A novel approach to study of the structural basis of enzyme polymorphism

Analysis of carboxylesterase B of *Escherichia coli* as model

Bertrand PICARD,* Philippe GOULLET* and Rajagopal KRISHNAMOORTHY†‡

*Laboratoire de Microbiologie, Faculté de Médecine Xavier Bichat (Université Paris VII), 16 Rue H. Huchard, 75018 Paris Cedex 14, and †Institut de Pathologie et de Biologie Cellulaires et Moléculaires, Unité 15 de l'Institut National de la Santé et de la Recherche Médicale, 24 Rue du Faubourg Saint Jacques, 75674 Paris Cedex 14, France

In order to understand the structural basis of charge differences among enzyme variants without undertaking purification and sequencing of the protein, an original approach was developed. The approach is applicable to any enzyme or protein provided that there is a specific staining procedure. This consists, as a first step, in the projection of electrophoretically obtained mobility values versus pI of all variants into a two-dimensional profile. In a second step, starting from the most common variant, various theoretical possibilities of substitutions are envisaged, taking into consideration the pH of the electrophoretic conditions, pI of the variants and range of variations of the pK values of several amino acid side chains. In a third step, verification of the theoretical data is obtained through comparative protein titration curves by combined isoelectrofocusing–electrophoresis of several pairs of relevant variants. The validity of this approach is tested on the highly polymorphic carboxylesterase B enzyme of *Escherichia coli* and is found to provide valuable information.

INTRODUCTION

Assessment of enzyme polymorphism by electrophoresis has been extensively employed to study the structure and evolution of populations and species (for review see Coyne, 1982). Since it has been recognized that the electrophoretic variants, called 'electromorphs', incompletely reflect the underlying allelic variation (Hubby & Lewontin, 1966; King & Ohta, 1975), attempts were made to improve the detection of cryptic enzyme variations (Bernstein *et al.*, 1973; Coyne, 1976; Singh *et al.*, 1976). From such studies various neutral models have been proposed, and one such model, the charge-state model, is considered to be appropriate for the available electrophoretic data (Ohta & Kimura, 1973). However, it is generally accepted that a considerable amount of structural data on the protein is required to enable the evolutionary mechanisms to be understood. Purification and sequencing of several protein variants from various species and populations require considerable effort in terms of time and material. In the present paper we propose a much simpler, novel, informative approach to understanding of the structural basis for differences among enzyme variants, taking carboxylesterase B from *Escherichia coli* (Goullet *et al.*, 1984) as a model system.

EXPERIMENTAL

Origin of strains

A total of 610 *E. coli* strains were examined for their esterase B pattern: 211 intestinal strains isolated from the faeces of various animal species, 142 human intestinal strains from healthy individuals and 257 human extra-intestinal invasive strains.

Preparation of bacterial extracts

Bacterial extracts were prepared essentially as described previously (Goullet, 1973). Briefly, bacteria were grown at 37 °C in broth with constant agitation for 18 h. After centrifugation the pellets were washed with 60 mM-Tris/glycine buffer, pH 8.7, resuspended in the same buffer and disrupted by intermittent exposure to ultrasonic oscillations (Branson Sonifier B 30 cell disrupter) for a total of 6 min at 4 °C, and the crude extract supernatants, containing 40–60 mg of protein/ml, were stored at –20 °C until use.

Conventional electrophoresis

Horizontal slab electrophoresis in a composite polyacrylamide/agarose gel (7% and 1.4% respectively) was performed according to the method described by Uriel (1966) in a discontinuous Tris/glycine buffer, pH 8.7, at constant value of 7 V/cm until the Bromophenol Blue marker had run 13 cm. Differences in relative mobilities of several strains were confirmed by multiple and comparative runs side by side on the same gel or through co-migration of a mixture of samples. The relative electrophoretic mobility (m_F value) is the running distance of the esterase band expressed as a percentage of that of the dye front, previously calibrated with several standard proteins.

Isoelectrofocusing

Thin-layer isoelectrofocusing in polyacrylamide gels (T = 5%, C = 3%) (Vesterberg & Svensson, 1966) was performed at 4 °C with carrier ampholytes of pH range 4–6.5 (Pharmalyte; Pharmacia, Uppsala, Sweden) in accordance with the manufacturer's instructions. Essentially, experiments were run at a constant power of 25 W

‡ To whom correspondence should be addressed.

for 2 h at 4 °C. Under these conditions the samples reach their equilibrium position (pI position) after 90 min and a further condensation of the zone after 30 min. The pH gradient was measured systematically for each gel after the elution of the carrier ampholytes from a zone where no sample was applied. The isoelectric points (pI) of esterases were determined by interpolation from this curve. After a large screening, several variants were run side by side in the same gel to confirm their pI differences.

Detection and identification of esterases

A discriminative esterase zymogram (Lawrence *et al.*, 1960; Uriel, 1961) with the synthetic substrates α -naphthyl acetate, β -naphthyl acetate, indoxyl acetate, α -naphthyl butyrate and β -naphthyl butyrate (Sigma Chemical Co., St. Louis, MO, U.S.A.) revealed that they all belong to esterase B group.

Furthermore, immunological identity, thermal stability and sensitivity to di-isopropyl phosphorofluoridate were examined as described previously (Goullet, 1973, 1980; Goullet *et al.*, 1984).

Titration curves by combined isoelectrofocusing-electrophoresis

The technique employed here has already been described (Righetti *et al.*, 1978). The essential features are that at the first step a pH gradient is formed extending from 3.5 to 10, the sample is applied in a groove running all along the gradient and electrophoresis is performed perpendicular to this gradient. Sigmoidal pH-mobility curves approximating to the titration curves of esterases are revealed by the specific staining procedures, with α -naphthyl acetate as substrate, as described previously (Uriel, 1961).

RESULTS

Among the 610 strains of *E. coli* analysed, 564 of them produced carboxylesterase B detectable both in conventional electrophoresis and in isoelectrofocusing.

Conventional electrophoresis

By conventional electrophoresis carboxylesterase B presents 12 allelic variants, whose mobility distribution extends from m_F 57 to m_F 75. The histogram in Fig. 1 reveals two modes of frequency distribution of these variants, designated groups B₁ and B₂. Group B₁ has a Gaussian distribution with maximal frequency at m_F 70 (292 strains) within the limits m_F 66 and 75, whereas group B₂, with a non-Gaussian distribution, trails off from m_F 57 to m_F 63. Group B₁ comprises 434 strains (77% of the total), and group B₂ 130 strains (23%). Both of these two subpopulations are categorized as carboxylesterase B by a combination of several criteria, namely substrate specificity, thermal stability, inhibition by di-isopropyl phosphorofluoridate, molecular size and immunospecificity (Goullet, 1973; Goullet *et al.*, 1984).

Isoelectrofocusing

On electrofocusing only ten allelic carboxylesterase B variants are detected, with isoelectric points varying from pH 4.5 to 5. More than 80% of the variants are confined to four different isoelectric points. The pattern of the frequency distribution revealed no distinct groups such as B₁ and B₂.

Fig. 1. Frequency distribution of carboxylesterase B variants from *E. coli*

A total of 610 isolates of *E. coli* were examined. The histogram plot of frequency of occurrence of variants against their m_F values allows the distinction of two groups, namely B₁ and B₂.

Two-dimensional electrophoretic profile

By plotting mobility values versus pI of all variants, a two-dimensional profile (Goullet & Picard, 1985) is obtained, as shown in Fig. 2(a). This representation allows discrimination of 41 electromorphs, of which 32 belong to group B₁ and nine to group B₂. Thus the effective number of variants increases at least 4-fold compared with when m_F and pI were considered alone. This is due to the fact that in most cases a variant defined by one parameter, say m_F , can have different values of pI, and vice versa. A three-dimensional projection of these data along with frequency values distinguishes not only groups B₁ and B₂ but also subdivides group B₁ into groups B_{1a} and B_{1b} (Fig. 2b).

Titration curves

Crude bacterial extracts were submitted to electrophoresis perpendicularly to a preformed stable carrier ampholyte pH gradient in order to obtain the mobility curves, which were revealed by specific staining procedures. Fig. 3(a) shows the comparative mobility curves obtained from the most frequent variant I (pI 4.6/ m_F 70) [see Fig. 3(a) legend and Fig. 2(a)] together with variant VI (pI 4.75/ m_F 70). The typical fork form corresponds to a substitution of a neutral amino acid residue in variant I by a histidine in variant VI (Righetti *et al.*, 1978). A qualitatively identical pattern was obtained by titrating both variants I and V (pI 4.5/ m_F 70) together (results not shown), which provided evidence for the opposite situation of substitution, namely a histidine residue of variant I replaced by a neutral amino acid in variant V. Fig. 3(b) represents the titration curves of variants I and VIII (pI 4.6/ m_F 72), and here the behaviour of the mobility curves is unequivocally explained by the substitution of an arginine residue of variant I by a histidine in variant VIII. A similar pattern was again observed for the variants I and VII (pI 4.6/ m_F 68) (results not shown). Titration curves of variants I and III (pI 4.75/ m_F 68) on the one hand and of variants I and IV (pI 4.9/ m_F 66) on the other revealed the substitution of an acidic amino acid residue by a neutral one and a basic one respectively (Righetti *et al.*, 1978) (results not shown).

Several other comparative titration curves were

Fig. 2. Two-dimensional profile and computer-derived three-dimensional projections of carboxylesterase B variants

(a) Projection of pI and m_F values of carboxylesterase B variants into a two-dimensional profile distinguished 41 alloenzymes; the number of isolates in which each alloenzyme was detected is indicated. Alloenzymes designated I to X, which were employed for comparative mobility curves, are shown in boxes (see the text). (b) Three-dimensional projection of the data in (a) reveals, apart from the B_1 and B_2 groups already observed in Fig. 1, the existence of two subgroups in group B_1 , namely B_{1a} and B_{1b} .

obtained to confirm the theoretical predictions proposed in Figs. 4 and 5, but all of these curves had qualitative mobility patterns already described previously (Righetti *et al.*, 1978) and it is of no interest to reproduce them here.

Furthermore, on titrating together five randomly chosen strain extracts from each high-frequency variant, we observed a single titration curve, providing evidence of their molecular homogeneity.

In addition, by analysing several titration curves, we were able to estimate that for all esterase B enzymes lying on one of the four straight lines (see Fig. 5) each protonation corresponds to a pI difference of 0.15 unit and a mobility jump of 2 units.

Fig. 3. Comparative protein mobility curves obtained by combined isoelectrofocusing-electrophoresis

(a) Comparison of variant I (pI 4.6/ m_F 70) and variant VI (pI 4.75/ m_F 70). The typical fork form oriented upwards with lines joining above the pH value of 7 corresponds to a substitution of a neutral amino acid residue in variant I by histidine in variant VI. Note that the pI value of each variant is the pH value at which they cross the groove. (b) Comparative titrations of variant I (pI 4.6/ m_F 70) and variant VIII (pI 4.6/ m_F 72). Here again there is a fork-like mobility pattern, but oriented downwards with lines joining well above their pI value; this is interpreted as representing a substitution of an arginine residue in variant I by histidine in variant VIII.

DISCUSSION

Each of the esterase B variants studied is considered to be the phenotype product of a single gene locus because only one molecular type was observed per bacterial strain. However, the wide dispersion of variants

Fig. 4. Theoretical substitutional relationship among variants of carboxylesterase B

Theoretical possibilities of several substitutions are depicted, assuming that they are all derived from the major variant I, and their difference in number of charges is mentioned in the upper and right-hand margins. Note that the substitutional relationships between variants I and VI on the one hand and between variants I and VIII on the other had already been determined (Fig. 3) by mobility curves. The variant numbering adopted in the text is derived from this Figure. Mutations not encountered in our investigations are shown as broken lines. The direction of arrows indicates the direction of substitutions. Key: N, neutral amino acid; B, basic amino acid; A, acidic amino acid; His, histidine; Arg, arginine.

observed in Fig. 2(a) on the one hand and their segregation into two subpopulations (see Fig. 1) on the other raised questions concerning their molecular relatedness.

In order to explain the observations we adopted the following method of analysis which is as follows. We made a theoretical projection of various possibilities of

Fig. 5. Distribution of the group-B₁ alloenzymes along the four theoretical straight lines

Definition of the line 1 through mobility curves and involvement of histidine in interrelating line 1 with lines 2 and 3 has already been shown in Figs. 3 and 4 (see also the Discussion section). If the major variant I is fixed and assumed to be the starting point of substitutions, then among the 32 variants of group B₁ at least eight variants (circled) must arise by serial substitutions. Group-B₂ variants lie apart from the group-B₁ enzymes.

substitutions in the major esterase B alloenzyme (designated I in Fig. 4). Changes in the values of mobilities or pI of esterase variants were estimated from the observed fact that for more than 50% of group-B₁ variants each difference of protonation results in uniform difference of 2 m_F units and 0.15 unit of pI.

In this way, as shown in Fig. 4, electrophoretic and isoelectrofocusing variants of the B₁ group can be individually classified into five different charge classes (+2, +1, 0, -1, -2). Three of the naturally occurring variants, namely II, III and IV shown in Fig. 2(a), belong to equivalent charge classes (-1, +1 and +2 respectively) in both electrophoresis and isoelectrofocusing.

Hence their substitutional differences with respect to variant I could define a linear relationship of pI versus mobility (thick line in Fig. 4). Whereas variants V and VI are defined as belonging respectively to -1 and +1 charge classes by isoelectrofocusing, both, along with variant I, are classified by electrophoresis into the 0 charge class. A similar discrepancy but with the inverse situation was observed for variants VII and VIII with respect to variant I. The anomalous but clearly defined behaviour of these variants suggested to us the involvement of histidine because the imidazolyl group of histidine should be in the protonated state in isoelectrofocusing (pI between 4 and 4.9) and in the deprotonated state in electrophoresis (pH 8.6). Also, the position of variant IX with respect to variant I in this profile can be explained by the substitution of an aspartic acid residue by histidine. Specifications of all substitutions are described in the legend to Fig. 4.

None of the distinct simple substitutional possibilities was able to explain the emergence of variant X from variant I. The completely contradictory behaviour of this variant in electrophoresis and isoelectrofocusing (-1 charge class in the former and +1 in the latter) from the theoretical profile of Fig. 4 suggests strongly the

involvement of serial substitutions, both implicating histidine residues, via either variant VI or VIII. These theoretical considerations are verified by comparative titration curves of various combinations of group-B₁ esterase variants. This established technique (Krishnamoorthy *et al.*, 1978; Righetti *et al.*, 1978, 1979) allowed us to confirm all the suggested substitutions unequivocally (see the Results section and Fig. 4).

Further extension of similar intervariant analysis to variants VI, VIII and IX shown in Fig. 4 both theoretically and through titration experiments allowed us to define a line along these variants extending from the +1 to the -1 charge class that had the same slope as that of the previous one. Along the same line of arguments and experimentally defined substitutions by titration curves, we were able to define for group-B₁ variants four straight lines of different levels obeying the rule $m = K \cdot M^{-\frac{1}{2}}$ (Offord, 1966) as shown in Fig. 5, all having the same slope.

The following points of physicochemical and biological interest emerge clearly from these data. Since all the group-B₁ variants are lying on or near the four charge-mobility-related straight lines with identical slopes, we conclude that all these variants are molecularly interrelated through simple and serial substitutions. All those lying on one of the lines (about 50% of the total) had the amino acid substitutions giving pK' values as predicted, whereas the others had pK' values that differed from those predicted, although within a limited range. Such restricted differences could be due to several factors: thus intrinsically the substitution may modify the conformation and hence the ionization of certain groups, or differential interaction of extrinsic factors such as salt ions and carrier ampholytes with the substituted groups can also alter the predictable pK' values. In this respect, we note from our data that the charge-state model (King & Ohta, 1975) is biochemically unrealistic, and other models may emerge, especially when extremely shallow immobilized pH gradients (Dossi *et al.*, 1983) become a regular isoelectrofocusing procedure for protein analysis.

Interestingly, the substitutions involving histidine are responsible not only for vertical and horizontal dispersions in the two-dimensional profile but also for relating one level to the other. Understanding of biological significance of this aspect will necessitate study of other enzyme variants in *E. coli*.

Some 25% of the variants must arise through serial substitutions of variant I. We stress this point because the serial substitutions must reflect a mutational rate of at least one order of magnitude higher than that of simple substitutions. This has important consequences in the analysis of the genetic diversity of a locus on the one hand and in consideration of this fact in terms of theories of molecular evolution on the other.

Although group-B₁ variants appeared to be widely diverse, physicochemical analysis showed them to be highly interrelated. However, no structural relationship

between group-B₁ variants and the clustered group-B₂ variants can be drawn on physicochemical grounds, even if one assumes multiple successive mutations, because no designable line will ever be parallel to that of B₁ group. Hence we believe that important changes of balance in charge have occurred in the molecules of group-B₂ variants. However, although enzymes of group B₂ differ from each other by charge changes, they nevertheless possess common electrophoretic features that separate them as a group from the variants of group B₁.

The most interesting biological aspect of group-B₂ esterases is that more than 92% of them are produced by *E. coli* human strains isolated from pathogenic sources. Among them, a substantial percentage of the strains not only possesses one of the two principal virulence factors, namely α -haemolysin and mannose-resistant haem-agglutinin, but also produces the two concomitantly, in contrast with the B₁ strains (Goulet & Picard, 1986).

In conclusion, we believe that the physicochemical approach employed here will be useful in the field of evolution, taxonomy and pathogenicity, through the understanding of molecular bases for phenotype differentiation.

REFERENCES

- Bernstein, S. C., Throckmorton, L. H. & Hubby, J. L. (1973) *Proc. Natl. Acad. Sci. U.S.A.* **70**, 3928–3931
- Coyne, J. A. (1976) *Genetics* **84**, 593–607
- Coyne, J. A. (1982) in *Isozymes, Current Topics in Biological and Medical Research* (Rattazzi, M. C., Scandalios, J. G. & Whitt, G. S., eds.), pp. 1–32, Alan R. Liss, New York
- Dossi, G., Celentano, F., Gianazza, E. & Righetti, P. G. (1983) *J. Biochem. Biophys. Methods* **7**, 123–142
- Goulet, Ph. (1973) *J. Gen. Microbiol.* **77**, 27–35
- Goulet, Ph. (1980) *J. Gen. Microbiol.* **117**, 493–500
- Goulet, Ph. & Picard, B. (1985) *Electrophoresis* **6**, 132–135
- Goulet, Ph. & Picard, B. (1986) *J. Gen. Microbiol.* **132**, 1853–1858
- Goulet, Ph., Picard, B. & Laget, P. F. (1984) *Ann. Microbiol. (Paris)* **134A**, 375–387
- Hubby, J. L. & Lewontin, R. C. (1966) *Genetics* **54**, 577–594
- King, J. L. & Ohta, T. (1975) *Genetics* **79**, 681–691
- Krishnamoorthy, R., Bosio, A. B., Labie, D. & Righetti, P. G. (1978) *FEBS Lett.* **94**, 319–323
- Lawrence, S. H., Melnick, P. J. & Weiner, H. E. (1960) *Proc. Soc. Exp. Biol. Med.* **105**, 572–575
- Offord, R. E. (1966) *Nature (London)* **211**, 591–593
- Ohta, T. & Kimura, M. (1973) *Genet. Res.* **22**, 201–204
- Righetti, P. G., Krishnamoorthy, R., Gianazza, E. & Labie, D. (1978) *J. Chromatogr.* **166**, 455–460
- Righetti, P. G., Krishnamoorthy, R., Lapoumeroulie, C. & Labie, D. (1979) *J. Chromatogr.* **177**, 219–225
- Singh, R. S., Lewontin, R. C. & Felton, A. A. (1976) *Genetics* **84**, 609–629
- Uriel, J. (1961) *Ann. Inst. Pasteur Paris* **101**, 104–119
- Uriel, J. (1966) *Bull. Soc. Chim. Biol.* **48**, 969–982
- Vesterberg, O. & Svensson, H. (1966) *Acta Chem. Scand.* **20**, 820–834