Enzyme catalysis: not different, just better

Jeremy R. Knowles

Where are we in our understanding of enzyme catalysis? The gloomier view is that protein structure and enzyme function are the finely balanced end-products of many weak interactions that can be summed only by massive computing power, and more precise parameterization than we enjoy at present. The cheerier position is that proteins are built on definable principles, and that enzymes use recognizable catalytic devices that will allow us to understand how existing enzymes work and to design new ones. To assess which interpretation is the more realistic, the simple reaction catalysed by triosephosphate isomerase is considered here. This examination illustrates some of the catalytic features of enzymes that are understood, and exposes a few that are not. But overall, the question turns out to have an optimistic answer.

TRIOSEPHOSPHATE isomerase catalyses the chemically simple interconversion of the two triose phosphates, dihydroxyacetone phosphate and *R*-glyceraldehyde 3-phosphate, through the intermediate formation of a *cis*-enediol(ate)[1-3] (Fig. 1). The enzyme merely mediates the protonations and deprotonations necessary to effect two enolizations, one from dihydroxyacetone phosphate to the enediol, and the other from glyceraldehyde 3-phosphate to the enediol. When we compare the catalytic power of the isomerase with that of a simple organic base having a $pK_a \sim 7$ (a notional catalyst from the primordial soup), the isomerase is 10^{10} times more effective[4] (Fig. 2). To improve this primitive catalyst, the kinetic barriers for each of the enolizations must be lowered (Fig. 2, solid arrows), and the enediol(ate) intermediate must be stabilized (Fig. 2, open arrow). How does triosephosphate isomerase do it?

There are, of course, many ways of summarizing the details of an enzyme's active site in terms of devices and motifs, yet in this case a reasonable description of the catalytic apparatus is possible with only three elements: a protein loop that binds and stabilizes the reaction intermediate (fulfilling the open arrow of Fig. 2), and a catalytic base and a catalytic acid to mediate the two enolizations (fulfilling the solid arrows of Fig. 2). What do we know about each of these features?

Intermediate sequestration

Many enzymes have flexible loops, floppy tails, mobile lids, or bivalve-like hinged domains, which appear to close down over the substrate(s) during the chemical changes of the catalytic act[5-10]. Reasons for such sequestration are usually not hard to find: for example, reactive intermediates should not be released into the solvent where they may decompose; stereoelectronics may require a particular substrate conformation to be maintained; water must be excluded; or regiospecificity or stereospecificity demands a tight grip on the reactants. Yet the catalytic value of these conformation changes is often presumed rather than proven.

In triosephosphate isomerase, a loop of ten or eleven amino acids between residues 166 and 176 moves about 7 Å from an 'open' position in the unliganded enzyme into a 'closed' position when there is substrate or inhibitor bound at the active site (Fig. 3*a*)[11-13]. From crystallographic work and molecular dynamics simulations, the conserved residues of this loop (those between residues 168 and 173) seem to form a rigid lid[14] that moves on hinges provided by the unconserved regions on each side. Within the conserved segment, residues 170 to 173 form a looplet that, when the lid is closed, provides a new hydrogen bond to the substrates' phosphate group. As the α carbon of residue 169 is less than 5 Å from that of 174, excision of the four residues of the looplet will not significantly distort the rest of the molecule[15]; but this deletion does make the active site lid too small to envelop the bound substrate (Fig. 3*b*). This mutant protein has been generated, and turns out to have lost nearly 10^5-fold in k_{cat} terms. Substrate binding is not seriously impaired, but the enediol phosphate intermediate analogue, phosphoglycolohydroxamate, is bound much more weakly[15]. It thus appears that the loop closure preferentially stabilizes the enediol phosphate intermediate, as well as the two transition states that flank it (Fig. 2).

Determination of the full energetics of this loopless isomerase confirms that the mutant enzyme has lost its grip on the enediol phosphate intermediate, this failure being vividly evident from the fact that now only one out of six substrate molecules 'makes it' through to product. In the absence of a proper lid, the enediol phosphate intermediate is lost from the truncated enzyme into solution, where it rapidly decomposes (to give methylglyoxal and P_i). (The ingenuity of the enzyme in binding substrate with the phosphate group in the plane of the enediol, so that this decomposition of the intermediate is stereoelectronically disfavoured in the active site, was first noted by Albert Eschenmoser in 1974). The flexible loop of triosephosphate isomerase thus has two catalytic functions: it ensures an efficient throughput of substrate to product, and it stabilizes the reaction intermediate (solid arrow in Fig. 2), which pulls down the free energy of the two flanking transition states. How the loop manages to be so discriminating, in stabilizing the reaction intermediate while affecting the bound substrate and product to a much lesser extent, will have to await a comparison of the high-resolution structures of the enzyme complexed with substrate and with the intermediate analogue phosphoglycolohydroxamate.

Significantly, the loop that we mutilated contains the amino-acid sequence -A-X-G-X-G-K-X-A- (single-letter notation), which has some similarity to the consensus turn that interacts with phosphate groups in some kinases, many dehydrogenases, *ras* p21, and other nucleotide-binding proteins[16-21]. It seems

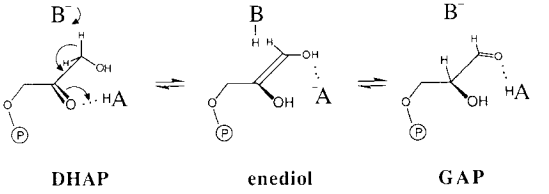

FIG. 1 The reaction catalysed by triosephosphate isomerase. A basic catalytic group B abstracts the pro-*R* proton on C-1 of dihydroxyacetone phosphate (DHAP), assisted by an acidic catalytic group HA, to produce the intermediate enediol. This intermediate then collapses to give the product glyceraldehyde 3-phosphate (GAP) and regenerates the enzyme. The same enediol is produced by the enolization of DHAP or of GAP, and the overall reaction simply puts the two enolizations back-to-back.

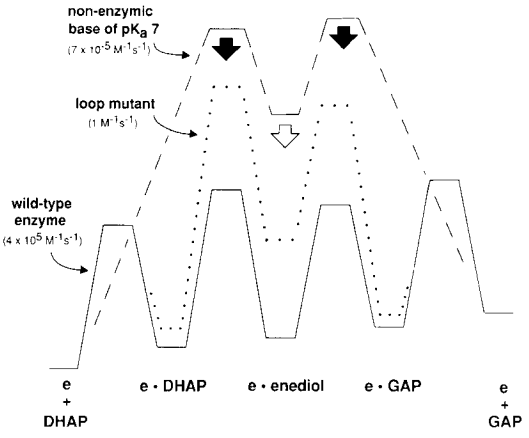

FIG. 2 Free-energy profiles for the isomerization of dihydroxyacetone phosphate (DHAP) and *R*-glyceraldehyde 3-phosphate (GAP), catalysed by wild-type triosephosphate isomerase[45] (solid line), by the loop mutant enzyme that lacks residues 170–173 (ref. 15) (dotted line), or by a non-enzymic base of $pK_a \sim 7$ (ref. 4) (dashed line). e, Enzyme.

that triosephosphate isomerase has simply recruited, for the catalytic purposes described above, the motif that we have (not altogether frivolously) called the 'phosphate gripper'.

A poised base

From experiments involving kinetics, stereochemistry, chemical modification and X-ray crystallography, we know that the carboxylate of glutamate 165 is the base responsible for the abstraction of the pro-*R* proton of the substrate dihydroxyacetone phosphate (or, for the reverse reaction, the proton at C-2 of *R*-glyceraldehyde 3-phosphate; Fig. 1). From the crystal structures of both the chicken and the yeast isomerases with either the substrate or a substrate analogue at the active site, it is clear that the oxygens of the carboxylate of Glu 165 are beautifully positioned for the proton abstractions required (Fig. 3*c*)[11–13,22,23]. This arrangement is exquisite for four reasons. First, the fit is snug (Fig. 3*c*), the distance from the nearer carboxylate oxygen to the substrate's C-1 or C-2 being only 2.8 or 3.4 Å. Second, the carboxylate group is bidentate, which means that the proton can be removed either from C-1 of dihydroxyacetone phosphate (from left to right in Fig. 1) or from C-2 of glyceraldehyde 3-phosphate (from right to left in Fig. 1), with minimal motion of the catalytic base. Third, the stereoelectronic requirement for rapid enolization[24] (that the abstracted proton lie orthogonal to the enolate plane defined by O-1, C-1, C-2 and O-2; see Fig. 4*a*) is nicely satisfied by the active site arrangement of triosephosphate isomerase. Fourth, the argument of Gandour[25] that a *syn* orbital of a carboxylate ion is much more basic than an *anti* orbital (Fig. 4*b*), has also been recognized in the configuration shown in Fig. 3*c*. So, with all the hindsight and all the physical-organic prejudice available, the choice and position of Glu 165 seem ideal. Indeed, when this group is moved by only ~1 Å (as in the mutant enzyme where Glu 165 is changed to Asp, for which the crystal structure shows no other significant alterations; E. Lolis and G. Petsko, unpublished results), the catalytic power of the enzyme falls by nearly a thousandfold[26]. Although it is too early to generalize, it is evident that in this case at least, the positioning of functionality at the active site of the enzyme needs to be quite precise if full catalytic potency is to be realized. Finally, as has been shown for a number of enzymes, complete removal of the catalytic group and its presumed replacement by water (as when Glu 165 is substituted by alanine or glycine) cuts the catalytic activity by more than a million times (K. Liu, unpublished work). The good news for catalyst engineers is that proper placement of appropriate groups in the right environment seems to be enough. The not-so-good news is that this placement must be very precise. (In this regard, it has been argued by several

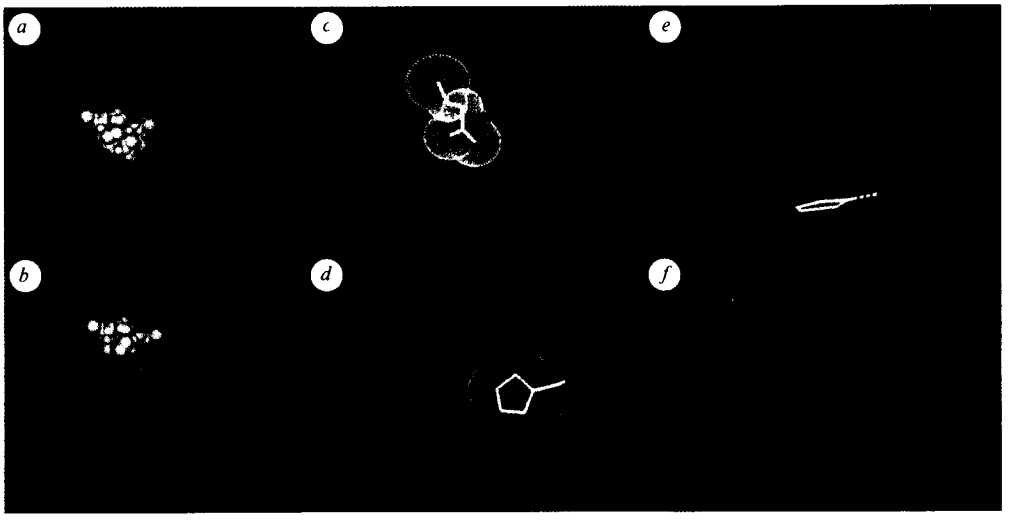

FIG. 3 Catalytic elements of triosephosphate isomerase. *a*, The substrate (orange) is sequestered from solvent by the complete mobile loop (turquoise) (the coordinates are from the crystal structure of the yeast enzyme containing bound phosphoglycolohydroxamate[22,23]); *b*, the substrate (orange) in the same position as in *a*, cannot be sequestered by a loop (turquoise) from which four amino acids (residues 170–173) have been removed (structure from *a*, after emendation and energy minimization); *c*, shows the snug fit of substrate (red) and the catalytic base of Glu 165 (yellow) (the dot surfaces are at the van der Waals' radii); *d*, shows the snug fit of substrate (red) and the catalytic electrophile of His 95 (green) (the dot surfaces are at the van der Waals' radii); *e*, showing the α helix (blue) that is aimed at the imidazole ring of His 95 (green); and *f*, shows the α helix (blue) that is aimed at the phosphate group of the substrate analogue phosphoglycolohydroxamate (red).

authors that the catalytic potential of the carboxylate group has rarely, if ever, been realized in model systems, because of the failure to exploit the *syn* orbitals of this group [27,28].)

An unexpected electrophile

Early experiments indicated that triosephosphate isomerase also uses electrophilic catalysis to speed the enolization steps, for the carbonyl group of the substrate is considerably polarized on binding to the enzyme[29,30]. A likely candidate for this electrophilic role is histidine 95, the Nε of which lies within 2.9 Å of each of the substrate oxygens on C-1 and C-2 (Fig. 3*d*). As with the catalytic base (the carboxyl group of Glu 165), the position of the imidazole ring of His 95 with respect to the substrate seems particularly appropriate, and provides for a strong hydrogen bond to the carbonyl oxygen of either of the substrates dihydroxyacetone phosphate or glyceraldehyde phosphate. There is even the seductive possibility (originally formulated by Jules Rebek) that the proton transfers are mediated merely by rotation of the appropriate side chain: the bidentate carboxyl group for the manipulation of carbon-bound protons, and the bidentate imidazole group for handling the oxygen-bound protons. Mutagenesis experiments have confirmed the importance of His 95 (conversion of this residue to glutamine gives an enzyme that is 100 times less active; conversion to asparagine cuts the activity by 10^4; refs 31–32), and the electrophilic and proton transfer roles of the imidazole ring have been further illuminated by infrared and X-ray crystallographic work on these mutant enzymes[33]. Indeed, when glutamine replaces His 95, the enzyme—deprived of its catalytic electrophile—finds a subtly different mechanistic pathway to achieve the interconversion of the two substrates at 1% of the wild-type rate[31]. This mutant apparently uses Glu 165 opportunistically to effect all the proton transfers, namely those to and from both carbon and oxygen centres.

On this basis, the reader would be forgiven for concluding that Fig. 3*c* and *d* adequately describes a precisely positioned general base (the carboxylate of Glu 165) and a well-placed general acid (the imidazolium of His 95), which are poised for 'push–pull' catalysis. Yet this is not how things are, and our understanding is not complete. The crystal structure of the isomerase indicates from the pattern of hydrogen bonds at the active site[11–13] that the imidazole ring of His 95 is not protonated at neutral pH; this implies that the enzyme uses imidazole rather than the much more electrophilic imidazolium as its acid component. We have investigated this using ^{15}N NMR, which is particularly suitable for such purposes[34]. First, all histidines in the yeast enzyme except that at position 95 were replaced with glutamine, and the resulting (fully active) enzyme was isolated, selectively enriched with ^{15}N-histidine. Using ^{15}N NMR, we found that the pK_a of His 95 is lowered from 6.7 in the denatured protein to less than 4.5 in the native enzyme, and we observed the formation of a new hydrogen bond (presumably to substrate) from the neutral imidazole ring (P. Lodi, unpublished), indicating that the enzyme does use the less acidic neutral imidazole as its electrophile. Why, when everything else about this enzyme seems chemically so reasonable, it should fail to use the more powerful electrophile, remains a puzzle. Yet, so as not to end in a minor key, let us look at one other feature of the active site.

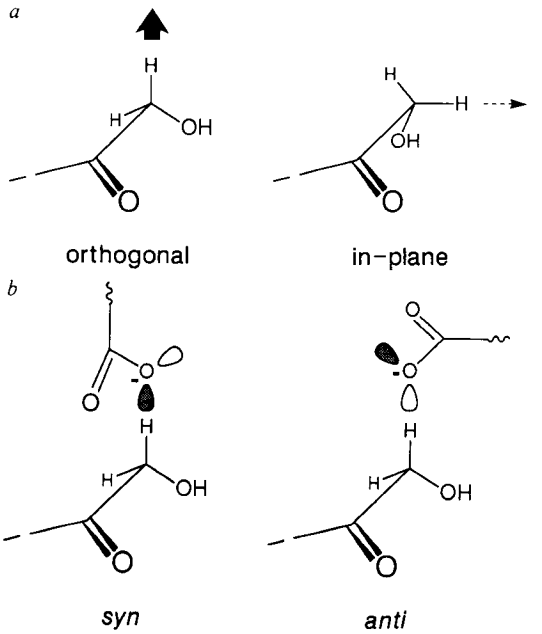

FIG. 4 Optimum geometry for enolization by a carboxylate base. *a*, The proton abstracted should lie orthogonal to the plane of the enolate[24]; and *b*, a carboxylate *syn* orbital (shaded) is believed to be more basic than an *anti* orbital (open)[25].

Well-aimed helices

Hol has published a splendid review[35] in which he develops the idea that the formal dipole associated with a classical α helix modulates the properties of groups at the helix termini. Whether the observed effects are due to the electric field of a helix dipole or merely to the local fields of the unsatisfied carbonyl groups and —NH— groups at each end of the helix[36], is less important than the effectiveness of the device, which is well established[37,38]. Triosephosphate isomerase has two conserved α helices that are directed towards the active site. One helix is trained on His 95, the imidazole ring of which is at the positive end (Fig. 3*e*). The aim is precise and, even if we cannot understand why the enzyme should want to lower the pK_a of His 95 to less than 4.5 and so ensure a neutral imidazole, the arrangement of this stretch of helix accounts for the perturbation very well.

The second helix that points into the active site has its positive

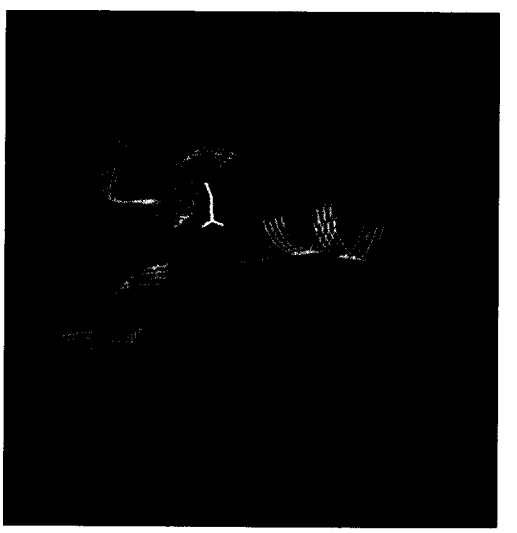

FIG. 5 One subunit of triosephosphate isomerase, with bound substrate (red), Glu 165 (yellow), His 95 (green), the two α helices that are aimed at the active site (pink), and the mobile loop that closes over the bound substrate (light blue).

end aimed at the phosphate group of the substrates (Fig. 3*f*). This arrangement too, is attractive, because there is only one positively charged group at the active site (Lys 12), and yet the substrates bind as their phosphate di-anions[39]. The arrangement shown in Fig. 3*f* provides two hydrogen bonds from main chain —NH— groups (of Gly 232 and Gly 233) to complete the neutralization of charge on the phosphate group.

If a single message emerges after all of the mechanistic and structural scrutiny of triosephosphate isomerase, it is one of precision[40], most evident from Fig. 3*c–f*. That the fit of substrate (or, more strictly, transition state[41–43]) to active site is exact has been clear for nearly a century[44], but there are other elements of precision that perhaps we have not fully appreciated. Indeed, one may speculate that the relatively large size of enzymes (compared with most of their substrates) may derive from the need for a matrix that positions their functional groups, focuses their helices and anchors the ends of their mobile loops (Fig. 5). And if our bioorganic effort and chemical mimicry continue to fall short of the catalytic potency of today's enzymes, at least we know that nature has been adjusting and refining for rather longer than we have[45]. □

Jeremy R. Knowles is in the Departments of Chemistry and Biochemistry, Harvard University, Cambridge, Massachusetts 02138, USA.

1. Rieder, S. V. & Rose, I. A. *J. biol. Chem.* **234,** 1007–1010 (1959).
2. Rose, I. A. *Brookhaven Symp. Biol.* **15,** 293–309 (1962).
3. Maister, S. G., Pett, C. P., Albery, W. J. & Knowles, J. R. *Biochemistry* **15,** 5607–5612 (1976).
4. Richard, J. P. *J. Am. chem. Soc.* **106,** 4926–4936 (1984).
5. Adams, M. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **70,** 1968–1972 (1973).
6. Clarke, A. R. *et al. Nature* **324,** 699–702 (1986).
7. Ludwig, M. L. *et al. Proc. natn. Acad. Sci. U.S.A.* **57,** 511–514 (1967).
8. Bennett, W. S. Jr & Steitz, T. A. *Proc. natn. Acad. Sci. U.S.A.* **75,** 4848–4852 (1978).
9. Pickover, C. A., McKay, D. B., Engelman, D. M. & Steitz, T. A. *J. biol. Chem* **254,** 11323–11329 (1979).
10. Winn, S. I., Watson, H. C., Harkins, R. N. & Fothergill, L. A. *Phil. Trans. R. Soc.* **B293,** 121–130 (1981).
11. Banner, D. W. *et al. Nature* **255,** 609–614 (1976).
12. Alber, T. C. *et al. Phil. Trans. R. Soc.* **B293,** 159–171 (1981).
13. Lolis, E. & Petsko, G. A. *Biochemistry* **29,** 6619–6625 (1990).
14. Joseph, D., Petsko, G. A. & Karplus, M. *Science* **249,** 1425–1428 (1990).
15. Pompliano, D. L., Peyman, A & Knowles, J. R. *Biochemistry* **29,** 3186–3194 (1990).
16. Möller, W. & Amons, R. *FEBS Lett.* **186,** 1–7 (1985).
17. Wierenga, R. K., Terpstra, P & Hol, G. J. *J. molec. Biol.* **187,** 101–107 (1986).
18. Seeburg, P. H., Colby, W. W., Capon, D. H., Goeddel, D. V. & Levinson, A. D. *Nature* **312,** 71–75 (1984).
19. Reinstein, J., Brune, M. & Wittinghofer, A. *Biochemistry* **27,** 4712–4720 (1988).
20. Redfield, A. G. & Papastavros, M. Z. *Biochemistry* **29,** 3509–3514 (1990).
21. Saraste, M., Sibbald, P. R. & Wittinghofer, A. *Trends biol. Sci.* **15,** 430–434 (1990).
22. Davenport, R. C. thesis, MIT, Cambridge, Massachusetts (1986).
23. Alber, T. C. *et al. Cold Spring Harbor Symp. quant. Biol.* **52,** 603–613 (1987).
24. Corey, E. J. & Sneen, R. A. *J. Am. chem. Soc.* **78,** 6269–6278 (1956).
25. Gandour, R. D. *Bioorg. Chem.* **10,** 169–176 (1987).
26. Raines, R. T., Sutton, E. L., Straus, D. R., Gilbert, W. & Knowles, J. R. *Biochemistry* **25,** 7142–7154 (1986).
27. Zimmerman, S. C. & Cramer, K. D. *J. Am. chem. Soc.* **110,** 5906–5908 (1988).
28. Huff, J. B., Askew, B., Duff, R. J. & Rebek, J. *J. Am. chem. Soc.* **110,** 5908–5909 (1988).
29. Webb, M. R. & Knowles, J. R. *Biochem. J.* **141,** 589–592 (1974).
30. Belasco, J. G. & Knowles, J. R. *Biochemistry* **19,** 472–477 (1980).
31. Nickbarg, E. B., Davenport, R. C., Petsko, G. A. & Knowles, J. R. *Biochemistry* **27,** 5948–5960 (1988).
32. Blacklow, S. C. & Knowles, J. R. *Biochemistry* **29,** 4099–4108 (1990).
33. Komives, E. A. *et al. Biochemistry* (in the press).
34. Bachovchin, W. W. *Biochemistry* **25,** 7751–7759 (1986).
35. Hol, W. G. J. *Prog. biophys. molec. Biol.* **45,** 149–195 (1985).
36. Tidor, B. & Karplus, M. *Biochemistry* (in the press).
37. Perutz, M. F., Gronenborn, A. M., Clore, G. M. Fogg, J. H. & Shih, D. T.-b. *J. molec. Biol.* **183,** 491–498 (1985).
38. Sali, D., Bycroft, M. & Fersht, A. R. *Nature* **335,** 740–743 (1988).
39. Belasco, J. G., Herlihy, J. M. & Knowles, J. R. *Biochemistry* **17,** 2971–2978 (1978).
40. Blow, D. M. & Smith, J. M. *Phil. Trans. R. Soc.* **B272,** 87–97 (1975).
41. Haldane, J. B. S. in *Enzymes* 182 (Longmans, London, 1930).
42. Pauling, L. *Chem. Eng. News* **24,** 1375–1377 (1946).
43. Leatherbarrow, R. J., Fersht, A. R. & Winter, G. *Proc. natn. Acad. Sci. U.S.A.* **82,** 7840–7844 (1985).
44. Fischer, E. *Ber. dtsch chem. Ges.* **27,** 2985–2993 (1894).
45. Albery, W. J. & Knowles, J. R. *Biochemistry* **15,** 5627–5631 (1976).

ACKNOWLEDGEMENTS. I thank, among many collaborators on the isomerase problem, R. T. Raines, A. Peyman, D. L. Pompliano, E. A. Komives, K. Liu, P. Lodi, E. Lolis and G. Petsko, upon whose contributions much of this account is based. Computer graphics were by P. Lodi and D. Pompliano. Supported by the General Medical Sciences Institute of the NIH and by the NSF.

Geodetic determination of tectonic deformation in central Greece from 1900 to 1988

H. Billiris*, D. Paradissis*, G. Veis*, P. England†, W. Featherstone†, B. Parsons†, P. Cross‡, P. Rands‡, M. Rayson‡, P. Sellers‡, V. Ashkenazi§, M. Davison§, J. Jackson‖ & N. Ambraseys¶

* Department of Surveying, National Technical University, Athens 15773, Greece
† Department of Earth Sciences, University of Oxford, Parks Road, Oxford OX1 3PR, UK
‡ Department of Surveying, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, UK
§ Institute of Engineering Surveying and Space Geodesy, University of Nottingham, University Park, Nottingham NG7 2RD, UK
‖ Department of Earth Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EQ, UK
¶ Department of Civil Engineering, Imperial College, Imperial College Road, London SW7 2BU, UK

The Global Positioning System has been used to measure the relative displacements of fifteen monuments in a hundred-year-old triangulation network spanning part of the Aegean extensional basin. These displacements reflect the tectonic deformation of the region over the past century, showing more than one metre of north–south extension across the network. The crust in this region appears to contain a few slowly deforming blocks separated by more rapidly deforming zones.

AT the same time that plate tectonics was being used successfully to describe a wide variety of geological observations, it was realized that the axioms of plate tectonics do not apply to parts of the continents (see, for example, ref. 1). Deformation is distributed over wide areas on the continents, rather than being concentrated on narrow boundaries as in the oceans. A key aim of tectonics is to determine the most appropriate and economical description of this distributed deformation; without such a description, it is difficult to make progress on studies of the dynamics of the deformation (for example, ref. 2). Most of the quantitative data on active continental deformation come from earthquake seismology and satellite imagery[3–6], but this approach is limited. Earthquake focal mechanisms cannot pro-