

Table 1. Prevalence of herpesvirus-like sequence in Kaposi's sarcoma (KS) tissue and blood of KS patients. Studies were conducted as described in text (2). Some of the tumor tissues and corresponding controls were provided by the University of California AIDS Specimen Bank under the direction of Dr. John Greenspan. Peripheral blood mononuclear cells from normal subjects came from healthy volunteers in the laboratory or were provided by Irwin Memorial Blood Centers, San Francisco, California.

Specimen	Subject's status		KS330Bam
	KS	HIV	
KS tumor	+	+	12/12
control tissue	+	+	4/11
KS tumor	+	—	1/1
control tissue	+	—	0/1
PBMC	+	+	7/7*
PBMC	+	—	3/3
PBMC	—	+	0/6
PBMC	—	—	0/14
PBMC subsets†	+	+	
CD8 ⁺			0/3
CD19 ⁺			3/3
Saliva	+	+	
Cells			0/4‡
Fluid			0/2
Semen	+	+	
Cells			0/5‡
Fluid			0/4

*Testing of samples obtained from at least two different visits of three of these patients yielded positive results. Samples from three of these HIV-infected KS patients were used for the subset analysis. †Cell populations with $\geq 90\%$ CD8⁺ or CD19⁺ cells. ‡Repeat samples from two patients on separate occasions were evaluated and also gave negative results.

semen of KS patients. Cells were removed by low-speed centrifugation and fluids were ultracentrifuged at 100,000g for 2 hours in a Beckman SW28 rotor. The cells and high speed pellets were resuspended, extracted, and analyzed as described (2). The KS330Bam sequence could not be detected in these body fluids (Table 1).

These results, showing the presence of the KS330Bam sequence in KS tissue and PBMC of both HIV-infected and uninfected individuals, provide further evidence of the specific association of a herpesvirus-like sequence with this malignancy. Its absence in the KS-derived cell lines may indicate the lack of cultivation of the KS cell itself, or an indirect effect of a herpes-like virus on endothelial cell proliferation, for example via cytokines as has been proposed (8). The preliminary data showing the absence of the KS330Bam sequence in saliva and semen may be a result of sampling error, and the potential route of transmission of this KS-associated agent remains unknown. Further evaluation is also needed to determine the distribution of this KS-associated agent in the population.

John A. Ambroziak
David J. Blackburn
Cancer Research Institute,
University of California,
San Francisco, CA 94143-0128, USA
Brian G. Herndier
Department of Pathology,
San Francisco General Hospital,
University of California,
San Francisco, CA 94143, USA
Richard G. Glogau
Department of Dermatology
University of California,
San Francisco, CA 94143, USA
John H. Gullett
Department of Medicine,
University of California,
San Francisco, CA 94143, USA
Alex R. McDonald
Department of Oral and
Maxillofacial Surgery,
University of California,
San Francisco, CA 94143, USA
Evelyn T. Lennette
Virolab,
Berkeley, CA 94710, USA
Jay A. Levy
Department of Medicine and
Cancer Research Institute,
University of California,
San Francisco, CA 94143-0128, USA

REFERENCES AND NOTES

1. Y. Chang *et al.*, *Science* **266**, 1865 (1994).
2. The PCR procedure specifically amplifies a 233-base pair portion of the KS330 Bam fragment using primers described by Chang *et al.* (1). This 233-base pair fragment is referred to as the KS330Bam sequence in the text, since its detection indicates the presence of the KS330Bam sequence. The conditions for PCR are: 94°C for 1 min (1 cycle); 94°C for 30 s, 58°C for 1 min, 72°C for 90 s (45 cycles); 72°C for 7 min (1 cycle). Each 100- μ l reaction contains 0.5 to 0.75 μ g of genomic DNA, 30 pmol of each of primer, 2.5 U of Taq polymerase, 200 μ M each of deoxynucleotide triphosphate, 10 mM tris-HCl, 1.5 mM MgCl₂, and 50 mM KCl (pH, 8.3). Reaction product (10 μ l) was electrophoresed on a 1.6% agarose-TBE gel containing ethidium bromide for 45 min (5 V/cm) to verify the presence or absence of the KS330Bam fragment by comparison with the amplified product of a positive control reaction. This control KS330Bam fragment, amplified from genomic DNA extracted from a KS tumor biopsy, was partially sequenced and shown to be identical to the KS330Bam sequence described by Chang *et al.* (1). PCR amplification of the human p53 tumor suppressor gene with primers P6-5 and P6-3 (7) was used to confirm DNA integrity. Total genomic DNA was isolated from tissue and PBMC by standard techniques.
3. B. G. Herndier *et al.*, *AIDS* **8**, 575 (1994).
4. J. A. Levy, in *AIDS: Pathogenesis and Treatment*, J. A. Levy, Ed. (Dekker, New York, 1989), pp. 159-229.
5. J. A. Levy *et al.*, *Virology* **178**, 113 (1990).
6. PBMC were obtained from heparinized blood by Ficoll-Hypaque separation per J. A. Levy *et al.* (5).
7. C. E. Mackewicz, H. W. Ortega, J. A. Levy, *J. Clin. Invest.* **87**, 1462 (1991).
8. J. A. Levy and J. L. Ziegler, *Lancet* **ii**, 78 (1983).

1 March 1995; accepted 31 March 1995

Building an Associative Memory Vastly Larger Than the Brain

Leonard M. Adleman proposes using the tools of molecular biology to solve computer science problems such as the directed Hamiltonian path problem (1). Richard J. Lipton (2) describes an improved biological procedure to solve directly any computational problem in the class NP and suggested approaches to speed up solution of some important practical problems. It is possible, at least in principle, to use these tools to produce an associative, or content addressable, memory of immense capabilities. A content addressable memory is one where a stored word may be retrieved from sufficient, partial, knowledge of its content, rather than needing to know a specific address as in standard computer memories. Content addressable memories are useful in a number of computer contexts and are widely thought to be an important component of human intelligence.

The memory is conceptually simple. It consists of a vessel containing DNA. It would be able to store binary words of a fixed length. One would write such a word in memory by placing in the vessel an appropriate single strand DNA molecule en-

coding it (3). In the simplest approach, two distinct DNA subsequences would be assigned to each component—the first coding for a “one” at that component and the other coding for a “zero.” The DNA molecule coding a particular word would be composed by concatenating the appropriate subsequences corresponding to its particular bits, in any order (4).

This memory would be “content addressable.” Given a “cue” consisting of a subset of the component values, one might retrieve any words in the memory consistent with these values as follows. For each component specified in the cue, one could introduce in turn the complement of the corresponding subsequence of DNA, affixed to a magnetic bead. This complement would then bond to DNA molecules in the memory having that subsequence, that is, coding for words containing that component value. A similar bonding procedure was used by Adleman in his computation (1). These molecules could then be extracted magnetically. After iteratively extracting on each component in the cue, one would retrieve molecules matching the cue exactly and

these could then be sequenced in order to read the stored word in its entirety.

Alternatively, it might be possible to introduce complementary subsequences for all the components specified in the cue simultaneously, with some marker attached to each complementary subsequence. In this implementation, the sequence with most marks would be most consistent with the cue and could be extracted. Such a parallel implementation would both save time and allow a full associative retrieve by finding the word most closely matching the cue even if the match is not exact. Such search and read operations would be logically analogous to the operation of a grandmother cell associative memory as described in (5), only realized biochemically instead of electronically. The biotin beads used by Adleman (1) are probably too large to perform this function, so an alternative marker technology would be necessary.

After the read operation, any DNA molecules removed would have to be reintroduced into the memory, so that stored words would not be lost. One way to do this simply would be to split these molecules back to single strand DNA. The strands without beads attached would then be the ones we wish to reintroduce. We might also wish to keep a copy of the entire memory, and return to the copy after several read operations.

By choosing a somewhat different encoding, and by care in choosing the DNA subsequences, one could achieve greater computational efficiency and ease in performing the relevant biochemical manipulations.

It would generally be more efficient to encode only the components for which the stored word has value "one." In this case we would assign only one DNA subsequence to each component. We would encode a word by concatenating all the subsequences corresponding to a component having value one. Thus relatively short DNA molecules could represent sparse vectors (that is, containing few ones) from a very large vector space. This encoding scheme would not encode any information in the order of the component subsequences on the molecule, but could otherwise use all potential bits of information.

There would be considerable freedom in choosing the DNA subsequences corresponding to components. The only constraint would be that none is a subsequence of another, nor of two others concatenated, as this would lead to ambiguities in the recall process. Typically subsequences between 10 and 100 nucleotides long would suffice to avoid such ambiguity (depending on the size of the memory). It might be convenient to end all subsequences with a distinctive "stop" subsequence. This would avoid ambiguities arising from concatena-

tion, and if the "stop" subsequence is chosen to be broken by a particular restriction enzyme, would facilitate chopping a DNA molecule into its component subsequences. One could use as subsequences, vectors chosen from an optimized code or alternatively naturally occurring DNA fragments. If the complement of one sequence is not reused to code for another sequence, then any complementary fragments which creep in during augmentation or reading would not corrupt the memory (since they would not be interpreted as stored words), and could be readily removed. Probably it would be convenient to choose the subsequences to facilitate the read operation.

Current technology is such that "automated oligonucleotide synthesizers [can] rapidly produce any DNA molecule containing up to about 100 nucleotides" (6, p. 319). These "can be joined together by repeated DNA cloning steps in various combinations to produce long custom designed DNA's of any sequence" (6, p. 319). One can also produce long DNA sequences even more easily by linking selected naturally occurring segments (6, p. 320).

Once representatives of component subsequences have been created, they could be cheaply copied. Writing would be relatively quick and inexpensive because we would already have test tubes containing each, previously created, component sequence. We would also have previously created sequences for use in "splinting" component subsequences together (7). If we begin each component subsequence with a start sequence A as well as a stop sequence B, the molecule BA (that is, the Watson-Crick complement of sequence B ligated to the complement of sequence A) would serve as a general purpose splint to join component subsequences in an order-independent manner. Alternatively, we could arrange to append subsequences in a specific order, using more specific splinting sequences.

The search for a stored molecule when we only encode the positions of "ones" is analogous to that described above.

When a DNA molecule is found containing the cue components, it must be sequenced to read out the rest of the word. DNA molecules may now be rapidly sequenced using a variety of techniques (6, chapter 7). This process might be expedited in our case by appropriate choice of component subsequences, as it is only necessary to identify which component subsequences are present, not to sequence them in detail. In fact it would be possible to choose the component subsequences to be unambiguously distinguishable using restriction mapping (8). In this case, reading could be accomplished much more rapidly than would be necessary to sequence an arbitrary molecule.

To expedite reading or to store more

information, a slightly more elaborate storage scheme would be possible. Consider storing DNA molecules consisting of an address portion and a data portion. The address portion would be created exactly as described above and addressed by content. When it is retrieved, the data portion would be sequenced and read. The information in the data portion might be identical, or different, from that stored in the address portion. The coding of the data portion might be identical, or different. One could imagine storing information in the data portion simply as a base four word, in which case it would be very compact, but would require full sequencing technology to read.

This technology could also be used as an ordinary random access memory in this fashion (9), where one must supply the full address portion in order to read a word. If the memory is used in this way, the read operation would be somewhat simplified, as it would be possible to attach a bead only to DNA molecules coding for the word desired to be recalled. This would be expedited if the address sequences were chosen from an error correcting code, making far less likely accidental bonding of retrieval sequences bonded to beads to molecules they only approximately match. Thus no other molecules would be removed from the memory during the read operation, and it would be much simpler to restore the memory after the read to its pre-read state.

Standard database operations such as delete are, of course, possible. To delete, for example, we would simply remove the appropriate molecules from the database.

With present technology, the search operations described above using beads would be performed on single-strand DNA, while sequencing and restriction mapping are done on double-strand DNA. It is possible but fairly slow to transform DNA from double to single strand, and vice versa. Thus it might be worthwhile to use an address portion that is single strand and a data portion that is double strand.

The storage that is in principle possible using these techniques is staggering. It is not completely implausible to imagine vessels (10) storing, say, 10^{20} words each vessel encoding several thousands, or even several tens of thousands, of bits. This compares to standard estimates of brain capacity as, perhaps, 10^{14} synapses each storing a few bits (11). With current technology, the read and write times would be on the order of hours. [Adelman's pathbreaking, manual computation took him 7 days (1).] But the current rate of technological progress in molecular biology is rapid, and there is no obvious fundamental physical limitation preventing achieving automated read and write operations on a much faster time scale. It is also worth noting that once a quantity of infor-

mation were encoded in DNA in this fashion, the whole vessel could be copied relatively easily by DNA replication; and likewise that the information in vessels could be readily merged (by physically merging them). DNA-based computing could conceivably provide a technological basis for superhuman intelligence.

Eric B. Baum
 NEC Research Institute,
 4 Independence Way,
 Princeton, NJ 08540, USA
 E-mail:eric@research.nj.nec.com

REFERENCES AND NOTES

1. L. M. Adleman, *Science* **266**, 1021 (1994).
2. R. J. Lipton, *ibid.* **268**, 542 (1995).
3. In principle a single molecule would suffice to encode the word. In practice it might be advisable to introduce many copies of the molecule encoding the word.
4. A similar encoding of Boolean vectors was previously used by Lipton (2).
5. E. B. Baum, J. Moody, F. Wilczek, *Biological Cybernetics* **59**, 217 (1988).
6. B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, J. D. Watson, *Molecular Biology of the Cell* (Garland, New York, 1994).

7. This splinting technique was used, for example, by Adleman (1).
8. A large number of restriction enzymes are known which cleave DNA molecules at particular locations. By subjecting a given DNA molecule to different restriction enzymes, one cleaves it into pieces of different, but well-defined lengths depending on the sequence. The lengths of the segments then provide constraints on the sequence. This is known as "restriction mapping" (6, p. 294). In our case, we need merely design sequences that can be readily distinguished one from another by the various lengths they are chopped into by various restriction enzymes. This would not be difficult.
9. J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular Cloning, A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, ed. 2, 1989).
10. A reasonable concentration for DNA is 0.06 g/liter of water (9). A millimole of molecules each 200 bases long would weigh about 50 g, hence occupy about 1000 liters. It is possible that practical limitations, for example, avoiding excessive annealing times in retrieval, would necessitate assigning a number of identical molecules to code each word.
11. One should, however, wonder whether perhaps huge numbers of bits are stored molecularly in each neuron or synapse, as techniques like those outlined here are available to biology.
12. I thank P. Kaplan and W. D. Smith for helpful comments.

10 January 1995; accepted 11 April 1995

Neutrophils and Drug Metabolism

In their report (1) Xiaoxia Jiang *et al.* studied the role of phagocytes in drug metabolism by measuring the anti-proliferative effect of products from drug-exposed neutrophils. They did not, however, present physicochemical evidence for reactive drug metabolites reportedly generated by these cells and did not provide evidence for an immunostimulatory effect of these products that might lead to drug-induced lupus. We and others demonstrated specific immune reactions to defined metabolites of lupus-inducing drugs. Unlike the parent drugs or unreactive metabolites, the reactive metabolites tested proved immunogenic for T cells and, in a T cell-dependent fashion, activated B cells. Thus, the parent drugs procainamide (PA) (2), propylthiouracil (PTU) (3) (both of them studied by Jiang *et al.*), and gold(I) thiomalate (3) all were shown not to be immunogenic for T cells, and the same results were obtained for the unreactive metabolite *N*-acetyl-PA (1). In contrast, the reactive intermediates of these drugs, hydroxylamino-PA (2), PTU-sulfonate (3), and gold(III) (4, 5) were shown to be immunogenic. Moreover, when the parent drugs PA, PTU, and gold(I) thiomalate were preincubated in vitro with either polymorphonuclear or mononuclear phagocytes, or myeloperoxidase/H₂O₂/Cl⁻, the respective immunogenic metabolites, hydroxylamino-PA, PTU-sulfonate, and gold(III)

were generated, as detected by chemical analysis (6) or specific responses of metabolite-sensitized T cells (2, 3, 5). Similarly, upon long-term administration to mice of the parent drugs PA (2), PTU (3), or gold(I) thiomalate (5), the respective immunogenic metabolites were demonstrable in phagocytic cells of these animals. In contrast to short-term treatment with the parent drugs, long-term treatment, which allows for more extensive generation of the reactive intermediates, also resulted in sensitization of T cells to the intermediates (3-5). Both the generation of and the T cell sensitization to these metabolites could be enhanced by stimulating the oxidative burst in phagocytes (2, 3), or, in the case of PA, by using mice that express the slow-acetylator phenotype (2). Finally, mice under long-term treatment with gold(I) thiomalate or PTU exhibited immunoglobulin G autoantibodies and signs of vasculitis (3).

Thus, the evidence weighs against the hypothesis of papers cited by Jiang *et al.* that reactive drug metabolites may induce lupus because they exert a mitogenic effect on lymphoid cells in vivo. Our findings support the view that the immunostimulation underlying drug-induced autoimmunity involves specific T helper cell responses to hitherto undefined self proteins altered by reactive drug metabolites. We do agree with Jiang *et al.* that generation of the respective

metabolites by cells of the immune system itself facilitates the development of adverse immune reactions to drugs.

Sherko von Schmiedeberg
 Department of Dermatology,
 Heinrich Heine University,
 D-40225 Düsseldorf, Germany
Carsten Goebel
Ernst Gleichmann
 Division of Immunology,
 Medical Institute of Environmental Hygiene,
 Heinrich Heine University
Jack Uetrecht
 Faculty of Pharmacy and Medicine,
 University of Toronto,
 Toronto, Ontario, M5S 2S2, Canada

REFERENCES

1. X. Jiang, G. Khursigara, R. L. Rubin, *Science* **266**, 810 (1994).
2. M. Kubicka-Muranyi, R. Goebels, C. Goebel, J. Uetrecht, E. Gleichmann, *Toxicol. Appl. Pharmacol.* **122**, 88 (1993); C. Goebel, M. Kubicka-Muranyi, R. Goebels, E. Gleichmann, *Immunobiol.* **189**, 114 (abstr.) (1993).
3. S. v. Schmiedeberg, H.-C. Schuppe, E. Csernok, W. Gross, E. Gleichmann, *Immunobiol.* **191**, 228 (abstr.) (1994); S. v. Schmiedeberg *et al.*, in preparation.
4. D. Schuhmann *et al.*, *J. Immunol.* **145**, 2132 (1990); J. Verwilghen, G. H. Kingsley, L. Gambling, G. S. Panayi, *Arthr. Rheum.* **35**, 1413 (1992).
5. C. Goebel, M. Kubicka-Muranyi, T. Tonn, J. Gonzales, E. Gleichmann, *Arch. Toxicol.*, in press.
6. B. Beverly and D. Couri, *Fed. Proc.* **46**, 854 (1987); J. P. Uetrecht, *Drug Metab. Rev.* **24**, 299 (1992); C. F. Shaw *et al.*, *Metal-Based Drugs* **1**, 351 (1994).

6 January 1995; accepted 6 April 1995

Response: We presented in our report (1) physicochemical evidence showing that all the tested lupus-inducing drugs were transformed by myeloperoxidase [figure 3 in (1)]. However, only the cytotoxic metabolite of procainamide, procainamide-hydroxylamine, has so far been identified. It should also be noted that we did reference the Kubicka-Muranyi paper (2) on the response of popliteal lymph node cells to procainamide-hydroxylamine. Similar studies on gold salts were not referenced because there have been no reports to our knowledge during the past 25 years of lupus related to gold therapy despite wide use of gold compounds in treatment of rheumatoid arthritis.

Adverse reactions to drugs, some of which are immune mediated, are common, but lupus-like disease related to more than 40 different medications is relatively rare. However, drug-induced lupus is provocative because it strongly resembles systemic lupus erythematosus (SLE), a disease for which there is no known cause. Influenced by the precedent set by penicillin-mediated allergic reactions, most studies have focused on the capacity of drugs or drug metabolites to bind components of the peripheral immune system, such as drug-altered self-proteins, T cells, or their contact sites on antigen pre-