Si '7 × 7' surface was determined[4] from transmission electron diffraction combined with scanning tunnelling microscopy data.

In determining the structure of crystals by X-ray diffraction, a major problem is that one knows only the intensities and not the phases of the diffracted beams. So direct inversion of the diffraction pattern by Fourier analysis to recover the electron density (which scattered the X-rays) is not possible. However, by recognizing that the electron density is not an arbitrary function but has special properties (for example, it is always positive and, as atomic electron density peaks strongly near the nucleus, is close to zero between atoms), relationships between the phases of diffracted beams can be established. These can then be factored into 'direct methods' of crystallographic analysis, which have made routine the determination of structures of moderately complex crystals (some hundreds of atoms in the unit cell). Electron crystallography is in principle similar — the scattering now occurs by the electrostatic potential of the atom, but this is simply related to the charge density by Poisson's equation — and direct methods have been used here too[5].

It is not straightforward to apply direct methods to the solution of surface structures by electron crystallography, but Erdman *et al.*[1] show that direct-methods techniques developed by Marks and colleagues[6] lead directly to a projection of the surface structure. They then use *ab initio* electronic-structure calculations to refine the coordinates normal to the surface to obtain a full three-dimensional structure.

In the $SrTiO_3$ structure (Fig. 1), there are alternating SrO and TiO_2 layers normal to the cubic axes. So one might expect a surface in that plane to terminate either with a TiO_2 layer or a SrO layer. In fact, in the observed structure[1] the crystal terminates with two TiO_2 layers, and with substantial reorganization in the outermost layer. The origin of the reorganization can be seen as follows. In the bulk structure each O atom is bonded to two Ti atoms. If this pattern were to persist in a surface of composition TiO_2, each Ti atom would have to be 4-coordinated (that is, forming four bonds to oxygen) in contrast to the octahedral 6-coordination in the bulk. The surface reconstruction is such that coordination polyhedra share edges (rather than just vertices as in the bulk) so that the Ti atoms in the surface layer are 5-coordinated.

Such behaviour might have been expected from crystal chemistry. In the tungsten oxide WO_3, for example, the structure is the same as the TiO_3 framework of $SrTiO_3$. WO_3 is easily reduced to phases WO_{3-x} in which, instead of having oxygen vacancies, the crystal reorganizes to eliminate the vacancies and maintain the octahedral coordination of the metal by edge-sharing between the octahedra. Similar behaviour is observed in the fantastically rich and highly organized structure of the many other oxides of the early transition metals (such as Ti, Nb, Mo and W)[7].

It should be emphasized that the work under discussion is the beginning of a long and surely very interesting story. Even for $SrTiO_3$ there is strong evidence[8] for other reconstructions of the surface; in addition, as well as the perovskites $A^{2+}B^{4+}O_3$ (such as $SrTiO_3$), there are other families $A^{+}B^{5+}O_3$ (such as $NaNbO_3$) and $A^{3+}B^{3+}O_3$ (such as $LaFeO_3$). And then, of course, there are many thousands of other oxide structures. ■

Michael O'Keeffe is in the Department of Chemistry and Biochemistry, Arizona State University, Tempe, Arizona 85287, USA.
e-mail: mokeeffe@asu.edu

1. Erdman, N. *et al. Nature* **419,** 55–58 (2002).
2. Mitchell, R. H. *Perovskites: Modern and Ancient* (Almaz, Thunder Bay, Ontario, 2002).
3. Marks, L. D. *Phys. Rev. Lett.* **51,** 1000–1002 (1983).
4. Takayanagi, K., Tanishiro, Y., Takahashi, S. & Takahashi, M. *Surf. Sci.* **164,** 367–392 (1985).
5. Dorset, D. L. *Structural Electron Crystallography* (Plenum, New York, 1995).
6. Marks, L. D., Erdman, N. & Subramanian, A. *J. Phys. Condens. Matter* **13,** 10677–10688 (2001).
7. Hyde, B. G. & Andersson, S. *Inorganic Crystal Structures* (Wiley, New York, 1989).
8. Jiang, Q. D. & Zegenhagen, J. *Surf. Sci.* **425,** 343–354 (1999).

Immunology

The roots of antibody diversity

Patricia J. Gearhart

When faced with foreign molecules our antibodies mutate, allowing them to bind to the intruders more strongly. In a story full of surprises, it looks as though the mechanism of mutation has finally been revealed.

One of the tactics our immune system uses to fight off intruders is the production of protective antibody proteins, which recognize and neutralize foreign molecules. Antibodies are secreted by the B cells of the immune system, and are unique among our proteins in their unlimited potential for diversity. Half of each antibody molecule is encoded by the so-called variable genes, and when a B cell recognizes and responds to a foreign molecule, these variable genes undergo mutation at a tremendously high frequency.

In a stunning set of papers, culminating in that on page 43 of this issue[1], Neuberger and colleagues seem to have discovered the elusive mechanism behind this 'hypermutation' — and the answer is totally unexpected. First they showed that the activation-induced cytosine deaminase (AID) protein, which is needed for hypermutation[2], leads to DNA mutations when expressed in bacteria[3]. More mutations of the cytosine 'letters' in DNA occur in bacteria that lack the enzyme uracil glycosylase, hinting that AID converts cytosines to uracils. (Normally, these erroneous uracils would be removed by the glycosylase before they caused a problem.) These findings are now followed[1] by a convincing demonstration that the generation of uracil from cytosine in B-cell DNA sparks hypermutation, which occurs as the cells try to fix the errors.

After vaccination or infection, the body first produces antibodies of relatively low affinity for the foreign molecule, the antigen. As an immune response progresses, these antibodies become hypermutated. The purpose of hypermutation is to create new protein sequences that can bind the antigen more strongly and specifically than their precursors. This allows us to respond quickly and effectively to pathogens that we have encountered previously. People who cannot mutate their antibodies suffer recurring bacterial and viral infections, and do not respond to vaccination.

Hypermutation occurs only in B cells that have been stimulated with antigen. During the development of B cells, before they encounter antigen, genetic shuffling has recombined blocks of DNA into long segments representing the basic antibody-encoding genes. Then, during hypermutation, many single mutations — and, rarely, small insertions and deletions — are introduced into the variable regions of these genes.

A breakthrough in understanding the mechanism behind hypermutation came two years ago, when Muramatsu *et al.*[2] discovered that the AID protein — which is expressed only in hypermutating B cells — is needed for the process. At the time it was thought that AID might convert ('deaminate') cytosine to uracil in the messenger RNA encoding an endonuclease enzyme, thereby enabling an active form of the enzyme to be produced. Endonucleases cut nicks in DNA strands, and it was proposed that nicks generated by this putative endonuclease in the variable regions of antibody genes could be the lesions that trigger hypermutation. This notion was attractive because numerous DNA breaks had been detected around variable genes. But the theory was questioned when similar numbers of breaks were found in mice lacking AID[4,5]. Researchers now suspect that these nicks are unrelated to hypermutation.

A second possibility, proposed by Neuberger and colleagues, was that AID converts cytosine to uracil in DNA (not RNA), as suggested by the increase in mutations of

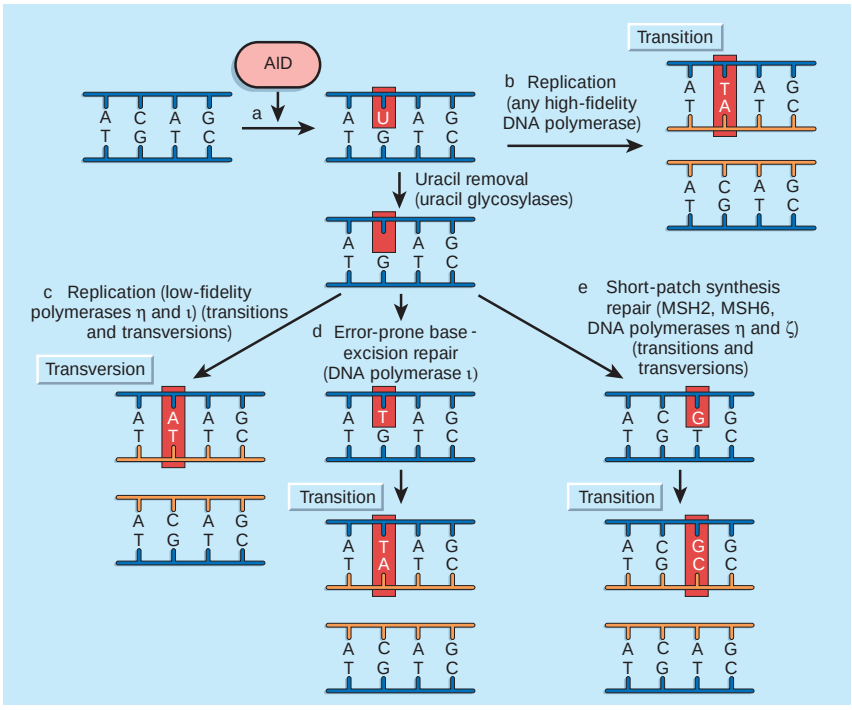

Figure 1 **A model for antibody diversity.** a, **The AID protein converts cytosines (C) in the antibody-gene variable region to uracils (U). These can be processed in several ways to generate mutations.** b, **High-fidelity DNA-replicating enzymes (polymerases) recognize U as thymine (T), and pair it with adenine (A) during replication.** c–e, **These pathways rely on the initial removal of U by uracil glycosylases.** c, **Low-fidelity polymerases bypass the resulting gap during replication, filling it more or less at random. One example is shown.** d, **In this repair pathway, a T is preferentially inserted in the gap, leading to errors after replication.** e, **In this repair process, the gap is filled in and extended, generating mutations that are mostly located opposite As and Ts. In chicken B cells lacking the protein XRCC2, the replication pathways predominate. All four pathways work in mice and humans. 'Transitions' swap one purine base (G or A) for the other; likewise for pyrimidines (C, T and U). 'Transversions' swap a pyrimidine for a purine, or vice versa. Blue lines, old DNA strands; orange lines, new strands produced after DNA replication.**

cytosines in AID-expressing bacteria[3]. In this model, AID produces a lot of uracils in the variable genes (Fig. 1a). The uracils are mistakes, as the usual bases in DNA are cytosine, guanine, adenine and thymine, and need to be corrected — but that often leads to more errors. One 'correction' mechanism involves the DNA polymerase enzymes that copy DNA, which mistakenly read uracil as thymine during replication, and introduce errors accordingly (Fig. 1b). Other correction mechanisms rely on the uracil first being removed by uracil glycosylase, leaving a gap. Mutations can then occur as the cell fills in the gap by one of three pathways — one DNA-replication and two DNA-repair pathways (Fig. 1c–e).

Di Noia and Neuberger[1] now provide support for the conversion of cytosine to uracil in DNA in eukaryotic cells. They took advantage of a chicken B-cell line engineered previously by Neuberger's group. These cells lack XRCC2, a protein involved in DNA recombination and repair, and so tend to deal with uracil in DNA by the replication pathways (Fig. 1b, c), not by DNA repair. Moreover, they mainly generate transversion-type mutations in the variable genes. In transversions, a pyrimidine base (cytosine, thymine or uracil) is swapped for a purine (adenine or guanine). In transitions, one pyrimidine is swapped for another, or one purine for another. The transversions in these B cells generally involve replacements of cytosine or guanine (rather than of adenine or thymine), and this probably occurs by the replication pathway that relies on the initial removal of uracil by the glycosylase (Fig. 1c).

The authors' experiments were simple yet effective. The uracil glycosylase was inhibited in the B cells, forcing the cells to deal with uracils through the other replication pathway. By its nature this pathway generates transitions of an initial cytosine–guanine base pair to a thymine–adenine pair (Fig. 1b). Di Noia and Neuberger reasoned that if the production of uracil in variable-region genes is the trigger for hypermutation, and if cells cannot deal with the uracil by removing it but instead copy over it, then the predominant mutation pattern would shift from transversions of cytosine and guanine to transitions. And this is exactly what they saw.

So what about hypermutation in mammalian B cells? Again, the generation of uracil could be the crucial starting point. In hypermutating mammalian B cells all four nucleotides are mutated with similar frequencies in variable genes, and transitions are favoured. All replication and repair pathways could be used to process AID-generated uracils, causing these patterns of hypermutation. After uracil is removed by either uracil glycosylase or another glycosylase, SMUG1, processing may be shunted into the repair pathways, allowing mutations of adenine and thymine — rather than just cytosine and guanine — to occur (Fig. 1d, e). During error-prone base-excision repair (Fig. 1d), the gap may be incorrectly filled in by DNA polymerase ι (ref. 6), which erroneously inserts thymine opposite guanine[7]. During short-patch repair synthesis (Fig. 1e), polymerase η could fill in the gap and synthesize several bases by strand displacement opposite adenines and thymines[8,9]. The mismatch-repair proteins MSH2 and MSH6, and polymerase ζ, may also be involved[10–12].

The identification of AID-generated uracils as the likely starting point is a major breakthrough in understanding hypermutation. Nonetheless, questions remain. The most important one is how AID targets the rearranged variable genes. Perhaps AID is delivered to these genes by interacting with transcription factors[13]. Also, why is the usually accurate base-excision repair using polymerase β subverted to mutation-generating repair with polymerase ι in variable genes? What is the exact role of MSH2 and MSH6? AID also initiates two other antibody-alteration processes: 'class switch recombination' in mammals, and variable-gene conversion in chickens. How can the DNA-deaminase model explain these very different events? Finally, before the model is totally convincing it must be shown that AID actually does deaminate cytosine in DNA, and that uracils are present in the variable genes. But these are exciting times for immunology, and it appears that one of the last black boxes has been opened. ■

Patricia J. Gearhart is in the Laboratory of Molecular Gerontology, National Institute of Aging, National Institutes of Health, 5600 Nathan Shock Drive, Baltimore, Maryland 21224, USA.
e-mail: gearhartp@grc.nia.nih.gov

1. Di Noia, J. & Neuberger, M. S. *Nature* **419,** 43–48 (2002); advance online publication, 31 July 2002 (doi:10.1038/nature00981).
2. Muramatsu, M. *et al. Cell* **102,** 553–563 (2000).
3. Petersen-Mahrt, S. K., Harris, R. S. & Neuberger, M. S. *Nature* **418,** 99–103 (2002).
4. Bross, L., Muramatsu, M., Kinoshita, K., Honjo, T. & Jacobs, H. *J. Exp. Med.* **195,** 1187–1192 (2002).
5. Papavasiliou, F. N. & Schatz, D. G. *J. Exp. Med.* **195,** 1193–1198 (2002).
6. Bebenek, K. *et al. Science* **291,** 2156–2159 (2001).
7. Frank, E. G. *et al. EMBO J.* **20,** 2914–2922 (2001).
8. Zeng, X. *et al. Nature Immunol.* **2**, 537–541 (2001).
9. Pavlov, Y. I. *et al. Proc. Natl Acad. Sci. USA* **99,** 9954–9959 (2002).
10. Phung, Q. H. *et al. J. Exp. Med.* **187,** 1745–1751 (1998).
11. Wiesendanger, M., Kneitz, B., Edelmann, W. & Scharff, M. D. *J. Exp. Med.* **191,** 579–584 (2000).
12. Diaz, M., Verkoczy, L. K., Flajnik, M. F. & Klinman, N. R. *J. Immunol.* **167,** 327–335 (2001).
13. Peters, A. & Storb, U. *Immunity* **4,** 57–65 (1996).