Ann. Rev. Microbiol. 1976. 30:409–25

ENZYME RECRUITMENT IN EVOLUTION OF NEW FUNCTION

◆1687

Roy A. Jensen

Department Biological Sciences, State University of New York at Binghamton, Binghamton, New York 13901

CONTENTS

INTRODUCTION

Homologous Proteins as Molecular Blueprints of Evolutionary Relationships

An immense, emergent base of data documenting primary sequences of proteins is appearing in the current literature. The use of the amino acid anatomy of proteins whose functions have been conserved intact through the evolutionary diversification of organisms is now a well-established approach to deciphering phylogeny. The problems of convergent evolution, a major hindrance at more general levels of analysis (especially with microorganisms), are largely avoided. The *Atlas* (23) already contains about 300 sequences of vertebrate proteins, allowing extensive lineage relationships to be drawn for such proteins as proteolytic enzymes, proteins of blood clotting, and a variety of hormones, hemoglobins, and immunoglobulins.

In microbial evolution, tenable formulations of phylogeny are based upon the substantial data in existence for cytochrome *c* as well as for the ferredoxins and other iron-sulfur proteins (24). Relatively general metabolic criteria, such as pathways of carbohydrate catabolism, energy metabolism, CO_2 fixation, cell wall structure, tetrapyrrole biosynthesis, alternative pathways of amino acid biosynthesis, and alternative patterns of enzyme control have proven useful in defining taxonomic relationships (24). Within such metabolic pathways, rigorous documentation of particular enzyme homologies undoubtedly offer impressive opportunities for more precise and refined evaluations of phylogenetic relationships.

Extraordinary insights into evolutionary relationships of proteins will be increasingly feasible as sequences of particular proteins become available. Substantial emphasis has been devoted to evolutionary relationships inferred from protein homologies in cases where identical function has been conserved in different organisms (14, 32, 41). This article emphasizes the prospects for defining evolutionary relationships between proteins that coexist in the same organism but that no longer serve the same function.

Primitive Expansion of Metabolic Capabilities

Pristine life must have been restricted to limited genetic information encoding a small number of proteins. The most attractive mechanism for acquisition of additional genetic information is that of gene duplication in tandem (J. Roth, manuscript in preparation). Divergence of the new gene copies via mutational modifications that altered enzymatic reactivities presumably allowed the expansion of metabolic capabilities and the evolution of new biochemical pathways.

Retrograde Evolution

Horowitz (42) addressed the conceptual difficulty of accounting for the successful recruitment of more than one enzyme in a multistep pathway, in view of the apparent lack of selective advantage prior to the establishment of an intact pathway. The hypothesis of retrograde revolution (42) invokes the stepwise and sequential recruitment of new enzymes in reverse order and is based on the assumption that

each intermediate of the backwardly evolving pathway was readily available in the primitive environment. This thesis is vulnerable to the uncertainty of the latter assumption (34) and is further weakened by the extreme chemical lability of many intermediary metabolites, as well as by the barriers to their transport that may have existed in the absence of specialized transport systems.

AN ALTERNATIVE FOR PATHWAY EVOLUTION BASED UPON SUBSTRATE AMBIGUITY

Substrate Specificity of Primitive Enzymes

Although modern enzymes (such as certain nucleases) can be extraordinarily specific, substrate specificities are perhaps broader than is generally appreciated. It is suggested that primitive enzymes possessed a very broad specificity, permitting them to react with a wide range of related substrates. This would maximize the catalytic versatility of an ancestral cell that functioned with limited enzyme resources. Subsequent elaboration of additional enzyme proteins following gene amplification (59) would allow the luxury of increased specialization and the improved metabolic efficiency that is thus permitted. Additionally, enzymic alteration from broad specificity to narrow specificity is consistent with expectations for the very early evolution of primitive proteins, because the change proceeds from a less restrictive state to a more restrictive one.

Substrate Ambiguity of Modern Enzymes

Many contemporary proteins exhibit considerable substrate ambiguity. For example, a strikingly low specificity of leucine pathway enzymes (13, 73), especially α-isopropylmalate synthetase, occurs in *Salmonella typhimurium* (54), *Pseudomonas aeruginosa* (73), *Neurospora crassa* (93), *Saccharomyces cerevisiae* (90), and *Hydrogenomonas* H16 (39). In some cases, as in isoleucine-valine biosynthesis, a balanced substrate ambiguity of a single series of enzymes is obligatory to the production in parallel of essential multiple products.

Sometimes both broad and narrow specificities are represented by coexisting enzyme species. Thus, *Aspergillus nidulans* possesses a general amidase in addition to acetamidase and formamidase (45). Permease multiplicity appears to be a common mechanism providing microorganisms with the capacity to deal appropriately with drastically different concentrations of exogenous metabolites (61). One transport component is frequently a high-affinity, low-capacity permease that allows cells to scavenge and concentrate low concentrations of extracellular amino acids for use in protein synthesis; a second component may be a low-affinity, high-capacity permease suitable for the utilization of larger quantities of amino acids as carbon and energy sources. Usually the high-capacity (general) system of transport is less specific than the high-affinity (specialized) system (3, 53).

Transaminase enzymes are notorious for their broad specificity, and this presumably accounts for the absence of transaminase mutants in well-defined systems of

biochemical genetics such as *Bacillus subtilis* (68). Generally, multiple transaminase activities of broadly variable and overlapping specificity can be separated from extracts of microorganisms such as *Pseudomonas aeruginosa* (47, 83) or *B. subtilis* (D. L. Pierson and R. A. Jensen, unpublished data), although a transaminase, imidazolylacetolphosphate: L-glutamate aminotransferase of *S. typhimurium,* appears to be quite specialized (35). A number of other enzymes (such as phosphatases) are generally reputed to possess broad substrate specificity.

The evolutionary exploitation of substrate ambiguity has been demonstrated with L-fucose isomerase, which also utilizes D-arabinose and L-xylose (12); with ribitol dehydrogenase, which can utilize xylitol and L-arabitol (66); with L-fucolokinase, which is reactive with D-ribulose (57); and with L-fuculose-1-phosphate aldolase, which can utilize D-ribulose-1-phosphate (57). In *P. fluorescens,* D-galactose and L-arabinose dehydrogenases each utilize both substrates, albeit differentially (8). Numerous examples of recruitment of substrate-ambiguous enzymes for new function in bacteria were described by Hegeman & Rosenberg (37).

Specializing Influence of Regulation Mechanisms

The in vitro potential of enzymes in modern metabolic pathways to carry out anomalous reactions is largely masked by regulatory controls. Erroneous reactions are maximized by high levels of enzyme, high concentration of abnormal substrate, and low concentration of normal substrate (e.g. 55). Repression of biosynthetic enzymes by end products, induction of catabolic enzymes by normal substrates, and relationships of feedback inhibition or substrate activation produce conditions that tend to reserve enzyme function for the production of the primary metabolites. In *Pseudomonas putida,* L-leucine feedback inhibits α-isopropylmalate synthetase activity, as well or better when 2-ketobutyrate, pyruvate, or 2-ketoisocaproate are substrate reactants as when the normal substrate, 2-ketoisovalerate, is used (R. Conrad and R. A. Jensen, and unpublished data).

It is noteworthy that experimental evolution in the laboratory is accompanied almost inevitably by loss of regulatory mechanisms as a first step (37). It is likely that regulatory mechanisms were undeveloped in primitive cells (21). Assuming the presence of substrate-ambiguous and unregulated enzymes, ancestral cell types may have produced a variety of minor products under particular conditions. Any fortuitously formed compounds that happened to be useful would have conferred a selective advantage, thereby providing a basis for increased and more specific production of that compound (by gene duplication and specialization via mutation).

Fortuitous Error and Gain of Multistep Pathways

Broad substrate specificity provides a kind of biochemical leakiness, especially in the absence of regulation, which is exploitable for gain of function. Nonenzymatic reactivities of some intermediary metabolites also provide a basis for a similar kind of leakiness. For example, the three-step conversion of chorismate to phenylalanine can occur via two nonenzymatic reactions and one nonspecific transaminase; a possible scheme for evolution of aromatic amino acid biosynthesis based upon these reactivities has been proposed (47). Thus, fortuitous reactivities, both enzymatic and

nonenzymatic, provide a background of biochemical heterogeneity against which individual enzyme recruitments improve the function of a slow, but already existing, multistep sequence. In some cases, the evolution of a single new enzyme may initiate a new multistep pathway. For example, one could imagine an ancestral sequence from pyruvate to L-valine. The evolutionary acquisition of threonine deaminase-producing α-ketobutyrate, a substrate analog of pyruvate, might be a plausible origin of the isoleucine-valine pathway.

The fortuitous error concept of enzyme recruitment dispenses with the principal assumption upon which the hypothesis of retrograde evolution depends. The apparent dilemma of how individual enzymes could have developed for a newly evolving pathway considering that they would be of no selective advantage until multiple recruitments established the intact pathway was resolved by assuming the presence of intermediary metabolites in the primitive environment (42). Alternatively, fortuitous error could have permitted individual alterations of enzyme capability to be of selective value through amplifications of an erroneous sequence. I suggest that many pathway sequences may have existed at a very low level of expression owing to fortuitous reactions, both enzymatic and nonenzymatic. A diverse background of quantitatively minor biochemical reactions is envisioned in which erroneous sequences are potentially available for recruitment to new metabolic function following the specialized development of each appropriate catalytic step.

Pathway Origin From One Ancestral Protein?

The concept of retrograde evolution, the likely significance of tandem gene duplication, and the existence of adjacent gene clusters encoding enzymes for certain pathways (operons) encouraged the suggestion (43) that each protein in a biochemical pathway is specified by a copy of a homologous gene. Hence, proteins coded by genes arranged in operons should be homologous in amino acid sequence. Where this has been specifically tested—with tryptophan synthetase subunits in *Escherichia coli* (19) and the muconate-lactonizing enzyme and muconolactone isomerase of *P. putida* (81)—homologies have not been found.

On the other hand, there is evidence that the analogous reactions of the converging β-ketoadipate pathway are catalyzed by homologous enzymes (65, 70). Hence, it seems likely that new enzyme functions are established most easily and most commonly by recruitment of proteins already catalyzing analogous reactions. Many proteins may be represented in modern cells by a number of homologous counterparts that have diverged considerably in function (e.g. see the possibilities outlined in Figure 1). It is intriguing in this context that when the gene for β-galactosidase was deleted from *E. coli,* a new β-galactosidase activity appeared following intensive selection (11, 30). The new enzyme may have evolved from a monosaccharide-reactive protein because it binds galactose, galactosamine, and γ-galactonolactone tightly (4). A priori the most ancient enzyme types would have had substantial opportunity to generate homologous proteins recruited to different function. Many of the dehydrogenases must have had a very ancient origin. For example, the ancient origin of glyceraldehyde phosphate dehydrogenases is suggested by the similarity of the enzyme in mammals, yeast, and *E. coli* (52).

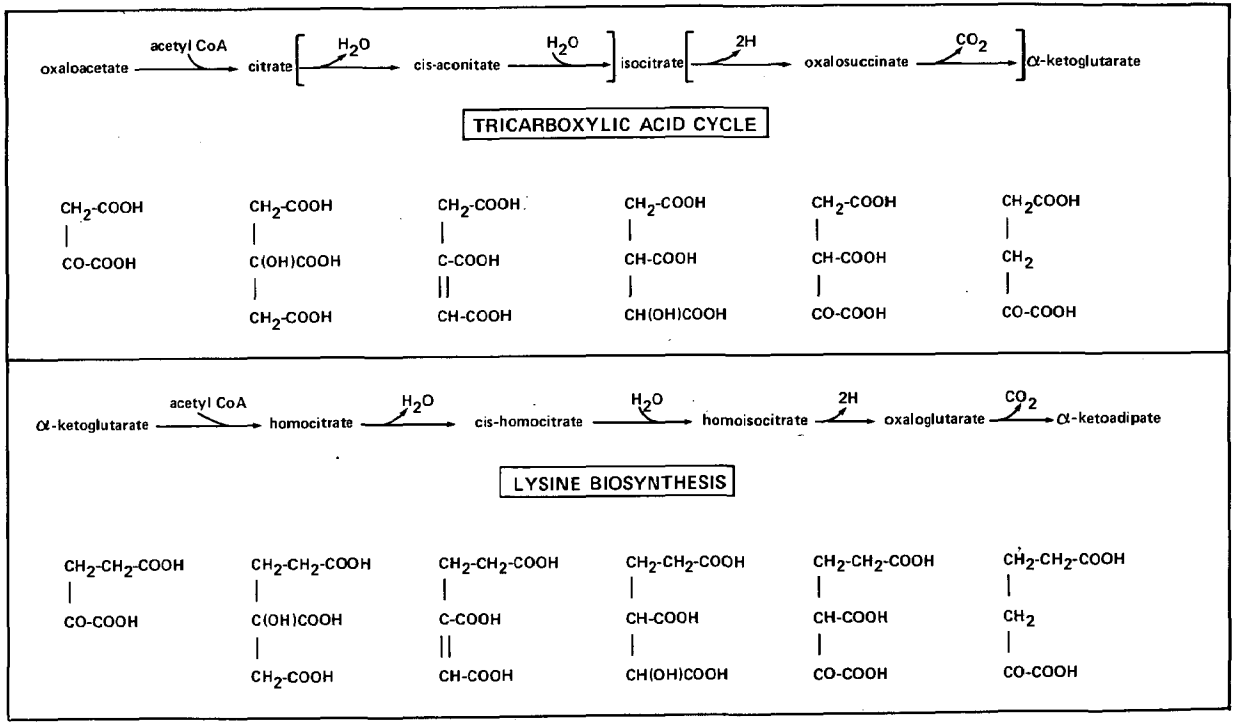
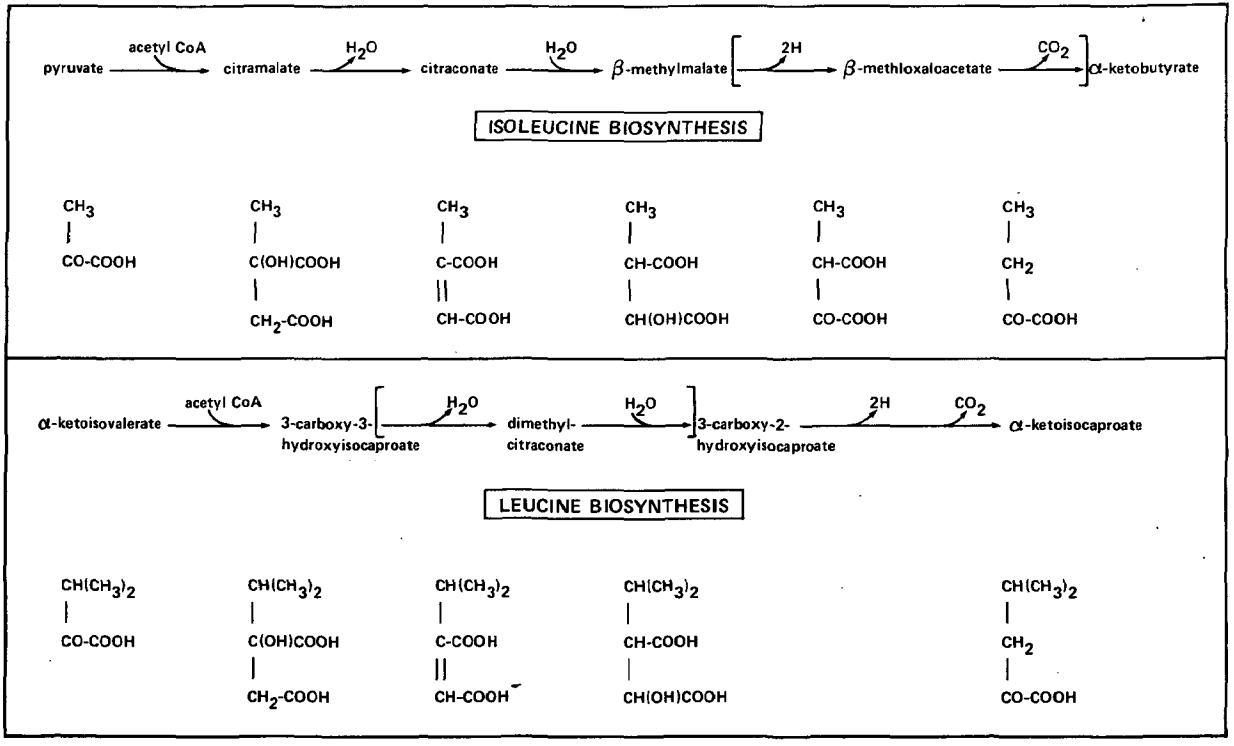


Figure 1 Analogous reaction sequences in major biochemical pathways. Each of the enzyme sequences shown begins with the condensation of acetyl CoA with an α-keto acid, followed by rearrangement, oxidation, and elimination of a carbon to produce an α-keto acid, which differs from the original keto acid because of the presence of an additional carbon. Instances where free intermediates may not be formed

are indicated with brackets. Each of the portions of the lysine and isoleucine pathways shown participate in one of several naturally occurring alternative pathways (see Table 1) for synthesis of these amino acids.



Pathway Evolution en Bloc

The concept of an ancestral repertoire of unregulated enzymes with broad reactivities allows the possibility that a new metabolite can form in minor amounts, even though it may be several enzymatic steps removed from the original substrate provided. If the presence of a new substrate produced minor amounts of a useful product through a multistep series of erroneous reactions, gene duplications would be selectively maintained since increased enzyme levels would increase the error level. Appropriate mutational modification of the gene copies could then complete the evolution of the new pathway.

Entire blocks of analogous enzyme sequences are commonplace in metabolism, and it seems likely that the recruitment of a new pathway from a preexisting and analogous sequence would have certain advantages. Exactly analogous reaction steps in four important biochemical pathways are illustrated in Figure 1. The tricarboxylic acid cycle is probably a pathway of ancient origin and contains many enzymes that are illustrative of commonly used reaction mechanisms in living cells.

It is suggestive that the modern leucine pathway enzymes exemplify extreme substrate ambiguity, permitting a number of substrate analogs to proceed through the entire reaction sequence (13, 73). In *Pseudomonas,* the metabolism of ethylmalic acids and probably *n*-propylmalic acids reflects the substrate ambiguities of leucine pathway enzymes (73). One can imagine the future evolution of an independent set of enzymes for ethylmalate metabolism following recruitment of the leucine pathway enzymes. It is not uncommon for parallel series of analogous reactions to be catalyzed by a single set of enzymes, an initial circumstance that could easily lead to the origin of a second homologous set of enzymes. A familiar example of a parallel pathway is the isoleucine-valine pathway of biosynthesis. Perhaps the existence of a second acetohydroxy acid synthetase in some enteric organisms and in *Pseudomonas* suggests an emergent evolution of isoleucine and valine independence now in progress.

Other examples of parallel series of analogous reactions mediated by common enzymes include the catabolism of branched-chain amino acids (63) and the conversion of D-mandelate and *p*-hydroxy-D-mandelate to benzoate and *p*-hydroxybenzoate, respectively (15). The parallel pathways of D-fucose and L-arabinose utilization in *Pseudomonas* are catalyzed by enzymes with sufficiently broad substrate specificites to accommodate 5-C or 6-C substrates (16). In *Aerobacter aerogenes,* L-mannose and L-rhamnose (6-deoxy-L-mannose) are metabolized by common isomerase, kinase, and aldolase enzymes (64). In species of *Pseudomonas* a single transaminase and a single dehydrogenase have dual reactivities that permit the conversion of prephenate to L-tyrosine by either the 4-hydroxyphenylpyruvate pathway or the pretyrosine pathway of synthesis (47). The *meta*-cleavage pathway in a number of pseudomonads utilizes a single set of ambiguous enzymes which can dissimilate analog derivatives of naphthalenes (94), cresols (5), benzoates (67), and catechols (75).

BIOCHEMICAL DIVERSITY IN NATURE

It is certainly true that major biochemical pathways exhibit a very substantial chemical uniformity, probably a reflection of their ancient origin and the limited feasible biochemical alternatives. However, the unity theme for biochemical pathways may have been overemphasized in the literature. The distribution in nature of two pathways for lysine biosynthesis has been of great interest because of the broad phylogenetic implications. The existence in nature of two alternative enzyme sequences leading to lysine initially appeared to be a unique instance of biochemical pathway diversity. Over the past few years, however, an accumulating body of evidence reveals more diversity than was previously appreciated in the means whereby various life forms synthesize essential cellular metabolites such as amino acids. Examples of alternative sequences for amino acid biosynthesis now emerging from the scientific literature are cited in Table 1. The prospects for presently known

Table 1 Alternate biochemical sequences for amino acid biosynthesis in nature[a]

Biosynthetic pathway	Sequence denotation	Occurrence	Reference
Arginine	acetylornithase	*Bacillus; Escherichia*	88
	transacetylase	*Micrococcus;* yeasts	
Isoleucine	threonine	major[b]	1
	β-methylaspartase	*Escherichia*	72
	citramalate	*Leptospira*	13
	ambiguous	*Escherichia;* yeast	90
Lysine	α-aminoadipate	higher fungi; euglenids	58
	diaminopimelate	bacteria, lower fungi, green algae, plants	89
Methionine	o-acetylhomoserine[c]	*Bacillus, Neurospora*	
	o-phosphorylhormoserine	green algae, green plants	22
	o-succinylhomoserine	blue-green bacteria; enterics	
Serine	ambiguous[d]	*Pseudomonas, Rhodopseudomonas*	38, 77
Tyrosine	4-hydroxyphenylpyruvate	*Bacillus;* enterics	47
	pretyrosine	blue-green bacteria, plants	48
	ambiguous	*Pseudomonas, Thiobacillus*[e]	82

[a] In addition, recent evidence (1975) suggests that reductive carboxylation participates in the formation of the keto acid precursors of leucine, valine, phenylalanine, and tyrosine in rumen microorganisms (76).

[b] The threonine pathway to L-isoleucine biosynthesis appears to be the major pathway in use so far (except for *Leptospira*), even in cases where an alternate sequence exists.

[c] Fungi possess two alternative (ambiguous) pathways to homocysteine, the direct precursor of L-methionine, providing that the sulfate assimilation pathway functions normally. The results of Paszewski & Grabski (69) indicate that the cystathionine pathway for methionine formation is the main one in *Aspergillus nidulans.* Another pathway involving homocysteine synthase is an alternative-conditional one, physiologically effective when the enzyme is derepressed.

[d] Containing both the phosphorylated (probably used by most microorganisms) and the non-phosphorylated pathways to serine.

[e] J. Shiveley and R. A. Jensen, unpublished data.

instances of pathway diversity offering a multifold basis for insight into evolutionary relationships are heightened by the likelihood that still other examples of alternative pathways will be uncovered as detailed work is done in a greater range of organisms.

It is common to repeat in other microorganisms assays for rigorously characterized enzyme activities of *E. coli* and to assume that positive results mean that the pathway in the test organism is identical with that of *E. coli.* Although this is certainly a reasonable first assumption, it is not proof. For example, the demonstration of prephenate dehydratase and prephenate dehydrogenase activities in *Pseudomonas* suggests that it contains the same aromatic pathway found in *E. coli.* However, L-tyrosine and L-phenylalanine can also be synthesized enzymatically from pretyrosine (the transamination product of prephenate) via pretyrosine dehydrogenase (47) and pretyrosine dehydratase (R. N. Patel and R. A. Jensen, unpublished data).

In molecular biology, the remarkable productivity of experimental systems such as *E. coli* and *B. subtilis* has tended to distort the generalized image of microbial characteristics. Relatively little information about biochemical pathways is available in many widely distributed microbial groups because they are relatively poor experimental systems for the efficient definition of gene-enzyme relationships. A true sense of the distribution of biochemical pathways in nature requires a balanced sampling of such groups. In this context, a novel biosynthetic pathway to L-isoleucine was found in the little-studied *Leptospira interrogans* (Table 1). The newly found pretyrosine pathway of L-tyrosine biosynthesis is utilized in many, if not all, species of blue-green bacteria, an exceedingly large and diverse group of procaryotic organisms (80). Yet, until recently, relatively few species were available in pure culture, and little enzymological characterization of biosynthetic or catabolic pathways had been done.

PLAUSIBLE PATHWAY HOMOLOGIES

Keto Acid Elongation Sequences

The analogous reactions illustrated in Figure 1 would be excellent candidates for homologous lineage. Such data could confirm or deny theories about the order of evolution of amino acid pathways (95). Amino acid sequence determination is the crucial and ultimate test of homology. Molecular weights of homologous proteins might differ for a variety of reasons. Even serological evidence is not necessarily reliable (71). For example, if one protein diverged from its homologous counterpart following fusion with another protein, molecular weight and antigenic properties could differ substantially. Screening for sequence homologies must be adequate to detect internal homologies reflecting the elongation mechanism that is instrumental in the evolution of some proteins, for example, human serum transferrin and bacterial ferredoxins (23, 62).

Sequential reactions of proteins in one pathway may be associated with a single protein in an analogous pathway (Figure 1). A newly arisen enzyme, fused with a preexisting enzyme, could act to channel an intermediate preferentially in a new

direction. In the evolutionary derivation of one pathway from another, the formation of multifunctional proteins as the result of gene fusion could channel metabolite flow preferentially into the new pathways, even prior to mutational divergence. The frequent occurrence of gene fusion in tryptophan biosynthesis—probably with methionine one of the last amino acid pathways to evolve (33, 49, 95)—may reflect an intermediate state of evolutionary development that might account for what is otherwise a rather confusing picture (18). Fused proteins may primarily be of selective value only for the intermediate stage of evolutionary change before sufficient divergence has occurred to permit reacquisition of independence.

Sugar Metabolism Sequences

The phosphoribose isomerase of the pentose phosphate pathway is analogous to phosphoglucose isomerase or triose phosphate isomerase in the Embden-Meyerhof pathway. The 6-phosphogluconate dehydrogenase of the pentose phosphate pathway is analogous to the reactions of isocitrate dehydrogenase and NADP-malate dehydrogenase, all being oxidative decarboxylations. The evolution of metabolism has been considered in the framework of energy formation, resulting in the following postulated order of evolution: Emden-Meyerhof, pentose phosphate, glyoxylate, and tricarboxylate (31, 44). Given the pentose phosphate sequence, the Entner-Doudoroff sequence would only require two additional enzymes. It is interesting that the anabolic hexose monophosphate and the catabolic Entner-Duodoroff pathways both share dehydrogenation of D-glucose-6-P and hydrolysis of 6-phospho-D-glucono-1,5-lactone. A rigorous search for possible homologies in the proteins of these pathways would be worthwhile.

Amino Acid Biosynthesis

Relationships whereby certain enzymes of the histidine biosynthetic pathway exhibit excellent reaction analogies with the three reactions of the phosphorylated pathway of L-serine synthesis and with three reactions of L-tryptophan synthesis are summarized in Table 2. A weaker analogy can be drawn between imidazole glycerolphosphate dehydratase (histidine) and indoleglycerolphosphate synthetase (trypto-

Table 2 Analogous reactions of histidine, tryptophan, and serine biosynthesis

Histidine pathway	Serine pathway[a]
histidinol dehydrogenase	phosphoglycerate dehydrogenase
histidinol phosphate transaminase	phosphoserine transminase
histidinol phosphatase	phosphoserine phosphatase
	Tryptophan pathway
phosphoribosyl formimino PRAIC ketol isomerase	phosphoribosyl anthranilate isomerase
L-glutamine amidotransferase	anthranilate synthetase
phosphoribosyl-ATP pyrophosphorylase	phosphoribosyl transferase

[a] The phosphorylated pathway of serine biosynthesis.

phan). If the analogous reactions shown prove to be homologous, it would suggest that many of the serine and tryptophan pathway enzymes were recruited from an ancestral pathway for histidine (or vice versa).

Peripheral Pathways

Peripheral pathways of secondary metabolism and other pathways that are not generally essential for growth are also a rich source of multistep pathway analogies. The use of these pathways is vulnerable to criticism on grounds that the cognate genes may be exchanged promiscuously among distantly related species (26). Widespread gene transfer creates uncertainty about whether a given protein evolved in the cell of current residence.

Nevertheless, these pathways exemplify the ubiquity of multistep analogies in nature. The converging β-ketoadipate pathway has been characterized extensively in *Pseudomonas* and *Acinetabacter* (70, 71). In *Streptomyces* two analogous sequences of five reactions each act in series in the biosynthesis of the diguanidinated inositol derivative, streptomycin (91). Ancestral strains presumably produced only the monoguanidinated inositol, the product of the first five enzymes.

In many leguminous plants, canavanine is produced by the canaline-urea cycle (74). These two pathways possess analogous reactions involving the sequential formation of canaline, O-ureidohomoserine, canavaninosuccinate, and canavanine in one series, while ornithine, citrulline, arginosuccinate, and arginine are produced in the other series. Undoubtedly, many antibiotics, antimetabolites, and other analog structures produced by microorganisms and plants are formed by reaction sequences analogous to those of other, more ubiquitous metabolic pathways. The eventual identification of enzyme homologies in these pathways would not be surprising.

PLAUSIBLE HOMOLOGIES OF INDIVIDUAL ENZYMES

Isoenzymes and Identical Enzyme Reactions in Different Pathways

Identical reactions may occur in different biochemical pathways, frequently biosynthetic on the one hand and catabolic, on the other hand. Such proteins probably are derivatives of isoenzymes that originated by gene duplication. A familiar precedent is the derivation of multiple lactate dehydrogenase genes from a single ancestral locus during vertebrate evolution (40). In some cases this is undoubtedly the origin of channeling phenomena in branched pathways whereby isoenzymes are complexed with different proteins (17, 83). The presence of an extra chorismate mutase in one strain of *B. subtilis,* but not in another (51), may exemplify an intermediary stage of evolution that led to the chorismate mutase complexes with prephenate dehydratase and prephenate dehydrogenase in *E. coli* (17). Regulatory isoenzymes such as those for 3-deoxy-D-*arabino* heptulosonate 7-phosphate synthetase (46) or for aspartokinase (78, 87) would be expected to exhibit homology in at least the portion of the protein essential for catalysis.

Recruitment for Divergence in Biosynthesis and Catabolism

Threonine deaminase has dual representation in many microorganisms as an inducible catabolic enzyme for threonine degradation and as a repressible biosynthetic enzyme for isoleucine biosynthesis. The appropriate physiological expression of these enzymes is dictated by wholly different regulatory controls.

The degradation of quinate and the biosynthesis of aromatic amino acids both generate 5-dehydroquinate as an intermediate. *Neurospora crassa* (28) and *Acinetobacter calcoaceticus* (85) possess inducible catabolic dehydroquinases in addition to the biosynthetic species of enzyme. *Pseudomonas fluorescens* possesses both a catabolic and a biosynthetic ornithine carbamyltransferase (79). In *Escherichia coli* only a repressible acetylornithine δ-transaminase is normally present in wild type for arginine biosynthesis. A single mutation apparently led to the appearance of a second arginine-inducible enzyme (7).

Recruitment for Reverse Catalytic Function

It seems likely that some homologous proteins have diverged following duplication of a parent gene to favor catalysis in opposite direction for reversible reactions. Thus one might expect degradative and biosynthetic transaminases for leucine, for example, to favor leucine and α-ketoisocaproate, respectively. In *Acetobacter calcoaceticus* (85) the reversible reaction between 5-dehydroshikimate and shikimate is mediated by a biosynthetic 5-dehydroshikimate reductase and a degradative shikimate dehydrogenase.

A clear example of evolution to reverse function in the laboratory is the evolution of propanediol catabolism through recruitment of enzymes normally involved in L-fucose and L-lactate metabolism; in *E. coli* an anaerobic L-lactaldehyde reductase was converted to an L-1,2-propanediol dehydrogenase (20). Tryptophanase from *Escherichia coli* can catalyze the reversal of its normal function, forming L-tryptophan from indole, pyruvate, and ammonia (92). The evolutionary potential of this enzyme is also suggested by its catalysis of a whole series of α,β-elimination reactions and β-replacement reactions. Of further interest is the analogy in reaction mechanism of tryptophanase with that of β-tyrosinase (92).

Homologous Proteins of Different Function

A number of families of homologous proteins with divergent function have been well characterized. These include the serine proteases (84), pyridoxal phosphate enzymes (25), enterosecretory proteins (2), and pyridine nucleotide-requiring proteins (6). In the latter case, six enzymes of diverse function (glyceraldehyde-3-phosphate dehydrogenase, alcohol dehydrogenase, glutamate dehydrogenase, dihydrofolate reductase, and lactate dehydrogenase) all exhibited partial homology of sequence.

The extensive possibilities for recruitment of new function are aptly illustrated by mammalian lactose synthetase, and enzyme that catalyzes the reaction between UDP-galactose and glucose to form lactose and UDP. Lactose synthetase consists of two proteins, one of which is a ubiquitous galactosyl transferase that catalyzes

the synthesis of glycoprotein in most tissues (9, 10). The presence of α-lactalbumen in the lactating mammary gland results in the ready utilization of glucose as an acceptor moiety. It is noteworthy that the potential of the transferase to act as lactose synthetase reflects the substrate ambiguity of the enzyme, because it has been shown (27) that the galactosyl transferase protein alone can form lactose if the glucose concentration is sufficiently high. The protein α-lactalbumen exhibits homology in the sequence with lysozyme. This complex relationship illustrates recruitment of reverse function in that lysozyme catalyzes the breaking of β-1,4-glucosidic linkages, while α-lactalbumen catalyzes their formation.

Another example of homologous proteins that participate in different functions includes the β-chain of human haptoglobin and the chymotrypsin family of serine proteases (56).

Shared Proteins

In *Bacillus subtilis,* anthranilate synthetase and *p*-aminobenzoate (PABA) synthetase compete for a common glutamine-binding subunit (50). A similar relationship exists in *Acetobacter calcoaceticus* (18). The two enzymes utilize the same substrates, but differ with respect to the position at which the amino group is placed on the aromatic ring. It seems probable that both enzymes originated from a common ancestral protein. The ancestral proteins were probably the ammonia-reactive forms of these enzymes (50, 96). The ability of both aminases to react as amidotransferases when complexed with subunit-X (the glutamine-binding subunit) also suggests a common origin of the aminases. A number of amidotransferases that may have originated as ammonia-utilizing aminases in primitive cells exist. It has been widely concluded that glutamine amidotransferase functions of enzymes arose from duplication and divergent evolution of an ancestral gene (29, 60, 86).

SUPPRESSOR MUTATIONS

Restoration of wild-type function by suppressor mutations offers many possibilities for insight into evolutionary versatilities (36). Such changes may equally well represent the recapture of some ancient enzyme relationship or the potential for a new evolutionary development. An especially apt example (55) is offered by a suppressor mutation in the biosynthetic pathway for proline from glutamate. A second partial blockade (i.e. the suppressor) in acetylornithine δ-transaminase causes partial starvation for L-arginine, derepression of N-acetylornithine deacylase, a limitation of N-acetylornithine, and an accumulation of N-acetylglutamic γ-semialdehyde behind the blockade. The deacylase has weak reactivity with N-acetylglutamic γ-semialdehyde. Under these conditions of high enzyme level, low concentration of normal substrate, and high concentration of abnormal substrate, the deacylase converts N-acetylglutamic γ-semialdehyde to glutamic γ-semialdehyde, thus bypassing the blockade in the proline pathway. A new pathway has thus evolved in which two common steps for proline and arginine biosynthesis precede the formation of N-acetylglutamic semialdehyde, the new branch point of the pathway.

SUMMARY

Ancient cells must have possessed small gene content. Primitive enzymes may have possessed broad specificity and undeveloped regulation mechanisms. The considerable substrate ambiguity of these enzymes resulted in the formation of minor amounts of erroneous products. Fortuitous formation of metabolites offered ancient cells maximum biochemical flexibility with minimal gene content. Gene duplication provided the opportunity for increased gene content and increased specialization of the diverging enzymes, the substrate specialization being further reinforced by the development of regulatory mechanisms. Recruitment of enzymes for new pathways did not necessarily require the sequential and backwardly evolving progression of evolutionary steps required by the hypothesis of retrograde evolution of biochemical pathways. Substrate ambiguity remains a conspicuous feature of many contemporary proteins, and evolutionary exploitation of substrate ambiguity in a variety of organisms is still apparent.

Literature Cited

1. Abramsky, T., Rowland, L. P., Shemin, D. 1965. *J. Biol. Chem.* 240:2971–75
2. Adelson, J. W. 1971. *Nature* 229: 321–25
3. Ames, G. F., Roth, J. R. 1968. *J. Bacteriol.* 96:1742–49
4. Arraj, J. A., Campbell, J. H. 1975. *J. Bacteriol.* 124:849–56
5. Bayly, R. C., Dagley, S., Gibson, D. T. 1966. *Biochem. J.* 101:293–301
6. Bennett, C. D. 1974. *Nature* 248:67–68
7. Billheimer, J. T., Jones, E. E. 1974. *Arch. Biochem. Biophys.* 161:647–51
8. Blachnitzky, E.-O., Wengenmayer, F., Kurz, G. 1974. *Eur. J. Biochem.* 47: 235–50
9. Brew, K., Vanaman, T. C., Hill, R. L. 1967. *J. Biol. Chem.* 242:3747–49
10. Brodbeck, U., Denton, W. L., Tanahashi, N., Ebner, K. E. 1967. *J. Biol. Chem.* 242:1391–97
11. Campbell, J. H., Lengyel, J. A., Langridge, J. 1973. *Proc. Natl. Acad. Sci. USA* 70:1841–45
12. Camyre, K. P., Mortlock, R. P. 1965. *J. Bacteriol.* 90:1157–58
13. Charon, N. W., Johnson, R. C., Peterson, D. 1974. *J. Bacteriol.* 117:203–11
14. Clarke, P. H. 1974. *Evolution in the Microbial World,* ed. M. J. Carlile, J. J. Skehel, pp. 183–217. London: Cambridge Univ. Press. 430 pp.
15. Clarke, P. H., Ornston, L. N. 1975. *Genetics and Biochemistry of Pseudomonas,* ed. P. H. Clarke, M. H. Richmond, pp. 191–261. London: Wiley. 366 pp.
16. Ibid, pp. 263–340
17. Cotton, R. G. H., Gibson, F. 1965. *Biochim. Biophys. Acta* 100:76–81
18. Crawford, I. P. 1975. *Bacteriol. Rev.* 39:87–120
19. Creighton, T. E. 1966. *Relationships and Interactions between Proteins of the Tryptophan Biosynthetic Pathway.* PhD dissertation, Stanford Univ., Palo Alto, Calif.
20. Cocks, G. T., Aguilar, J., Lin, E. C. C. 1974. *J. Bacteriol.* 118:83–88
21. Cove, D. J. 1974. *Nature* 251–256
22. Datco, A. H., Giovanelli, J., Mudd, S. H. 1974. *J. Biol. Chem.* 249:1139–55
23. Dayhoff, M. O., ed. 1972. *Atlas of Protein Sequence and Structure,* Vol. 5, Suppl. I, 1973. Washington DC: Natl. Biomed. Res. Found.
24. DeLey, J., Kersters, K. 1975. *Comprehensive Biochemistry,* ed. M. Florkin, E. H. Stotz, pp. 1–77. Amsterdam:Elsevier. 284 pp.
25. Dunathan, H. C., Voet, J. G. 1974. *Proc. Natl. Acad. Sci. USA* 71:3888–91
26. Dunn, N. W., Gunsalus, I. C. 1973. *J. Bacteriol.* 114:974–79
27. Fitzgerald, D. K., Brodbeck, U., Kiyosawa, I., Mawal, R., Colvin, B., Ebner, K. E. 1970. *J. Biol. Chem.* 245:2103–8
28. Giles, N. H., Partridge, C. W. H., Ahmed, S. I., Case, M. E. 1967. *Proc. Natl. Acad. Sci. USA* 58:1453–60
29. Greishaber, M., Bauerle, R. 1972. *Nature New Biol.* 236:232–35
30. Hall, B. G., Hartl, D. L. 1974. *Genetics* 76:391–400

31. Hall, J. B. 1971. *J. Theoret. Biol.* 30:429–54
32. Hartley, B. S. 1970. *Phil. Trans. R. Soc. London Ser. B* 257:77–87
33. Hartman, H. 1974. *Origins Life* 6: 423–27
34. Hartman, H. 1973. *J. Mol. Evol.* 4:359–70
35. Hartman, P. E., Hartman, Z., Stahl, R. C. 1971. *Adv. Genet.* 16:1–34
36. Hartman, P. E., Roth, J. R. 1973. *Adv. Genet.* 17:1–105
37. Hegeman, G. D., Rosenberg, S. L. 1970. *Ann. Rev. Microbiol.* 24:429–62
38. Heptinstall, J., Quayle, J. R. 1970. *Biochem. J.* 117:563–72
39. Hill, F., Schlegel, H. G. 1969. *Arch. Mikrobiol.* 68:1–17
40. Holmes, R. S. 1972. *FEBS Lett.* 28: 51–55
41. Holmes, R. S., Scopes, R. K. 1974. *Eur. J. Biochem.* 43:167–77
42. Horowitz, N. H. 1945. *Proc. Natl. Acad. Sci. USA* 31:153–57
43. Horowitz, N. H. 1965. *Evolving Genes and Proteins,* ed. V. Bryson, H. J. Vogel. New York: Academic. 629 pp.
44. Horvath, R. S. 1974. *J. Theoret. Biol.* 47:361–71
45. Hynes, M. J. 1975. *J. Gen. Microbiol.* 91:99–109
46. Jensen, R. A., Nasser, D. S. 1968. *J. Bacteriol.* 95:188–96
47. Jensen, R. A., Pierson, D. L. 1975. *Nature* 254:667–71
48. Jensen, R. A., Stenmark, S. L. 1975. *J. Mol. Evol.* 4:249–59
49. Jukes, T. H. 1973. *Nature* 246:22–26
50. Kane, J. F., Holmes, W. M., Jensen, R. A. 1972. *J. Biol. Chem.* 247:1587–96
51. Kane, J. F., Stenmark, S. L., Calhoun, D. H., Jensen, R. A. 1971. *J. Biol. Chem.* 246:4308–16
52. Kaplan, N. O. 1970. *Homologies in Enzymes and Metabolic Pathways.* ed. W. J. Whelan, J. Schultz, pp. 66–88. Amsterdam:North Holland. 463 pp.
53. Kay, W. W. 1971. *J. Biol. Chem.* 246:7373–82
54. Kohlhaw, G. B., Leary, T. R. 1970. *Methods in Enzymology,* ed. H. Tabor, C. W. Tabor, p. 771. New York: Academic
55. Kuo, T.-T., Stocker, B. A. D. 1969. *J. Bacteriol.* 98:593–98
56. Kurosky, A., Barnett, D. R., Rasco, M. A., Lee, T.-H., Bowman, B. H. 1974. *Biochem. Genet.* 11:279–92
57. LeBlanc, D. J., Mortlock, R. P. 1971. *J. Bacteriol.* 106:90–96
58. LéJohn, H. B. 1971. *Nature* 231:164–68
59. Lewis, E. B. 1951. *Cold Spring Harbor Symp. Quant. Biol.* 16:159–63
60. Li, H.-C., Buchanan, J. M. 1971. *J. Biol. Chem.* 246:4713–19
61. Lin, E. C. C. 1970. *Ann. Rev. Genet.* 4:225–261
62. MacGillivray, R. T. A., Brew, K. 1975. *Science* 190:1306–7
63. Martin, R. R., Marshall, V. D., Sokatch, J. R., Unger, L. 1973. *J. Bacteriol.* 115:198–204
64. Mayo, J. W., Anderson, R. L. 1969. *J. Bacteriol.* 100:948–55
65. Meagher, R. B., Ornston, L. N. 1973. *Biochemistry* 12:3523–30
66. Mortlock, R. P., Fossitt, D. D., Wood, W. A. 1965. *Proc. Natl. Acad. Sci. USA* 54:572–79
67. Murray, K., Duggleby, C. J., Sala-Trepat, J. M., Williams, P. A. 1972. *Eur. J. Biochem.* 28:301–10
68. Nester, E. W., Jensen, R. A., Nasser, D. S. 1969. *J. Bacteriol.* 97:1500–2
69. Paszewski, A., Grabski, J. 1974. *Mol. Gen. Genet.* 132:307–20
70. Patel, R. N., Meagher, R. B., Ornston, L. N. 1973. *Biochemistry* 12:3531–37
71. Patel, R. N., Meagher, R. B., Ornston, L. N. 1974. *J. Biol. Chem* 249:7410–19
72. Phillips, A. T., Nuss, J. I., Moosic, J., Foshay, C. 1972. *J. Bacteriol.* 109: 714–19
73. Rabin, R., Salamon, I. I., Bleiweis, A. S., Carlin, J., Ajl, S. J. 1968. *Biochemistry* 7:377–87
74. Rosenthal, G. A. 1972. *Plant Physiol.* 50:328–31
75. Sala-Trepat, J. M., Murray, K., Williams, P. A. 1972. *Eur. J. Biochem.* 28:347–56
76. Sauer, F. D., Erfle, J. D., Mahadevan, S. 1975. *Biochem. J.* 150:357–72
77. Schmidt, L. A., Sojka, G. A. 1973. *Arch. Biochem. Biophys.* 159:357–72
78. Stadtman, E. R., Cohen, G. N., Le Bras, G., De Robichon-Szulmajster, H. 1961. *Cold Spring Harbor Symp. Quant. Biol.* 26:319–21
79. Stalon, V., Ramos, F., Pierard, A., Wiame, J.-M. 1972. *Eur. J. Biochem.* 29:25–35
80. Stanier, R. Y., Kunisawa, R., Mandel, M., Cohen-Bazire, C. 1971. *Bacteriol. Rev.* 35:171–205
81. Stanier, R. Y., Wachter, D., Gasser, C., Wilson, A. C. 1970. *J. Bacteriol.* 105:351–62
82. Stenmark, S. L., Pierson, D. L., Glover, G. I., Jensen, R. A. 1974. *Nature* 247:290–92

83. Stenmark-Cox, S. L., Jensen, R. A. 1975. *Arch. Biochem. Biophys.* 167: 540–46
84. Stroud, R. M. 1974. *Sci. Am.* 231:74–88
85. Tresguerres, M. E. F., Ingledew, W. M., Canovas, J. L. 1972. *Arch. Mikrobiol.* 82:11–19
86. Trotta, P. P., Birt, M. E., Haschemeyer, R. H., Meister, A. 1971. *Proc. Nat! Acad. Sci. USA* 68:2599–2603
87. Truffa-Bachi, P., Cohen, G. N. 1968. *Ann. Rev. Biochem.* 37:79–108
88. Udaka, S. 1966. *J. Bacteriol.* 91:617–21
89. Vogel, H. J. 1960. *Biochim. Biophys. Acta* 41:172–73
90. Vollbrecht, D. 1974. *Biochim. Biophys. Acta* 362:382–89
91. Walker, J. B. 1974. *J. Biol. Chem.* 249:2397–404
92. Watanabe, T., Snell, E. E. 1972. *Proc. Natl. Acad. Sci. USA* 69:1086–90
93. Webster, R. E., Gross, S. R. 1965. *Biochemistry* 4:2309–17
94. Williams, P. A., Catterall, F. A., Murray, K. 1975. *J. Bacteriol.* 124:679–85
95. Wong, J. T.-F. 1975. *Proc. Natl. Acad. Sci. USA* 72:1909–12
96. Zalkin, H., Murphy, T. 1975. *Biochem. Biophys. Res. Commun.* 67:1370–77

CONTENTS