they possess variants that might lead to an altered protein sequence or gene expression. Examples of such variants are *APOE* (in Alzheimer's disease) and *HLA-DQB1* (in type 1 diabetes).

Crohn's disease is characterized by perturbed control of inflammation in the gut and in its interaction with bacteria, and damage to the gut wall[5,6]. In studies of how the immune system recognizes bacteria, one of the groups — Ogura *et al.* — had previously identified[4] two intracellular versions of Toll-like receptors; these are molecules on immune cells that detect bacterial components known as lipopolysaccharides[7]. The new receptors, dubbed NOD1 and NOD2, both regulate NF-κB, which is a master regulator of the genes involved in inflammation and has been implicated in gut–microbe interactions and inflammatory bowel disease[5,6,8]. *NOD2* was therefore a strong 'functional candidate gene' for Crohn's disease. There are, however, hundreds of plausible candidate genes in the immune system. But the real clincher for *NOD2* came from its location in the genome[4]: it lies on chromosome 16, bang in the area of the strongest linkage to disease. Ogura *et al.*[2] had then only to sequence the gene from patients to find the disease-associated alleles.

The other group, Hugot *et al.*[1], did not know anything about the *NOD2* gene when they started their quest. Rather, they began to search the 20 megabases of chromosome 16 for SNPs that were most frequently inherited with the disease. This is called association mapping. Linkage analysis is referred to as the positional approach to gene identification because it does not rely on guessing the identity of a functional candidate gene. Association mapping extends linkage analysis to map the position of the disease gene to a much higher resolution, on average to within about 60,000 base pairs. This still means that many polymorphisms must be analysed in many families, a process that is still severely limited because of the unfinished nature of the human genome sequence, the lack of complete knowledge of SNPs, and the steep cost of genotyping. Furthermore, many statistical tests were, or could have been, carried out in the search, so there is the problem of obtaining positive-looking results that are actually false, just by chance.

Hugot *et al.* did not let these limitations stop them. They took their chance, a slim one, and won. They sequenced some genomic DNA cloned from chromosome 16 that showed a sniff of disease association, identified a gene, characterized the polymorphisms in it and discovered that these were strongly associated with Crohn's disease. Their gene turned out to be *NOD2*. In both reports[1,2], the linkage result narrowed the odds.

The new work provides strong evidence that a defective *NOD2* gene makes its bearer susceptible to Crohn's disease. But it confers risk only, not certain diagnosis, and disease occurrence is dependent on the presence of bacteria in the gut and probably many other environmental factors. Moreover this is not *the* gene for Crohn's disease, nor even necessarily the only relevant gene in this region of chromosome 16. Other chromosomes have been linked to the disease[9], and the genes concerned have yet to be discovered and fitted into the puzzle.

More broadly, however, these studies will encourage geneticists in their search for disease causation. It is currently assumed that common diseases might stem from many common variants with allele frequencies in the general population exceeding 15%. However, in the case of this Crohn's disease gene, *NOD2*, many rare variants (with frequencies of 5% or less) are involved. My guess is that the genetic basis of common disease will be a mixture of common and rare alleles, the expression of which is dependent on many environmental factors.

What about help for sufferers from Crohn's disease? Better diagnosis is a good bet. From the new findings[1,2] it seems that up to about 15% of cases (compared with 5% in the general population) may have alleles that alter the function of *NOD2*. These disease-associated *NOD2* alleles tend not to be found in patients with ulcerative colitis, a related form of inflammatory bowel disease. For therapy, given that normal *NOD2* protects against the gut inflammation seen in Crohn's disease, mimicking that effect might reduce the risk of disease. Sulfasalazine and glucocorticoids are among the most effective drugs in treating inflammatory bowel disease currently available and they target the NF-κB complex[10]. Restoring the normal function of *NOD2* in people in which the gene is defective might help prevent disease in the first place. Better diagnosis and (especially) treatments are not immediate prospects. But the new work constitutes tangible progress towards those ends. ■

*John A. Todd is at the Juvenile Diabetes Research Foundation/Wellcome Trust Laboratory, Cambridge Institute for Medical Research, University of Cambridge, Wellcome Trust/MRC Building, Hills Road, Cambridge CB2 2XY, UK.*
*e-mail: john.todd@cimr.cam.ac.uk*

1. Hugot, J.-P. *et al. Nature* **411,** 599–603 (2001).
2. Ogura, Y. *et al. Nature* **411,** 603–606 (2001).
3. Hugot, J. *et al. Nature* **379**, 772–773 (1996).
4. Ogura, Y. *et al. J. Biol. Chem.* **276**, 4812–4818 (2001).
5. Fiocchi, C. *Gastroenterology* **115**, 182–205 (1998).
6. Schreiber, S., Nikolaus, S. & Hampe, J. *Gut* **42**, 477–484 (1998).
7. Brown, P. *Nature* **410**, 1018–1020 (2001).
8. Mahida, Y. R. & Johal, S. *Clin. Exp. Immunol.* **123**, 347–349 (2001).
9. Cho, J. *Curr. Gastroenterol. Rep.* **2**, 434–439 (2000).
10. Wahl, C., Liptay, S., Adler, G. & Schmid, R. M. *J. Clin. Invest.* **101**, 1163–1174 (1998).

Chemistry

# A new twist on molecular shape

Frank Weinhold

What are the forces that control the twisting and folding of molecules into complex shapes? Don't look for the answers in your organic chemistry textbook.

The forces that hinder internal rotations (twisting) about single bonds are arguably the most important factor controlling the shape and dynamics of complex polymeric molecules, including proteins, yet these forces remain widely misunderstood. Although it is relatively easy to determine the skeletal sequence of bonded atoms in a chain polymer, it is far more difficult to determine how the twists and turns at each skeletal bond conspire to produce the overall macromolecular shape. Understanding how proteins fold so quickly and correctly is widely recognized as the greatest existing challenge to structural biologists. But as the US National Institutes of Health observed last year[1]: "for 50 years they've tried — and failed — to crack the code that governs folding."

Like a three-armed turnstile, a single carbon–carbon bond is known to exert a force that opposes initial twisting from the resting state, but reverses after a 60° displacement to drive the system towards an alternative resting state rotated by 120° from the first. The classic example of a simple molecule with a central carbon-carbon bond is ethane ($C_2H_6$). Kemp and Pitzer[2] first established the shape of the internal energy profile for ethane in 1936, showing that it has three energy wells (minima) corresponding to the most stable internal rotation state (conformer) shown in Fig. 1 (overleaf). The big theoretical question ever since has been: what causes the internal rotation barriers (roughly 3 kcal $mol^{-1}$) between one well and the next? More specifically, what causes the internal energy to rise in passing from the stable 'staggered' conformation to the unstable 'eclipsed' conformation? The emerging answer, as pointed out by Pophristic and Goodman[3] on page 565 of this issue, contrasts sharply with the stock explanation presented in almost every textbook of organic chemistry.

Generations of chemistry students have

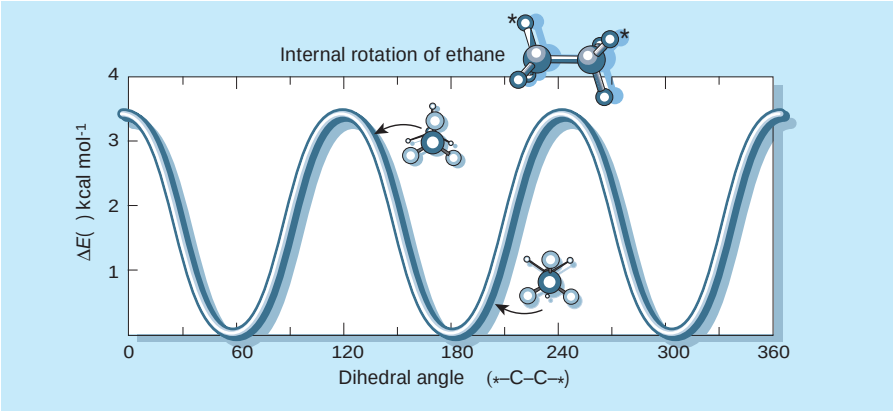



Figure 1 **Energy profile for internal rotations in ethane. As one of the methyl groups ($CH_3$) rotates about the central carbon–carbon bond, the molecule switches three times between a high-energy 'eclipsed' conformation and the preferred low-energy 'staggered' conformation. When the molecule is viewed end-on, the near-side CH bonds appear to interleave the far-side bonds in the staggered conformations ($\phi = 60°$, 180° or 300°), but cover the far-side bonds in the eclipsed conformations ($\phi = 0°$, 120° or 240°). The dihedral twisting angle, $\phi$, is defined as the angle between a specific pair of opposing C–H bonds (asterisks), shown here in the *gauche* conformation.**

been taught that energy barriers like those found in ethane arise from spatial (steric) effects related to repulsive interactions between hard-sphere-like atoms. This explanation is intuitively satisfying, but to address the ethane problem fully, one must adopt the wave-like imagery of quantum mechanics, particularly of the orbitals that 'house' the electron pairs that form bonds between atoms. According to the Pauli exclusion principle of quantum mechanics, each bond orbital can be occupied by a maximum of two electrons. As pointed out in 1975 by Weisskopf[4], two filled (doubly occupied) orbitals cannot be crowded into the same spatial region except by developing additional ripples to preserve their mutual orthogonality (a kind of 'perpendicularity' in the quantum space of electron orbitals). These unfavourable (energy-raising) contortions give rise to a 'quantum pressure' that opposes steric overlap of filled orbitals. According to the textbook picture, such steric repulsion between bonds (due to the overlap of two filled bond orbitals) is the cause of the ethane energy barrier.

Pophristic and Goodman[3] use a computational method that is closely related to Weisskopf's picture in order to evaluate quantitatively the orbital interactions in ethane. They find that the steric repulsions in ethane are only of secondary importance (in accordance with other analysis methods) and actually favour the eclipsed conformer, whereas the stable state is the staggered conformer. It is no longer valid to assume, as the textbooks do, that steric considerations play the leading role in deciding the equilibrium structure of ethane. So what controls ethane's preference for the staggered conformation?

The true origin of ethane-like barriers can be found in another quantum-mechanical effect arising from the interactions of electron orbitals. There is growing awareness[5] of the importance of hyperconjugative interactions between filled and unfilled orbitals in barrier phenomena. In the case of ethane, the hyperconjugative interaction involves partial electron transfer from a nearly doubly occupied (bonding) orbital to a nearly vacant (antibonding) orbital[6,7]. As chemistry students also learn, the bonding orbital between two atoms arises from constructive mixing of two atomic orbitals, and this low-energy combination is naturally chosen for occupancy by the two electrons of the bond. But the wave-mixing rules of quantum mechanics dictate that each constructive (in-phase) combination must be accompanied by the corresponding destructive (out-of-phase or antibonding) orbital.

Figure 2 shows two-dimensional and three-dimensional depictions of the orbital interactions in the staggered and eclipsed conformations of ethane. The figures show the bonding and antibonding patterns (labelled $\sigma_{CH}$ and $\sigma_{CH}$*) for two adjacent carbon–hydrogen single bonds in each conformation. The solid and dotted contours (in two dimensions) or blue and yellow lobes (three dimensions) correspond to the peaks and troughs, respectively, of the orbital waveforms. The mixing of adjacent $\sigma_{CH}$–$\sigma_{CH}$* orbitals is more favourable (for example, blue with blue, and yellow with yellow) when ethane adopts the staggered (Fig. 2a) rather than the eclipsed (Fig. 2b) conformation.

The hyperconjugative interaction in Fig. 2a leads to delocalization of negative charge, whereby the electrons in the filled orbital are able to move partially into the unoccupied orbital. It may seem counterintuitive that the electron pair in a carbon–hydrogen bond would delocalize into an adjacent high-energy antibonding orbital, the $\sigma_{CH}$*

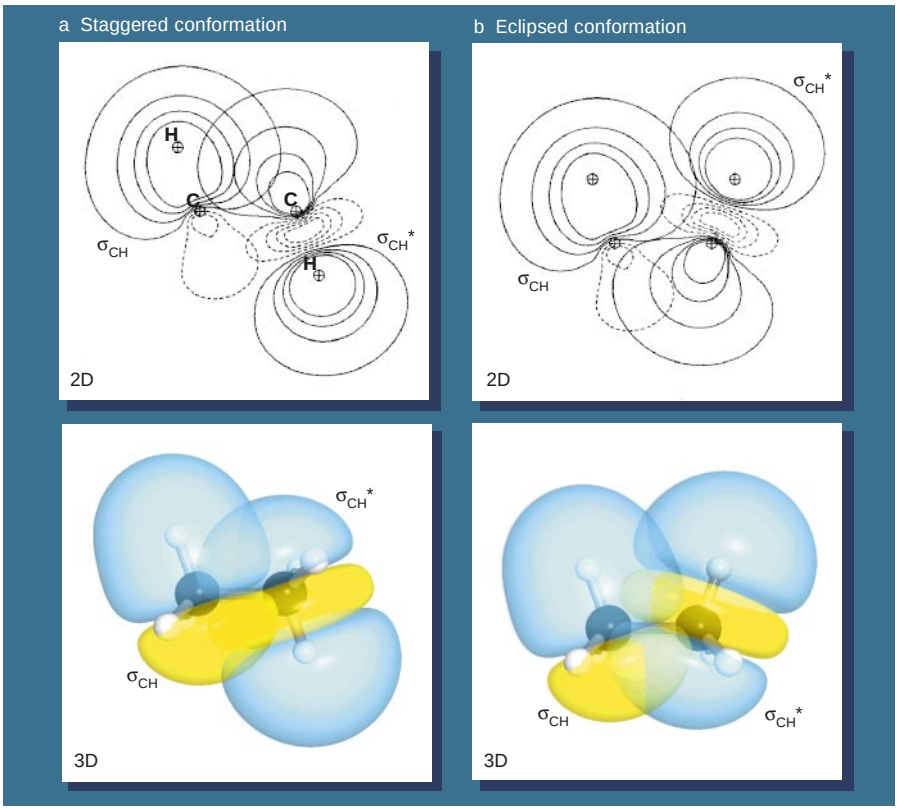
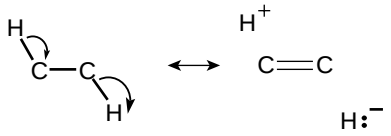



Figure 2 **Contour plots (two-dimensional, 2D) and surface plots (3D) of adjacent carbon–hydrogen bond orbitals ($\sigma_{CH}$ and $\sigma_{CH}$*) of ethane.** a, **In the staggered conformation the more favourable orbital overlap leads to a stronger (more stabilizing) interaction.** b, **There is a less favourable interaction in the eclipsed conformation, leading to a higher-energy state.**

orbital. Yet according to the famously weird rules of quantum superposition[8] such hyperconjugative $\sigma_{CH}$–$\sigma_{CH}$* interactions are indeed rewarded with an additional energy lowering of a few kilocalories per mole. The enhanced strength of hyperconjugative delocalizations in the staggered arrangement is the quantum mechanical origin of the energy wells in Fig. 1.

When bonding electrons delocalize they can lead to alternative bonding patterns (resonance structures). For example, the electron delocalization associated with the $\sigma_{CH}$–$\sigma_{CH}$* interaction corresponds to a partial double-bond resonance structure, as represented in chemical notation below. Twisting about a double carbon–carbon bond is difficult, so rotation is inhibited.




So ethane's energy barriers can be viewed as a form of hyperconjugative 'resonance stabilization' for the electrons of σ-type single bonds, weaker than, but closely related to, the π-electron resonance stabilization in systems with alternating single and double bonds. From this viewpoint, ethane-like molecules adopt staggered conformations not to relieve steric congestion, but to achieve optimal resonance stabilization.

The hyperconjugative picture can also explain other conformational effects that arise in barrier phenomena. These include the '*gauche* effect' (the tendency of lone electron pairs or polar bonds to adopt isomeric arrangements that are non-aligned but not *trans*), the 'anomeric effect' (*gauche* effect for carbohydrates), and various other 'stereoelectronic effects' on molecular shape and reactivity[9,10]. Hyperconjugative interactions also account for the striking long-range effects (far beyond the range of plausible steric influences) that are observed in the barrier potentials of toluene and related aromatic species[11]. As shown by Pophristic and Goodman, hyperconjugative interactions are also largely responsible for inducing the skeletal distortions, such as carbon–carbon bond elongation, that accompany internal rotations. At the orbital level, a host of apparently unconnected effects appear to be due to the same electronic interaction. Electrons, after all, know only a few tricks.

The hyperconjugative interactions in Fig. 2 offer new electronic opportunities for controlling the conformational destiny of complex biomolecules. Broader awareness of these interactions among molecular scientists should inspire new chemical and biochemical strategies to control twisting processes — for example, through cooperative coupling to other delocalizing interactions. Improved understanding of the electronic origins of molecular twisting may also spur overdue efforts to incorporate more realistic quantum-mechanical effects into molecular modelling and protein folding calculations. The most pressing question raised by Pophristic and Goodman is: when will chemistry textbooks begin to serve as aids, rather than barriers, to this enriched quantum-mechanical perspective on how molecular turnstiles work? ■

*Frank Weinhold is in the Theoretical Chemistry Institute and Department of Chemistry, University of Wisconsin, Madison, Wisconsin 53706, USA.*
*e-mail: weinhold@chem.wisc.edu*

1. *The Structures of Life* NIH Publication No. 01-2778 (National Institutes of Health, Washington DC, 2000). http://www.nigms.nih.gov
2. Kemp, J. D. & Pitzer, K. S. *J. Chem. Phys.* **4,** 749 (1936).
3. Pophristic, V. & Goodman, L. *Nature* **411,** 565–568 (2001).
4. Weisskopf, V. F. *Science* **187,** 605–612 (1975).
5. Michl, J. & West, R. *Acc. Chem. Res.* **33,** 821–824 (2000).
6. Brunck, T. K. & Weinhold, F. *J. Am. Chem. Soc.* **101,** 1701–1709 (1979).
7. Reed, A. E. & Weinhold, F. *Isr. J. Chem.* **31,** 277–285 (1991).
8. Weinhold, F. *J. Chem. Ed.* **76,** 1141 (1999).
9. Petillo, P. & Lerner, L. in *The Anomeric Effect and Associated Stereoelectronic Effects: ACS Symposium Series* (ed. Thatcher, G. R. J.) 539 (Am. Chem. Soc., New York, 1993).
10. Bretscher, L. E. *et al. J. Am. Chem. Soc.* **123,** 777–778 (2001).
11. Lu, K.-T. *et al. J. Chem. Phys.* **102,** 6787–6805 (1995).

Immunology

# Brief encounter

Raymond M. Welsh

To a certain extent, the response of immune cells to infection is preset. T cells seem to be activated early on or not at all, and, once stimulated, embark on a complete programme of division and specialization.

When a doctor palpates a patient's abdomen or the glands under the jaw, the intent is often to assess the size of the spleen or the lymph nodes, which swell in response to infection. This swelling results from the proliferation of antigen-specific T and B cells, which also become specialized in different ways to control intruding pathogens. When the infection is quelled, the immune response is silenced, and many of the T and B cells die. Others lurk as slowly dividing 'memory' populations in lymphoid and non-lymphoid organs, until activated by the same or a related pathogen. Puzzlingly, the timing of the T-cell response is often more or less the same regardless of the pathogen, and T cells may continue multiplying for days after the antigen is cleared. In contrast, the magnitude of the response can vary markedly, depending on the antigen and its form of delivery (a point not lost on vaccinologists). Writing in *Nature Immunology*, van Stipdonk and colleagues[1] and Kaech and Ahmed[2] resolve the puzzle: they show that T cells are programmed to multiply and differentiate after only a brief encounter with a stimulating antigen.

In a typical immune response, a pathogen is first processed by antigen-presenting cells (APCs), which display on their surface short peptides from the pathogen, in complex with 'major histocompatibility' (MHC) proteins. Naive T cells may recognize this peptide–MHC complex, and this interaction starts a series of events inside the T cells, which is amplified by the interaction of 'co-stimulatory molecules' on the T cells and the APCs[3]. Thereafter, the T cells both secrete and respond to growth and differentiation factors, and take on functions that enable them to ward off the pathogens, for example by killing infected cells.

Van Stipdonk *et al.*[1] questioned how long T cells must be exposed to antigen for this response to be stimulated and maintained. They found that $CD8^+$ T cells (so called because they express a molecule called CD8 on their surface) required as little as 2 hours of exposure to APCs to initiate several rounds of division and differentiation into cell-killing (cytotoxic) T cells. To show this, the authors engineered mouse fibroblast cells to express a co-stimulatory molecule, as well as a peptide from the ovalbumin protein that could stimulate a population of naive, ovalbumin-specific $CD8^+$ T cells. They then treated the T cells with a fluorescent dye, the intensity of which decreases by half with each cell division, and exposed them to the fibroblasts (the APCs) for different time periods. They studied the T cells to see how much they had proliferated, how much dye they had lost (an indicator of how many times they had divided), and whether they had differentiated.

Cell division did not begin until at least 24 hours after stimulation, but by 48 hours the cells were proliferating rapidly, dividing every 5 to 6 hours. This remarkably fast generation time is consistent with previous studies[4–6]. Moreover, when the authors separated the T cells from the APCs after as little as 2 hours, the entire proliferation and differentiation programme continued[1]. The authors obtained similar results *in vivo*.

One implication is that T cells 'decide' very early during an infection whether to respond to the pathogen. This sheds light on a recent report by Mercado *et al.*[4], who showed that the magnitude of the response