ABSTRACT

Title of Document: CHARACTERIZATION OF MULTIPLE SURFACE LOOPS IN ESCHERICHIA COLI BIOTIN PROTEIN LIGASE FUNCTIONS

Poorni Ranmali Adikaram, Doctor of Philosophy, 2012

Directed By: Dr. Dorothy Beckett, Department of Chemistry and Biochemistry

Multifunctional proteins utilize several strategies to interact with different partners, resulting in diverse cellular outcomes. Structural, thermodynamic and kinetic features of these proteins influence the switch between functions. The *Escherichia coli* biotin protein ligase, BirA, is a bifunctional protein essential for biotin homeostasis. It transfers biotin to the biotin carboxyl carrier protein (BCCP) subunit of acetyl-CoA carboxylase in its metabolic role or dimerizes and binds the biotin biosynthetic operator as a transcriptional repressor. Each function involves forming a protein-protein interaction, and because a single surface of BirA is used to form both interactions, the two are mutually exclusive. The BirA interaction surface contains several loops, two with highly conserved sequences, and the remaining with variable sequences. In this work the roles of four loops in facilitating BirA function were investigated. Amino acids from surface loops were replaced with alanine to obtain 18 alanine substituted variants. Homodimerization energetics measured using sedimentation equilibrium yielded an 8

kcal/mol range for variants from all loops. Steady-state and stopped-flow kinetic assays yielded 7 of 18 variants that exhibited slower rates than wild-type in biotin transfer to BCCP. The majority of alanine substituted variants are from constant loops. These results indicate that the biotin transfer reaction is mediated primarily through the constant loop and homodimerization is facilitated by all surface loops. The energetics of transcription repression complex assembly, which comprises contributions from dimerization and DNA binding, was assessed using DNaseI footprint titrations. Although variants exhibit a broad range in total assembly energetics, all dimers bind with similar affinities to DNA, implying independence between DNA binding and dimerization domains. The switch between functions was also investigated using inhibition DNaseI footprint titrations. A direct correlation between inhibition of repression complex assembly and rates of BirA-BCCP association was observed, reinforcing a kinetic mechanism for the switch between BirA functions. These studies indicate that multiple surface loops form the structural basis for bifunctionality, and BirA switches between protein-protein interactions through a kinetically controlled mechanism. Elucidation of structural and mechanistic aspects of the BirA functional switch enhances our understanding of how multifunctionality evolves and the mechanism of switching between biological functions.

CHARACTERIZATION OF MULTIPLE SURFACE LOOPS IN ESCHERICHIA COLI BIOTIN PROTEIN LIGASE FUNCTIONS.

By

Poorni Ranmali Adikaram

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2012

Advisory Committee:
Professor Dorothy Beckett, Chair
Professor Catherine Fenselau
Professor Douglas Julin
Assistant Professor Nicole LaRonde-LeBlanc
Professor Kevin McIver

Acknowledgements

I would like to thank my advisor, Dr. Beckett, for her guidance during my time in graduate school, especially these last few months. I also thank my committee members for their encouragement and advice.

I thank previous lab members for their help, friendship and good times. I would especially like to thank Maria, who can be summed up in one word: awesome! Thanks also to Chris, for keeping me sane and entertained.

I thank my fellow Biochemistry friends Urszula, Katie, Jai, Joe and Mark and all the other friends I have made throughout the years. I also thank Jim Watson, for being an amazing instructor to TA for and for his support.

I thank my friends and relatives both here and abroad for their love and support. I would also like to thank my husband, parents, sister and mother-in-law for their patience, understanding and never-ending love as I worked to finish my degree.

This work was funded, in whole or in part, by the National Institutes of Health Grants R01GM46511 and S10RR15899 to Dr. Beckett.

Table of Contents

List of Tables

List of Figures

Chapter 1: Background and Introduction

1.1 Multiple interactions facilitate cellular processes

A combination of interactions, including protein-protein, protein-nucleic acid and protein-small ligand interactions, give rise to almost all cellular processes. While most of these interactions lead to favorable outcomes, some unexpected and aberrant interactions lead to disease. Characterizing these macromolecular interactions provides information that can be used to predict binding partners, design new interactions and design drugs. This information, in turn, can be used to understand essential cellular pathways, improve drug target affinities and efficiencies, and to cure disease.

Multifunctional proteins can interact with other proteins and with nucleic acids to affect different cellular outcomes, such as metabolism, immune response, apoptosis and development (1-4). For example, the pyruvate kinase M2 enzyme is an essential component of the glycolytic pathway, catalyzing the conversion of phosphoenolpyruvate to pyruvate. In addition, this isoform interacts with signaling proteins, transcription factors and different oncoproteins in tumor cells leading to signal transduction, transcription regulation, glycolysis and tumor growth (5, 6). Protein-nucleic acid interactions result in transcription, translation, DNA repair, replication, and mRNA processing. The tumor suppressor p53 is one such protein capable of interacting with both proteins and nucleic acids to regulate cellular function. Under conditions of DNA damage and oxidative stress, p53 is activated through interaction with protein kinases (7, 8). This interaction, in turn, activates the transcriptional regulatory role of p53 and it binds to

specific DNA sequences involved in DNA repair, apoptosis and senescence (9). The unique ability of p53 to interact with both proteins and DNA gives it a multifunctional role in tumor suppression.

Protein-protein and protein-nucleic acid interactions do not necessarily result in favorable outcomes (10). Many known human diseases arise from mutations at protein interaction interfaces. For example, the interaction of replication protein A, essential for DNA repair, replication and recombination, with BRCA2, a protein involved in DNA double-strand break repair through the homologous recombination pathway, is inhibited due to a mutation in the *BRCA2* gene (11), leading to accumulation of DNA breaks that can result in breast cancer. Likewise, infection by the papilloma virus can lead to warts, cervical cancer and other tumors in humans (10). The E2 viral protein is important for replication of the virus in human cells through several interactions (Fig. 1). The E2 C-terminal domain dimerizes and binds viral DNA, while the N-terminal transactivation domain binds the E1 helicase protein, resulting in regulation of viral protein- and DNA-binding interactions. The transactivation domain also interacts with the human Brd4 protein, forming a bond between virus and host cells, leading to infection in humans. Thus, studying protein interfaces will lead to a better understanding of the mechanisms behind these interactions, which can be applied in the design of new drugs and to cure disease.

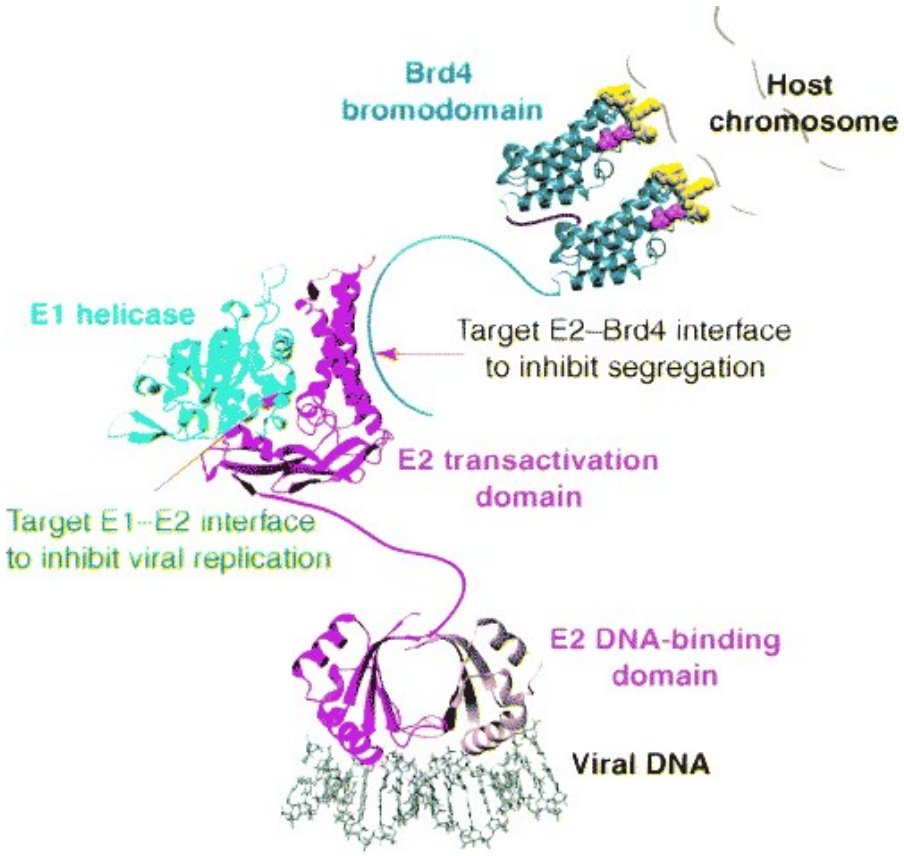


Figure 1. Interactions of the E2 papilloma virus protein. The E2 DNA-binding domain interacts with viral DNA, while the transactivation domain interacts with the E1 helicase. The transactivation domain also interacts with human Brd4 protein in the first steps of human infection. The figure was obtained from (10).

The inhibition of protein-protein interactions has emerged as the latest approach in the drug discovery field (12). Traditionally, antibodies have been used to target protein-protein interactions because these molecules bind their targets with high

selectivity and affinity (13). Also many drugs are synthesized to target the active site or binding site of the ligand to inhibit the reaction (14). Inhibiting protein-protein interactions using small compounds result in the inhibition of subsequent steps in the pathway, though challenges exist in this method. Protein-protein interaction interfaces are usually large and flat with an average contact surface of approximately 1500-3000Å^2 (15). Since very few small molecules naturally bind to a protein interaction surface, designing and developing new drugs de novo that efficiently disrupt or mimic an interaction using such a small fraction of the interface is a daunting task. However, "hotspots," which are residues within certain regions on protein surfaces that contribute most to binding energetics have been identified, and are used as targets in drug design. In more than 30% of human cancers, RAS, a GTPase important in signaling, has been mutated. RAS binds RAF, a kinase that activates the MAP kinase signaling pathway, which regulates cell survival, proliferation and apoptosis (16). Small molecule inhibitors of the RAS-RAF interaction have been identified, and these drugs could inhibit this protein-protein interaction, thereby inhibiting tumor growth (17). Thus, though targeting a protein interaction surface to inhibit a protein-protein interaction is challenging, it can result in the inhibition of unfavorable reactions in the cell.

Multifunctional proteins are believed to contribute to one-third of all protein-protein interactions (18) and the structural basis for multispecifity varies for different proteins (19). Loops, due to their conformational flexibility, are important in mediating interactions (20). Intrinsically disordered proteins, or intrinsically disordered regions of a protein, are flexible and can undergo various conformational changes to accommodate

different interacting partners (21). Binding pockets in proteins also facilitate direct interactions with their partners through specific interactions (22). Imperfect packing, flat surfaces and bridging water molecules at binding sites also lead to protein-protein interactions (23). Hotspots are involved in forming interactions, and often two hotspot regions on each protein partner interact with each other to form the protein complex (24). In addition, multifunctional proteins can bind to their numerous partners through distinct interaction interfaces, thereby utilizing different residues for binding (25). The human 14-3-3 protein family is comprised of 7 isoforms (β, ϵ, η, γ, τ, σ and ζ) and is important for the regulation of metabolism, cell signaling, apoptosis, stress response and tumor formation (26). These proteins bind phosphopeptides in a specific binding site to regulate activation, inhibition, stabilization or orientation of its protein partners (27) (Fig. 2). Each protein can homodimerize or form a heterodimer with another isoform, and there is evidence that dimerization is used to control the cellular activities of the protein. Specificity of the seven isoforms for their various binding partners is facilitated through flexibility of the molecule, especially of the αG-αI helices and the CD and HI loop regions, which allow different peptides to enter the binding site. Thus, a combination of factors, including conformational flexibility, binding pocket specificity and protein assembly state all contribute to the proteins' ability to act as a multifunctional enzyme.

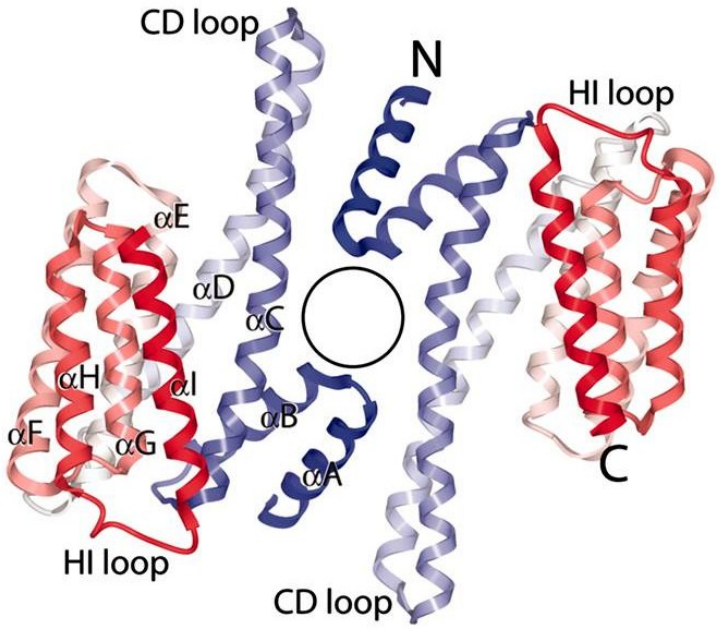


Figure 2. Crystal structure of the 14-3-3 protein dimer. Each monomer is colored blue to red from the N to C-terminal domain. The flexible CD and HI loops that are involved in ligand binding are labeled. The figure was obtained from (26).

One feature of multifunctional proteins is their ability to switch between functions to produce different cellular outcomes. The mechanism of switching between two or more functions can be under kinetic or equilibrium control. If under kinetic control, the rates of association and dissociation of protein partners controls the switch, while if under thermodynamic control, it is the affinities, or equilibrium dissociation constants of the interactions that favors one protein-protein interaction over the other. For example, two proteins, thrombomodulin (TM) and a monoclonal antibody (mAb), both bind to thrombin, a serine protease involved in blood clotting (28). Both proteins bind thrombin

in the same binding site and though this is not a biologically relevant system, the thermodynamic and kinetic profiles have been determined for these two protein-protein interactions. They have similar binding affinities, but their association and dissociation rates are very different. The TM-thrombin reaction is fast and reversible, while the thrombin-mAb reaction is slow and almost irreversible. The difference in rates for these two interactions implies a kinetically driven mechanism for thrombin functional switching.

The *Escherichia coli* biotin protein ligase, BirA, is an integral part of the bacterial proteome and is an excellent model to study protein multispecifity. BirA utilizes several structural elements to switch between two competing protein-protein interactions that result in either fatty acid synthesis or transcription repression. The study of the structural basis for multispecificity, as well as the mechanism of the switch, has enabled us to better understand the evolution of BirA and its central role in the cell.

1.2 Biotin regulatory system in E. coli

The biotin regulatory system is centered on biotin protein ligase, called BirA in *E. coli* (Fig. 3). BirA is a bifunctional protein involved in both protein-protein and protein-DNA interactions in metabolic and transcription repression roles (29, 30). The ligase is activated through the sequential binding of two ligands, biotin and ATP, to form the intermediate bio-5'-AMP (31). In its metabolic role, holoBirA associates with the biotin carboxyl carrier protein (BCCP) subunit of acetyl-CoA carboxylase (ACC) to transfer biotin. ACC utilizes biotin as a cofactor to make malonyl-CoA. Two holoBirA monomers

can also homodimerize and bind to the biotin biosynthetic operator to act as a transcription repressor (32, 33). The partitioning of BirA between its metabolic and transcription regulatory roles is governed by the amount of apoBCCP in the cell, which is in turn governed by cell growth rate (34, 35). Since ACC catalyzes the first committed step of fatty acid biosynthesis, the faster cells grow, the greater the requirement for membrane synthesis. Hence, there is a greater requirement for apoBCCP to be biotinylated, partitioning holoBirA towards its metabolic role. During periods of slow cell growth, when apoBCCP levels are low, holoBirA dimerizes, binds to DNA and represses transcription of biotin biosynthetic genes. In this way, BirA acts as a bifunctional enzyme to regulate biotin levels in the cell.

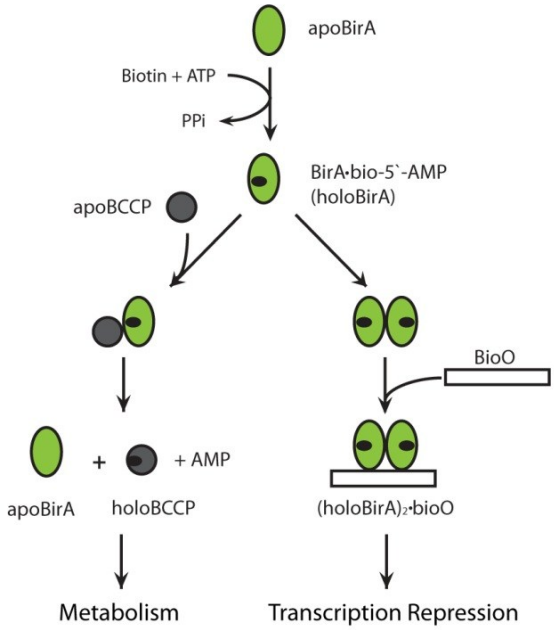


Figure 3. A schematic representation of the biotin regulatory system in *E. coli*. ApoBirA binds biotin and ATP to form bio-5'-AMP, and the resulting holoBirA binds one of two protein partners in either a transcription repression or metabolic role.

1.3 The BirA structure

BirA is a monomer in its unliganded form with a molecular weight of 35.3 kDa, and the 321 amino acids are organized into three domains (Fig. 4). The N-terminal domain, comprising residues 1-60, contains three α-helices and two antiparallel β-strands (36). Two of the helices, a three residue loop and the two β-strands form a winged helix-turn-helix motif, which is involved in DNA binding. The N-terminal domain is connected to the central domain through a linker. The central domain, compromising residues 80-269, contains a 7-stranded mixed beta-sheet packed against five helices. This domain also contains 5 loops and the active site of the protein. The C-terminal domain contains two three-stranded antiparallel beta-sheets and comprises residues 271-320.

The crystal structures of apoBirA (36, 37), BirA bound to biotin (38), as well as the dimer bound to biotinol-5'-AMP (btnOH-AMP), a non-hydrolysable analog of bio-5'-AMP, have been obtained (39). The apoBirA crystal structure contains four disordered loops, three of which, including residues 116-124, 140-146 and 193-199, become ordered in the biotin-bound structure. The 116-124 loop is known as the biotin binding loop (BBL) as the residues in this loop fold over the biotin molecule, thereby positioning it in the active site. In the btnOH-AMP bound crystal structure, the fourth loop, from residues 212-234, is ordered. This loop is known as the adenylate binding loop (ABL) as it folds over the adenylate moiety and positions both ligands to synthesize the physiological corepressor, bio-5'-AMP.

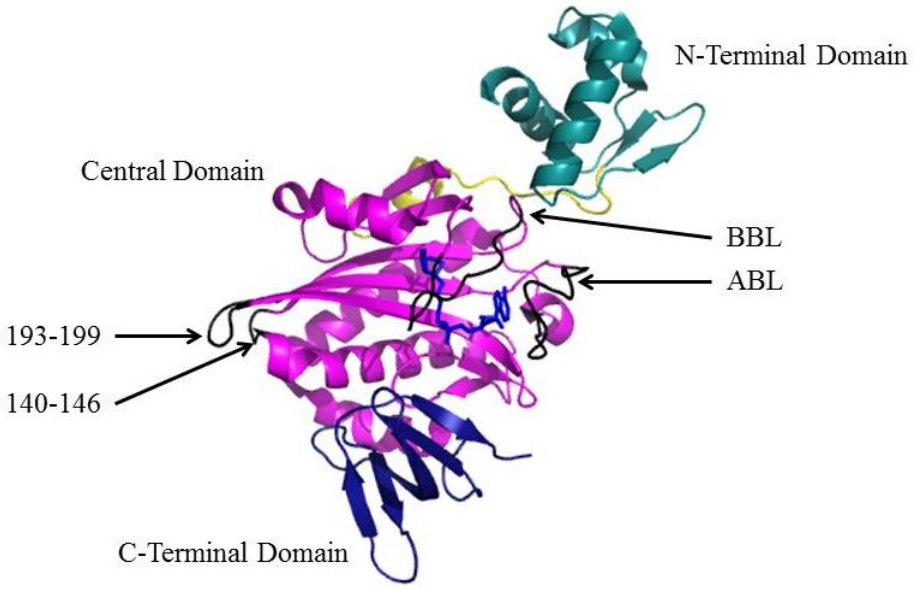


Figure 4. A model of the biotinol-5'-AMP-bound BirA monomer structure. The domains are colored: N-terminal domain, cyan; central domain, magenta; C-terminal domain, blue. The linker region is colored in yellow. Loops that are disordered in apoBirA are colored black and are labeled. The figure was generated using PyMOL (40) with input file 2EWN (39).

Several surface loops have previously been characterized for their roles in various BirA functions. The BBL, which contains the glycine rich sequence GRGRXG is conserved in all biotin protein ligases (41). Characterization of three variants from this loop, including G115S, R118G and R119W indicate that this loop is important for biotin and bio-5'-AMP binding, homodimerization, and DNA binding through defects in self-assembly (42, 43). The characterization of two variants from two other loops, including A146Δ and D197Y, indicated that these two loops are involved in homodimerization and DNA binding, again due to defects in protein assembly (43). Further characterization of single amino acid substituted variants in the BBL and ABL indicated that there is overlap

in use of structural information for homo- and heterodimerization, and that residues specific for each interaction exist (44). These results indicate that multiple disordered surface loops are required for ligand binding, homodimerization and subsequent DNA binding and in the interaction with BCCP. Thus, these multifunctional loops facilitate various BirA processes and enable BirA to act as a bifunctional protein.

1.4 BirA as a metabolic enzyme

BirA transfers biotin onto a conserved lysine residue on a β-turn of BCCP, forming an amide linkage between the carboxyl group of biotin and the ϵ-NH_2 group of lysine. The structure of the *E. coli* BCCP87 domain, the C-terminal BCCP fragment shown to be functionally equivalent to the full length protein, has been solved, and contains two sets of four antiparallel beta-sheets (45-48). The structure of the BirA-BCCP from the *E. coli* complex has not been solved, but has been obtained for the *Pyrococcus horikoshii* system (49). A model of the *E. coli* complex was constructed with known structures of holoBirA monomer and apoBCCP87, using the PhBPL-PhBCCP complex as template (Fig. 5) (39, 45, 49, 50). While the PhBPL-PhBCCP association is believed to occur through main-chain interactions (49), the BirA-BCCP interaction is mediated, at least in part, by BirA surface loops (44).

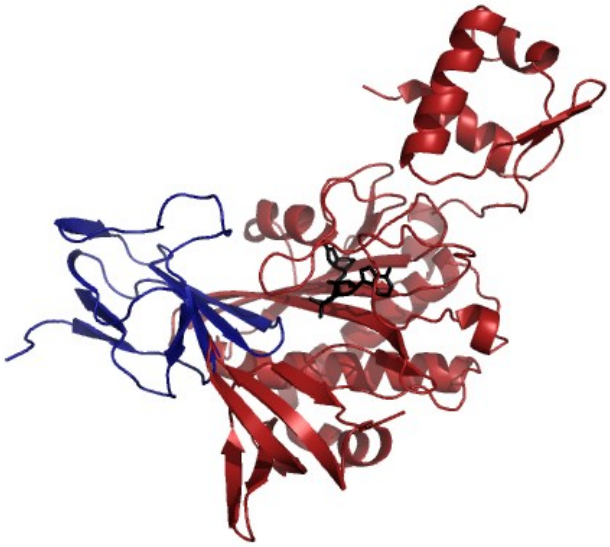


Figure 5. Model of the BirA-BCCP87 complex from *E. coli*. BirA is colored red, and BCCP87 is blue. The figure was constructed using PyMOL (40) with a file constructed by Zachary Wood comprising the PDB files 2EWN and 1BDO for BirA and BCCP87, respectively, using the PhBPL-PhBCCP file 2EJG as a template.

E. coli contains only one biotin-dependent carboxylase, ACC, which catalyzes the conversion of acetyl-CoA to malonyl-CoA in the first step of fatty acid synthesis (Fig. 6). The *E. coli* ACC is encoded by four independent peptides (51). The BCCP subunit contains the site of biotin attachment. The biotin carboxylase (BC) subunit catalyzes the carboxylation of biotin through the formation of a carboxyphosphate intermediate from bicarbonate and ATP. This intermediate is then attacked by free electrons on the N1 atom of biotin to form carboxybiotin (52). The carboxyltransferase (CT), which comprises two distantly related CT-α and CT-β subunits (53), transfers the carboxyl group to acetyl-CoA to form malonyl-CoA. The interactions between these subunits occur through a

"swinging-subunit" model, in which the BCCP subunit translocates during catalysis so biotin can interact with the BC and CT subunits (54).

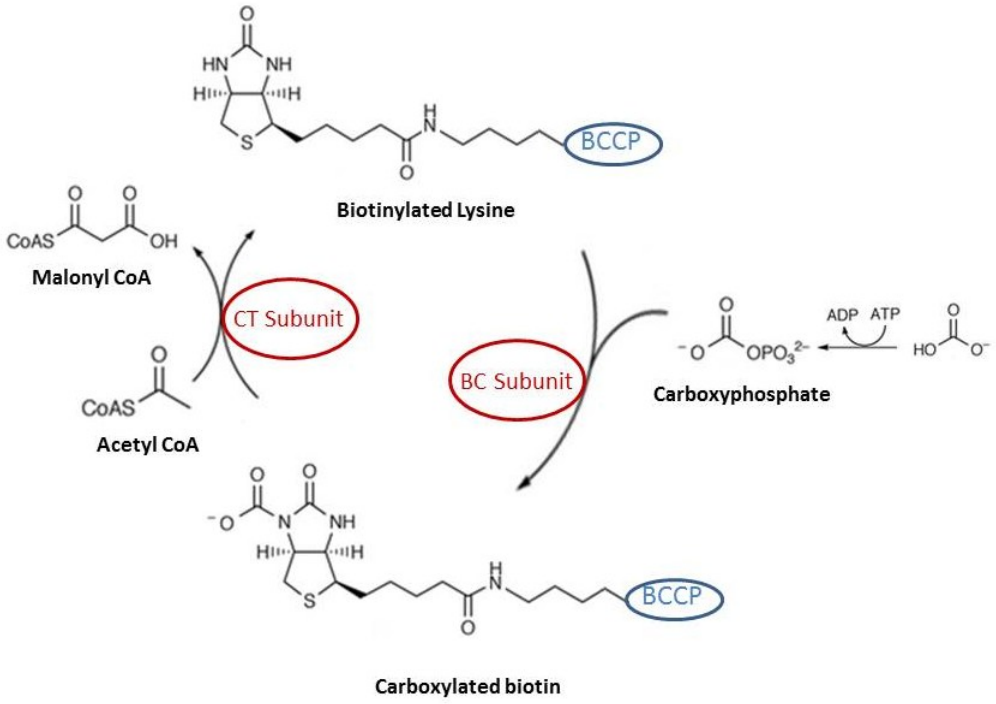


Figure 6. Reactions catalyzed by the subunits of acetyl-CoA carboxylase. Biotin is carboxylated by the BC subunit, and the carboxyl group is transferred by the CT subunit to acetyl CoA to form malonyl CoA (51, 52).

The reaction catalyzed by ACC is the first committed step in fatty acid biosynthesis and is the rate-limiting step of this process (55). Long-chain fatty acids synthesized from malonyl-CoA are used to produce cellular components, including the plasma membrane. All subunits of ACC are important for cell viability, as evidenced by the inhibition of cell growth due to mutations in BCCP and CT genes (56). ACC is the

target of assay development and drug design due to its integral role in metabolism and cell growth (57).

1.5. BirA as a transcription repressor

BirA transcription repression complex assembly involves corepressor, bio-5'-AMP, binding to apoBirA, followed by holoBirA dimerization and site-specific binding to the biotin biosynthetic operator, bioO (32, 33, 58, 59). The biotin operator consists of a 40-base pair inverted palindrome with two overlapping divergent promoters, P_a and P_b (Fig. 7a) (58, 60). A crystal structure of the BirA-bioO complex is not available, but DNaseI, hydroxyl radical and dimethyl sulfate footprinting have revealed contact regions between the protein and DNA (61, 62). The winged helix-turn-helix (wHTH) motif of the BirA N-terminal domain binds DNA through several interactions (Fig. 7b). This motif in each monomer interacts with the two 12-base pair termini of the operator site. The recognition helix of the HTH motif contacts the major groove of DNA, while the wing interacts and fits into the minor groove. Mutations in the recognition helix sequence of BirA result in loss of the DNA binding function of the repressor (29, 30, 63). The DNA binding domain is relatively mobile compared to the remainder of the molecule, while poor packing of the wing region lends flexibility to this domain (39). All these factors contribute to BirA's exceptional specificity in binding to the biotin operator sequence.

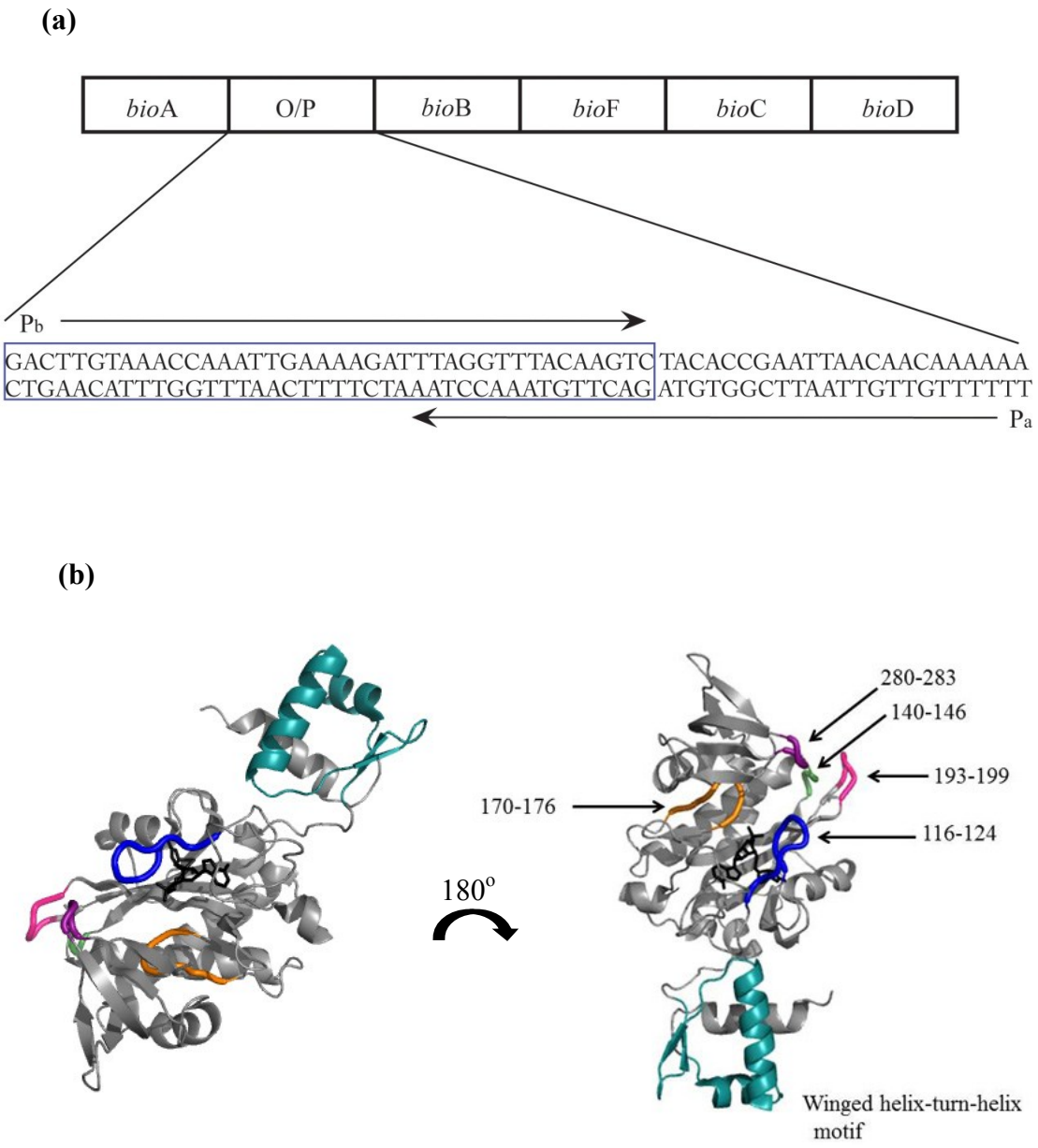


Figure 7. (a) Schematic representation of the *E. coli* biotin operon. The transcription regulatory region is expanded to show the sequence of the biotin operator (boxed region) and the two promoters, P_a and P_b. (b) The crystal structure of BirA bound to btnOH-AMP highlighting the winged helix-turn-helix motif in cyan. The left figure was rotated 180° to the right and rotated upside-down to obtain the figure on the right. The colored loops represent the flexible loops at the dimer interface. The figure was generated using PyMOL (40) with PDB file 2EWN.

The assembly of many transcription repression complexes, including BirA, involves coupling of protein-protein and protein-DNA interactions. Several factors contribute to the assembly of a repression complex, including the self-assembly of protein monomers, the binding of protein oligomers to DNA, and the structural rearrangement of the DNA, the protein or the protein-DNA complex to achieve optimal arrangement for interaction (64). The bacteriophage lambda cI repressor, for example, dimerizes and binds one or more of three lambda operator sites to control the switch between lysogenic and lytic phage cycles (65). The step-wise process of repression complex assembly involves lambda repressor dimerization followed by DNA binding, while further self-assembly of the lambda repressor on DNA can result in tetramers or octamers. The dissociation constant for dimer formation is 10^{-8} M, while 100-fold higher concentrations are required to form the tetramer and octamer (65). The higher order species are involved in cooperative binding to multiple lambda operator sites to mediate repression. In this manner, the lambda repressor complex participates in multiple protein-protein and protein-DNA interactions, resulting in different developmental states.

1.6 The BirA functional switch

Once holoBirA is formed, the enzyme can partition in one of two directions: it can dimerize and bind to the biotin biosynthetic operator, or it can interact with BCCP and transfer biotin. The structural, thermodynamic and kinetic bases for the switch have been extensively studied. Structurally, both protein-protein interactions utilize the same surface of the holoBirA monomer for interaction, rendering the two interactions mutually

exclusive. Several loops at the dimer interface have been found to be important in mediating the two interactions (Fig. 8) (43, 44).

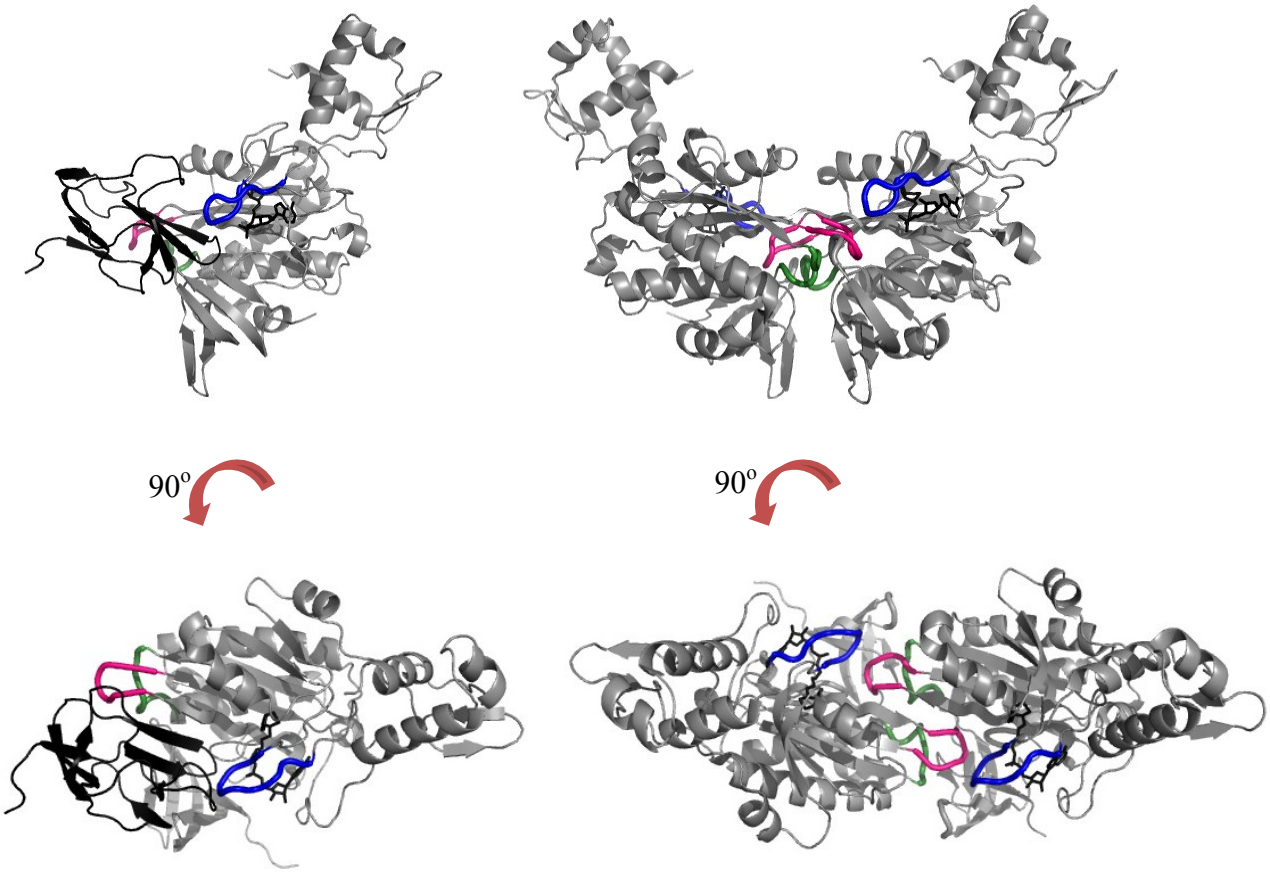


Figure 8. Models of the BirA-BCCP heterodimer (left) and BirA homodimer (right). The top figures were rotated 90° to display cross-sections of the interfaces. Loops that were previously studied are colored: blue, 116-124; green, 140-146; pink, 193-199. All models were created using PyMOL (40) with the PDB file 2EWN for the homodimer and a file constructed by Zachary Wood comprising the PDB files 2EWN and 1BDO for BirA and BCCP87, respectively, using the PhBPL-PhBCCP file 2EJG as a template.

The thermodynamic basis for the functional switch has been investigated. Homodimerization occurs with an equilibrium dissociation constant of 5-10 μM (ΔG = -7.0 kcal/mol) (59), while heterodimerization occurs at 2-3 μM (ΔG = -7.4 kcal/mol). Although the energetics are similar, several experimental results led to the conclusion that there were distinct chemistries behind these two interactions (44). HoloBirA dimerizes when bound to two bio-5'-AMP analogs, biotinol-5'-AMP and biotin sulfamoyl 5'-AMP, but apoBCCP does not interact with these ligand-bound BirA molecules (44, 66). Further, the structures of the two interfaces indicate that the homodimer forms more electrostatic interactions than the heterodimer. Investigation of the temperature and salt dependencies of these two interactions indicated that while homodimerization is characterized by a large unfavorable enthalpy (41 ±3 kcal/mol) and a modest dependence on salt, heterodimerization has a modest unfavorable enthalpy (7 ± 2 kcal/mol) and no salt dependency (44, 67). The calculated entropies (-TΔS) of homo and heterodimerization are -47 and -14 kcal/mol, respectively. Thus, although the entropic driving forces for both interactions are favorable, the larger entropy for homodimerization has been attributed to solvent release upon dimer formation (67). Since the two interactions are isoenergetic, kinetic control for the switch was proposed.

The kinetic basis for the switch has been investigated by measuring the effect of the presence of apoBCCP on the initial rate of holoBirA binding to bioO (68). In the presence of apoBCCP, holoBirA can interact with this substrate and transfer biotin to form holoBCCP and apoBirA, thus decreasing the pool of holoBirA that interacts with bioO. In the experiment, the extent of inhibition increased with increasing acceptor

protein concentration. Analysis of sedimentation velocity measurements to obtain the rates of homodimerization indicates slow kinetics for the monomer-dimer interconversion, with k_{on} and k_{off} values of 20 (± 8) $M^{-1}s^{-1}$ and 2.7 (±0.5) x 10^{-4} s^{-1}, respectively (69). These rates are considerably slower than the measured rate of heterodimerization of 14,700 (± 200) $M^{-1}s^{-1}$(46) which suggests that at sufficiently high apoBCCP concentrations, holoBirA partitions towards its metabolic role in a kinetically controlled manner.

1.7 Biotin protein ligases from other organisms

Biotin protein ligase is wide-spread in all kingdoms of life, with the exception of some bacterial species that do not contain biotin-dependent carboxylases, including *Buchnera* sp., *Borrelia burgdorferi, Aeropyrum pernix,* thermoplasmas, and mycoplasmas (70). The *birA* genes from many species of eubacteria and archaea were identified using a similarity search (70). This gene is widely distributed in these two domains of life, and based on gene sequence, two types of proteins could be identified, those that act as transcription regulators, and those that do not. Presence of an N-terminal DNA binding domain with a helix-turn-helix motif was used as the criterion to identify ligases with transcription regulatory functions. The DNA binding property of ligases is widely distributed in the *Bacillus/Clostridium* group, gamma-proteobacteria, and archaea. The extensive conservation in sequence among eubacteria and archaea makes this protein the only transcriptional regulator with a conserved binding signal. Since bifunctional ligases are found in both archaea and eubacteria, the oldest ancestor was believed to be bifunctional (71).

Many bacterial ligases and all eukaryotic ligases are monofunctional and contain no N-terminal DNA binding motif in their structure. The C-terminal domains of both mono- and bifunctional ligases are conserved. For example, the full-length enzymes from *E. coli* and humans show a 10.5% identity and 19% similarity in sequences (72). By contrast, alignment of the C-terminal domains only yields 20% identity and 36% similarity in sequence, emphasizing the conservation of the catalytic domain among ligases. Two of the five surface loops identified in BirA are exceptionally conserved among all ligases (Fig. 9). The residues in these two conserved loops are important for every BirA function including ligand binding, dimerization and biotin transfer. Due to high sequence conservation, similar roles in ligand binding and biotin transfer for these constant loops in other organisms are also anticipated.

	116-124	140-146	170-176	193-199	280-283
E. coli	GRGRRGRKWF	EQGPAA	RVKWPND	GKTGDAA	IGDK
S. typhi	GRGRRGRKWF	EQGPAA	RVKWPND	GKTGDAA	IGDK
B. subtilis	GRGRMSRVWH	DIPLQK	DIKWPND	AEEDRVR	TLNG
P. horikoshii	GHGRLNRKWE	KVPQKD	RIKWPND	GDK----	--DG
S. cerevisiae	GRGRGGNTWI	PVTNRN	RIKWPND	HFINNKY	DHGN
A. thaliana	GRGRTKNVWE	VVPLIQ	KIKWPND	TYRSKKF	DQVV
H. sapiens	GKGRGGNVWL	QLGQRI	RVKWPND	TLMGETF	SAEG
	*:** . *		:*****		

Figure 9. Multiple sequence alignment of biotin protein ligase from different species. The 116-124 and 170-176 loop sequences show high sequence identity, while the other loops do not. *Escherichia coli* and *Salmonella typhi* are gamma-proteobacteria, *Bacillus subtilis* belongs to the Bacillus/Clostridium group, *Pyrococcus horikoshii* is an archaea bacteria, *Saccharomyces cerevisiae*, *Arabidopsis thaliana* and *Homo sapiens* are eukaryotes. The sequence alignment was performed using ClustalW (72).

1.8 The experimental problem

BirA participates in two protein-protein interactions resulting in different biological outcomes. Although the structural basis for BirA multispecifity has been studied previously for a few loops, a comprehensive analysis of all surface loops in their roles in mediating the two interactions was required. Further, assembly of the transcription repression complex and the switch between the two protein functions was previously studied, but a complete study of the roles of surface loops in facilitating these processes was needed. This work focuses on investigating the roles of surface loops in various BirA functions. To this end, 18 amino acids from four surface loops were replaced with alanine and the variants were purified and characterized.

Chapter 2 focuses on the thermodynamic and kinetic characterization of the surface loop variants. The energetics of homodimerization for all variants were measured using sedimentation equilibrium. A range of free energies of dimerization, from -2.4 to -10.6 kcal/mol was observed for alanine-substituted variants across all loops. The kinetic parameters of the total reaction, comprising both bio-5'-AMP synthesis and subsequent biotin transfer to BCCP87 indicated that three variants in the conserved loop 170-176 were incapable of performing the entire reaction. To determine which step the variants were incapable of catalyzing, and to determine the bimolecular association rate constants for the BCCP interaction for all variants, stopped flow studies of the second half reaction were performed. Seven of the 18 variants were slower than wild-type BirA in biotin transfer, indicating that a subset of variants, primarily from the conserved loop, are more important for biotin transfer. Overall, the conserved loops are important for the essential

biotin transfer reaction, and all surface loops co-evolved to support the non-essential homodimerization reaction.

Characterization of the surface loop variants in the functional switch is described in Chapter 3. The effects of dimer interface mutations on the total assembly energetics of the repression complex were assessed by DNaseI footprint titrations. While a range of total assembly energetics was observed, the affinity of each dimer for DNA was very similar to wild-type holoBirA dimer, indicating that dimerization energetics determines total assembly energetics. Additionally, the impacts of surface loop alanine substitutions on the switch between homo- and heterodimerization were investigated using inhibition DNaseI footprint titrations. A direct correlation between inhibition of repression complex assembly with hetero-association rates was observed, reinforcing the mechanism of the switch as a kinetic partitioning between the two protein-protein interactions.

This work investigated several surface loops involved in mediating BirA multispecificity. These loops evolved to support the essential biotin transfer reaction as well as the non-essential homodimerization reaction. Moreover, the direct correlation observed between inhibition of repression complex assembly and rates of heterodimer association reinforces a mechanism of kinetic control for the switch between BirA functions. Combined, these results explain how BirA functions as a metabolic enzyme in the essential biotin transfer reaction while also regulating levels of biotin by acting as a transcription repressor in *E. coli*.

Chapter 2: Structural basis for BirA multispecificity

The work presented in this chapter was originally published in the Journal of Molecular Biology: Poorni R. Adikaram and Dorothy Beckett. Functional Versatility of a Single Protein Surface in Two Protein:Protein Interactions. Journal of Molecular Biology. 2012; 419, 223-233.

2.1 Abstract

The ability of the *Escherichia coli* protein BirA to function as both a metabolic enzyme and a transcription repressor relies on use of a single surface for two distinct protein:protein interactions. BirA forms a heterodimer with the biotin acceptor protein of acetyl-coenzyme A carboxylase and catalyzes post-translational biotinylation. Alternatively, it forms a homodimer that binds sequence-specifically to DNA to repress transcription initiation at the biotin biosynthetic operon. Several surface loops on BirA, two of which exhibit sequence conservation in all biotin protein ligases and the remainder of which are highly variable, are located at the two interfaces. The function of these loops in both homodimerization and biotin transfer was investigated by characterizing alanine-substituted variants at 18 positions of one constant and three variable loops. Sedimentation equilibrium measurements reveal that 11 of the substitutions, which are distributed throughout conserved and variable loops, significantly alter homodimerization energetics. By contrast, steady-state and single-turnover kinetic measurements indicate that biotin transfer to biotin carboxyl carrier protein is impacted by 7 substitutions, the majority of which are in the constant loop. Furthermore, constant loop residues that

function in biotin transfer also support homodimerization. The results reveal clues about the evolution of a single protein surface for use in two distinct functions.

2.2 Introduction

Multispecificity in protein:protein interactions provides a mechanism for integration of seemingly unrelated cellular processes. For example, through its ability to bind to many distinct protein partners, ubiquitin functions in protein turnover, DNA repair, cellular trafficking and chromatin remodeling (73). Although general mechanisms, including conformational flexibility and intrinsic disorder (19), for achieving multispecific protein binding have been identified, the evolution of multispecificity in protein interactions remains obscure. One pressing question is how multispecificity is achieved without severely impacting function.

The *Escherichia coli* biotin protein ligase BirA provides an ideal system for investigation of a multispecific protein binding surface (29, 30). BirA binds biotin and ATP to catalyze synthesis of bio-5`-AMP, and the resulting enzyme intermediate complex (holoBirA) partitions to one of two functions in response to cellular biotin demand (Fig. 10). During rapid growth, when demand is high, holoBirA transfers biotin to the biotin carboxyl carrier protein (BCCP) subunit of acetyl-coenzyme A (CoA) carboxylase, the enzyme that catalyzes the first committed step of fatty acid synthesis (35). ApoBCCP provides the direct signal for biotin demand (34) and only upon its depletion can holoBirA homodimerize and bind sequence-specifically to the biotin operator DNA, bioO, to repress transcription initiation at the biotin biosynthetic operon

(33). The enzyme intermediate, bio-5'-AMP, binding is required for tight homodimerization (59, 74). Although posttranslational biotinylation is essential for viability, only a subset of biotin protein ligases are also repressors (70).

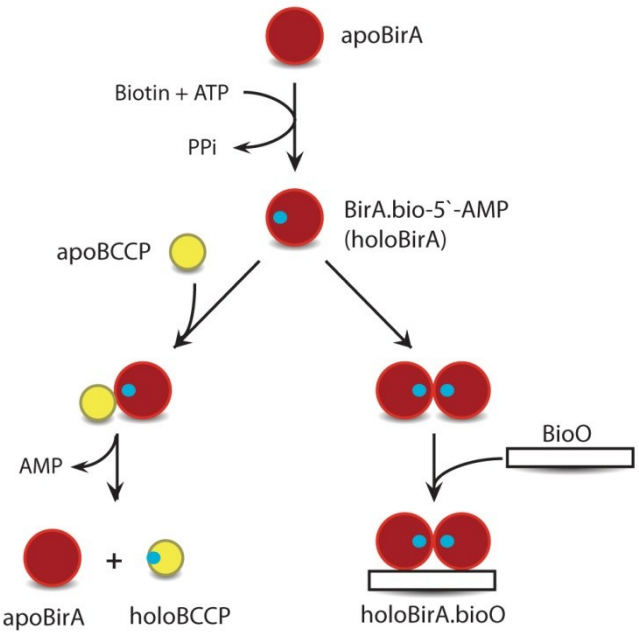
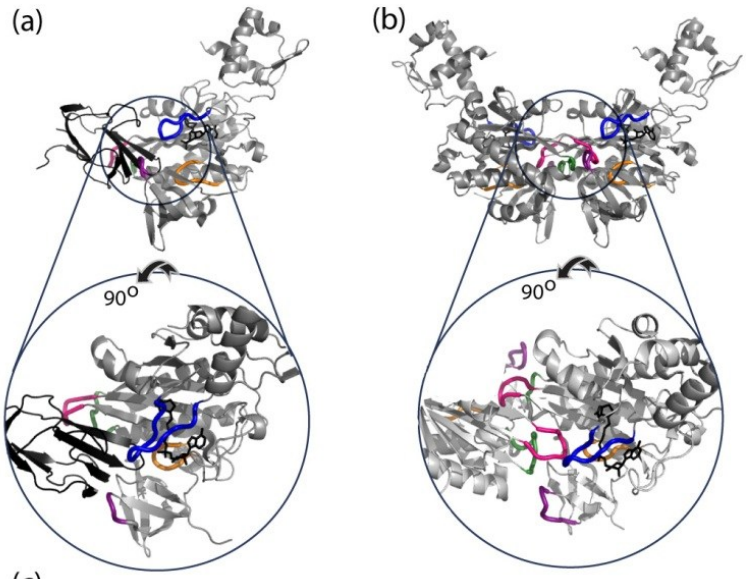


Figure 10. The *E. coli* biotin regulatory system. BirA catalyzes bio-5`-AMP synthesis from biotin and ATP, and the resulting holoBirA either channels biotin to metabolism by linking it to BCCP or homodimerizes and binds sequence-specifically to the biotin operator to repress transcription of the biotin biosynthetic operon.

A single surface on BirA is used for both heterodimerization and homodimerization (Fig. 11a and b) with both dimers forming by extension of a β-sheet in the BirA central domain (38, 39). Additionally, several BirA surface loops are located in the two interfaces. Previous studies reveal that substitution or deletion of residues in a subset of these loops results in both decreased repression *in vivo* and impaired DNA

binding and homodimerization *in vitro* (29, 30, 42, 43). A number of these sequence changes also affect the interaction with BCCP (44). Due to their limited scope, these results were insufficient to draw conclusions about the relationship between the structure of the bispecific BirA surface and its two functions.

Alignment of sequences of bifunctional bacterial ligases provides clues to the evolution of bispecific protein:protein interactions in BirA (Fig. 11c). The average sequence identity among these ligases ranges from 90 to 21%. However, alignment of several interface loop sequences reveals that they fall into two classes designated "conserved and variable." The sequences of conserved loops 116-124 and 170-176 are similar in biotin ligases from bacteria to humans, including monofunctional ligases, which either do not form homodimers or do so using surfaces distinct from that used by *Ec*BirA (49, 75, 76). Consequently, this conservation likely reflects the critical roles of these loops in heterodimerization and/or biotin transfer. This does not, however, preclude roles for some of the constant loop residues in homodimerization (29, 30, 42-44). By contrast, the variable loop sequences, comprising residues 140-146, 193-199, and 280-283, likely evolved to support homodimerization while not compromising the essential heterodimerization function. Furthermore, the sequence variability in these loops for the bifunctional ligases indicates degenerate solutions to evolution of homodimerization function.



(c)

	116-124	140-146	170-176	193-199	280-283
E. coli	RGRRGRKWF	EQGPAAA	RVKWPND	GKTGD--AA	IGD-K
S. typhi	RGRRGRKWF	EQGPAAA	RVKWPND	GKTGD--AA	IGD-K
N. europaea	RGRRGRSWS	QCPVNTL	ALKWPND	GDMLS--LG	FAEGS
M. flagellatus	RGRRGRSWQ	QLGASAL	QLKWPND	GDMEG--PS	FPDGR
B. subtilis	RGRMSRVWH	DIPLQKT	DIKWPND	AEEDR--VR	RTLNG
L. lactis	HGRRGRKFY	SHDILKI	KLKWVND	LDLESSSAS	KLGTS
T. thermophilus	RGRRGRPWE	GLPLSAL	GLKWPND	AEGEE--VS	HTGKG
	:** .* :		:** **		

Figure 11. Models of holoBirA-BCCP87 heterodimer (a) and holoBirA homodimer (b). Rotation of the molecules by 90° displays cross-sections of the heterodimer interface (bottom left) and homodimer interface (bottom right), highlighting the surface loops involved in the two interactions. All models were created using PyMOL(40) with input file 2EWN for the homodimer and a file constructed by Zachary Wood for the heterodimer. This model, which uses the experimentally determined structure of the *Pyrococcus horikoshii* biotin protein ligase:BCCP complex as a template (2EJG) (49), was constructed using the holoBirA monomer coordinates from 2EWN (39) and the coordinates for apoBCCP87 (1BDO) (45). (c) Multiple sequence alignment of several bifunctional microbial biotin protein ligases performed using ClustalW (40). The color codes for loop sequences correspond to the coloring of loops in the models in (a) and (b).

The hypothesized roles of four of the BirA surface loops, including three variable and one constant loop, in bispecific dimerization were investigated by measuring the functional consequences of single alanine substitutions for the two reactions. Sedimentation equilibrium measurements of alanine variants reveal that substitutions in any of the four loops can impact homodimerization and yield dimerization free energies spanning a range of 8.0 kcal/mol. By contrast, steady-state and single turnover kinetic measurements of biotin transfer indicate large perturbations primarily for alanine substitutions in the conserved loops. Additionally, the constant loop residues that function in biotin transfer to BCCP also support the homodimerization reaction. These results reveal a strategy for acquisition of additional protein interaction potential in a protein surface that is evolutionarily constrained to function in an essential metabolic reaction.

2.3 Experimental procedures

2.3.1 Chemicals and biochemicals

All chemicals used in buffer preparation were at least reagent grade. The isopropyl-β-D-thiogalactoside (IPTG), phenylmethylsulfonyl fluoride (PMSF), polyethyleneimine (PEI), unlabeled biotin and ATP were purchased from Sigma. The 1,4-dithio-DL-threitol (DTT) was purchased from Research Organics. Biotin [8, 9- ^{3}H] was purchased from Perkin Elmer and stored under argon at -70^{o}C. ATP was prepared by dissolving the disodium salt in water and adjusting the pH to 7.5 and its concentration was determined using UV absorbance with an extinction coefficient of 15,400 M cm^{-1} at

259 nm. The unlabeled biotin stock solution was prepared by dissolving biotin in 10 mM Tris HCl, 200 mM KCl, 2.5 mM $MgCl_2$ and adjusting the pH to 7.5 at 20^{o}C. The bio-5`-AMP was synthesized and purified as previously described (31, 58). The 87-amino acid C-terminal fragment (BCCP87) of BCCP from acetyl-CoA carboxylase was purified as previously described (46).

2.3.2 Mutagenesis and expression and purification of BirA variants

Site directed mutagenesis was performed using the QuikChange II XL Site-Directed Mutagenesis kit (Stratagene) using a pBtac2 (Boehringer Mannheim) derivative that contains the BirA coding sequence under transcriptional control of the tac promoter. A C-terminal $(His)_6$ tag, which has no effect on function, enabled purification of the variant proteins away from the chromosomally encoded wtBirA. After DpnI digestion of the mutagenized plasmid, the DNA was transformed into XL10-Gold cells (Stratagene), and several single colonies were screened for expression. The sequences were confirmed for the entire gene. Proteins were purified as previously described (77), with the exception that a final Q-Sepharose column step was added. The absorbance at 280 nm was used to calculate the concentration of each protein using the molar extinction coefficient of 47,510 $M^{-1}cm^{-1}$ calculated from the amino acid composition (78). Each protein was approximately 95% pure, as assessed by SDS-PAGE analysis, and all proteins were >90% active as determined by stoichiometric titrations with bio-5`-AMP monitored by fluorescence spectroscopy (58).

2.3.3 Sedimentation equilibrium measurements

Homodimerization of each protein bound to bio-5'-AMP was determined by sedimentation equilibrium using a Beckman Coulter Optima XL-1 analytical ultracentrifuge with a four-hole An-60 rotor. Standard 12-mm six-hole or 3-mm two-hole cells with charcoal-filled Epon centerpieces and sapphire windows were used in all experiments. Prior to centrifugation, proteins (~ 400 µl) were exhaustively dialyzed against assay buffer [10 mM Tris-HCl (pH 7.5 at 20°C), 200 mM KCl and 2.5 mM $MgCl_2$]. The bio-5`-AMP, which was diluted into the dialysis buffer immediately before sample preparation, was added to protein at a 1.5:1 molar ratio. Samples prepared at three protein concentrations were centrifuged at three rotor speeds (79) (ranging from 18,000 to 26,000 rpm) at 20°C. After reaching equilibrium at each speed, the absorbance at 295 or 300 nm was measured with a step size of 0.001 cm and five replicates per step.

2.3.4 Data analysis of sedimentation equilibrium measurements

The individual absorbance *versus* radius profiles were analyzed separately and globally using a single species model to obtain the reduced molecular weight, σ, using the program WinNonLin (80), from which the molecular weight was calculated using the following relationship:

$$\sigma = \frac{M(1-\bar{\nu}\rho)}{RT}\omega^2 \qquad (1)$$

where M is the molecular weight, $\bar{\nu}$ is the partial specific volume of the protein, ρ is the density of the buffer, ω is the angular velocity of the rotor, R is the gas constant and T is

the absolute temperature. The nine data sets were then globally analyzed using a monomer-dimer model (81) to obtain an association constant, K_a, using the equation:

$$c_t(r) = \delta + c_m(r_o)e^{\sigma_m(\frac{r^2-r_o^2}{2})} + K_a(c_m(r_o)^2)e^{2\sigma_m(\frac{r^2-r_o^2}{2})} \quad (2)$$

where c_t is the total concentration at position r, δ is the baseline offset, $c_m(r_o)$ is the concentration at the reference radial position r_o and σ_m is the reduced molecular weight of the monomer. The association constants, which are obtained in absorbance units from the analysis, are reported as molar equilibrium dissociation constants or K_{dim}. The quality of each analysis was assessed by the square root of variance of the fit and the distribution of residuals.

2.3.5 Steady-state kinetic measurements

Kinetics of biotin transfer as a function of BCCP87 concentration was measured by monitoring the incorporation of [^{3}H]-biotin into BCCP87. Reactions containing 5 μM cold biotin, 60 nM [^{3}H]biotin (specific activity, 32.6 Ci/mmol), 2.5 mM ATP, 0.1 mg/ml acetylated bovine serum albumin carrier protein (Invitrogen) and a range of apoBCCP87 concentrations were performed at 20^{o}C in assay buffer. After initiation by addition of BirA (wild type or variant) that had been preincubated at 20^{o}C to a final concentration of 100 nM (G193A, 50 nM), 20 μl volumes of the reaction were quenched at various time points by spotting onto Whatman filter paper that had been pretreated with a solution containing 1 mM biotin and 10% trichloroacetic acid (Alfa Aesar) in water and washed as previously described (56, 82). Each dried filter paper was transferred to a scintillation

vial containing Ready Protein^{+} cocktail (Beckman Coulter), and radioactivity was measured by scintillation counting for 10 min.

2.3.6 Data analysis of steady-state kinetic measurements

The initial velocity, obtained as the slope of the counts per minute *versus* time plots, *versus* BCCP87 concentration data were subjected to nonlinear least-squares analysis using the Michaelis-Menten equation in GraphPad Prism 4 (83) to obtain V_{max} and K_m. The V_{max} and k_{cat} values were converted from counts per minute to micromolar biotin using the relation between counts per minute and micromoles obtained from a control reaction performed at conditions in which the [^{3}H]biotin is quantitatively incorporated into BCCP87.

2.3.7 Initial rate measurements of biotin transfer from the intermediate bio-5'-AMP

The rates of BirA-catalyzed biotin transfer from bio-5`-AMP to the BCCP87 acceptor substrate were measured using a Kintek SF-2001 stopped-flow instrument equipped with fluorescence detection. A solution of the enzyme, wild type or alanine variant, at a concentration of 1μM was first preincubated for at least 15 min with either 800 nM biotin and 500 μM ATP or 800 nM bio-5`-AMP at 20$^{o}$C. The resulting enzyme·intermediate complex was rapidly mixed at a 1:1 vol:vol ratio with BCCP87 to achieve final concentrations of 400 nM and 10-90 μM for BirA·bio-5'-AMP and apoBCCP87, respectively. Bio-5`-AMP was used for variants K172A, N175A and D176A, which are incapable of bio-5`-AMP synthesis. The excitation wavelength was

295 nm, and emission was measured above 340 nm using a cutoff filter (Corion Corp.). At least four traces were obtained at each BCCP87 concentration.

2.3.8 Data analysis of single-turnover kinetic measurements

The fluorescence intensity *versus* time traces, obtained for each variant at a range of BCCP87 concentrations, were analyzed using either a single-exponential or a double-exponential model to obtain apparent rate(s). Further analysis of the rate *versus* concentration profiles was performed as described in Results.

2.4 Results

2.4.1 Homodimerization is sensitive to alanine substitution in variable and constant loops

The effects of alanine substitutions in the variable loops 140-146, 193-199 and 280-283 and constant loop 170-176 on homodimerization were measured using equilibrium analytical ultracentrifugation. The conserved loop comprising residues 116-124 was not included in the present work because previously published data indicate roles for this loop in both homodimerization and the interaction with BCCP (42-44). The total number of variants in the four remaining loops was limited to 18 either because only non-alanine residues were replaced or the protein could not be purified. Since stable dimerization requires binding of the allosteric effector, bio-5'-AMP, all measurements were performed on complexes of proteins bound to the adenylate. Saturation was ensured by preparing solutions at a 1.5:1 molar excess of bio-5'-AMP:protein in the micromolar range of concentration, at which adenylate binding is stoichiometric. Samples prepared at

three protein concentrations were centrifuged at three rotor speeds (Fig. 12), and global analysis of the resulting data using a monomer-dimer model provided the equilibrium dimerization constant. For each measurement, the agreement between the data and the model was assessed from the distributions of the residuals of the fits and the magnitude of the square root of the variance of the fit (Fig. 12).

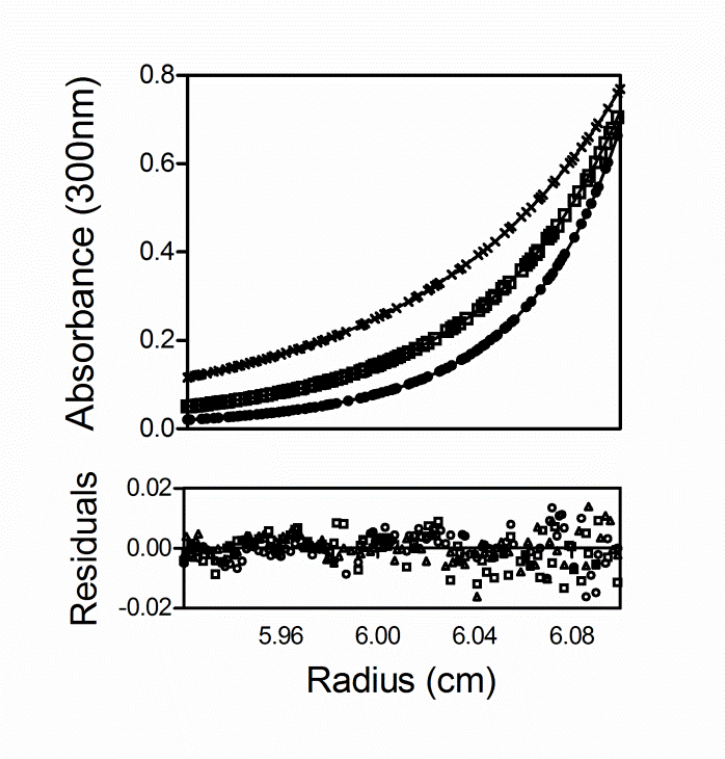


Figure 12. Absorbance *versus* radius profiles of E140A-bio-5`-AMP obtained at 40 μM protein concentration and centrifuged at 18,000 (×), 22,000 (□) and 26,000 (○) rpm. The lines represent the best-fit curves obtained from global analysis of three data sets to a monomer-dimer model. The residuals of the fits are shown in the lower panel.

Dimerization measurements for all alanine-substituted variants in the four loops indicate a broad range of equilibrium constants (Table 1). With the exception of the G142A, D176A, I280A, K172A and K194A variants, the sedimentation equilibrium data for the proteins are well described by a monomer-dimer model. The equilibrium constants for the reaction range from $17x10^{-3}$ to $1x10^{-8}$ M (Table 1) corresponding to Gibbs free energies of -2.5 to -10.6 kcal/mol (Fig. 13a). For the G142A, K172A, D176A, and I280A variants, which exhibit weak dimerization (Table 1, Fig. 13a), the concentration *versus* radius profiles yielded only the monomer at the highest loading concentration of 150 μM. Hence, the reported dimerization constants are lower limits. The K194A variant shows the tightest dimerization with a K_{dim} of approximately 10 nM, which corresponds to an energetic enhancement to dimerization of approximately -3.5 kcal/mol. The large errors in the equilibrium constant and free energy of dimerization for this variant (Table 1) reflect the necessity, due to the detection limits of the instrument's absorption optics, of using protein concentrations at which the monomer population was low. Therefore, the reported equilibrium constant and free energy are upper limits.

Table 1. Equilibrium parameters for dimerization of BirA variants.

BirA variant	K_{dim} (M)[a,b]	ΔG^o_{dim} (kcal/mol)[c]
Wild Type	**7 (± 3) x 10^{-6}**	**-6.8 (± 0.3)**
E140A	**22 (± 5) x 10^{-6}**	**-6.2 (± 0.2)**
Q141A	**3 (± 2) x 10^{-6}**	**-7.5 (± 0.3)**
G142A	**17 (± 52) x 10^{-3}**	**-2.4 (± 0.9)**
P143A	**4 (± 1) x 10^{-5}**	**-5.9 (± 0.2)**
R170A	**4 (± 1) x 10^{-5}**	**-6.0 (± 0.2)**
V171A	**3 (± 2) x 10^{-7}**	**-8.8 (± 0.6)**
K172A	**7 (± 4) x 10^{-4}**	**-4.2 (± 0.5)**
N175A	**6 (± 2) x 10^{-5}**	**-5.7 (± 0.2)**
D176A	**6 (± 6) x 10^{-3}**	**-3.0 (± 0.4)**
G193A	**5 (± 4) x 10^{-7}**	**-8.4 (± 0.5)**
K194A	**1 (± 7) x 10^{-8}**	**-10.6 (± 2.4)**
T195A	**10 (± 4) x 10^{-5}**	**-5.3 (± 0.2)**
G196A	**5 (± 4) x 10^{-6}**	**-7.1 (± 0.4)**
D197A	**5 (± 3) x 10^{-4}**	**-4.5 (± 0.3)**
I280A	**4 (± 3) x 10^{-3}**	**-3.3 (± 0.7)**
G281A	**9 (± 4) x 10^{-5}**	**-5.4 (± 0.3)**
D282A	**2 (± 2) x 10^{-6}**	**-7.8 (± 0.6)**
K283A	**12 (± 7) x 10^{-6}**	**-6.6 (± 0.4)**

The equilibrium constants were measured in standard buffer [10 mM Tris-HCl (pH 7.50 ± 0.02 at 20°C), 200 mM KCl and 2.5 mM $MgCl_2$] as described in Materials and Methods. In all cases, the protein was saturated with bio-5'-AMP. The Gibbs free energies were calculated using the equation $\Delta G^o = RT\ln K_{dim}$

[a] The errors represent the standard error of two independent experiments

[b] The K_{dim} values reported for G142A, K172A, D176A and I280A are lower limits and that for K194A is an upper limit.

[c] The reported uncertainties for each variant were calculated using standard error propagation methods.

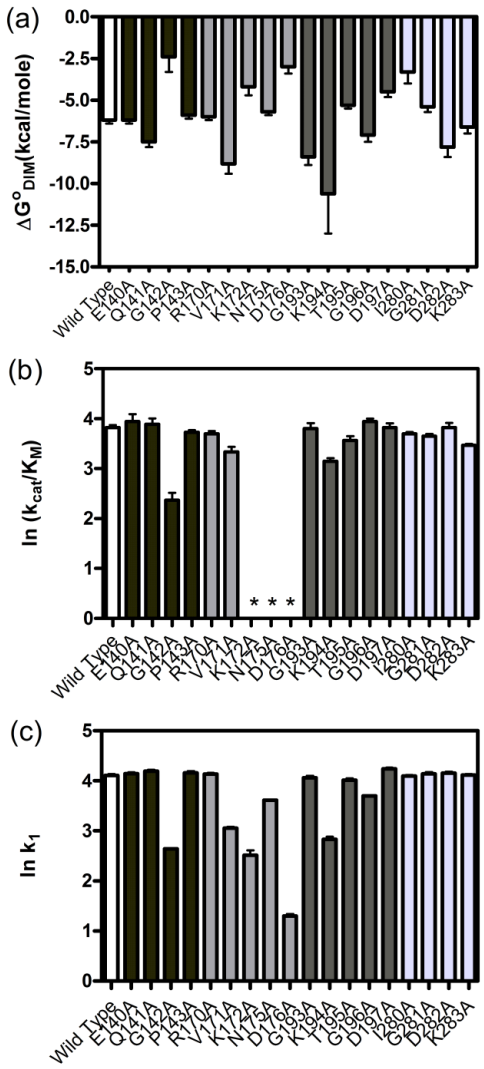


Figure 13. Homodimerization energetics and kinetic parameters for biotin transfer to BCCP. (a) Gibbs free energies of homodimerization. (b) The k_{cat}/K_m values for biotin transfer to BCCP87 obtained from steady-state measurements. *, not determined. (c) The bimolecular rates of association, k_1, of variants with BCCP87 obtained from single-turnover kinetic measurements. The kinetic parameters obtained for (b) and (c) are shown on a natural logarithmic scale to place them on an energy scale. Error bars for each graph represent the 68% confidence intervals. The bar color coding in each panel is wtBirA (white) and variable loop 140-146 (black), constant loop 170-176 (medium gray), variable loop 193-199 (dark gray) and variable loop 280-283 (light gray).

2.4.2 Two-step biotin transfer to BCCP is sensitive to alanine substitutions in primarily the constant loops

BirA-catalyzed biotin transfer is a two-step reaction in which the adenylated intermediate, bio-5'-AMP, is first synthesized from biotin and ATP, and the biotin is then transferred from the adenylate to BCCP. Since any alanine substitution in the loops can, in principle, affect either or both of these steps, two assays, one of the overall reaction and the second of biotin transfer from the intermediate alone, were employed to evaluate the variants. The impact of each substitution was first measured using steady-state kinetic measurements of BirA-catalyzed [^{3}H]biotin incorporation into a C-terminal fragment of the *E. coli* acetyl-CoA carboxylase BCCP, BCCP87, which is functionally equivalent to the full-length BCCP in BirA-catalyzed biotin transfer (46). In the assay, [^{3}H]biotin incorporation into the acceptor protein is monitored as a function of time in reactions in which ATP and total biotin (labeled plus unlabeled) concentrations were constant and in which BCCP87 concentration was varied. Biotin and ATP concentration were at saturating and K_m, respectively (84). For each variant, the dependence of the initial rate, obtained from the linear portion of the product versus time curves (Fig. 14a, inset), on BCCP87 concentration was analyzed using the Michaelis-Menten formalism to obtain the values of K_m and k_{cat} (Fig. 14a). The parameters obtained for the variant D197A are 12±2 μM and 0.08±0.01 s^{-1} for K_m and k_{cat}, respectively, identical with those obtained for wtBirA of 9±2 μM and 0.060±0.006 s^{-1}, respectively.

Figure 14. Kinetics of biotin transfer to BCCP. (a) Steady-state kinetic analysis of the two-step BirA-catalyzed biotin transfer reaction. Inset: initial rates of biotin incorporated at 5 μM (■), 25 μM (▲), 75 μM (▼) and 100 μM (♦) BCCP concentrations. For clarity, only four of seven lines are shown. Each data point is an average of two independent measurements with error bars representing the standard error. Initial rates *versus* substrate concentration were subjected to nonlinear least-squares analysis using the Michaelis-Menten model. Each initial rate represents the average of values measured in two independent experiments, with error bars representing the standard error. (b) Stopped-flow measurements of biotin transfer to BCCP in the second half-reaction. The fluorescence intensity *versus* time trace for D197A at 90 μM BCCP is shown with residuals for a single-exponential model (middle panel) or a double-exponential model (lower panel). (c) The dependence of the apparent rate of biotin transfer on apoBCCP concentration. The plots obtained for wtBirA (■), T195A (▲) and V171A (▼) are shown. Each data point is the average of rates obtained from at least four stopped-flow traces, and the lines represent the best fits of the rate *versus* concentration profiles to a linear equation.

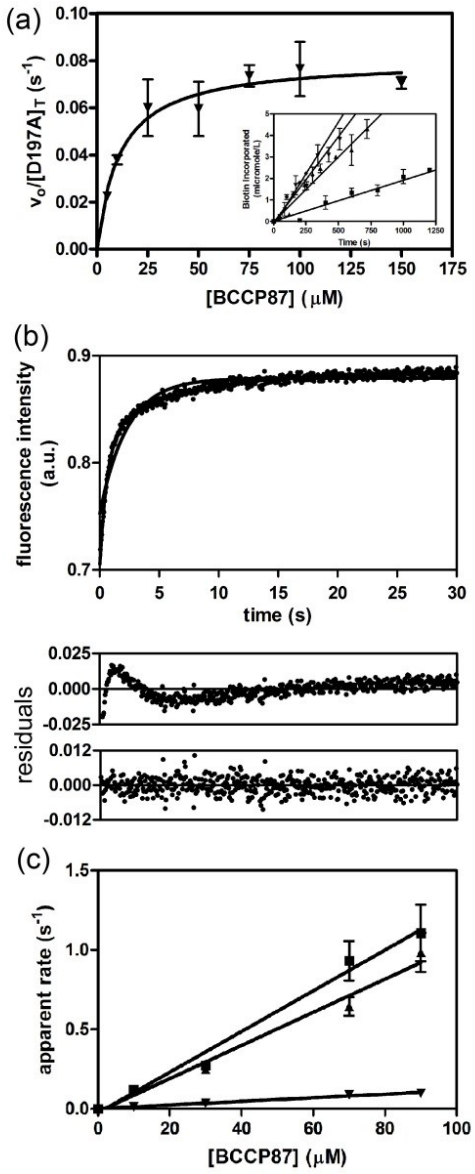


Figure 14. (Description on previous page)

Steady-state measurements of biotin transfer catalyzed by the entire set of variants indicate relative insensitivity of the variable loop sequences to amino acid replacement (Table 2 and Fig. 13b). By contrast, alanine substitutions in the 170-176 loop yielded three proteins, K172A, N175A and D176A, that have no activity in catalyzing biotin transfer to BCCP in the steady-state assay. Additional measurements indicate that these

proteins can, when provided with synthetic intermediate, bio-5'-AMP, catalyze biotin transfer to BCCP87 (see below). Among the alanine-substituted variants in the variable loops, only two, K194A and G142A, yielded K_m values that differ significantly from that measured for the parent wild-type protein. The perturbations to K_m for the two proteins are roughly 5-fold and 20-fold, respectively. The k_{cat} values for biotin transfer catalyzed by all variable loop variants are identical to that measured for wtBirA.

Effects of alanine substitutions on the heterodimer interaction were also assessed using single-turnover measurements of the second half-reaction. This assay allows separation of effects on the interaction of the enzyme intermediate complex with BCCP in the second half-reaction from those on the overall reaction, which is particularly important for this system because the rate-determining step in the overall reaction corresponds to the first half-reaction, bio-5'-AMP synthesis (46, 84, 85). In these measurements, preformed BirA-adenylate is rapidly mixed with apoBCCP87, and the resulting time-dependent increase in the intrinsic BirA fluorescence signal is monitored (Fig. 14b) (46). This increase, which accompanies the conversion of adenylate-bound BirA to apoBirA that occurs upon biotin transfer, has been shown to parallel product, bioBCCP product, bioBCCP, accumulation (46). Analysis of each transient was performed using either a single-exponential or a double-exponential model (Fig. 14b). For variants that exhibit double-exponential behavior, the apparent rate of the first, faster phase increases with acceptor protein concentration and that of the second phase is substrate concentration independent (46). Apparent rate *versus* BCCP concentration data show a linear dependence with no leveling off (Fig. 14c). The simplest interpretation of

this behavior is that it reflects collision of the enzyme intermediate complex with apoBCCP, which, in the kinetic scheme shown below, is governed by the bimolecular rate constant k_1:

$$\text{BirA·bio-5'-AMP} + \text{BCCP} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \text{BirA·bio-5'-AMP·BCCP} \xrightarrow{k_2} \text{BirA} + \text{bioBCCP} + \text{AMP} \qquad (3)$$

The slope obtained from linear regression of the rate *versus* acceptor protein concentration profiles yields the bimolecular association constant governing ternary complex formation, k_1 (Fig. 14c), from BirA·bio-5'-AMP and BCCP (85). The concentration-independent phase, which occurs with a rate of approximately 0.2 s^{-1}, is assigned to the product dissociation step governed by k_2 (46, 85).

Rate *versus* concentration profiles for all 18 variants indicate that the majority of the proteins behave identically to wild type BirA (Table 2, Fig. 13c). Of the seven variants that deviate, four are proteins with substitutions in the constant loop. Only three variants with substitutions in variable loops including G142A, K194A and G196A differ from wtBirA in biotin transfer from the adenylate. All variants exhibit the same linear dependence of apparent rate on substrate concentration that is observed for wild type BirA. However, the magnitudes of k_1 values obtained for these proteins are 2.5-600 fold slower than that of the wild type protein (Table 2), indicating that they are compromised in the rate at which they associate with apoBCCP. Moreover, with the exception of G196A, for variants that could be assayed using both steady-state and single turnover methods, decreases in k_1 correlate with increases in K_M., consistent with the Michaelis constant reflecting, in part, binding of the enzyme intermediate complex to the acceptor protein.

Table 2. The steady-state and single-turnover kinetic parameters for BirA-catalyzed biotin transfer to BCCP87

BirA variant	K_m (BCCP) (μM)[a]	k_{cat} (s^{-1})[a]	k_{cat}/K_m ($M^{-1}s^{-1}$)[b]	k_1 ($M^{-1}s^{-1}$)[c]
Wild Type	**9 ± 1**	**0.060 ± 0.003**	**6700 ± 800**	**12,800 ± 800**
E140A	**8 ± 3**	**0.07 ± 0.01**	**8800 ± 3500**	**13,900 ± 800**
Q141A	**9 ± 1**	**0.07 ± 0.02**	**7800 ± 2400**	**16,000 ± 1000**
G142A	**214 ± 9**	**0.05 ± 0.02**	**230 ± 90**	**440 ± 10**
P143A	**13 ± 1**	**0.070 ± 0.005**	**5400 ± 600**	**14,500 ± 1200**
R170A	**16 ± 2**	**0.080 ± 0.002**	**5000 ± 600**	**13,700 ± 800**
V171A	**23 ± 6**	**0.050 ± 0.003**	**2200 ± 600**	**1100 ± 100**
K172A	**N.D.[d]**	**N.D.**	**N.D.**	**320 ± 50**
N175A	**N.D.**	**N.D.**	**N.D.**	**4100 ± 100**
D176A	**N.D.**	**N.D.**	**N.D.**	**20 ± 2**
G193A	**11 ± 3**	**0.070 ± 0.005**	**6400 ± 1800**	**12,000 ± 1000**
K194A	**64 ± 9**	**0.090 ± 0.005**	**1400 ± 200**	**700 ± 100**
T195A	**19 ± 3**	**0.07 ± 0.01**	**3700 ± 800**	**10,400 ± 800**
G196A	**8 ± 1**	**0.070 ± 0.005**	**8700 ± 1300**	**5000 ± 100**
D197A	**12 ± 2**	**0.08 ± 0.01**	**6700 ± 1400**	**17,000 ± 1000**
I280A	**12 ± 1**	**0.0600 ± 0.0003**	**5000 ± 400**	**12,500 ± 600**
G281A	**20 ± 2**	**0.09000 ± 0.00005**	**4500 ± 500**	**14,000 ± 1000**
D282A	**9 ± 2**	**0.060 ± 0.003**	**6700 ± 1500**	**14,400 ± 800**
K283A	**17 ± 1**	**0.050 ± 0.001**	**2900 ± 200**	**10,100 ± 500**

The k_{cat}, K_m and k_1 values were measured as described in "Materials and Methods" in standard buffer [10 mM Tris-HCl (pH 7.50 ± 0.02 at 20°C), 200 mM KCl and 2.5 mM $MgCl_2$].

[a] The errors represent the standard error of two independent experiments.

[b] The reported uncertainties were calculated using standard error propagation.

[c] The errors represent the uncertainties in the linear regression of the k_{obs} *versus* [apoBCCP87] profiles.

[d] N.D., not determined because these variants were incapable of catalyzing bio-5`-AMP synthesis

2.5 Discussion

The roles of BirA loop sequences in its two functions were assessed using combined sedimentation equilibrium and kinetic analysis of alanine substituted proteins. Consistent with the hypothesis that the variable loop sequences support the homodimerization required for transcription repression, alanine substitutions in the 140-146, 193-199, and 280-283 loops have large energetic consequences for homodimerization but only minor impact on heterodimerization. Contrary to the hypothesized role of constant loops in biotin transfer only, the 170-176 loop sequence function in both interactions.

Previous studies of the thermodynamics of BirA homodimerization and heterodimerization indicate that although the Gibbs free energies measured for the two interactions at 20^{o}C are similar, the underlying enthalpic and entropic contributions are very different (44), consistent with distinct structural determinants for the two interactions. The measured effects of alanine substitutions in the four surface loops on homodimerization and biotin transfer allow mapping of these structural differences (Fig. 15). Highlighting of loop residues at which alanine substitution results in large perturbations to the homointeraction or heterointeraction indicates distinct gross features of the two interfaces. Data obtained previously on the 116-124 loop, which functions in both homodimerization and biotin transfer, are included in this analysis (42-44). The homodimer interface consists of two dyad symmetric sub-interfaces, each of which is characterized by continuous interaction between the variable loop residues of one monomer with the constant loop residues of the second. By contrast, the heterodimer

interface is bipartite with the constant loops, which constitute the enzyme active site forming one locus of interaction and the 193-199 loop forming the second.

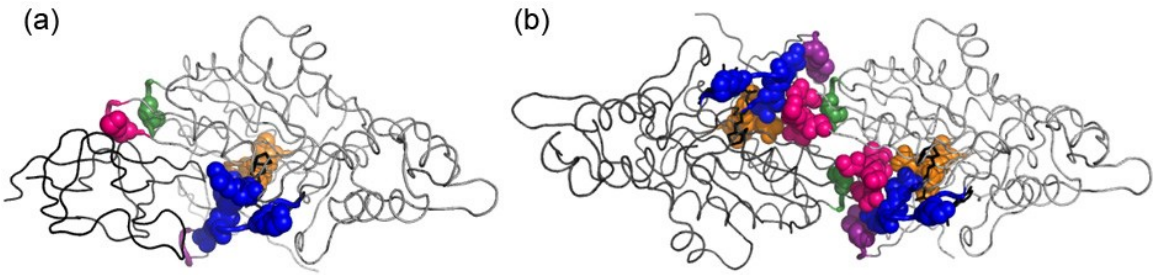


Figure 15. The BirA heterodimer (a) and homodimer (b) interfaces with residues that exhibit sensitivity to alanine substitution (with k_1 rates at least 10 times less than wtBirA, or the $\Delta\Delta G_{dim} \geq$ 1kcal/mol) highlighted. The color coding is identical with that used for the loops in Fig. 11. The data for the 116-124 loop (in blue) were obtained from previously published results (42-44).

Detailed structural interpretation of the energetic consequences of alanine substitutions for homodimerization is complicated by the known inherent flexibility of variable loops 140-146 and 193-199 and of the constant loop 116-124, which are disordered in the apoBirA monomer structure (36). Thus, alanine substitutions can potentially result in altered conformations of any of these loops. Furthermore, for several variants, the functional effects may reflect a disruption of intermolecular and/or intramolecular interactions. For example, replacement of V171 with alanine renders dimerization more favorable by -2 kcal/mol. In the wild-type BirA dimer structure, the V171 side chain is packed in the interior of each protein monomer far from the interface. Disruption of another apparently purely intramolecular interaction between I280 and

W310 leads to a 3.5 kcal/mole penalty for dimerization when the isoleucine is replaced by alanine. The D197 side chain interacts electrostatically with R119 in an intermolecular interaction, and substitution of D with A results in an energetic penalty to dimerization of 2.3 kcal/mol. The stabilizing effect of the K194 replacement reflects the relief of charge repulsion between the positively charged lysine and multiple arginine side chains in the 116-124 loop of the opposing monomer. Relative to homodimerization, alanine replacement of a small number of residues impacts biotin transfer. The inability of constant loop variants K172A, N175A and D176A to synthesize bio-5`-AMP and their slow rates of biotin transfer to BCCP likely reflects roles for these side chains in positioning of biotin and ATP in the active site, as well as associating with BCCP to transfer the biotin from the adenylate. Interpretation of the slow rate of biotin transfer observed for the K194A variant is complicated by the fact that it exists as a dimer, which competes with apoBCCP binding, at the 500nM concentration required for detection in the stopped-flow fluorescence measurements. Nevertheless, in the holoBirA·BCCP model, the distance between the epsilon amino group of K194 on BirA and the carboxylate of E128 of BCCP is consistent with formation of a salt bridge. This interaction may, in part, form the basis of the selectivity that is observed in the biotin transfer reaction (85).

The results reported in this work when combined with those of previous studies suggest a model for how the surface loops on BirA support bispecific protein:protein interactions. First, structural information in constant loops is multifunctional. Alanine substitutions in the 170-176 loop indicate perturbations to bio-5'-AMP synthesis, biotin

transfer to the acceptor protein and homodimerization. Previous studies of the second constant loop comprising residues 116-124 indicate similar complexity in its function with variants exhibiting defects in biotin and bio-5'-AMP binding, heterodimerization and homodimerization (42-44). The sequence conservation in these loops reflects their critical roles in the multistep posttranslational biotin addition reaction that is required for viability. Second, in the homodimer interface, the variable loop sequences complement the constant loops. This is illustrated in Fig. 15, which shows that, in the homodimer interface, the majority of the variable loop side chains that are functionally sensitive to alanine substitution contact constant loop residues. Thus, it is the complementarity of variable loop side chains to the constrained residues of the constant loops that enables use of a single surface on BirA for two distinct protein:protein interactions.

The range of homodimerization energetics measured for the loop variants illustrates the ease with which the interaction can be significantly altered through minor structural changes. These large energetic changes reflect, in part, the 2-fold symmetry of the homodimer interface, with each alanine substitution perturbing at least two interactions. The large impact on homodimerization coupled with the absence of impact of several amino acid changes in the variable loops on heterodimerization/biotin transfer prompts consideration of why wild-type BirA did not evolve to homodimerize more tightly. From the standpoint of transcription repression alone, tighter dimerization would lead to energetically more favorable assembly of the BirA·bioO complex and resulting greater repression. It is possible that, if homodimerization were too tight, heterodimerization could not compete, with resulting deleterious consequences for

viability. Comparison of the results of steady state and single-turnover kinetic measurements for the K194A variant supports this possibility. The lack of an effect on k_{cat} in the steady-state measurements reflects the experimental design in which reactions are initiated by enzyme addition. Upon bio-5'-AMP synthesis, the significantly higher apoBCCP concentration ensures that the enzyme intermediate complex collides with acceptor protein rather than itself. By contrast, in the single-turnover measurements preincubation of the K194A variant in its complex with the adenylate results in significantly slower biotin transfer to BCCP than is observed for the wild-type enzyme. This slower transfer may be limited by homodimer dissociation. An alternative explanation for the relatively weak homodimerization is that very tight dimerization would render the repression sensitive to biotin concentrations that are insufficient to support viability. Assembly of the repression complex occurs *via* coupled dimerization of the adenylate-bound monomer and binding of the preformed dimer to bioO (33). Simulations of the dependence of transcription repression on biotin concentration indicate that as the equilibrium dimerization constant for holoBirA becomes tighter, repression occurs at lower biotin concentrations (86). These predictions are currently being tested.

In general, protein surface loops can serve important roles in mediating interactions with other proteins. Structure-based sequence alignments of common domains have led to the proposal that surface loops evolve to either allow (enable) or prevent (disable) an interaction (20). A surface loop on the *E. coli* guanylate kinase has been experimentally demonstrated to promote its oligomerization (87). In some cases loop length, and not sequence, is the most important property for self-association (88).

Alternatively, as observed in BirA, the particular loop sequence is important for tuning dimerization energetics (89). However, the variability in 140-146, 193-199 and 280-283 loop sequences among the bifunctional ligases suggests that many sequences can support homodimerization. This apparent tolerance of homodimerization to sequence variation, which may be facilitated by the intrinsic structural flexibility of the interface loops, is intriguing in light of the complementarity between variable and constant loop side chains in the *Ec*BirA homodimer interface (Fig. 15).

This work demonstrates that variable surface loops evolved to support BirA function in the homodimerization required for transcription repression by complementing the constant loops, which are constrained because of their essential role in posttranslational biotin addition. The results further illustrate that relatively subtle changes in surface loop sequence can elicit large energetic changes in protein:protein interactions. The observations have implications for both evolution of protein:protein interactions and the design of novel interactions.

Chapter 3: Protein-protein interactions in the BirA functional switch

3.1 Abstract

Many multifunctional proteins switch between different functions in response to cellular signals, but the mechanism of control of the switch is not known in most cases. The biotin protein ligase from *Escherichia coli*, BirA, is a bifunctional enzyme involved in biotin homeostasis. It transfers biotin onto an acetyl-CoA carboxylase subunit, biotin carboxyl carrier protein (BCCP), in its role as a metabolic enzyme, or dimerizes and binds to the biotin biosynthetic operator in its role as a transcription repressor. BirA forms two mutually exclusive protein-protein interactions using the same protein surface. Previous experiments revealed that the interaction surface contains several loops, all of which are important for homodimerization, and only a few of which are impacted in their rates of association with BCCP. In this work, the role of surface loops in transcription repression complex assembly and in the switch between the two interactions was investigated using DNaseI footprint titrations. Direct titrations reveal that the energetics of repression complex assembly is determined by homodimerization energetics. Inhibition footprint titrations reveal a direct correlation between inhibition of repression complex assembly and rates of heterodimer association. These results emphasize the central role that protein:protein interactions play in regulatory complex assembly and reinforce a kinetic control mechanism for the BirA functional switch.

3.2 Introduction

Many proteins are known to carry out more than one cellular function. This multifunctionality can provide a mechanism for linking seemingly distinct cellular processes such as gene transcription and metabolism. The ability of a protein to switch functions frequently reflects swapping of one protein partner for another. For example, under normal growth conditions the σ-like (SL) domain of the general stress response protein PhyR of *Caulobacter crescentus* interacts intramolecularly with the PhyR receiver domain (90). In stressful conditions the receiver domain is phosphorylated, leading to its dissociation from the SL domain, and binding of PhyR to NepR, an anti σ^{EcfG} protein. Swapping of these interactions results in transcription activation of stress response genes. Understanding the mechanisms of these biological switches requires elucidation of the structural, kinetic and thermodynamic rules that govern switching of protein partners.

The *Escherichia coli* biotin protein ligase, BirA, provides an example of a protein that undergoes a functional switch in response to cellular metabolic state (29, 30). BirA binds biotin and ATP to synthesize bio-5'-AMP forming holoBirA (Fig. 16). This complex either partitions towards biotin transfer to the biotin carboxyl carrier protein (BCCP) subunit of acetyl-coA carboxylase (ACC), which catalyzes the first committed step of fatty acid synthesis, or it associates with another holoBirA monomer and binds site-specifically to DNA to represses transcription of biotin biosynthetic genes (31, 33, 58). The holoBirA functional switch depends on intracellular apoBCCP concentration, which is, in turn, governed by growth rate (34, 35). In rapid growth conditions, when

fatty acid requirement is high, holoBirA partitions towards its metabolic role. In slow growth, with its accompanying lower requirement for fatty acids, holoBirA dimerizes and acts as a transcription repressor, thereby inhibiting biotin biosynthesis when it is not needed. The BirA functional switch occurs *via* formation of two distinct protein-protein interactions. Homodimerization, which is a prerequisite of DNA binding, is essential for transcription repression. Likewise, the hetero interaction with BCCP, a requirement for biotin transfer, is essential for BirA's enzymatic role.

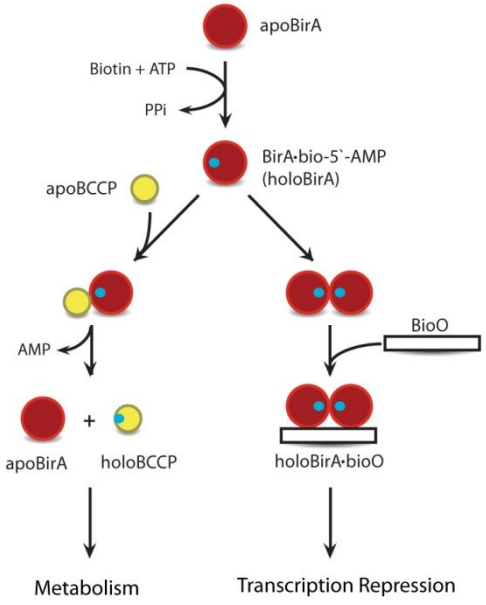


Figure 16. The biotin regulatory system in *E. coli*. BirA binds biotin and ATP, and the resulting BirA·bio5'-AMP (holoBirA) either transfers biotin to apoBCCP or dimerizes and binds site-specifically to the biotin operator (bioO).

Much is known about the structural basis of the BirA functional switch. The structure of holoBirA dimer bound to biotinol-5'-AMP, a non-hydrolysable analog of

bio-5'-AMP, has been determined by X-ray crystallography (39). Although the structure of the *E. coli* BirA-BCCP complex has not been experimentally obtained, a model has been constructed from the known structures of holoBirA and apoBCCP87, a C-terminal 87 biotin-accepting fragment of BCCP (Fig. 17) (45, 46, 50). The crystal structure of the BirA-BCCP complex from *Pyrococcus horikoshii* confirms the validity of the model (49). A single BirA surface, which is characterized by several loops (43, 44, 91), is utilized for both the hetero- and homodimer complexes, rendering the two interactions mutually exclusive. BirA variants with alanine substitutions in four of the loops, comprising residues 140-146, 170-176, 193-199 and 280-283, possess a broad range in homodimerization energetics relative to wild-type BirA (91). A subset of these variants also associates more slowly with BCCP than does wild-type BirA (91). Although informative about the structural requirements for the two protein-protein interactions, these results do not report on the impact of the loop perturbations on repression complex assembly or the functional switch.

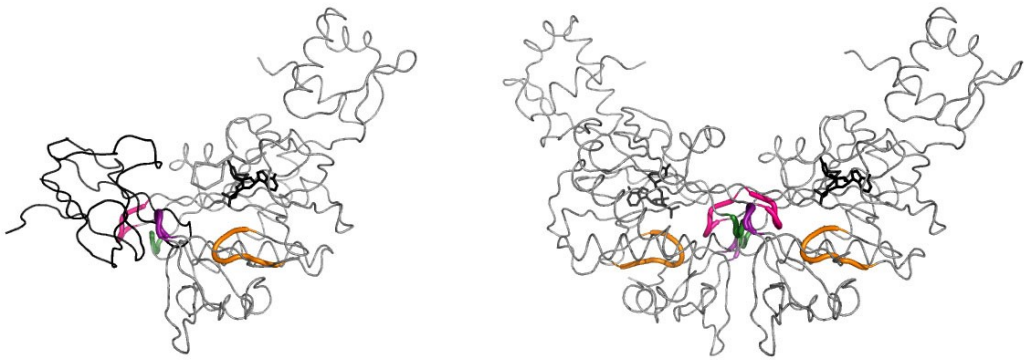


Figure 17. Models of the holoBirA·BCCP87 heterodimer (left) and holoBirA homodimer (right). Amino acids that were replaced with alanine are from loops 140-146 (green), 170-176 (orange), 193-199 (pink) and 280-283 (purple). Models were created using PyMOL (40) with the input file 2EWN for the homodimer and a file constructed by Zachary Wood for the heterodimer.

Mechanistically, the holoBirA switch can theoretically be subject to kinetic or equilibrium control. However, results of multiple experiments support kinetic control. The two interactions are roughly isoenergetic with equilibrium dissociation constants for homodimerization and heterodimerization of 5-10 μM and 2-3 μM, respectively (59, 68, 69). Measurements of the effect of apoBCCP on the initial rate of holoBirA binding to bioO indicated that, depending on its concentration, the acceptor protein decreases the final equilibrium occupancy of bioO by holoBirA dimer (68). This is consistent with the notion that apoBCCP depletes the pool of holoBirA that is available for homodimerization, thereby lowering bioO occupancy. Sedimentation velocity measurements of the dynamics of homodimerization indicate very slow interconversion between the two species with values for k_{on} and k_{off} of 20 (± 8) $M^{-1}s^{-1}$ and 2.7 (±0.5) x 10^{-4} s^{-1}, respectively (69). Furthermore, the bimolecular association rate constant for the BirA-BCCP87 interaction of 14,700 (± 200) $M^{-1}s^{-1}$ is considerably faster than the rate constant for homodimerization (46). Thus, due to the faster kinetics of heterodimerization, at sufficiently high apoBCCP concentrations holoBirA preferentially partitions towards biotin transfer.

In this work, control of the BirA switch was further investigated by measuring the effects of perturbations on homo and heterodimerization resulting from alanine substitutions in the interacting surface loops on the process. The diverse set of surface loop variants at our disposal were used to study the switch. Direct footprinting experiments on wild-type and the 18 BirA surface loop variants indicate that the energetics of total assembly of repression complex is dictated by homodimerization

energetics. Results of inhibition footprinting experiments indicate a direct correlation between inhibition of repression complex assembly and heterodimer association rates. This relationship between heterodimerization rates and inhibition of repression complex assembly reinforces the mechanism of the BirA functional switch as a kinetic partitioning between alternate protein-protein interactions.

3.3 Experimental Procedures

3.3.1 Chemicals and biochemicals

All chemicals were at least reagent grade. The DNaseI was purchased from Sigma-Aldrich. The OneTaq HotStart DNA Polymerase and T4 Polynucleotide Kinase were purchased from New England Biolabs. The γ^{32}P-ATP used for labeling DNA was obtained from Perkin Elmer and the dNTPs were purchased from Promega. Biotin and ATP were purchased from Sigma-Aldrich and stock solutions were prepared as described in (91) and the bio-5'-AMP was synthesized and purified as previously described (31, 58). The 87-fragment biotin carboxyl carrier protein (BCCP87) was purified as described in (46), with the addition of a Hydroxyapatite ULTROL (PALL) column chromatography step to remove any trace amounts of DNase in the protein preparation (68).

3.3.2 Expression and purification of BirA variants

Mutations in the BirA coding sequences were introduced using the QuikChange II XL Site-Directed Mutagenesis Kit (Stratagene) in the pBTac2 plasmid, which contains the BirA coding sequence under transcriptional control of the tac promoter. The variant proteins, which were produced with a carboxyl terminal $(His)_6$ tag to enable separation of

the variant from the chromosomally encoded wild-type protein, were purified as previously described (77, 91). After purification, the proteins were dialyzed against storage buffer [10 mM Tris-HCl (pH 7.50 ± 0.02 at 4^{o}C), 200 mM KCl, 0.1 mM dithiothreitol and 5% (v/v) glycerol] and stored at -70^{o}C. Protein concentrations were determined spectrophotometrically at 280 nm using the extinction coefficient of 47,510 $M^{-1}cm^{-1}$, calculated from the amino acid composition (78). The proteins were >95% pure as assessed by SDS-PAGE analysis, and >90% active in bio-5'-AMP binding, as determined by stoichiometric titrations monitored using steady state fluorescence emission spectroscopy (58).

3.3.3 Preparation of bioO DNA for DNaseI footprint titrations

Labeling of the DNA fragment for footprinting measurements was accomplished by polymerase chain reaction amplification of a fragment in which one of the primers was ^{32}P-labeled. The two primers were designed to flank the HindIII and PstI restriction sites of plasmid pBioZ (58, 60). First the HindIIIpBioZ primer (4 μM) was radiolabeled using γ-^{32}P ATP (0.8 μM) with T4 polynucleotide kinase. The labeling reaction, which also contained 10X kinase buffer and unlabeled ATP (0.8 μM) was allowed to proceed for 30 min at 37^{o}C, and then quenched with the addition of 10 μl formamide dye [80% deionized formamide, 1X TBE buffer, 0.02 % (w/v) bromophenol blue, and 0.02 % xylene cyanol (w/v)]. The resulting sample was incubated at 95^{o}C for 3 minutes (92) and the labeled DNA was separated from unincorporated nucleotides by electrophoresis on a 20% denaturing polyacrylamide gel (29:1 acrylamide/bisacrylamide in 8 M urea) at 300 V. The labeled oligonucleotide was removed from the gel by electro-elution, purified

through an Elutip-D column (Schleicher & Schuell), and concentrated by ethanol precipitation (93). The PCR for production of the labeled DNA fragment contained 1 nM pBioZ plasmid, 100-150 nM of each of the two primers, 200 nM deoxynucleotide triphosphates, and 0.05 units of Taq DNA Polymerase in reaction buffer in a total volume of 50 µl. The reaction was performed for 30 cycles of 95°C for 2 min (1x), 95°C for 1 min, 54°C for 1 min, 72°C for 1 min (30x) and a final 72°C extension for 5 min (1x) and the resulting DNA fragment was separated from template and primers by electrophoresis on a 1% agarose gel. After locating the band by phosphorimaging, the DNA was electro-eluted from the gel, purified through an Elutip-D column and concentrated by ethanol precipitation. The final labeled DNA was stored in TE buffer at a final concentration of ≤ 20,000 cpm/µl at 4°C for up to one month.

3.3.4 DNaseI footprint titrations

The DNaseI footprinting reactions were performed as described in (94), with a few modifications. The labeled DNA at a final concentration of 60 pM was combined with BirA protein, wild type or variant, at a range of concentrations in binding buffer containing 10 mM Tris-HCl (pH 7.5 at 20°C), 200 mM KCl, 2.5 mM $MgCl_2$, 1 mM $CaCl_2$, 2 µg/ml sonicated calf-thymus DNA, 100 µg/ml bovine serum albumin, 50 µM biotin and 500 µM ATP (or bio-5'-AMP at twice the highest final BirA concentration for those variants incapable of bio-5'-AMP synthesis). After incubating at 20°C for at least 45 minutes DNA cleavage was initiated by the addition of 5 µl DNase I (diluted to 0.0028 mg/ml in binding buffer minus the ligands, calf-thymus DNA and BSA), the sample was mixed by gently vortexing, and cleavage was allowed to proceed for 2 min.

The reaction was quenched by addition of 33 µl of 50 mM Na_2EDTA, the sample was rapidly vortexed, and the DNA was precipitated by addition of 700 µl of solution containing 0.4 M NH_4OAc and 50 µg/ml yeast phenylalanyl tRNA in ethanol. After a 20 minute incubation in a dry ice/ethanol bath the DNA was pelleted by centrifugation at 13,000 rpm for 20 min and the resulting pellets were washed with 80% ethanol twice. The dried pellets were resuspended in 7 µl formamide dye [80% deionized formamide, 1X TBE buffer, 0.02 % (w/v) bromophenol blue, and 0.02 % xylene cyanol (w/v)], samples were heated to 90^{o}C for 10 minutes, quick cooled on ice, loaded onto a 10% denaturing polyacrylamide gel (19:1 acrylamide/bisacrylamide in 8 M urea), and electrophoresed at 93 W. The gel was dried and exposed to a phosphoimager screen for at least 36 hours prior to scanning using a STORM phosphorimaging system (Perkin Elmer).

3.3.5 Inhibition DNase I footprint titrations

The inhibition DNase I footprinting titrations were performed as described in (68, 94). The holoBirA concentration used in each titration, which was determined from the direct DNase I footprinting titrations, is the concentration at which the bioO is just saturated with a particular holoBirA variant. All reactions were carried out in binding buffer containing 50 µM biotin and 500 µM ATP, or, for variants that are incapable of adenylate synthesis, bio-5`-AMP at twice the total BirA concentration. Each reaction contained approximately 10,000 cpm of radiolabeled DNA plus BCCP87 at varying concentrations in a total reaction volume of 45 µl. These samples, plus a sample of wild-type or variant BirA at 10X the final desired concentration, were preincubated at 20^{o}C for

at least 45 minutes. All proteins, with the exception of K194A and G142A, were pre-incubated in binding buffer containing biotin and ATP or bio-5'-AMP (please see Results for explanation). Binding reactions were initiated by addition of 5 µl BirA to a tube containing BCCP87 and operator DNA. After mixing by gentle vortexing the sample was incubated at 20°C for 80 seconds at which time bioO occupancy was probed by addition of 5 µl of DNase I at 0.003 mg/ml and the cleavage reaction was allowed to proceed for 30 s. All subsequent steps were performed as described above for the direct footprints.

3.3.6 Data analysis

For all footprinting images the intensities of the DNA bands in the bioO site were quantified relative to a standard band outside of the site (ImageQuant). For the direct footprint titrations, the optical densities were used to produce binding isotherms (94), which were subjected to non-linear least square analysis (GraphPad Prism 4.0), using the following equation

$$\bar{Y} = \frac{K_T\,[holoBirA]^2}{1 + K_T\,[holoBirA]^2} \qquad (4)$$

to obtain values for K_T, the equilibrium constant governing both repressor dimerization and dimer binding to bioO. The equilibrium constant for dimer binding to bioO, K_{bioO}, is obtained from the total equilibrium constant using the expression $K_T = K_{DIM} \times K_{bioO,}$ where K_{DIM} is the equilibrium association constant for holoBirA dimerization measured using equilibrium analytical ultracentrifugation.

The inhibition footprint images were processed as described above with the exception that the isotherms were analyzed using the following equation:

$$\bar{Y} = \frac{K_T\,([holoBirA]\left(1 - \frac{(K_{a,I}[apoBCCP])}{1 + K_{a,I}[apoBCCP])}\right))^2}{1 + K_T\,([holoBiA]\left(1 - \frac{(K_{a,I}[apoBCCP])}{1 + K_{a,I}\,[apoBCCP]}\right))^2} \quad (5)$$

to obtain $K_{a,I}$, the inhibition association constant governing the interaction between holoBirA and apoBCCP. The inverse of $K_{a,I}$, the inhibition dissociation constant, K_I, is reported in Results. The K_T value used for analysis of each inhibition footprint titration was obtained from the direct DNase I footprint titrations as described above.

3.4 Results

3.4.1 The total repression complex assembly energetics for each BirA variant is dictated by the dimerization energetics

The effects of amino acid replacements at the dimer interface on the two-step assembly of the biotin repressor on bioO were measured using quantitative DNase I footprinting. Alanine replacement in the residues of four surface loops, comprising residues 140-146, 170-176, 193-199 and 280-283, yields proteins with dimerization free energies ranging from -2.4 to -10.6 kcal/mol (91). DNaseI footprint titrations were performed to assess how these perturbations to dimerization influence assembly of the repressor on DNA.

Repression complex assembly occurs in two-steps including dimerization of two holoBirA monomers followed by binding of this dimer to bioO. The DNaseI footprinting experiment reports on the energetics of both steps and the equilibrium association constant governing dimer binding to bioO, K_{bioO}, can be calculated using the equilibrium dimerization constant, $K_{DIM,}$ obtained from sedimentation equilibrium experiments. In the footprinting experiment, each protein is saturated with bio-5'-AMP either through inclusion of biotin and ATP or, for those variants incapable of intermediate synthesis, direct addition of bio-5'-AMP, and incubated with approximately 60 pM of radiolabeled bioO. This low DNA concentration allows for the assumption that the total protein concentration in the reaction is equal to the free protein.

As observed in the DNaseI footprint titration obtained for the G196A variant (Fig. 18a), with increasing protein concentration, increased occupancy of the bioO site is obtained. Quantitation of the band intensities in the operator site as a function of holoG196A concentration (Materials and Methods) yields a binding isotherm (Fig. 18b), which was subjected to non-linear least squares analysis using Equation 4 to obtain a K_T value of 8.0 (± 1.0) x $10^{15}M^{-2}$. Using this total equilibrium constant and the known equilibrium dimerization constant for the variant (5 (± 4) x 10^{-6} M), the calculated value of K_{bioO}, the equilibrium constant for dimer binding to bioO, for the variant is 3.4 (± 0.7) x $10^{10}M^{-1}$, identical, within error, to the value of 2.0 (± 1.0) x 10^{10} M^{-1} obtained for wild type BirA.

The K_T and ΔG_T values for wild-type and all 18 variants are shown in Table 3, and the calculated K_{bioO} and ΔG_{bioO} values for all proteins are shown in Table 4. Although

a range of K_T and ΔG_T values are observed, the calculated K_{bioO} and ΔG_{bioO} values for all proteins are similar in magnitude. An approximately 1 kcal/mol difference in ΔG_{bioO} is seen for the K194A variant relative to wild-type. However, the dimerization constant obtained for this variant was at the upper limit of measurement due to limitations in the instrument's absorption optics (91) and has a large error associated with it. Thus, the values of K_{bioO} and ΔG_{bioO} have this same large uncertainty associated with them.

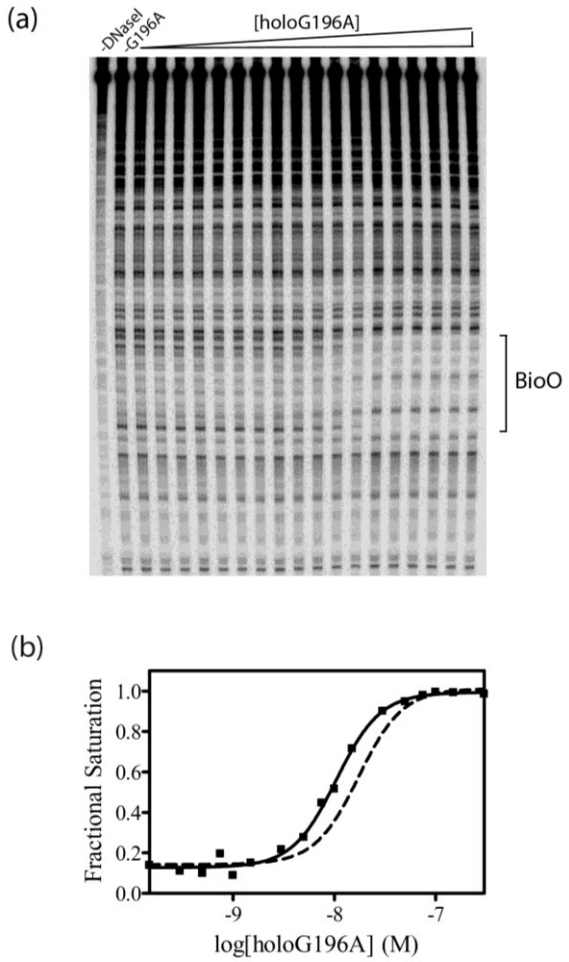


Figure 18. (a) The direct DNase I footprint titration of bioO obtained for BirA G196A (b) The isotherm obtained from quantitation of footprinting data obtained for G196A. The solid line represents the best-fit curve obtained from non-linear least squares analysis of the data. The results obtained for wild-type BirA are shown as a dashed line.

Table 3. Total assembly energetics of repression complex assembly

BirA Variant	K_T (M^{-2})[a]	ΔG^o_T (kcal/mol)[b, c]
Wild Type	**4.0 (± 2.0) x 10^{15}**	**-21.0 (± 0.4)**
E140A	**1.1 (± 0.5) x 10^{15}**	**-20.4 (± 0.3)**
Q141A	**2.3 (± 0.9) x 10^{15}**	**-20.8 (± 0.2)**
G142A	**3.0 (± 0.8)[c] x 10^{12}**	**-16.9 (± 0.2)**
P143A	**1.86 (± 0.01) x 10^{14}**	**-19.400 (± 0.002)**
R170A	**9.0 (± 4.0) x 10^{14}**	**-20.3 (± 0.3)**
V171A	**5.3 (± 2.3) x 10^{16}**	**-22.7 (± 0.3)**
K172A	**7.3 (± 2.1) x 10^{13}**	**-18.8 (± 0.2)**
N175A	**1.13 (± 0.09) x 10^{15}**	**-20.40 (± 0.05)**
D176A	**1.2 (± 0.6) x 10^{13}**	**-17.7 (± 0.3)**
G193A	**1.4 (± 0.3) x 10^{16}**	**-21.9 (± 0.1)**
K194A	**2.8 (± 0.1) x 10^{17}**	**-23.70 (± 0.02)**
T195A	**2.5 (± 0.4) x 10^{14}**	**-19.60 (± 0.08)**
G196A	**8.0 (± 1.0) x 10^{15}**	**-21.6 (± 0.1)**
D197A	**2.2 (± 0.2) x 10^{13}**	**-18.10 (± 0.04)**
I280A	**1.4 (± 0.4) x 10^{13}**	**-17.8 (± 0.1)**
G281A	**6.0 (± 2.0) x 10^{14}**	**-20.1 (± 0.2)**
D282A	**1.15 (± 0.05) x 10^{16}**	**-21.80 (± 0.02)**
K283A	**6.52 (± 0.05) x 10^{14}**	**-20.100 (± 0.005)**

The equilibrium constant, K_T, was measured in standard buffer [10 mM Tris-HCL (pH 7.50 ± 0.02 at 20°C), 200 mM KCl, 2.5 mM $MgCl_2$, 1 mM $CaCl_2$, 2 µg/ml sonicated calf-thymus DNA, 100 µg/ml bovine serum albumin, 50 µM biotin and 500 µM ATP (or bio-5'-AMP)] as described in "Materials and Methods."

[a] The errors represent the standard error of two independent experiments.

[b] ΔG^o_{KT} values were calculated using $\Delta G^o_{KT} = RT\ln K_{KT}$

[c] The reported uncertainties for each variant were calculated using error propagation.

Table 4. Energetics of binding of BirA variants to the biotin operator

BirA Variant	K_{bioO} (M^{-1})[a]	ΔG^{o}_{bioO} (kcal/mol)[b,c]
Wild Type	**2.0 (± 1.0) x 10^{10}**	**-13.8 (± 0.4)**
E140A	**2.0 (± 0.9) x 10^{10}**	**-13.7 (± 0.3)**
Q141A	**7.0 (± 3.0) x 10^{9}**	**-13.1 (± 0.2)**
G142A	**4.0 (± 1.0) x 10^{10}**	**-14.2 (± 0.2)**
P143A	**9.26 (± 0.03) x 10^{9}**	**-13.313 (± 0.002)**
R170A	**3.0 (± 1.0) x 10^{10}**	**-13.9 (± 0.3)**
V171A	**1.4 (± 0.6) x 10^{10}**	**-13.5 (± 0.3)**
K172A	**5.0 (± 1.0) x 10^{10}**	**-14.2 (± 0.2)**
N175A	**6.4 (± 0.5) x 10^{10}**	**-14.43 (± 0.05)**
D176A	**6.0 (± 3.0) x 10^{10}**	**-14.4 (± 0.3)**
G193A	**7.0 (± 1.0) x 10^{9}**	**-13.1 (± 0.1)**
K194A	**3.3 (± 0.1) x 10^{9}**	**-12.71 (± 0.02)**
T195A	**2.6 (± 0.3) x 10^{10}**	**-13.9 (± 0.1)**
G196A	**3.4 (± 0.7) x 10^{10}**	**-14.1 (± 0.1)**
D197A	**9.7 (± 0.7) x 10^{9}**	**-13.34 (± 0.04)**
I280A	**5.0 (± 1.0) x 10^{10}**	**-14.3 (± 0.1)**
G281A	**6.0 (± 2.0) x 10^{10}**	**-14.3 (± 0.2)**
D282A	**1.55 (± 0.06) x 10^{10}**	**-13.61 (± 0.02)**
K283A	**7.58 (± 0.06) x 10^{9}**	**-13.197 (± 0.005)**

[a] The K_{bioO} value was calculated using the equation $K_T = K_{DIM} \times K_{bioO}$, where K_{DIM} was determined previously (Table 1).
[b] ΔG^{o}_{KbioO} values were calculated using $\Delta G^{o}_{KbioO} = RT\ln K_{KbioO}$
[c] The reported uncertainties for each variant were calculated using error propagation.

3.4.2 The inhibition of biotin repression complex assembly is directly correlated with BirA-BCCP association rates

The extent to which transcription repression complex assembly by the BirA variants is inhibited by apoBCCP87 was determined using inhibition DNaseI footprinting. Stopped-flow kinetic measurements indicate that several of the variants with alanine substitutions in surface loops are slower than wild-type BirA in their rates of association with apoBCCP (91). It is also known that holoBirA homodimerization and holoBirA-apoBCCP heterodimerization are roughly isoenergetic, implying that the switch between these two protein-protein interactions is kinetically controlled (68, 69). The inhibition footprinting experiments were performed to determine the relationship of inhibition of BirA repression complex assembly to rates of hetero-association for the variants.

Repression complex assembly involves dimerization of two holoBirA monomers and subsequent binding of this dimer to bioO. Addition of apoBCCP87 to this system perturbs homodimerization through the association of holoBirA with apoBCCP87 and resulting biotin transfer with concomitant conversion of holoBirA to apoBirA, a species that dimerizes very weakly (68). By depleting the dimer pool, occupancy of the bioO site is also decreased. In the inhibition footprinting experiment, a range of apoBCCP87 concentrations is used to obtain an inhibition isotherm. Nonlinear least squares analysis of the data using Equation 5 (Materials and Methods) yields the inhibition constant, K_I, which reports on the energetics of the holoBirA-BCCP87 interaction.

The design of the inhibition footprinting experiment takes into account information obtained from direct footprints described above by using a total BirA concentration at which the bioO site is just saturated with BirA, a concentration at which the system is poised for inhibition by apoBCCP87. As indicated in the inhibition footprint obtained for I280A (Fig. 19a.), a loss of the bioO footprint is observed with increasing apoBCCP87 concentration. The inhibition isotherm (Fig. 19b) obtained from the data was analyzed using non-linear least squares analysis to obtain a K_I of 12.0 (± 2.1) x 10^{-6} M, a value that is approximately four-fold larger than that obtained for wild type BirA. In contrast to wild type BirA, the lower plateau of the I280A inhibition curve is greater than zero. As previously shown (68), this plateau in these isotherms depends on the total holoBirA concentration employed with higher concentration yielding higher lower plateau values. Due to its weaker dimerization, the I280A concentration used in this experiment was 1 μM, significantly higher than the 7.5 x 10^{-8} M concentration employed for wild type BirA.

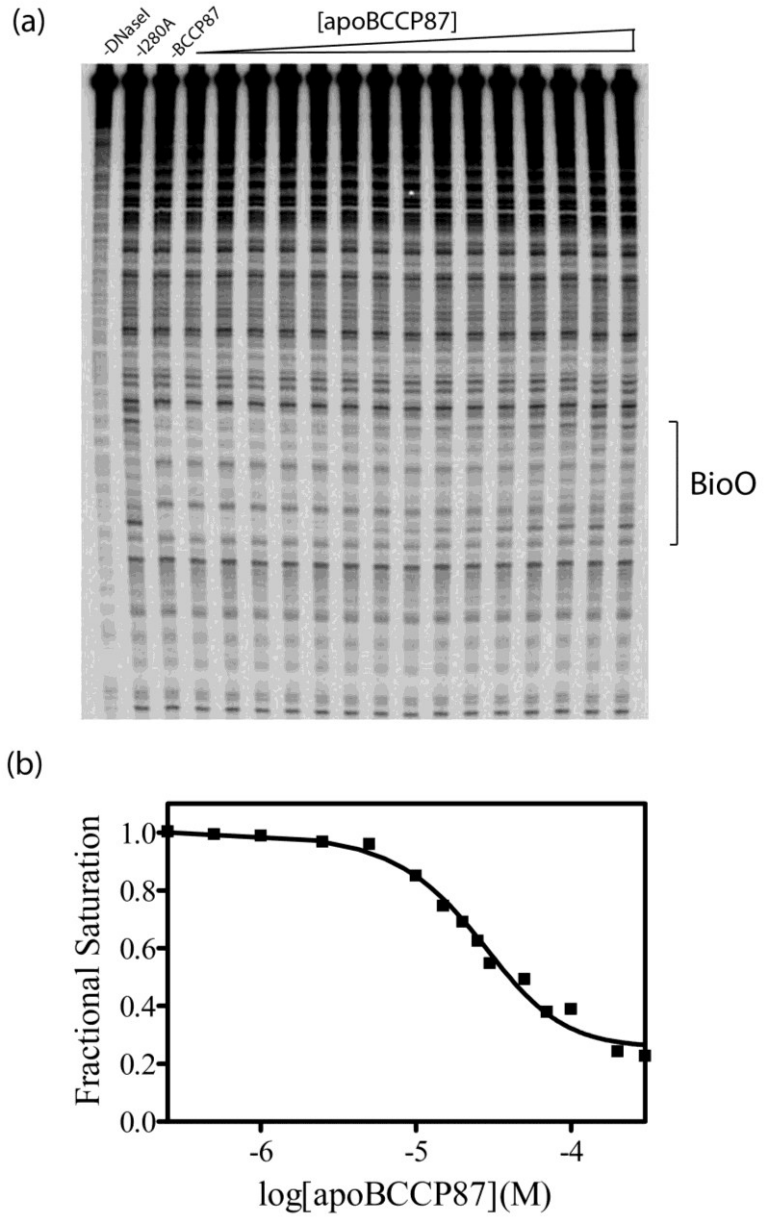


Figure 19. (a) The inhibition DNase I footprint titration of bioO obtained for BirA I280A. (b) The inhibition isotherm obtained from quantitation of footprinting data for I280A. The solid line represents the best-fit curve obtained from non-linear least squares analysis of the data.

The inhibition DNaseI footprint titrations were performed on all variants and data were analyzed as described above. In the experiment, all variants with the exceptions of K194A and G142A were preincubated with sufficient biotin and ATP to obtain saturation with bio-5'-AMP, or directly with bio-5'-AMP for those variants incapable of intermediate synthesis, at 20^{o}C for 45 min before initiation of the inhibition reaction. However, when G142A and K194A variants are pre-incubated with ligand, very high BCCP87 concentrations, up to 1 mM, were required to observe any inhibition of repression complex assembly. However, when preincubated in the absence of ligand, inhibition of assembly by these same variants was observed in the same apoBCCP87 concentration range as observed for wild-type BirA. The K_I and ΔG_I values obtained for all variants are shown in Table 5. While the parameters for many of the variants are similar to that obtained for wild-type BirA, some variants including G142A, K172A, D176A, K194A and G196A, differ from wild type BirA by approximately 1 kcal/mol or more in ΔG_I.

Table 5 Energetics of heterodimerization of BirA variants with BCCP87

BirA Variant	K_I (M)[a]	ΔG^o_I (kcal/mol)[b,c]
Wild Type	**2.8 (± 0.9) x 10^{-6}**	**-7.4 (± 0.2)**
E140A	**4.0 (± 1.0) x 10^{-6}**	**-7.3 (± 0.2)**
Q141A	**4.0 (± 1.0) x 10^{-6}**	**-7.3 (± 0.2)**
G142A	**15 (± 5) x 10^{-6}**	**-6.5 (± 0.2)**
P143A	**7.0 (± 1.0) x 10^{-6}**	**-6.9 (± 0.1)**
R170A	**3.8 (± 0.2) x 10^{-6}**	**-7.24 (± 0.02)**
V171A	**9.0 (± 2.0) x 10^{-6}**	**-6.8 (± 0.1)**
K172A	**20 (± 0.7) x 10^{-6}**	**-6.21 (± 0.02)**
N175A	**4.0 (± 0.4) x 10^{-6}**	**-7.2 (± 0.1)**
D176A	**200 (± 1) x 10^{-6}**	**-5.022 (± 0.003)**
G193A	**5.0 (± 0.5) x 10^{-6}**	**-7.1 (± 0.1)**
K194A	**15 (± 0.4) x 10^{-6}**	**-6.44 (± 0.02)**
T195A	**9.0 (± 1.0) x 10^{-6}**	**-6.8 (± 0.1)**
G196A	**22 (± 5.0) x 10^{-6}**	**-6.2 (± 0.1)**
D197A	**9.0 (± 0.1) x 10^{-6}**	**-6.735 (± 0.005)**
I280A	**12 (± 2) x 10^{-6}**	**-6.6 (± 0.1)**
G281A	**6.0 (± 0.8) x 10^{-6}**	**-6.9 (± 0.1)**
D282A	**4.0 (± 1.0) x 10^{-6}**	**-7.3 (± 0.2)**
K283A	**4.85 (± 0.02) x 10^{-6}**	**-7.09 (± 0.03)**

The inhibition constants and Gibbs free energies were determined as described in "Materials and Methods."

[a] K_I was obtained through analysis of the data to Equation 5. The value obtained from the fit, $K_{A,I}$ was inversed to obtain the inhibition dissociation constant reported here. The errors represent the standard error of two independent experiments.

[b] ΔG^o_I values were calculated using $\Delta G^o_I = RT\ln K_I$

[c] The reported uncertainties for each variant represent the standard error of two independent experiments.

3.5 Discussion

The effects of perturbations to holoBirA homo and heterodimerization on repression complex assembly and the switch between the two functions were assessed through direct and inhibition DNase I footprinting titrations. The results reveal that the total energetics of repression complex assembly can be tuned by homodimerization energetics. Furthermore, a direct correlation between heterodimer association rates and inhibition of repression complex assembly was observed, reinforcing a mechanism of BirA functional switching that is governed by kinetic control.

Repression complex assembly is a two-step process of holoBirA dimerization followed by site-specific binding of the dimer to the biotin operator (33), and the total equilibrium constant, K_T, obtained from DNaseI footprint titrations provides a measure of both steps. Using previously measured values of homodimerization equilibrium constant, K_{DIM}, (Table 1) (91), the affinity of the dimer for DNA, K_{bioO}, is calculated from the expression $K_T=K_{DIM}$ x K_{bioO}. The variants are characterized by a range of K_T values (Table 3, Fig. 20a), but the calculated K_{bioO} values are all similar to that obtained for wild-type BirA (Table 4, Fig. 20b), consistent with control of total assembly energetics by dimerization energetics. The correlation between K_T and K_{DIM} in repression complex assembly, for which the correlation coefficient is 0.9556, illustrates the control that dimerization energetics exerts on repression complex assembly (Fig. 21a).

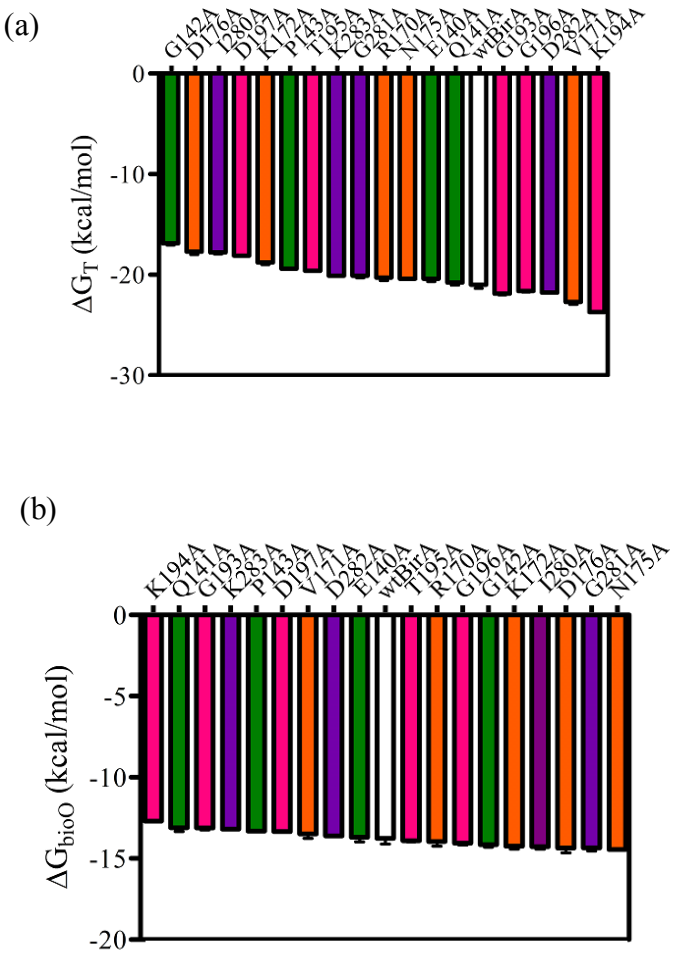


Figure 20. Gibbs free energies of wild-type BirA and all variants for (a) total transcription repression complex assembly, ΔG_T (b) the protein dimer binding to bioO, ΔG_{bioO}

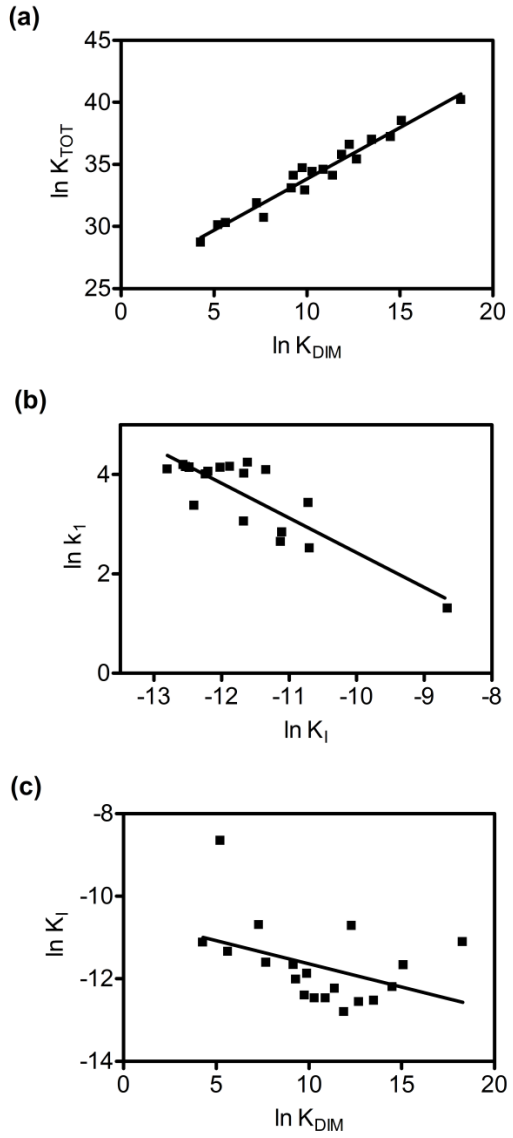


Figure 21. Correlation plots of (a) ln K_{TOT} versus ln K_{DIM} ($r^2 = 0.9556$). (b) ln k_1 versus ln K_I ($r^2 = 0.7190$). (c) ln K_I versus ln K_{DIM} ($r^2 = 0.1702$).

The similarity of the K_{bioO} values for the variants indicates that amino acid substitutions in the dimer interface either do not perturb the structure of this interface or that the N-terminal domain functions independently of the remainder of the molecule. The introduction of alanine substitutions into the homodimer interface may alter its

conformation, and this altered conformation could potentially change the affinity of the dimer for bioO. However, the observed lack of impact of K_{bioO} values observed for the dimer interface variants suggests that the N-terminal domain functions independently. This functional independence is consistent with the known ability of the DNA binding domain to adopt different positions relative to the dimerization domain. In the btnOH-AMP bound dimer structure, two DNA-binding domain orientations that differ by 6.6^{o} relative to the central domain are observed, indicating that this domain is mobile relative to the rest of the molecule (36, 39). Additionally, the winged helix-turn-helix motif DNA-binding domain itself is structurally heterogeneous. The wing region, which fits into the minor groove of DNA, is poorly packed or disordered in the absence of DNA (39), lending flexibility to the DNA binding region. Thus, the variable arrangement of the entire DNA binding domain in the crystal structure suggests independence of the DNA binding domain from the dimerization domains which renders the affinity of the holoBirA dimer for bioO insensitive to alterations in the dimer interface.

The role of heterodimerization kinetics in regulating the BirA functional switch was investigated by measuring the abilities of the alanine variants to inhibit assembly of the repression complex. In the inhibition footprint titrations the presence of apoBCCP decreases the final equilibrium bioO occupancy by depleting the pool of holoBirA available for homodimerization (68). Using a range of apoBCCP87 concentrations to disrupt the holoBirA-bioO interaction, an inhibition constant, K_I, was determined for each variant. Of the several variants that exhibit inhibition constants greater than wild-type (Table 5), most also exhibit slower rates of biotin transfer to apoBCCP87 (91) (Fig.

22a and 22b). Variants with K_I values greater than wild type BirA (less negative ln K_I values) include G142A, K172A, D176A, K194A and G196A, which are all slower in hetero association with apoBCCP (Fig. 22a and 22b). The correlation between inhibition constant and the bimolecular association rate constant for holoBirA with apoBCCP87 (Fig. 21b), which is characterized by a correlation coefficient of 0.7190, emphasizes the central role that kinetics plays in determining the BirA functional switch.

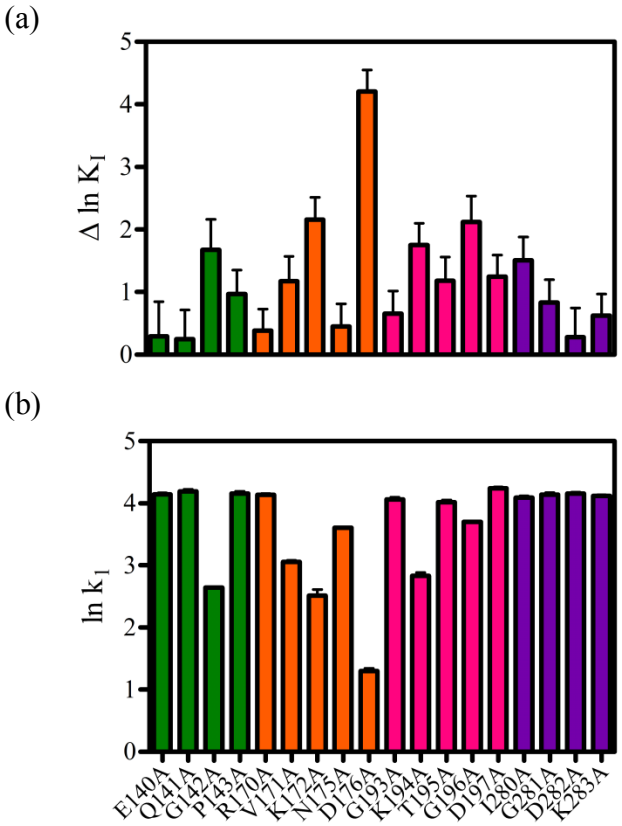


Figure 22. (a) $\Delta \ln K_I = \ln K_{I(variant)} - \ln K_{I(wt)}$. (b) The bimolecular rates of association, k_1, of the BirA-BCCP87 interaction (1). The errors for graph (a) were calculated using standard error propagation. The errors in graph (b) represent the standard error of two independent experiments.

Theoretically, in the BirA functional switch, a kinetic or thermodynamic mechanism could be responsible for one biological outcome over the other. Previous studies have shown that homo and heterodimerization are characterized by isoenergetic equilibrium constants (59, 68, 69). Further evidence that the switch is not under thermodynamic control is presented in Fig. 21c, which shows the absence of a correlation between the equilibrium constant for homodimerization and the ability of apoBCCP to inhibit binding of a variant to bioO (correlation coefficient = 0.1702). Thus the magnitude of the equilibrium constant governing holoBirA homodimerization has no influence on the switch.

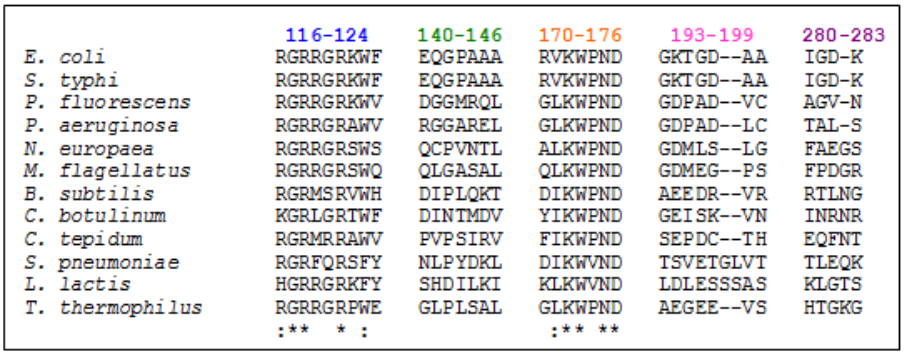
A strong relationship between the results obtained from this work and the biology of the biotin regulatory system is observed. The rate of any bimolecular process increases with increasing component concentrations, and under conditions of fast cell growth and high apoBCCP concentrations, the probability that holoBirA binds this substrate is higher, resulting in biotin transfer. Under slow growth conditions due to lower apoBCCP concentrations, holoBirA partitions towards its transcription repression role, thereby switching between its two functions in a kinetically controlled manner (35). Other multifunctional proteins employ different structural, kinetic or thermodynamic mechanisms to interact with their partners (28, 95), and elucidation of the mechanism of switching for these systems requires measurements of the reaction that are relevant to the process. In each case, the thermodynamic or kinetic rules that govern the interactions determine the final biological output of the system.

Chapter 4: Summary and Future Directions

The work presented in this thesis focuses on characterizing the roles of surface loop variants in a variety of BirA functions including homodimerization, biotin transfer, assembly of the transcription repression complex and in controlling the bifunctional switch. All surface loops were found to be important for homodimerization and only a subset of residues on some surface loops was found to be essential for the biotin transfer reaction. Study of repression complex assembly revealed that homodimerization energetics determines total assembly energetics. Finally, a direct correlation between heterodimer association rates and inhibition of repression complex assembly was observed, reinforcing a mechanism of kinetic control in the BirA bifunctional switch.

The structural basis for BirA bifunctionality is achieved through surface loops. The flexibility of these loops and their location at the dimer interface enables them to participate in two protein-protein interactions. Two different types of loops termed constant and variable were identified based on protein sequence conservation. Constant loops have highly conserved sequences among ligases from all organisms and in BirA the essential biotin transfer reaction is facilitated by these loops. Based on the high sequence conservation and their presence in all biotin protein ligases, we can assume that these constant loops facilitate the biotin transfer reaction in all ligases. The variable loop sequences show very little homology, even among bifunctional ligases (Fig. 23). These loops participate in the homodimerization reaction in BirA, and we can assume that they facilitate dimer association in other bifunctional ligases as well. However, because

variable loop sequences from bifunctional ligases are so diverse, it is uncertain how different loop sequences can evolve to perform the same function as in *E. coli* BirA. The ideal method to investigate this is to perform alanine scanning mutagenesis on loop residues from other bifunctional ligases and measure the effects of loop perturbations on homodimerization. By studying surface loops from other bifunctional ligases, we can understand the evolution of variable loops and determine if, regardless of the sequence, these loops facilitate homodimerization as they do for BirA. If not, an investigation of other structural strategies used by these bifunctional ligases to enable homodimerization without compromising the essential biotin transfer reaction would allow us to understand the evolution of this bifunctional protein in general.



	116-124	140-146	170-176	193-199	280-283
E. coli	RGRRGRKWF	EQGPAAA	RVKWPND	GKTGD--AA	IGD-K
S. typhi	RGRRGRKWF	EQGPAAA	RVKWPND	GKTGD--AA	IGD-K
P. fluorescens	RGRRGRKWV	DGGMRQL	GLKWPND	GDPAD--VC	AGV-N
P. aeruginosa	RGRRGRAWV	RGGAREL	GLKWPND	GDPAD--LC	TAL-S
N. europaea	RGRRGRSWS	QCPVNTL	ALKWPND	GDMLS--LG	FAEGS
M. flagellatus	RGRRGRSWQ	QLGASAL	QLKWPND	GDMEG--PS	FPDGR
B. subtilis	RGRMSRVWH	DIPLQKT	DIKWPND	AEEDR--VR	RTLNG
C. botulinum	KGRLGRTWF	DINTMDV	YIKWPND	GEISK--VN	INRNR
C. tepidum	RGRMRRAWV	PVPSIRV	FIKWPND	SEPDC--TH	EQFNT
S. pneumoniae	RGRFQRSFY	NLPYDKL	DIKWVND	TSVETGLVT	TLEQK
L. lactis	HGRRGRKFY	SHDILKI	KLKWVND	LDLESSSAS	KLGTS
T. thermophilus	RGRRGRPWE	GLPLSAL	GLKWPND	AEGEE--VS	HTGKG
	:** * :		:** **		

Figure 23. Multiple sequence alignment of bifunctional bacterial biotin protein ligases using ClustalW (72). The constant loop sequences are well conserved and the variable loop sequences are very diverse in these bifunctional ligases.

Like BirA, many other multifunctional proteins utilize loops to mediate several interactions (96, 97). Surface loops do not disrupt folding of the protein and can be used in ligand binding and partner recognition as well. Apart from loops, other structural strategies can be employed by proteins to interact with different targets including binding pockets, intrinsically disordered regions of the protein, hotspots and/or distinct interaction interfaces (19). A multispecific protein often co-evolves with its partners to carry out specific cellular functions, implying that the structural components used by these proteins to facilitate interactions also co-evolved. The structural basis for functional evolution has not been studied for most multifunctional proteins, and could provide valuable insight into the mechanisms of protein multispecificity. The structural components used to form each protein-protein interaction should be identified, and overlap in use of structural information would illustrate the structural versatility of the protein in facilitating numerous interactions. These studies would provide information about the structure-function relationship and evolution of multifunctional proteins.

One feature of multifunctional proteins is their ability to switch between functions depending on cellular requirements. Functional switching can be mechanistically under equilibrium or kinetic control. A kinetic mechanism implies that the rates of association govern protein assembly, and an equilibrium mechanism is governed by affinities for one protein-protein interaction over the other. For BirA, a kinetic mechanism of switching between two protein-protein interactions was observed. However, for many multifunctional proteins the switching mechanism is not known. Studying the basis for the switch in multifunctional proteins enables an understanding of the rules governing

protein association and the biological outcomes that result from one protein-protein interaction over the other. For example, ubiquitin is a protein involved in multiple homo and hetero interactions, yet the kinetic or thermodynamic basis for switching between partners is not known (98). This information can be used to understand ubiquitin assembly, preference for binding partners and ubiquitin functions, including protein degradation, cell-cycle regulation, DNA repair and apoptosis (99). Understanding the mechanism of functional switching in multifunctional proteins can add to our knowledge of cellular processes and explain different biological outcomes. However, complications arise when studying the switch mechanism in eukaryotic proteins versus prokaryotic proteins due to compartmentalization in eukaryotic systems (100). For example, the human biotin protein ligase, holocarboxylase synthetase, or HCS, interacts with five different biotin-dependent carboxylases in human cells (51). The kinetic preference of HCS for three substrates over the other two was found to be based on the location of the carboxylase in the cell (101). Three carboxylases that function in the mitochondria are biotinylated faster than two carboxylases found in the cytosol, presumably because the cytosolic HCS has less access to mitochondrial carboxylases. Thus, when studying the switching mechanism in eukaryotic cells the compartmentalization of protein partners should also be considered.

This work has laid the groundwork for several future studies. The loops studied in this work represent only four of eight interfacial loops identified previously (39). The study of other loops will provide a complete understanding of the roles of all surface loops in BirA functions. The loops 292-298 and 309-313 located in the C-terminal

domain (Fig. 24) are of particular interest because similar loops were identified in the *P. horikoshii* biotin protein ligase (PhBPL) by sequence alignment and are within contact distance to PhBCCP (49). The BirA 309-313 loop is within contact distance to the other BirA monomer in the homodimer structure and to BCCP in the heterodimer model (39, 45, 50). Based on the importance of these loops in mediating the PhBPL-PhBCCP interaction, we can speculate on their importance for facilitating BirA homo and heterodimerization. The best way to determine this is to perform alanine scanning mutagenesis on residues of these loops to determine their roles in BirA functions, and thereby also learn more about the role of the C-terminal domain. Also, since only four residues from the 116-124 loop were previously studied, a more comprehensive analysis of the roles of the residues in this loop can provide information on the importance of this conserved loop in BirA functions. Further, obtaining structures of variants already characterized will provide a clearer understanding of the effect of alanine substitutions at the dimer interface. These structures may provide insight into the structural rearrangements that occur due to amino acid replacements that cannot be explained from existing structures of the wild-type protein. This may, in turn, help to explain some unexpected results obtained from our work, such as the more favorable Gibbs free energy of homodimerization observed for the V171A variant relative to wtBirA.

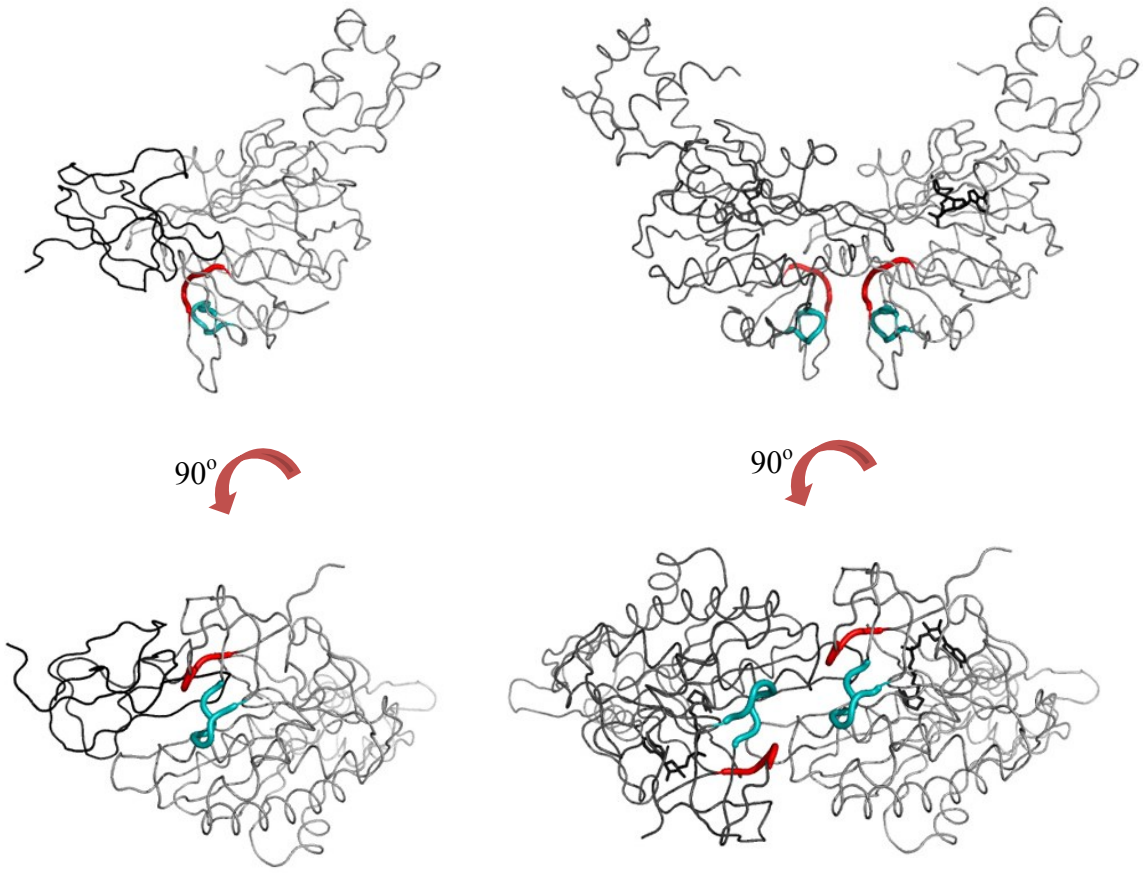


Figure 24. Models of the BirA-BCCP heterodimer (left) and BirA homodimer (right). The top figures were rotated 90° to display cross-sections of the interfaces. Two BirA loops located at the dimer interface are highlighted: 292-298, cyan; 309-313, red. All models were created using PyMOL (40) with the PDB file 2EWN for the homodimer and a file constructed by Zachary Wood comprising the PDB files 2EWN and 1BDO for BirA and BCCP87, respectively, using the PhBPL-PhBCCP file 2EJG as a template.

In this work, we assumed that the alanine substituted variants from the four surface loops bind biotin, ATP and bio-5'-AMP in a manner identical to wild-type BirA.

This is a reasonable assumption, since the biotin-binding and ATP-binding sites identified from crystal structures are located far from the dimer interface loops. Yet, a study of the ligand binding properties of these variants may yield interesting results, especially because in the dimer structure the 193-199 loop from the second monomer and the 170-176 loop from the same monomer are within close proximity to the btnOH-AMP moiety. It is possible that residues from these loops are required for both ligand binding and to position the adenylate in the active site. The ligand binding properties of all alanine substituted surface loop variants should be measured to understand the roles of these loops in ligand binding and catalysis.

A diverse set of variants that display very different dimerization properties are at our disposal and can be used to study BirA properties *in vivo*. Through recombineering (102), the chromosomal DNA of *E. coli* can be mutated to code for variants with altered dimerization constants, which in turn can be used to study the relationship between *in vitro* measurements and BirA properties *in vivo*. The wtBirA *in vivo* concentration is estimated to be 40 nM and we can assume that variants transcribed from chromosomal DNA will be present at similar concentrations, but this will have to be experimentally measured. Variants with very weak dimerization constants, in the upper micromolar to millimolar range, will assumedly be unable to dimerize and repress transcription at these low *in vivo* concentrations. These variants can be used to measure the effect of uncontrolled biotin production on cell growth/death rates and the biotin transfer reaction. Conversely, the chromosome can be mutated to produce BirA variants with very tight dimerization constants, in the nanomolar range, and these variants will theoretically

always be bound to the biotin operator. These variants can be used to determine the effect of over-repression of the biotin biosynthetic operator on the essential biotin transfer reaction and the switch between functions. Comparison of *in vivo* and *in vitro* measurements is essential to determine the validity of experiments we perform in the test tube.

In this work the significance of BirA surface loops for homodimerization and biotin transfer were investigated. The results provide insight into the evolution of these loops to mediate multiple functions and explain a mechanism of switching between two functions. We hope this work will be the starting point for future studies of BirA surface loops, the study of the evolution of biotin protein ligase, and for studying the mechanisms of multispecifity in other multifunctional proteins.

Bibliography

1. Hohenauer T & Moore AW (2012) The Prdm family: expanding roles in stem cells and development. *Development* 139(13):2267-2282.
2. Valenta T, Hausmann G, & Basler K (2012) The many faces and functions of [beta]-catenin. *EMBO Journal* 31(12):2714-2736.
3. Harris HE, Andersson U, & Pisetsky DS (2012) HMGB1: A multifunctional alarmin driving autoimmune and inflammatory disease. *Nature Reviews Rheumatology* 8(4):195-202.
4. Hinz M, Arslan SC, & Scheidereit C (2012) It takes two to tango: IκBs, the multifunctional partners of NF-κB. *Immunological Reviews* 246(1):59-76.
5. Gupta V & Bamezai RNK (2010) Human pyruvate kinase M2: A multifunctional protein. *Protein Science* 19(11):2031-2044.
6. Zwerschke W*, et al.* (1999) Modulation of type M2 pyruvate kinase activity by the human papillomavirus type 16 E7 oncoprotein. *Proceedings of the National Academy of Sciences* 96(4):1291-1296.
7. Zhang C*, et al.* (2010) P53 and p38 MAPK pathways are involved in MONCPT-induced cell cycle G2/M arrest in human non-small cell lung cancer A549. *Journal of Cancer Research and Clinical Oncology* 136(3):437-445.
8. Shieh S-Y, Ahn J, Tamai K, Taya Y, & Prives C (2000) The human homologs of checkpoint kinases Chk1 and Cds1 (Chk2) phosphorylate p53 at multiple DNA damage-inducible sites. *Genes & Development* 14(3):289-300.

9. Vousden KH & Lu X (2002) Live or let die: the cell's response to p53. *Nature Reviews Cancer* 2(8):594-604.

10. Ryan DP & Matthews JM (2005) Protein–protein interactions in human disease. *Current Opinion in Structural Biology* 15(4):441-446.

11. Wong JMS, Ionescu D, & Ingles CJ (2003) Interaction between BRCA2 and replication protein A is compromised by a cancer-predisposing mutation in BRCA2. *Oncogene* 22(1):28-33.

12. Jubb H, Higueruelo AP, Winter A, & Blundell TL (2012) Structural biology and drug discovery for protein-protein interactions. *Trends in Pharmacological Sciences.* 33(5):241-248.

13. Hudson PJ & Souriau C (2003) Engineered antibodies. *Nature Medicine* 9(1):129-134.

14. Surade S & Blundell Tom L (2012) Structural Biology and Drug Discovery of Difficult Targets: The Limits of Ligandability. *Chemistry & Biology* 19(1):42-50.

15. Wells JA & McClendon CL (2007) Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* 450(7172):1001-1009.

16. Walker K & Olson MF (2005) Targeting Ras and Rho GTPases as opportunities for cancer therapeutics. *Current Opinion in Genetics & Development* 15(1):62-68.

17. White AW, Westwell AD, & Brahemi G (2008) Protein–protein interactions as targets for small-molecule therapeutics in cancer. *Expert Reviews in Molecular Medicine*

18. Pagel P*, et al.* (2005) The MIPS mammalian protein–protein interaction database. *Bioinformatics* 21(6):832-834.

19. Erijman A, Aizner Y, & Shifman JM (2011) Multispecific Recognition: Mechanism, Evolution, and Design. *Biochemistry* 50(5):602-611.

20. Akiva E, Itzhaki Z, & Margalit H (2008) Built-in loops allow versatility in domain–domain interactions: Lessons from self-interacting domains. *Proceedings of the National Academy of Sciences* 105(36):13292-13297.

21. Fuxreiter M, Simon I, & Bondos S (2011) Dynamic protein–DNA recognition: beyond what can be seen. *Trends in Biochemical Sciences* 36(8):415-423.

22. Graves B*, et al.* (2012) Activation of the p53 pathway by small-molecule-induced MDM2 and MDMX dimerization. *Proceedings of the National Academy of Sciences* 109(29):11788-11793.

23. Keskin O & Nussinov R (2007) Similar Binding Sites and Different Partners: Implications to Shared Proteins in Cellular Pathways. *Structure* 15(3):341-354.

24. Moreira IS, Fernandes PA, & Ramos MJ (2007) Hot spots—A review of the protein–protein interface determinant amino-acid residues. *Proteins: Structure, Function, and Bioinformatics* 68(4):803-812.

25. Schreiber G & Keating AE (2011) Protein binding specificity versus promiscuity. *Current Opinion in Structural Biology* 21(1):50-61.

26. Yang X*, et al.* (2006) Structural basis for protein–protein interactions in the 14-3-3 protein family. *Proceedings of the National Academy of Sciences* 103(46):17237-17242.

27. Kleppe R, Martinez A, Døskeland SO, & Haavik J (2011) The 14-3-3 proteins in regulation of cellular metabolism. *Seminars in Cell & Developmental Biology* 22(7):713-719.

28. Baerga-Ortiz A, Bergqvist S, Mandell JG, & Komives EA (2004) Two different proteins that compete for binding to thrombin have opposite kinetic and thermodynamic profiles. *Protein Science* 13(1):166-176.

29. Barker DF & Campbell AM (1981) The birA gene of Escherichia coli encodes a biotin holoenzyme synthetase. *Journal of Molecular Biology* 146(4):451-467.

30. Barker DF & Campbell AM (1981) Genetic and biochemical characterization of the birA gene and its product: Evidence for a direct role of biotin holoenzyme synthetase in repression of the biotin operon in Escherichia coli. *Journal of Molecular Biology* 146(4):469-492.

31. Lane MD, Rominger KL, Young DL, & Lynen F (1964) The Enzymatic Synthesis of Holotranscarboxylase from Apotranscarboxylase and (+)-Biotin. *Journal of Biological Chemistry* 239(9):2865-2871.

32. Streaker ED, Gupta A, & Beckett D (2002) The Biotin Repressor: Thermodynamic Coupling of Corepressor Binding, Protein Assembly, and Sequence-Specific DNA Binding. *Biochemistry* 41(48):14263-14271.

33. Streaker ED & Beckett D (2003) Coupling of Protein Assembly and DNA Binding: Biotin Repressor Dimerization Precedes Biotin Operator Binding. *Journal of Molecular Biology* 325(5):937-948.

34. Li SJ & Cronan JE (1993) Growth rate regulation of Escherichia coli acetyl coenzyme A carboxylase, which catalyzes the first committed step of lipid biosynthesis. *Journal of Bacteriology* 175(2):332-340.

35. Cronan JE (1988) Expression of the biotin biosynthetic operon of Escherichia coli is regulated by the rate of protein biotination. *Journal of Biological Chemistry* 263(21):10332-10336.

36. Wilson KP, Shewchuk LM, Brennan RG, Otsuka AJ, & Matthews BW (1992) Escherichia coli biotin holoenzyme synthetase/bio repressor crystal structure delineates the biotin- and DNA-binding domains. *Proceedings of the National Academy of Sciences* 89(19):9257-9261.

37. Brennan RG, Vasu S, Matthews BW, & Otsuka AJ (1989) Crystallization of the bifunctional biotin operon repressor. *Journal of Biological Chemistry* 264(1):5.

38. Weaver LH, Kwon K, Beckett D, & Matthews BW (2001) Corepressor-induced organization and assembly of the biotin repressor: A model for allosteric activation of a transcriptional regulator. *Proceedings of the National Academy of Sciences* 98(11):6045-6050.

39. Wood ZA, Weaver LH, Brown PH, Beckett D, & Matthews BW (2006) Co-repressor Induced Order and Biotin Repressor Dimerization: A Case for Divergent Followed by Convergent Evolution. *Journal of Molecular Biology* 357(2):509-523.

40. DeLano WL (2002) The PyMOL Molecular Graphics System.

41. Tissot G, Douce R, & Alban C (1997) Evidence for multiple forms of biotin holocarboxylase synthetase in pea (Pisum sativum) and in Arabidopsis thaliana: subcellular fractionation studies and isolation of a cDNA clone. *Journal of Biochemistry* 323(1):179-188.

42. Kwon K & Beckett D (2000) Function of a conserved sequence motif in biotin holoenzyme synthetases. *Protein Science* 9(8):1530-1539.

43. Kwon K, Streaker ED, Ruparelia S, & Beckett D (2000) Multiple Disordered Loops Function in Corepressor-induced Dimerization of the Biotin Repressor. *Journal of Molecular Biology* 304(5):821-833.

44. Zhao H, Naganathan S, & Beckett D (2009) Thermodynamic and Structural Investigation of Bispecificity in Protein–Protein Interactions. *Journal of Molecular Biology* 389(2):336-348.

45. Athappilly FK & Hendrickson WA (1995) Structure of the biotinyl domain of acetyl-coenzyme A carboxylase determined by MAD phasing. *Structure* 3(12):1407-1419.

46. Nenortas E & Beckett D (1996) Purification and Characterization of Intact and Truncated Forms of the Escherichia coli Biotin Carboxyl Carrier Subunit of Acetyl-CoA Carboxylase. *Journal of Biological Chemistry* 271(13):7559-7567.

47. Roberts EL*, et al.* (1999) Solution Structures of Apo and Holo Biotinyl Domains from Acetyl Coenzyme A Carboxylase of Escherichia coli Determined by Triple-Resonance Nuclear Magnetic Resonance Spectroscopy. *Biochemistry* 38(16):5045-5053.

48. Yao X, Wei D, Soden C, Summers MF, & Beckett D (1997) Structure of the Carboxy-Terminal Fragment of the Apo-Biotin Carboxyl Carrier Subunit of Escherichia coli Acetyl-CoA Carboxylase. *Biochemistry* 36(49):15089-15100.

49. Bagautdinov B, Matsuura Y, Bagautdinova S, & Kunishima N (2008) Protein Biotinylation Visualized by a Complex Structure of Biotin Protein Ligase with a Substrate. *Journal of Biological Chemistry* 283(21):14739-14750.

50. Weaver LH, Kwon K, Beckett D, & Matthews BW (2001) Competing protein:protein interactions are proposed to control the biological switch of the E coli biotin repressor. *Protein Science* 10(12):2618-2622.

51. Jitrapakdee S & Wallace JC (2003) The biotin enzyme family: conserved structural motifs and domain rearrangements. *Current Protein and Peptide Science* 4: 217-229.

52. Knowles JR (1989) The Mechanism of Biotin-Dependent Enzymes. *Annual Review of Biochemistry* 58(1):195-221.

53. Cronan Jr JE & Waldrop GL (2002) Multi-subunit acetyl-CoA carboxylases. *Progress in Lipid Research* 41(5):407-435.

54. Tong L (2012) Structure and function of biotin-dependent carboxylases. *Cellular and Molecular Life Sciences*:1-29.

55. Davis MS, Solbiati J, & Cronan JE (2000) Overproduction of Acetyl-CoA Carboxylase Activity Increases the Rate of Fatty Acid Biosynthesis in Escherichia coli. *Journal of Biological Chemistry* 275(37):28593-28598.

56. Chapman-Smith A, Morris TW, Wallace JC, & Cronan JE (1999) Molecular Recognition in a Post-translational Modification of Exceptional Specificity. *Journal of Biological Chemistry* 274(3):1449-1457.

57. Soriano A, *et al.* (2006) Escherichia coli acetyl-coenzyme A carboxylase: Characterization and development of a high-throughput assay. *Analytical Biochemistry* 349(2):268-276.

58. Abbott J & Beckett D (1993) Cooperative binding of the Escherichia coli repressor of biotin biosynthesis to the biotin operator sequence. *Biochemistry* 32(37):9649-9656.

59. Eisenstein E & Beckett D (1999) Dimerization of the Escherichia coli Biotin Repressor: Corepressor Function in Protein Assembly. *Biochemistry* 38(40):13077-13084.

60. Otsuka A & Abelson J (1978) The regulatory region of the biotin operon in Escherichia coli. *Nature* 276(5689):689-694.

61. Streaker ED & Beckett D (1999) Ligand-linked Structural Changes in the Escherichia coli Biotin Repressor: The Significance of Surface Loops for Binding and Allostery. *Journal of Molecular Biology* 292(3):619-632.

62. Streaker ED & Beckett D (1998) A map of the biotin repressor-biotin operator interface: binding of a winged helix-turn-helix protein dimer to a forty base-pair site. *Journal of Molecular Biology* 278(4):787-800.

63. Buoncristiani MR, Howard PK, & Otsuka AJ (1986) DNA-binding and enzymatic domains of the bifunctional biotin operon repressor (BirA) of Escherichia coli. *Gene* 44(2-3):255-261.

64. Rusinova E, Ross JBA, Laue TM, Sowers LC, & Senear DF (1997) Linkage between Operator Binding and Dimer to Octamer Self-Assembly of Bacteriophage λ cI Repressor. *Biochemistry* 36(42):12994-13003.

65. Senear DF, *et al.* (1993) The primary self-assembly reaction of bacteriophage cI repressor dimers is to octamer. *Biochemistry* 32(24):6179-6189.

66. Brown PH, Cronan JE, Grøtli M, & Beckett D (2004) The Biotin Repressor: Modulation of Allostery by Corepressor Analogs. *Journal of Molecular Biology* 337(4):857-869.

67. Zhao H, Streaker E, Pan W, & Beckett D (2007) Protein−Protein Interactions Dominate the Assembly Thermodynamics of a Transcription Repression Complex. *Biochemistry* 46(47):13667-13676.

68. Streaker ED & Beckett D (2006) The Biotin Regulatory System: Kinetic Control of a Transcriptional Switch. *Biochemistry* 45(20):6417-6425.

69. Zhao H & Beckett D (2008) Kinetic Partitioning Between Alternative Protein–Protein Interactions Controls a Transcriptional Switch. *Journal of Molecular Biology* 380(1):223-236.

70. Rodionov DA, Mironov AA, & Gelfand MS (2002) Conservation of the Biotin Regulon and the BirA Regulatory Signal in Eubacteria and Archaea. *Genome Research* 12(10):1507-1516.

71. Lombard J & Moreira D (2011) Early evolution of the biotin-dependent carboxylase family. *BMC Evolutionary Biology* 11(1):232.

72. Larkin MA, *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21):2947-2948.

73. Winget JM & Mayor T (2010) The Diversity of Ubiquitin Recognition: Hot Spots and Varied Specificity. *Molecular Cell* 38(5):627-635.

74. Prakash O & Eisenberg MA (1979) Biotinyl 5'-adenylate: corepressor role in the regulation of the biotin genes of Escherichia coli K-12. *Proceedings of the National Academy of Sciences* 76(11):5592-5595.

75. Tron CM*, et al.* (2009) Structural and Functional Studies of the Biotin Protein Ligase from Aquifex aeolicus Reveal a Critical Role for a Conserved Residue in Target Specificity. *Journal of Molecular Biology* 387(1):129-146.

76. Gupta V*, et al.* (2010) Structural Ordering of Disordered Ligand-Binding Loops of Biotin Protein Ligase into Active Conformations as a Consequence of Dehydration. *PLoS ONE* 5(2):e9222.

77. Naganathan S & Beckett D (2007) Nucleation of an Allosteric Response via Ligand-induced Loop Folding. *Journal of Molecular Biology* 373(1):96-111.

78. Gill SC & von Hippel PH (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Analytical Biochemistry* 182(2):319-326.

79. Roark DE (1976) Sedimentation equilibrium techniques: Multiple speed analyses and an overspeed procedure. *Biophysical Chemistry* 5(1–2):185-196.

80. Johnson ML, Correia JJ, Yphantis DA, & Halvorson HR (1981) Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques. *Biophysical Journal* 36(3):575-588.

81. Laue TM (1995) Sedimentation equilibrium as thermodynamic tool. *Methods in Enzymology* 259:427-452.

82. Ingaramo M & Beckett D (2009) Distinct Amino Termini of Two Human HCS Isoforms Influence Biotin Acceptor Substrate Recognition. *Journal of Biological Chemistry* 284(45):30862-30870.

83. Motulsky H & Christopoulos A (2004) *Fitting Models to Biological Data Using Linear and Nonlinear Regression. A Practical Guide to Curve Fitting* (Oxford University Press, New York).

84. Xu Y & Beckett D (1994) Kinetics of Biotinyl-5'-adenylate Synthesis Catalyzed by the Escherichia coli Repressor of Biotin Biosynthesis and the Stability of the Enzyme-Product Complex. *Biochemistry* 33(23):7354-7360.

85. Ingaramo M & Beckett D (2011) Biotinylation, a Post-translational Modification Controlled by the Rate of Protein-Protein Association. *Journal of Biological Chemistry* 286(15):13071-13078.

86. Tungtur S, Skinner H, Zhan H, Swint-Kruse L, & Beckett D (2011) In vivo tests of thermodynamic models of transcription repressor function. *Biophysical Chemistry* 159(1):142-151.

87. Hible G*, et al.* (2005) Calorimetric and Crystallographic Analysis of the Oligomeric Structure of Escherichia coli GMP Kinase. *Journal of Molecular Biology* 352(5):1044-1059.

88. Sato K*, et al.* (2009) Metal-binding loop length and not sequence dictates structure. *Proceedings of the National Academy of Sciences* 106(14):5616-5621.

89. Sakurai K & Goto Y (2002) Manipulating Monomer-Dimer Equilibrium of Bovine beta-Lactoglobulin by Amino Acid Substitution. *Journal of Biological Chemistry* 277(28):25735-25740.

90. Herrou J, Rotskoff G, Luo Y, Roux B, & Crosson S (2012) Structural basis of a protein partner switch that regulates the general stress response of α-

proteobacteria. *Proceedings of the National Academy of Sciences* 109(21):E1415–E1423.

91. Adikaram PR & Beckett D (2012) Functional Versatility of a Single Protein Surface in Two Protein:Protein Interactions. *Journal of Molecular Biology* 419(3–4):223-233.

92. Iannello R (1995) DNase I footprinting using PCR-generated end-labeled DNA probes. *In Vitro Transcription and Translation Protocols,* Methods in Molecular Biology 37:379-391.

93. Wood EJ (1983) Molecular cloning. A laboratory manual by T Maniatis, E F Fritsch and J Sambrook. pp 545. Cold Spring Harbor Laboratory, New York. 1982. ISBN 0-87969-136-0. *Biochemical Education* 11(2):82-82.

94. Brenowitz M (1986) Quantitative DNase footprint titration: a method for studying protein-DNA interactions. *Methods in Enzymology* 130:132-181.

95. Suzuki F*, et al.* (1998) Functional Interactions of Transcription Factor Human GA-binding Protein Subunits. *Journal of Biological Chemistry* 273(45):29302-29308.

96. Pearson JT, Dabrowski MJ, Kung I, & Atkins WM (2005) The central loop of Escherichia coli glutamine synthetase is flexible and functionally passive. *Archives of Biochemistry and Biophysics* 436(2):397-405.

97. Kovalevskiy OV, Solonin AS, & Antson AA (2010) Structural investigation of transcriptional regulator HlyIIR: Influence of a disordered region on protein fold and dimerization. *Proteins: Structure, Function, and Bioinformatics* 78(8):1870-1877.

98. Kimura Y & Tanaka K (2010) Regulatory mechanisms involved in the control of ubiquitin homeostasis. *Journal of Biochemistry* 147(6):793-798.

99. Mukhopadhyay D & Riezman H (2007) Proteasome-Independent Functions of Ubiquitin in Endocytosis and Signaling. *Science* 315(5809):201-205.

100. Martin W (2010) Evolutionary origins of metabolic compartmentalization in eukaryotes. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365(1541):847-855.

101. Ingaramo M & Beckett D (2012) Selectivity in Post-translational Biotin Addition to Five Human Carboxylases. *Journal of Biological Chemistry* 287(3):1813-1822.

102. Sharan SK, Thomason LC, Kuznetsov SG, & Court DL (2009) Recombineering: a homologous recombination-based method of genetic engineering. *Nature Protocols* 4(2):206-223.