RESEARCH COMMUNICATION

Large changes of transition-state structure during experimental evolution of an enzyme

Srinivasan K,* Alexandros KONSTANTINIDIS,* Michael L. SINNOTT*‡ and Barry G. HALL†

*Department of Chemistry (M/C 111), University of Illinois at Chicago, Chicago, IL 60607-7061, U.S.A. and †Department of Biology, University of Rochester, Rochester, NY 14627, U.S.A.

The question of whether, during the evolution of an enzyme, the transition state of the catalysed reaction is largely unchanged, or whether transition state and protein change together, was examined using the *egb* β-galactosidases of *Escherichia coli*. Charge development at the first chemical state was assumed [Konstantinidis and Sinnott (1991) Biochem. J. **279**, 587–593] to be proportional to $\Delta\Delta G^{\ddagger}$, the ratio of second-order rate constants for the hydrolysis of β-D-galactopyranosyl fluoride and 1-fluoro-D-galactopyranosyl fluoride, expressed as a free-energy difference. $\Delta\Delta G^{\ddagger}$ (kJ · mol^{-1}) falls from 10.4 for wild-type enzyme to 6.8 and 7.2 as a consequence of two different single amino-acid changes (which arise from single evolutionary events), to 6.3 as a consequence of the two amino-acid changes together, and then increases slightly to 7.3 as a consequence of a third single evolutionary change involving three further amino-acid changes.

INTRODUCTION

Application of transition-state theory to enzymic reactions yields the fruitful idea that enzymes work by stabilizing transition states (e.g. [1]). Enzymes are the result of natural selection, and there is an extensive literature on the changes in the energies of bound intermediates and transition states to be expected as a consequence of selection for increased catalytic competence [2–6]. To the best of our knowledge, however, the question of whether transition-state structure is essentially invariant during evolution, or whether transition states are sufficiently plastic as to be altered as the protein in which they are embedded undergoes evolutionary changes, has never been addressed. We now present results showing that single evolutionary events can cause large changes in the amount of positive charge on the glycone at the first transition state of reactions catalysed by the *ebg* β-galactosidase of *Escherichia coli*.

Experimental selection pressure can be unequivocally directed towards the catalytic competence of this enzyme [7]. The wild-type enzyme, *ebg*^{o}, is too catalytically feeble to permit growth on lactose when it is the only β-galactosidase in the bacterium, but selection on lactose as sole carbon source results, 90% of the time, in a (mutant) Class I enzyme, and 10% of the time in a (mutant) Class II enzyme, both of which permit growth on lactose. (Class II enzymes can also be produced by selection on lactulose.) A second round of selections, of Class I enzymes on lactulose or Class II on galactosylarabinose, results in Class IV enzymes [8].

Each round of selections can result in simultaneous spontaneous changes in the repressor gene, *ebgR*, and the two structural genes for the enzyme, *ebgA* and *egbC* (the active form of the enzyme being an $\alpha_4\beta_4$ heterooctamer [9]). The system is one in which directed evolution takes place [10]. The kinetic characteristics of the Class I, Class II, and Class IV enzymes are, however, determined, at least to the limit of resolution of our present kinetic probes, by only two amino-acid changes in the *ebgA* peptide [11], Class I enzymes containing an Asp^{92}→Asn change, Class II enzymes a Trp^{977}→Cys change, and Class IV enzymes both [12]. In accordance with previous usage, we denote the Asp^{92}→Asn change by the superscript *a*, and the Trp^{977}→Cys change by the superscript *b*. The simplest Class I enzyme is thus *ebg*^{a}, the simplest Class II enzyme is *ebg*^{b}, and the simplest Class IV enzyme is thus *ebg*^{ab}. However, another type of Class IV enzyme possesses, in addition to the Asp^{92}→Asn and Trp^{977}→Cys changes in the large subunit, two additional changes, Ser^{979}→Gly in the large subunit and Glu^{122}→Gly in the small subunit. We designate the large subunit Ser^{979}→Gly change by a superscript *c* and the small subunit Glu^{122}→Gly change by superscript *d*. Thus, this second type of Class IV enzyme is *ebg*^{abcd}. Selection of Class IV-containing *ebg* strains on lactobionate has resulted in one Class V-containing strain, in which there are now four amino-acid changes in the large subunit (Asp^{92}→Asn, Glu^{93}→Lys, Trp^{977}→Cys, and Ser^{979}→Gly) and one in the small subunit (Glu^{122}→Gly) [12]. As we indicate the Glu^{93}→Lys change by superscript *e*, this Class V enzyme is *ebg*^{abcde}.

We have applied a new high-resolution mechanistic probe to reactions catalysed by the wild-type enzyme, *ebg*^{o}, *ebg*^{a}, *ebg*^{b}, *ebg*^{ab} and *ebg*^{abcde}. The mechanistic probe is the ratio of second-order rate constants for the hydrolysis of β-D-galactopyranosyl fluoride (**I**) and 1-fluoro-D-galactopyranosyl fluoride (**II**). The initially non-departing (α) fluorine of the difluoride exerts a strong cation-destabilizing inductive effect, as demonstrated by the 4×10^4-fold slower rate of spontaneous hydrolysis of 1-fluoro-D-glucopyranosyl fluoride than β-D-glucopyranosyl fluoride [13]. The anomeric hydrogen of a glycosidase substrate necessarily executes substantial motions during the reaction, and is unlikely to make critical contacts with the enzyme protein. The small difference in size of the C–H and C–F groups {0.045 nm (0.45 Å) [14]: the previous [13] estimate of 0.03 nm (0.3 Å) for this difference is based on an older data set [15]} means therefore that the effect of substitution of the anomeric hydrogen of the substrate with a fluorine atom is overwhelmingly electronic. The ratio of rate constants for hydrolysis of monofluoride and difluoride, expressed as differences in free energies of activation ($\Delta\Delta G^{\ddagger}$ values), reflect the change in charge between the unbound substrates and the first irreversible transition state. If the usual

‡ To whom correspondence should be addressed.

Brønsted-type assumptions are made, then $\Delta\Delta G^{\ddagger}$ values can be taken to be proportional to the charge development at the transition state: $\Delta\Delta G^{\ddagger}$ corresponds to a Brønsted slope without normalization by a reference reaction (see e.g. [16]). The reference reaction could be the spontaneous hydrolysis, for which $\Delta\Delta G^{\ddagger}$ is estimated to be 28.3 kJ·mol^{-1} from results in the *gluco* series. Spontaneous hydrolysis is a reaction in which close to, but somewhat less than, a full positive charge will be developed at the transition state [13].

EXPERIMENTAL

ebg enzymes were isolated as described previously [9,11]. The sources of all enzymes were also as described [11], except for *ebg*^{abcde}, which was isolated from *Salmonella typhimurium* LT2 containing the pUF 17 plasmid bearing genes for the Class V *E. coli ebg* enzyme [12].

Liberation of fluoride ions was followed electrochemically as described earlier [13], and $k_{cat.}$ values were calculated on the assumption, broadly justified by pre-steady-state measurements [9,17], that the most active isolated protein was completely pure and all four active sites were fully active. Kinetic parameters refer to 0.125 M potassium phosphate buffer, pH 7.5, containing 5 mM $MgCl_2$ and 25 μM 2,2′-dipyridyl.

lacZ β-Galactosidase was purchased from Cambrian Chemicals, lot no. 46J224: $k_{cat.}$ values, which refer to 0.05 M sodium phosphate buffer, pH 7.0, containing 1.0 mM $MgCl_2$, were calculated on the assumption that $k_{cat.}$ for *p*-nitrophenyl galactoside was 156 s^{-1} [18].

Kinetic parameters were estimated by fitting ten initial rates of fluoride liberation directly to the Michaelis–Menten equation using the non-linear least-squares program 'Kaleidagraph' (Synergy Software, Reading, PA, U.S.A.). Standard deviations of parameters ($V_{max.}$, K_m or $V_{max.}/K_m$) were always less than 10%.

β-D-Galactopyranosyl fluoride was made by literature procedures [19].

1-Fluoro-D-galactopyranosyl fluoride was synthesized analogously to the *gluco* compound [13,20]. Penta-*O*-acetyl-β-D-galactopyranose (1.85 g) was dissolved in dry chloroform and anhydrous $AlCl_3$ (0.35 g) was added. The solution was stirred for 20 min at room temperature and then poured into hexane (700 ml). The aluminium salts precipitated out and were filtered off, and the hexane solution was evaporated. The crude solid tetra-*O*-acetyl-β-D-galactopyranosyl chloride (0.8 g) was dissolved in CCl_4 (50 ml), *N*-bromosuccinimide (1.6 g) was added and the solution was heated under reflux for 20 min, with irradiation by a household 150 W tungsten-filament light-bulb. The solvent was then evaporated, and the residue was dissolved in ether, which was extracted with water to remove succinimide. Crude tetra-*O*-acetyl-1-bromo-D-galactopyranosyl chloride (0.4 g) was obtained as a syrup after drying with Na_2SO_4 and evaporation, and was converted into the tetra-*O*-acetyl-1-fluoro-D-galactopyranosyl fluoride by stirring with anhydrous AgFl (0.9 g) in freshly dried acetonitrile (25 ml) for 3 days. The precipitated silver salts were filtered off (Celite), the acetonitrile was evaporated, and the tetra-*O*-acetyl-1-fluoro-D-galactosyl fluoride was isolated by chromatography on silica gel with ethyl acetate/light petroleum (3:7, v/v) as eluent. Twofold recrystallization from ether/light petroleum (b.p. 40–60 °C) gave the pure material, m.p. 59.6 °C (0.2 g, 11% from pentacetyl glucose). Zemplén deacetylation as for the *gluco* compound [13] gave 1-fluoro-D-galactopyranosyl fluoride (92%) m.p. 149 °C (decomp.), $[\alpha]_D^{25}$ +92° (*c* 0.17, in water):(Found: C, 36.05; H, 5.12; $C_6H_{10}O_5F_2$ requires C, 36.0; H, 5.35%). ^{19}F-n.m.r. (H_2O): δ (p.p.m., with respect to trifluoroacetic acid) $F_{ax.}$ −13.7, $F_{eq.}$ −4.6, $J_{F_{ax.}F_{eq.}}$ 149 Hz, $J_{F_{ax.}H_2}$ 13.6 Hz.

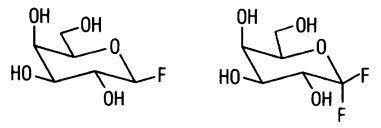
RESULTS AND DISCUSSION

In Table 1 are contrasted Michaelis–Menten parameters for the hydrolyses of β-D-galactopyranosyl fluoride (I) and 1-fluoro-D-galactopyranosyl fluoride (II) by *lacZ* β-galactosidase, *ebg*^{o}, *ebg*^{a}, *ebg*^{b}, *ebg*^{ab}, and *ebg*^{abcde}. The *ebg* enzymes are simple double-displacement glycosidases [21], in which one or other of the two chemical steps, depending on the leaving group ability of the aglycone, determines $k_{cat.}$ [9,17,21,22]. For *lacZ* β-galactosidase, *egb*^{a}, *ebg*^{b} and *egb*^{ab} it is clear from the similarity of the $k_{cat.}$ value for (I) to the rate of hydrolysis of the galactosyl-enzyme intermediate [9,17,22,23], which is rate-determining for substrates with far worse leaving groups than fluoride [9,18,22], that the second chemical step determines $k_{cat.}$ The similarity of $k_{cat.}$ for hydrolysis of (I) by *ebg*^{o} to that for 3,4-dinitrophenyl galactoside (186 s^{-1} [21]), the dependence of $k_{cat.}$ for aryl galactosides on leaving group ability [21], and the much lower intrinsic nucleofugacity of 3,4-dinitrophenolate compared with fluoride [24], suggest that for *egb*^{o}, $k_{cat.}$ for (I) represents hydrolysis of the glycosyl enzyme. In the light of the behaviour of other *ebg* enzymes, and its low absolute value, it is likely that $k_{cat.}$ for *ebg*^{abcde}-catalysed hydrolysis of (I) represents hydrolysis of the galactosyl-enzyme intermediate. Since as yet, though, we do not know whether $k_{cat.}$ values for (II) represent formation or hydrolysis of the fluorogalactosyl enzymes, comparison of monofluoride and difluoride must be confined to $k_{cat.}/K_m$, which reflects only the first chemical step and binding.

There are significant changes in $\Delta\Delta G^{\ddagger}$ values calculated from these second-order rate constants as a consequence of evolutionary changes in the enzyme, though the changes are not additive: either the Asp^{92}→Asn or the Trp^{977}→Cys change causes a comparable decrease in the charge at the first transition state, which is only slightly decreased by the additional introduction of the other amino-acid change. On the assumption of linearity between $\Delta\Delta G^{\ddagger}$ and transition-state charge development, there is an attenuation of transition-state charge by 35% as a consequence of the *egb*^{o}→*ebg*^{a} change, 31% as a consequence of the

Table 1 Hydrolysis of β-D-galactopyranosyl fluoride (I) and 1-fluoro-D-galactopyranosyl fluoride (II) by *E. coli* β-galactosidases at 25.0 °C

Parameters are reproducible (with different batches of substrate) to ± 10%. The *ebg* enzyme data refer to results obtained in 0.125 M potassium phosphate buffer, pH 7.5, containing 5 mM $MgCl_2$ and 25 μM 2,2′ bipyridyl, the *lacZ* β-galactosidase data to 0.05 M sodium phosphate buffer, pH 7.0, containing 1.0 mM $MgCl_2$.



I II

Enzyme	$k_{cat.}$ (s^{-1})	K_m (mM)	$k_{cat.}$ (s^{-1})	K_m (mM)	$\Delta\Delta G^{\ddagger}$ (kJ·mol^{-1})
lacZ	1185	1.82	1.74	1.26	15.2
ebg^{o}	172	2.18	0.80	0.67	10.4
ebg^{a}	22	0.45	0.105	0.033	6.8
ebg^{b}	10	0.54	0.22	0.22	7.2
ebg^{ab}	3.3	0.21	0.138	0.112	6.3
ebg^{abcde}	2.7	0.18	0.262	0.335	7.3

$ebg^{o} \rightarrow ebg^{b}$ change, but only 39% as a consequence of both changes. The third evolutionary change (ebg^{ab} to ebg^{abcde}) results in a slight increase of effective charge.

The decreases in charge may arise from a closer approach of the nucleophilic Glu^{382} residue to the anomeric centre as a consequence of the $Asp^{92} \rightarrow Asn$ and $Trp^{977} \rightarrow Cys$ changes. Site-directed mutagenesis of the nucleophilic Glu^{358} of *Agrobacterium* β-glucosidase to the shorter aspartic acid residue produces an increase in oxocarbonium ion character of the first transition state that can be detected by α-deuterium kinetic isotope effects [25].

There is 34% amino-acid identity between the *ebgA* and *lacZ* gene products [12]. The $\Delta\Delta G^{\ddagger}$ value for the *lacZ* enzyme, though, is greater than for any of the *ebg* enzymes, indicating that it has a larger transition-state charge development. This is in accord with higher α-deuterium kinetic isotope effects, both for the formation of galactosyl-enzyme intermediates from galactosyl pyridinium salts [21,26] and for the hydrolysis of the galactosyl-enzyme intermediates [9,22,23] for the *lacZ* enzyme than for *ebg* enzymes. The estimate of charge development from $\Delta\Delta G^{\ddagger}$ for the *lacZ* enzyme is, moreover, probably an underestimate, since non-covalent events may govern $k_{cat.}/K_m$ for the monofluoride as they do for other good, uncharged substrates [18,27].

The first two evolutionary changes in *ebg* protein result in a lowering of $\Delta\Delta G^{\ddagger}$. Therefore, at least for these changes, movement in evolutionary space, although towards *lacZ* in terms of efficiency of lactose hydrolysis, is away from it in terms of transition-state structure. These changes appear to be exploring an evolutionary dead end, though, as the third evolutionary change appears to result in a slight increase of transition-state charge development, indicating a small movement in evolutionary space towards *lacZ*.

The plasticity of a transition state for glycosyl transfer is, with hindsight, reasonable in the light of its commonly preassociative character: for such reactions the free-energy surface in the region of the transition state can be envisaged as a plateau [28], such that comparatively small changes in, for instance, nucleophile positioning can cause significant changes in the position of the transition state [29]. The small effect from the second mutation may be because one mutation has already brought the system to the edge of the plateau. Indeed, a third evolutionary change appears to take the transition state slightly away from the edge of this plateau.

Financial support from Grants GM 46663 and GM 42469 of the National Institutes of Health of the USPHS is gratefully acknowledged.

REFERENCES

1 Jencks, W. P. (1976) Adv. Enzymol. **43**, 219–410
2 Albery, W. J. and Knowles, J. R. (1976) Biochemistry **15**, 5631–5640
3 Burbaum, J. J., Raines, R. T., Albery, W. J. and Knowles, J. R. (1989) Biochemistry **28**, 9283–9305
4 Chin, J. (1983) J. Am. Chem. Soc. **105**, 6502–6503
5 Pettersson, G. (1989) Eur. J. Biochem. **184**, 561–566
6 Pettersson, G. (1991) Eur. J. Biochem. **195**, 663–670
7 Hall, B. G. (1976) J. Mol. Biol. **107**, 71–84
8 Hall, B. G. (1982) Evol. Biol. **15**, 85–150
9 Elliott, A. C., K, S., Sinnott, M. L., Smith, P. J., Bommuswamy, J., Guo, Z., Hall, B. G. and Zhang, Y. (1992) Biochem. J. **282**, 155–164
10 Hall, B. G. (1990) Bioessays **12**, 551–558
11 Hall, B. G. (1981) Biochemistry **20**, 4042–4049
12 Hall, B. G., Betts, P. W. and Wotton, J. C. (1989) Genetics **123**, 635–648
13 Konstantinidis, A. and Sinnott, M. L. (1991) Biochem. J **279**, 587–593
14 Withers, S. G., Street, I. P. and Percival, M. D. (1988) Am. Chem. Soc. Symp. Ser. **374**, 59–77
15 Trotter, J. P. (1973) in The Chemistry of the Carbon–Halogen Bond (Patai, S., ed.), pp. 49–62, John Wiley and Sons, New York, Toronto and Chichester
16 Williams, A. (1984) Acc. Chem. Res. **17**, 425–430
17 Li, B. F. L., Osborne, S., Sinnott, M. L. and Yates, D. W. (1983) J. Chem. Soc. Perkin Trans. **2**, 365–369
18 Sinnott, M. L. and Souchard, I. J. L. (1973) Biochem. J. **133**, 89–98
19 Micheel, F. and Klemer, A. (1961) Adv. Carbohydr. Chem. **16**, 85–97
20 Praly, J.-P. and Descotes, J.-P. (1987) Tetrahedron Lett. **28**, 1405–1408
21 Burton, J. and Sinnott, M. L. (1983) J. Chem. Soc. Perkin Trans. **2**, 359–364
22 Li, B. F. L., Holdup, D., Morton, C. A. J. and Sinnott, M. L. (1989) Biochem. J. **260**, 109–114
23 Sinnott, M. L. and Viratelle, O. M. (1973) Biochem. J. **133**, 81–87
24 Jones, C. C. and Sinnott, M. L. (1977) J. Chem. Soc. Chem. Commun. 767–768
25 Withers, S. G., Rupitz, K., Trimbur, D. and Warren, R. A. J. (1992) Biochemistry **31**, 9979–9985
26 Sinnott, M. L. and Withers, S. G. (1974) Biochem. J. **143**, 751–762
27 Selwood, T. and Sinnott, M. L. (1990) Biochem. J. **268**, 317–323
28 Sinnott, M. L. and Jencks, W. P. (1980) J. Am. Chem. Soc. **102**, 2026–2032
29 Jencks, W. P. (1985) Chem. Rev. **85**, 511–527

Received 8 December 1992/15 January 1993; accepted 18 January 1993