RESEARCH: PROTEIN STRUCTURE

Sieves in Sequence

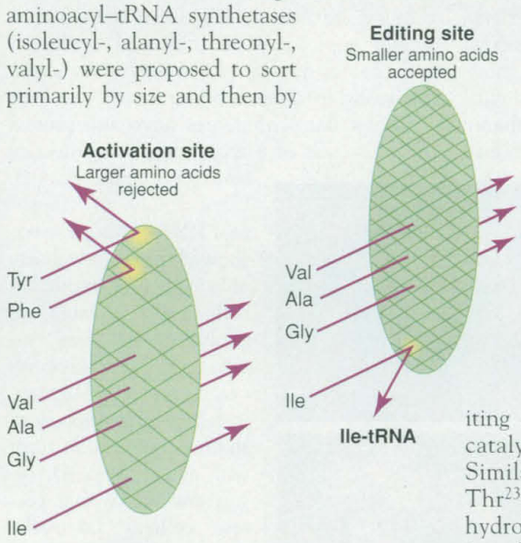
Alan R. Fersht

In an obscure publication in 1958, Linus Pauling put his finger on the fundamental problem in molecular recognition: A protein must keep ligands that are smaller than the correct one out of its binding site (*1*). His classic example, translated into modern terms, was the following: How does the isoleucine-selective enzyme, isoleucyl–tRNA synthetase (IleRS), exclude valine, which is smaller than isoleucine by only one methylene group? The chemical process of selection by this enzyme requires two steps. Isoleucine is first converted into the rather unstable isoleucyl-adenylate, which remains bound to IleRS, and is then transferred to its specific transfer RNA (tRNA) to yield Ile-$tRNA^{Ile}$. In this issue, the structure of the isoleucyl–tRNA synthetase, now revealed in molecular detail, provides the resolution of this classic problem (*2*).

The dimensions of the dilemma were originally highlighted by R. B. Loftfield who showed (*3*, *4*) that IleRS is remarkably accurate and misincorporates valine for isoleucine into the proteins ovalbumin and globin in only 1 out of 3000 positions. Yet the enzyme activates both isoleucine and valine to the adenylated form with similar turnover numbers, and the activated valine remains bound to the enzyme with only a 150-fold weaker affinity. This paradox was resolved rapidly by A. N. Baldwin and P. Berg (*5*), who found that addition of $tRNA^{Ile}$ to the IleRS·valyl–adenylate complex led to hydrolysis of the valyl adenylate rather than synthesis of Val-$tRNA^{Ile}$. This experiment gave birth to the concept of "editing" or "proofreading" mechanisms.

Maintaining an accurate flow of information from DNA, through replication, to protein synthesis is beyond the limits of conventional specificity, and so the cell has introduced proofreading mechanisms at several steps. The cell checks the accuracy of DNA replication (*6*), proofreads the work of other aminoacyl-tRNA synthetases that have smaller or isosteric competing substrates, and monitors translation from RNA to protein (*7*).

The "strong" force in specificity is steric repulsion: Whereas a smaller substrate can always rattle around in a larger cavity, it is energetically very difficult to cram a larger substrate into a cavity built for a smaller one. On the basis of this idea, I proposed the "double-sieve" editing mechanism for sorting amino acids (*8*): Objects can be sorted into size ranges by running them over a coarse sieve that allows the objects smaller than the mesh to fall through onto a finer grid, which in turn allows the objects smaller than its mesh to fall through. Some of the aminoacyl–tRNA synthetases (isoleucyl-, alanyl-, threonyl-, valyl-) were proposed to sort primarily by size and then by specific chemical features (see the figure). The active site for aminoacylation acts as the coarse sieve, activating at a significant rate only those amino acids that are the same size as or smaller than the desired one. The active site for hydrolysis is the fine sieve, which destroys the products of those amino acids that are smaller than the correct ones. There are of course gradations of specificity due to variations in binding energies, and specific chemical binding features take care of isosteric amino acids. An alternative proposal, "kinetic proofreading," put forward by J. J. Hopfield (*9*), uses the same features of the coarse sieve twice over.



Checking it twice. The double-sieve mechanism for the isoleucyl–tRNA synthetase (*8*). Hydrolytic editing reduces the error rate for the misactivation of valine from an expected value between 1 in 10 and 1 in 100 to 1 in 40,000.

Now, 40 years after the original posing of the problem, Nureki and colleagues have beautifully demonstrated that the isoleucyl–tRNA synthetase operates at high specificity by double-sieve editing (*2*). They solved the structure of the isoleucyl–tRNA synthetase from *Thermus thermophilus* and its complexes with isoleucine and valine. The protein contains a nucleotide-binding fold that binds ATP. The fold has two characteristic ATP-binding motifs: His^{54}-Val^{55}-Gly^{56}-His^{57} and Lys^{591}-Met^{592}-Ser^{593}-Lys^{594}. In the L-Ile·IleRS complex, a single Ile is bound at the bottom of the ATP cleft, with the hydrophobic side chain in a hydrophobic pocket, surrounded by Pro^{46}, Trp^{518}, and Trp^{558}. L-Leucine cannot fit into this pocket because of the steric hindrance of one of its terminal methyl groups. Larger amino acids are similarly excluded from this site. In the L-Val·IleRS complex, valine is bound to the same site, but the contact area with Pro^{46} and Trp^{558} is lower. This site is the coarse sieve. A long polypeptide sequence is inserted into the nucleotide-binding fold. This forms a four-domain structure, which has been shown by mutagenesis to contain the hydrolytic editing site (*10*). In the L-Val·IleRS complex, a second molecule of Val is bound here, whereas no electron density was observed for isoleucine in the L-Ile·IleRS complex. The binding cavity is surrounded by Trp^{32} and Tyr^{386} and is too small to accommodate isoleucine. Deletion of the 47 residues that constitute this site in the IleRS from *Escherichia coli* abolishes the editing activity, and the mutant efficiently catalyzes the formation of Val-$tRNA^{Ile}$. Similarly, mutation of the equivalent of Thr^{230}, the putative nucleophile in the hydrolytic activity, to Ala abolishes editing. The inserted sequence is the fine sieve. The hydrolytic editing activity is induced by the addition of $tRNA^{Ile}$ to the IleRS·Val-AMP complex. Model building suggests that one of the domains in the insert rotates on the addition of the tRNA so that it and the two active sites form a closed cavity in which the aminoacylation and editing clefts face each other.

The author is in the Department of Chemistry and Centre for Protein Engineering, Lensfield Road, Cambridge CB2 1EW, UK. E-mail: arf10@cam.ac.uk

References

1. L. Pauling, in *Festschrift Arthur Stoll* (Birkhäuser Verlag, Basel, Switzerland, 1958), p. 597.
2. O. Nureki *et al.*, *Science* **280**, 578 (1998).
3. R. B. Loftfield, *Biochem. J.* **89**, 82 (1963).
4. _____ and D. Vanderjagt, *J. Mol. Biol.* **71**, 1353 (1972).
5. A. N. Baldwin and P. Berg, *J. Biol. Chem.* **241**, 839 (1966).
6. D. Brutlag and A. Kornberg, *ibid.* **247**, 241 (1972).
7. R. C. Thompson and P. J. Stone, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 198 (1977).
8. A. R. Fersht, *Enzyme Structure and Mechanism* (Freeman, San Francisco, 1977), p. 283.
9. J. J. Hopfield, *Proc. Natl. Acad. Sci. U.S.A.* **71**, 4135 (1974).
10. L. Lin, S. P. Hale, P. Schimmel, *Nature* **384**, 33 (1996).