specificity of proteins for DNA.

Ten years after James Watson and Francis Crick published their model of the DNA double helix[4], Karst Hoogsteen reported a crystal structure[5] of a complex in which analogues of A and T formed a base pair that had a

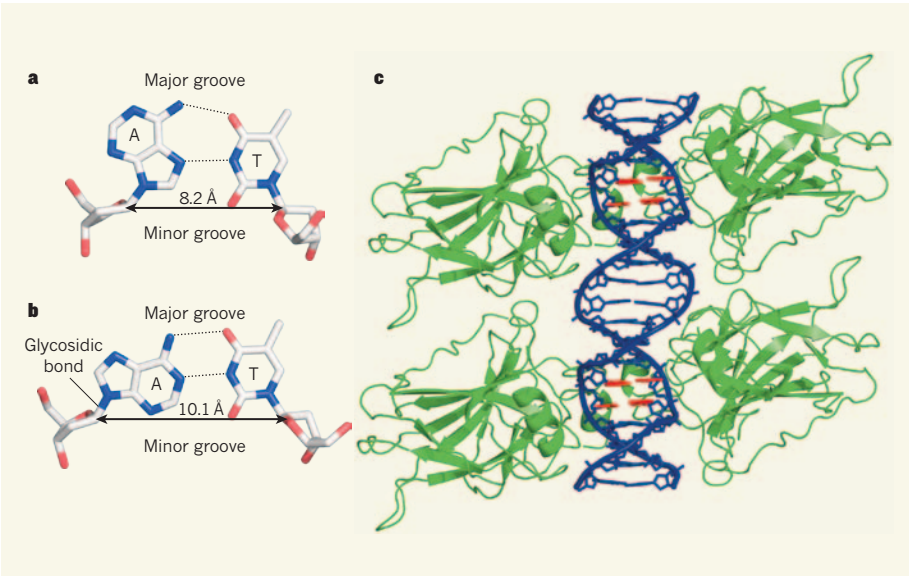


**Figure 1 | Hoogsteen and Watson–Crick base pairing.** Base pairs in nucleic acids can adopt different geometries. **a**, This A•T Hoogsteen base pair forms in a complex of DNA with the p53 protein. Structural data are from ref. 9. **b**, This A•T Watson–Crick base pair forms at the same position in an analogous p53–DNA complex, although the DNA has a different overall sequence. Structural data are from ref. 11. The adenine is rotated by approximately 180° about the glycosidic bond compared with the Hoogsteen base pair, so that the lower hydrogen bond involves a different chemical group in the adenine from that used in the lower hydrogen bond in **a**. The pattern of chemical groups in each groove of the DNA differs between **a** and **b**, as does the distance across the base pair. **c**, This complex[9] consists of a tetramer of p53 DNA-binding domains (green) bound to DNA that contains Hoogsteen base pairs (red). The distortion to the double-helix structure caused by the Hoogsteen base pairs is clearly visible. Nikolova *et al.*[1] report that Hoogsteen base pairs form transiently in free DNA.

different geometry from that described by Watson and Crick (Fig. 1a, b). Similarly, an alternative base-pairing geometry can occur for G•C pairs. Hoogsteen pointed out that if the alternative hydrogen-bonding patterns were present in DNA, then the double helix would have to assume a quite different shape. Hoogsteen base pairs are, however, rarely observed.

Nikolova and colleagues' key finding[1] is that, in some DNA sequences, especially CA and TA dinucleotides, Hoogsteen base pairs exist as transient entities that are present in thermal equilibrium with standard Watson–Crick base pairs. The detection of the transient species required the use of NMR techniques that have only recently been applied to macromolecules[6].

Why is this finding important? Hoogsteen base pairs have, after all, previously been observed in protein–DNA complexes[7–9] (Fig. 1c). But it has not been possible to determine whether Hoogsteen base pairs are present in free DNA. Nikolova and colleagues' study[1] reveals that the ability to flip between Watson–Crick and Hoogsteen base pairing in free DNA is an intrinsic property of individual sequences. This implies that some proteins have evolved to recognize only one base-pair type, and use intermolecular interactions to shift the equilibrium between the two geometries[9].

DNA has many features that allow its sequence-specific recognition by proteins. This recognition was originally thought to primarily involve specific hydrogen-bonding interactions between amino-acid side chains and bases. But it soon became clear that there was no identifiable one-to-one correspondence — that is, there was no simple code to be read. Part of the problem is that DNA can undergo conformational changes that distort the classical double helix. The resulting variations in the way that DNA bases are presented to proteins can thus affect the recognition mechanism.

More significantly, it has become evident that distortions in the double helix are themselves dependent on base sequence. This enables proteins to recognize DNA shape in a manner reminiscent of the way that they recognize other proteins and small ligand molecules. For example, stretches of A and T bases can narrow the minor groove of DNA (the narrower of the two grooves in the double helix), thus enhancing local negative electrostatic potentials and creating binding sites for appropriately placed, positively charged arginine amino-acid residues[3].

Nikolova and colleagues' discovery[1] that DNA base pairs can so easily leave their favoured Watson–Crick conformation comes as a surprise. But viewed in another way, the phenomenon is not so different from protein side chains undergoing a conformational change so as to optimize binding with another protein. The real surprise where DNA is concerned is that the constraints of the double helix don't preclude this possibility. The presence of Hoogsteen base pairs in detectable amounts — even in free DNA — therefore provides a notable example of the remarkable plasticity of the canonical double helix. It also implies that, if DNA bases are regarded as letters, each letter potentially has two meanings that determine both hydrogen-bonding patterns and structural variations in the double helix.

Structural biologists have long recognized that there is no second code in which certain amino acids recognize complementary DNA bases in protein–DNA interactions. Nevertheless, protein–DNA binding is still commonly thought of purely in terms of codes and sequence motifs, rather than as the binding of two large macromolecules that have complex shapes and considerable conformational flexibility[10]. Nikolova and colleagues' discovery reminds us that DNA offers proteins not only an enticing linear alphabet, but also a set of conformations that can be recognized in a sequence-dependent way. Understanding how the linear sequence of bases in DNA is recognized by proteins is therefore a problem that must be solved in three dimensions. This will require structural, biochemical, genomic and computational studies on both naked double helices and protein–DNA complexes. ■

**Barry Honig** *is at the Howard Hughes Medical Institute, Center for Computational Biology and Bioinformatics, Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032, USA.* **Remo Rohs** *is in the Molecular and Computational Biology Program, Department of Biological Sciences, University of Southern California, Los Angeles, California 90089, USA.*
*e-mails: bh6@columbia.edu; rohs@usc.edu*

1. Nikolova, E. N. *et al. Nature* **470,** 498–502 (2011).
2. Parker, S. C. J., Hansen, L., Abaan, H. O., Tullius, T. D. & Margulies, E. H. *Science* **324,** 389–392 (2009).
3. Rohs, R. *et al. Nature* **461,** 1248–1253 (2009).
4. Watson, J. D. & Crick, F. H. C. *Nature* **171,** 737–738 (1953).
5. Hoogsteen, K. *Acta Crystallogr.* **16,** 907–916 (1963).
6. Palmer, A. G. & Massi, F. *Chem. Rev.* **106,** 1700–1719 (2006).
7. Aishima, J. *et al. Nucleic Acids Res.* **30,** 5244–5252 (2002).
8. Nair, D. T., Johnson, R. E., Prakash, S., Prakash, L. & Aggarwal, A. K. *Nature* **430,** 377–380 (2004).
9. Kitayner, M. *et al. Nature Struct. Mol. Biol.* **17,** 423–429 (2010).
10. Rohs, R. *et al. Annu. Rev. Biochem.* **79,** 233–269 (2010).
11. Chen, Y., Dey, R. & Chen, L. *Structure* **18,** 246–256 (2010).

CLIMATE CHANGE

# Old droughts in New Mexico

**A long climate record reveals abrupt hydrological variations during past interglacials in southwestern North America. These data set a natural benchmark for detecting human effects on regional climates.** SEE LETTER P.518

JOHN WILLIAMS

Southwestern North America is a pretty dry place, and is likely to get drier this century because of anthropogenic climate warming. On page 518 of this issue[1], Fawcett *et al.* provide a climate record from deep in the past that will help in assessing the future hydrological regime for the area.

Climate models consistently project declines in winter precipitation for the southwest, in response to rising greenhouse gases as the subtropical dry zones expand polewards[2,3]. This precipitation decline, combined with expected increases in evaporation rates and reduced snowpack, would severely strain the region's capacity to adapt to climate change. Moreover, the southwest is historically prone to droughts, with six multi-year droughts in the nineteenth and twentieth centuries, including the infamous 1930s Dust Bowl[4]. These droughts are linked to yearly-to-decadal variations in sea surface temperatures in the tropical Pacific, enhanced by local soil-moisture feedbacks[5]. Further understanding of the mechanisms of hydrological variability, together with efforts to limit societal vulnerability to climate change, are priorities in global-change research[4].

Palaeoclimatic studies have made an essential, if not particularly reassuring, contribution to this effort, by providing insight into the natural behaviour of hydrological systems and a longer-term context for historical and projected changes. Tree-ring records offer sobering evidence of widespread and decades-long 'megadroughts' in the western United States over the past several millennia that dwarfed recorded historical droughts[6]. Records spanning the past 11,000 years (the Holocene interglacial) demonstrate long-term shifts in southwestern monsoon