Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Chemical Biology

Advances in synthesis of biotin and assembly of lipoic acid

John E Cronan[1,2]

Although biotin and lipoic acid are two universally conserved cofactors essential for intermediary metabolism, their synthetic pathways have become known only in recent years. Both pathways have unusual features. Biotin synthesis in *Escherichia coli* requires a methylation that is later removed whereas lipoic acid is assembled on the enzymes where it is required for activity by two different pathways.

Addresses
[1] Department of Microbiology, University of Illinois, Urbana, IL 61801, USA
[2] Department of Biochemistry, University of Illinois, Urbana, IL 61801, USA

Corresponding author: Cronan, John E (jecronan@illinois.edu)

Current Opinion in Chemical Biology 2018, **47**:60–66

This review comes from a themed issue on **Mechanistic biology**

Edited by **John A Gerlt** and **Tobias J Erb**

https://doi.org/10.1016/j.cbpa.2018.08.004

Introduction

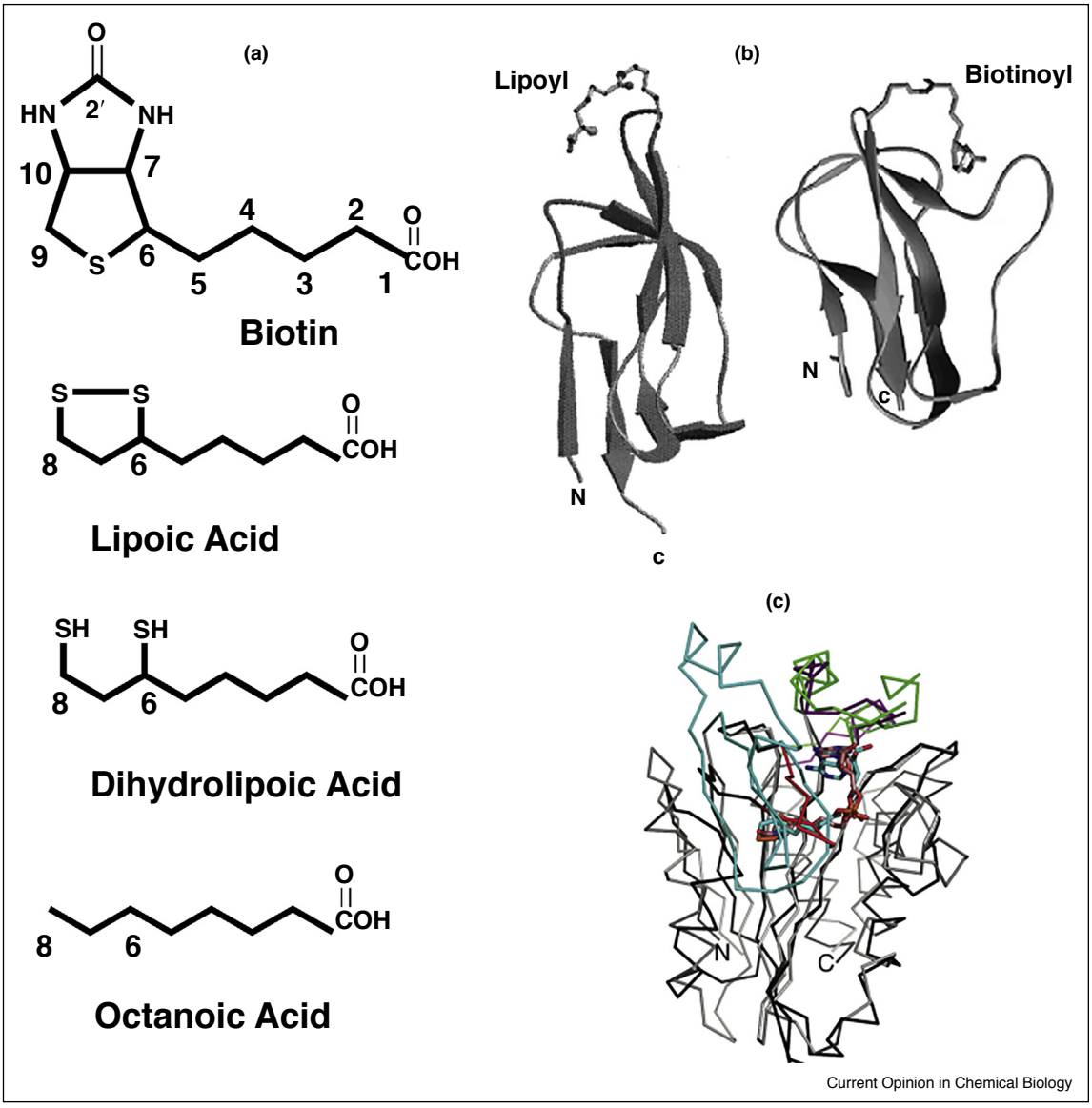
Biotin and lipoic acid (Figure 1a) share many similarities [1]. Both cofactors must be covalently attached to their cognate enzyme subunits to play their roles in central metabolism. Attachment is via an amide linkage to a specific lysine residue. Once attached, they act as part of 'swinging arms' that shuttle intermediates between active sites in a covalent form of substrate channeling. The protein species modified by biotin or lipoate attachment are rare. *Escherichia coli* has a single biotinylated protein and three lipoylated proteins whereas mammals have four each of biotinylated and lipoylated proteins.

The *ca*. 60–80 residue protein domains to which biotinoyl and lipoyl moieties are attached have very similar structures [1] (Figure 1b). Both biotinoyl and lipoyl domains are flattened β-barrels, comprised of two 4-stranded antiparallel β-sheets. The lysine residue earmarked for cofactor attachment is found in an exposed position in a tight type I β-turn. At the opposite end of the domains the N-terminal and C-terminal residues of the domain are found in close together in space. Indeed, depending on the pair of domains chosen for overlay, the backbone atoms can superimposed to 1 Å and these proteins define a protein family (PF00364). Indeed, mutation of residues close to the target lysine are sufficient to allow a biotinoyl domain to accept lipoic acid [2]. Both biotinoyl and lipoyl domains are tethered to their cognate enzyme subunits by long (25–30 residue) polypeptide chains, characteristically rich in alanine, proline and charged amino acids that form flexible but extended linkers. The biotinoyl domains are at the C-termini of biotinylated proteins whereas the lipoyl domains are found at the N-termini of lipoylated proteins. Although only lipoyl linkers have been studied in detail, a lipoyl linker can be substituted for a biotinoyl linker [3]. In the original swinging arm concept only the cofactor plus the lysine side chain was thought to swing. However, has become clear that entire domains swing between active sites [4]. A useful way to think of the domain-linker arrangement is as a tadpole: the modified lysine being the nose and the linker, the tail. Note that these domain structures are highly conserved throughout biology. Foreign domains often are modified when expressed in diverse host cells. Biotin and lipoic acid can be attached by ligases called biotin–protein ligase and lipoate ligase, respectively [1]. Both ligase reactions proceed via an active site-bound acyl-adenylate intermediate that is attacked by the ε-amino group of the lysine to be modified. All biotin proteins are modified via a ligase whereas lipoate ligases function only when environmental lipoate is available [1,5]. The backbones of the *E. coli* biotin and lipoate ligases can be superimposed to about 2.8 Å (Figure 1c). Together with the enzymes of lipoate assembly (discussed below) they comprise a new protein family (PFAM 03099.13). *In vivo* assembly of lipoyl moieties proceeds via pathway in which the lipoate precursor, octanoate, is first attached to the target proteins and then converted to lipoate via sulfur insertion [5]. This will be discussed below.

Recent progress in biotin synthesis

Biotin is composed of two fused heterocyclic rings plus a valerate side chain (Figure 1). The cofactor is an essential vitamin for mammals and birds; only bacteria, archaea, plants and some fungi synthesize biotin. Commercial biotin is made by chemical synthesis, but industrial interest in obtaining biotin by bacterial synthesis has resulted in detailed studies of the assembly of the biotin heterocyclic rings by DuPont and Shiseido [1]. Both companies determined the origin of the biotin carbon atoms by feeding ^{13}C-labeled precursors to *E. coli* [6,7]. These analyses showed that the biotin carbon atoms are

Figure 1



Biotin and lipoic acid. **(a)** Shown are the structures of biotin, lipoic acid, reduced lipoic acid (dihydrolipoic acid) and the octanoate precursor. Note that the biotin structure is larger to aid legibility of the numbering system. **(b)** The structures of the lipoyl and biotinoyl domains of *E. coli*. Note that the 'thumb' structure that protrudes from the *E. coli* biotinoyl domain is only found in acetyl-CoA carboxylase subunits, the biotinoyl domains of other enzymes lack this protrusion and thus more closely resemble lipoyl domains. **(c)** Superimposition of the *E. coli* BirA biotin protein ligase (black, PDB entry 2EWN) superimposed on the *E. coli* LplA lipoate ligase (gray, PDB entry 2ART) with insertions corresponding to the BirA biotin and biotinoyl adenylate binding loops colored cyan and green, respectively. Lipoyl adenylate (cyan) and biotinoyl adenylate are depicted in stick representation. The lipoyl adenylate binding loop is not visible due to its disorder. The C-termini have been deleted for clarity.

derived from acetate, alanine and CO_2. The valerate side chain plus the first two carbons (carbons 6 and 7) of the first ring were known to come from pimelic acid, a seven-carbon α, ω-dicarboxylic acid (which can replace biotin in some bacteria). The ^{13}C-labeling studies showed that six of the seven carbon atoms (carbons 2–7, Figure 1a) are derived by head-to-tail condensations as in fatty acid synthesis whereas the valerate carboxyl (C1) is derived from CO_2. The differing labeling patterns of the pimelate carboxyl carbons indicated that free pimelate was not a pathway intermediate and argued that pimelate was made by a fatty acid synthesis pathway in which the acetate-derived carboxyl (now C7 of the ring) was linked to a thiol, such as that of acyl carrier protein (ACP). Straightforward use of the membrane lipid fatty acid synthesis pathway seemed precluded by the requirement that the hydrophobic enzyme active sites must tolerate a carboxyl group. *E. coli* genetics had defined the genes responsible