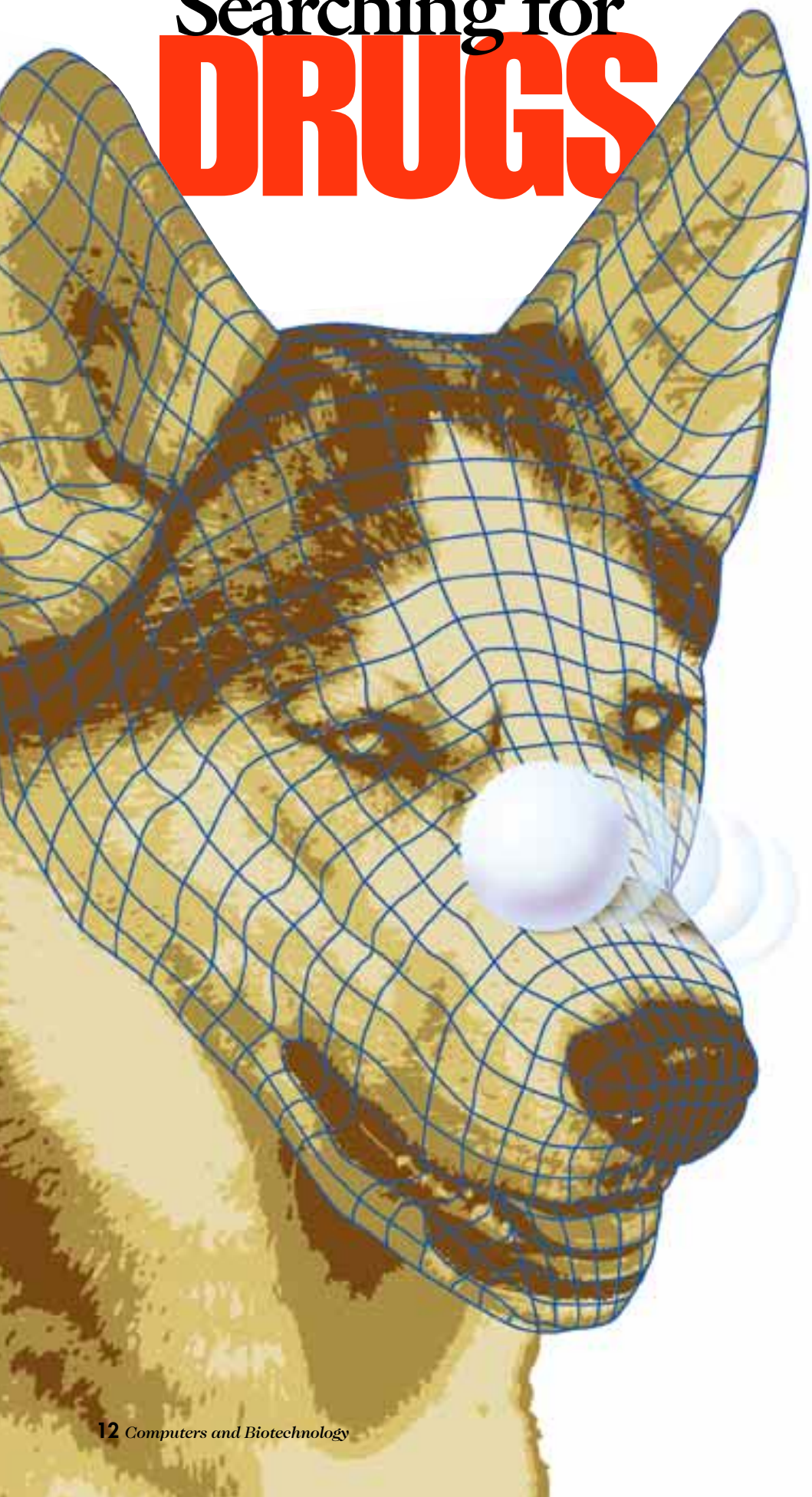
Searching for DRUGS



Finding Similar Molecules

Suppose you know of a molecule that binds to a protein in a way that partially treats a disease. Perhaps, however, that molecule doesn't make a good drug. (It may interact weakly, or can't get into a cell, or has unwanted side effects, or also interacts with another protein in a harmful way.) Perhaps another molecule could bind to the same protein but work better. How do we find it? Traditionally, scientists tried to change the molecule's structure to mimic other drugs – or they used the old hit-or-miss, trial-and-error method. Sometimes this works. But surely it also misses molecules.

One way to expand the search is to teach a computer to analyze complicated shapes. We humans recognize shapes very quickly. We immediately know the difference between a cat and a dog. But can a computer tell the difference? Both animals have a head, two eyes, a nose, a body, four legs, and a tail. How could a computer distinguish two breeds of dogs – a task that is closer to the problems of drug discovery? Here are some methods scientists have developed.

Ping Pong Ball

Let's pretend you want to analyze a particular dog's shape. You can use a ping pong ball as a probe and roll it over the entire surface of the dog. At each point on the surface, a computer records the exact location of the ball on a three-dimensional grid. The entire set of coordinates (X, Y, Z) of the ball in the grid represents the dog's shape.

In the same way, a computer traces a molecule in a three-dimensional grid, measures the location of a small probe molecule when it bumps up against the larger molecule and records the coordinates. This is one method called "quantitative structure activity relationship" or QSAR. We can use it to record the shape of several molecules that we know "lock" to a certain protein and then ask the computer to search for a "key" feature that all those molecules share.

Then scientists can search for other molecules with the same feature. Scientists used a QSAR approach in developing the new antibacterial drug norfloxacin, the Alzheimer's treatment donepezil (Aricept™), and the migraine drug zolmitriptan (Zomig™).

The Animal's Den

Another approach allows us to guess the shape of the protein's binding site by examining the molecules that bind to it. What shapes do these molecules have in common? What can that tell us about the protein's "lock" (binding site)? Scientists call this approach the "pharmacore" method, and it is like trying to define the shape of a den by examining the animals that fit in it. If they are all long and thin, the den is probably narrow and deep. Once you have modeled the shape of the protein's "den" (or "lock"), you can look for molecules that have a "key" that would fit it. These molecules may all have quite different chemical properties, but they share a similar "key" for the same protein.

Finding Different Molecules

The ping pong ball and animal's den approach only get us so far in our search for new molecules, though. Suppose you want to find molecules that are different from those you have studied? Perhaps none of the known molecules interacts in just the right way with a protein involved in a deadly disease. In that case, you may want to combine different molecules together to make new compounds.

This method, called "combinatorial chemistry," allows scientists to synthesize millions of new compounds. It is like choosing all possible combinations from a Chinese restaurant menu: an entry from column A combined with each meat option from column B; the second entry from A combined with each option from B, and so on. If each column had ten thousand options, there are 100 million (10^4 x 10^4) possible combinations. Maybe one of them is just the new drug you want to find...

Now the trick is to screen those compounds to see if any interact with the disease protein. But there are so many compounds to test – even using the robots that are now common in labs! And many compounds have almost identical shapes and biological activity. Scientists would like to test only those with significantly different shapes and – probably – different activities. Computers are helping weed out the similar shapes, so scientists can concentrate on representatives from different families of shapes. Then, if they find one promising molecule, they can study its other family members.

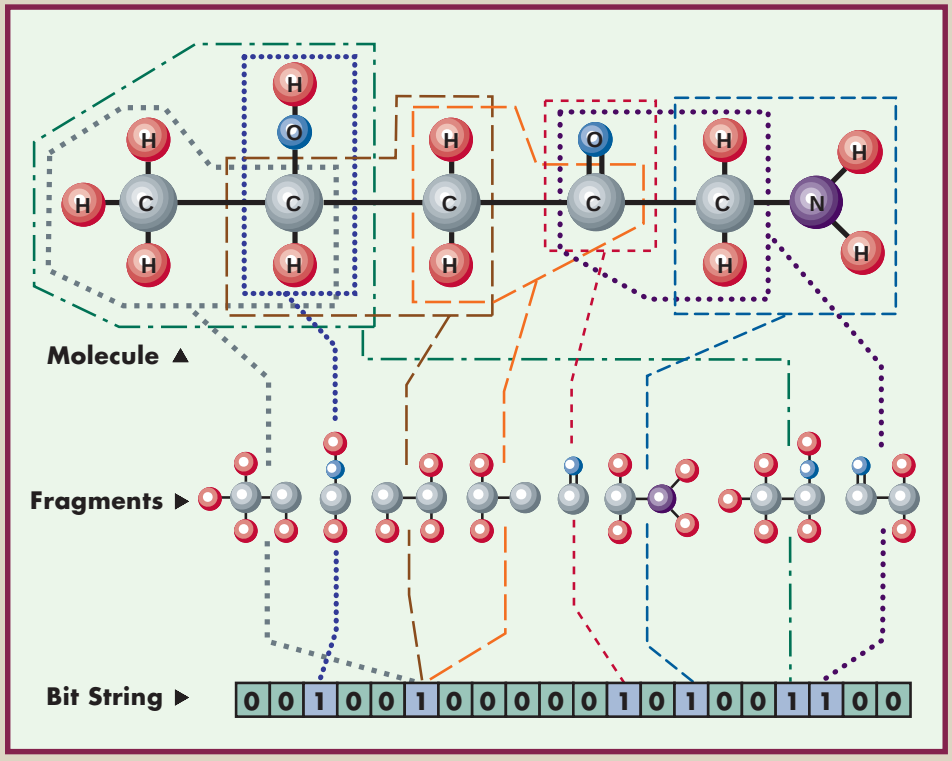
Alternatively, they can use these methods to look for molecules with shapes similar to one they know, in order to develop variations on a successful drug theme. That's how new alternatives to the antidepressant drug fluoxetine (Prozac®) were created; scientists looked for molecules with similar "keys" to that which allows the drug molecule to interact with chemicals in our brains.

Fragments

The most common way to sort molecules is to break them into fragments, like breaking the shape of a dog down into nose, tail, ears, legs, etc. Then, a computer represents the pattern of those fragments as "bit strings." For each possible fragment, it assigns a "1" or a "0" to show if the fragment is present to create a string like this: 0001001111011010. Then the computer compares millions of bit streams. Molecules with similar strings may have similar shapes and binding properties, so computers can sort molecules into families. Scientists can use these families either to concentrate on differently shaped molecules or to isolate members of one family.

Virtual Libraries

Computerized "virtual libraries" store the shapes of larger molecular fragments. Computers then "cook up" all the possible shape combinations in the Chinese menu columns and point out those "meals" that are similar and different. Once scientists identify molecules of interest, they can synthesize and test them. Alternatively, the computer can model the molecule's structure and superimpose it on the protein to see if it fits.



Fragments: The molecule on the top row is broken down into some of the possible overlapping fragments, which are shown in the second row. The third row shows a bit string where these fragments are represented as a "1". (The "H" (hydrogen) atom doesn't affect the activity of the molecule, so H_3C-CH, CH-CH_2, and CH_2-C all produce the same fragment, C-C.)