

# Coordinate regulation of energy transduction modules in *Halobacterium* sp. analyzed by a global systems approach

Nitin S. Baliga\*, Min Pan, Young Ah Goo, Eugene C. Yi, David R. Goodlett, Krassen Dimitrov, Paul Shannon, Ruedi Aebersold, Wailap Victor Ng, and Leroy Hood

Institute for Systems Biology, 1441 North 34th Street, Seattle, WA 98103-8904

Contributed by Leroy Hood, September 16, 2002

The extremely halophilic archaeon *Halobacterium NRC-1* can switch from aerobic energy production (energy from organic compounds) to anaerobic phototrophy (energy from light) by induction of purple membrane biogenesis. The purple membrane is made up of multiple copies of a 1:1 complex of bacterioopsin (Bop) and retinal called bacteriorhodopsin that functions as a light-driven proton pump. A light- and redox-sensing transcription regulator, Bat, regulates critical genes encoding the biogenesis of the purple membrane. To better understand the regulatory network underlying this physiological state, we report a systems approach using global mRNA and protein analyses of four strains of *Halobacterium* sp.: the wild-type, *NRC-1*; and three genetically perturbed strains: *S9* (*bat+*), a purple membrane overproducer, and two purple membrane deficient strains, *SD23* (a *bop* knockout) and *SD20* (a *bat* knockout). The integrated DNA microarray and proteomic data reveal the coordinated coregulation of several interconnected biochemical pathways for phototrophy: isoprenoid synthesis, carotenoid synthesis, and bacteriorhodopsin assembly. In phototrophy, the second major biomodule for ATP production, arginine fermentation, is repressed. The primary systems level insight provided by this study is that two major energy production pathways in *Halobacterium* sp., phototrophy and arginine fermentation, are inversely regulated, presumably to achieve a balance in ATP production under anaerobic conditions.

The *Halobacterium NRC-1* lives in a harsh environment, 4.5 molar salt solution and intense sunlight, and exhibits phototrophy, the ability to synthesize ATP by using sunlight under anaerobic conditions. Under aerobic conditions, ATP synthesis in *Halobacterium* sp. predominantly occurs via the trichloroacetic acid (TCA) cycle (1). In contrast, under anaerobic conditions, ATP synthesis can occur via at least two pathways: fermentation of arginine (2) and photosynthesis (3). The induction of phototrophic growth triggers the synthesis of bacteriorhodopsin (bR); gas vesicles, which make the cells buoyant and move them toward the surface and light; and the sensory rhodopsin I/halobacterial transducer I signal transduction machinery, which confers selective phototaxis for certain wavelengths of light (4, 5). This metabolic switch is the hallmark of the organism. Phototrophy induces bR, a complex of a protein, bacterioopsin (Bop), and retinal, creating a light-driven proton pump, multiple copies of which are then assembled into a two-dimensional lattice called the purple membrane. This proton pump lattice works in conjunction with a membrane ATP synthetase to generate ATP (3). This simple description of phototrophy belies the complexity of the process. We have applied global genomic and proteomic approaches to its systems analysis.

On the basis of consideration of metabolic pathways and known regulatory networks, a minimum of three biomodules can be proposed to participate in the synthesis of bR (6–8). These include the isoprenoid and carotenoid biomodules, which in tandem synthesize the retinal precursor lycopene, and the bR biomodule, in which lycopene is converted to retinal and assem-

bled with Bop into bR. A biomodule is a group of proteins that execute a particular function. Metabolic biomodules can be bounded by branch points in their pathways. Four genes responsible for aspects of bR synthesis, *bop*, *brp*, *bat*, and *crtB1*, constitute a regulon, a linked cluster of genes that are coregulated (6). *CrtB1* catalyzes a key step in the carotenoid biomodule, and *Brp* and *Bop* participate in key steps in the bR biomodule (6–8). The transcriptional regulator *Bat* coordinates the expression of these three genes and itself through interactions with a common putative transcription factor-binding motif on each gene. *Bat* contains signaling domains for responding to environmental changes in cellular redox-potential (PAS/PAC) and light intensity (GAF) (5, 6, 9), critical parameters for the regulation of phototrophy.

The *Halobacterium NRC-1* genome is organized as a 2,014-kb chromosome and two minichromosomes: pNRC200 and pNRC100, 365 and 191 kb in size, respectively. It contains ≈2,413 predicted protein-coding genes (1). The global discovery approach analyzes a particular system such as phototrophy by comparing various system states (wild type, knockout, or overexpression of individual genes in the system) with regard to global patterns of gene or protein expression. From these global patterns of expression, one can deduce systems-level hypotheses that can then be tested by further genetic or environmental perturbations and subsequent global analyses.

The initial systems approach we use here focuses on three genetic perturbations of genes encoding a regulator and a structural protein for bR synthesis: *bat*, the transcriptional activator that mediates the bR regulatory network; and *bop*, the structural gene for Bop. The *Halobacterium S9* strain has four amino acid substitutions in the PAS/PAC domain of *Bat* leading to *Bat* and purple membrane (bR) overproduction (6). The *S9* mutant overrides the normal physiological control by redox potential of *bat*, and operationally its metabolism is shifted to permanent phototrophy. Purple membrane or bR production is abolished in two different mutants derived from the *S9* strain, by transposable insertion sequence element (ISH1) insertional inactivation of the *bat* and *bop* genes in the *SD20* and *SD23* strains, respectively (10). We will denote the *S9* strain *bat+*, the *SD20* strain *bat-*, the *SD23* strain *bop-*, and the *NRC-1* strain wild type. By microarray analyses, we have analyzed the expression patterns for all 2,413 genes in the wild type and mutants. We have compared the relative changes in patterns of protein expression for ≈11% of the proteome using isotope-coded affinity tag (ICAT) analyses (11) on the simultaneously analyzed *bat+* and *bat-* mutants.

The global mRNA and protein analyses have verified the existence of at least 496 of 1,024 hypothetical genes. The proteomics analyses have revealed that 65% of the protein

Abbreviations: TCA, trichloroacetic acid; bR, bacteriorhodopsin; Bop, bacterioopsin; ICAT, isotope-coded affinity tag; UAS, upstream activator sequence.

\*To whom correspondence should be addressed. E-mail: nbaliga@systemsbiology.org.

concentration changes observed in the *bat+* and *bat-* strains are not reflected at the mRNA levels and, accordingly, many of these may be regulated by posttranscriptional control mechanisms, stressing the potential importance of posttranscriptional control mechanisms.

The most significant outcomes from these global discovery approaches were two important systems-level insights that have led to new hypotheses. First, the transcription factor Bat coordinately and directly controls key enzymes in the carotenoid and bR biomodules and indirectly regulates the isoprenoid and other biomodules. Second, the arginine fermentation biomodule, on the one hand, and the carotenoid and bR biomodules, on the other, are inversely regulated, presumably to provide an optimal steady-state level of ATP under anaerobic conditions.

## Materials and Methods

**Strains and Culturing.** *Halobacterium NRC-1* is the wild-type strain. The *bat+* strain overproduces purple membrane because of mutations in the *bat* gene. The *bat-* and *bop-* strains are knockouts of *bat* and *bop*, respectively, and neither produces purple membrane. Culturing of all strains was done at 37°C in complex halobacterial medium containing trace metals in the presence of full-spectrum white light, as described (10).

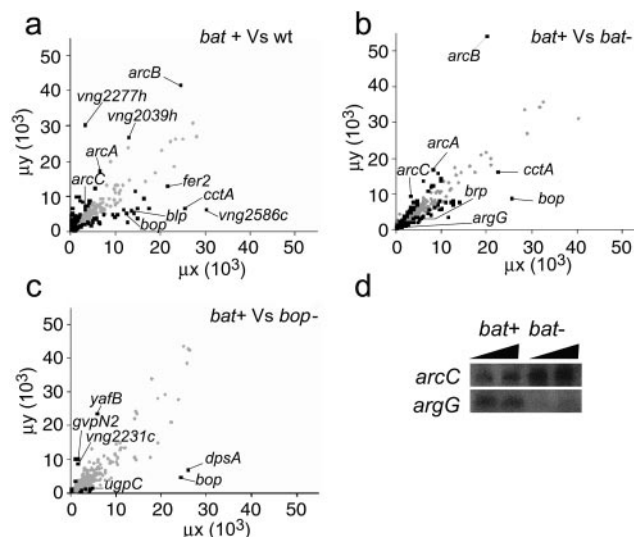
**Microarray Analysis.** All of the 2,413 protein-coding genes encoded in *Halobacterium NRC-1* genome were PCR-amplified; the purified PCR products were resuspended in 50% DMSO at 150 ng/μl and spotted by using a Virtek robot (Virtek Vision, Waterloo, ON, Canada) in quadruplicate on to polyamine glass slides (TeleChem, Sunnyvale, CA).

Ten micrograms each of total DNA-free RNA from two samples (RNA1 and RNA2), prepared by using the Absolutely-RNA kit (Stratagene, La Jolla, CA), were labeled with Alexa 594 (dye1) and Alexa 660 (dye2) (Molecular Probes) and purified through a Micro Bio-Spin P-30 column (Bio-Rad). A flip-dye experiment, wherein RNA1 and -2 were labeled with dye2 and -1, respectively, was also performed to rule out a bias in dye incorporation. The hybridization was performed overnight at 42°C in 40 μl of hybridization buffer (Amersham Pharmacia Biosciences) under coverglass, in sealed chambers (Corning). After hybridization, the slides were processed through three successive washes with 2× SSC + 0.2% SDS, 2× SSC, and 0.2× SSC, 15 min each at 42°C. Slides were scanned with ScanArray (Perkin-Elmer), and spots in microarray images were extracted by using AnalyzerDG (MolecularWare, Cambridge, MA).

The raw intensity data were processed, and the statistical significance of differential expression of each gene was calculated by maximum likelihood analysis (12). To estimate the level of differential expressions of genes due to inherent biological variation, total RNA from four identically grown cultures of *Halobacterium NRC-1* were interrogated against each other. The statistical significance for differential expression due to biological variation was calculated by using maximum likelihood analysis, and the maximum value of the test statistic  $\lambda$  for this control was determined to be 18. Therefore, differential expression of genes in experimental analysis was considered significant for  $\lambda > 18$ .

**Northern Analysis.** Northern blot analysis was performed on 5 and 10 μg of total RNA from *bat+* and *bat-* cultures (OD<sub>600</sub> = 1.4) by using the Northernmax kit (Ambion, Austin, TX). The probes were gene-specific PCR fragments <sup>32</sup>P-labeled using the Prime-It II Random Primer Labeling Kit (Stratagene).

**Differential Proteomics.** Total protein extracts were prepared from *bat+* and *bat-* cultures (OD<sub>600</sub> = 1.4), as described (13). ICAT labeling and mass spectrometer analysis on total proteins were conducted as described by Han *et al.* (14), with minor modifi-



**Fig. 1.** Microarray analysis of wild-type (*NRC-1*) *Halobacterium* sp. and three mutants: *S9* (*bat* overexpression), *SD20* (*bat* knockout), and *SD23* (*bop* knockout). Scatter plots on average median-normalized intensities ( $\mu_x$  and  $\mu_y$ ) for three microarray comparisons: (a) *bat+* vs. wild type; (b) *bat+* vs. *bat-*; and (c) *bat+* vs. *bop-*. Representative data points (black dots) with significant differential expression ( $\lambda > 18$ ) are labeled. (d) Northern blot analysis on 5 (lanes 1 and 3) and 10 μg (lanes 2 and 4) of total RNA for two representative genes involved in arginine degradation (*arcC*) and arginine synthesis (*argG*).

cations. Total protein (2.5 mg) was denatured with 6 M urea and 0.05% SDS and immediately reduced with 5 mM tributylphosphine. Cysteine residues were selectively labeled with a 2-fold molar excess of either light (d0) (*bat-*) or heavy (d8) ICAT (*bat+*) (Applied Biosystems). d0- and d8-ICAT labeled proteins were mixed in 1:1 ratio and digested with trypsin.

The tryptic peptides were analyzed by liquid chromatography–electrospray ionization–tandem mass spectrometry. Tandem mass spectra for selected peptide ions were searched against the *Halobacterium NRC-1* sequence database by using the SEQUEST algorithm, and relative expression ratios were calculated by using the EXPRESS software tool (14). A second experiment wherein the labeling of the two protein samples was flipped with respect to the normal and heavy ICAT reagents was performed to rule out a bias in isotope incorporation.

**Promoter Analysis.** Upstream regions (–500 to +1 bp) for selected genes were scanned for conserved motifs with MEME software (15). Motif alignments were further analyzed by Sequence Logo Analysis (16, 17).

## Results

**Transcriptome Analysis.** The effects of the genetic perturbations in *bat* (overexpression and knockout) and *bop* (knockout) on the transcription of 2,413 genes encoded in the *Halobacterium NRC-1* genome were analyzed by DNA microarray analysis. Each sample pair was analyzed four times independently, so that reliable statistical analyses could be carried out. As little as a 50% change in expression patterns could be detected in a statistically significant manner by this method (12). The wild-type and the *bat-* and *bop-* mutant total RNAs were compared against the *bat+* mutant total RNA, because the *bat-* and *bop-* strains were derived from *bat+* strain. The median-normalized signal intensities for the two RNA populations in each of the three comparisons are plotted (Fig. 1 a–c).

**Comparisons of *bat+* and Wild-Type Strains.** The mutations in Bat, which result in overproduction of the purple membrane in the

**Table 1. Subtractive mRNA analysis for various biomodules relevant to phototrophy and ATP production**

Biomodule	Gene	<i>bat</i> <sup>+</sup> / <i>wt</i>	<i>bat</i> <sup>-</sup> / <i>bat</i> <sup>+</sup> *	<i>bop</i> <sup>-</sup> / <i>bat</i> <sup>+</sup>
bR regulon and linked genes	<i>bat</i>	+	-	0
	<i>bop</i>	+	-	-
	<i>brp</i>	+	-	0
	<i>crtB1</i>	+	-/-	0
	<i>blp</i>	+	-	0
bR synthesis	<i>1459H</i>	0	0/-	0
	<i>bop</i>	+	-	-
Isoprenoid synthesis	<i>brp</i>	+	-	0
	<i>dmd</i>	0	0/+	0
	<i>mvaA</i>	0	0/+	0
	<i>mvk</i>	0	0/+	0
Carotenoid synthesis	<i>idi</i>	0	0/+	0
	<i>crtB1</i>	+	-/-	0
	<i>crtl2</i>	+	-/-	0
Halorhodopsin synthesis	<i>hop</i>	0	-	0
Arginine synthesis	<i>argG</i>	0	-/-	0
	<i>argH</i>	0	0/-	0
Arg-tRNA synthesis	<i>argS</i>	-	+/+	0
Arginine fermentation	<i>arcA</i>	-	+/+	0
	<i>arcB</i>	-	+/+	0
	<i>arcC</i>	-	+/+	0
AA transporter	<i>yhdG</i>	0	0/+	0
Glutamate metabolism	<i>gdhB</i>	0	0/+	0
	<i>glnA</i>	0	0/+	0
Pyrimidine metabolism	<i>carA</i>	-	0	0
	<i>carB</i>	0	0/+	0
	<i>pyrB</i>	0	+	0
	<i>pyrE2</i>	0	+	0
	<i>nrdB2</i>	0	0/+	0
	<i>trxA2</i>	0	0/+	0
	<i>acc</i>	0	0/+	0
TCA cycle	<i>fer2</i>	+	0/+	0
	<i>icd</i>	0	0/+	0
	<i>korA</i>	+	0/+	0
	<i>korB</i>	+	0/0	0
	<i>porB</i>	0	-	0
	<i>sdh</i>	0	+/+	0
	<i>sdhB</i>	0	+/+	0
	<i>sdhC</i>	0	+	0
	<i>1524C</i>	0	0/+	0

0, unchanged; +, increased; -, decreased.

\*Comparison mRNA/protein.

*bat*<sup>+</sup> strain, were associated with induction of 66 genes and repression of 85 genes relative to the wild-type *NRC-1* strain (Fig. 1a; Table 2, which is published as supporting information on the PNAS web site, www.pnas.org). Of these 151 genes, 98 were of unknown function. The activities of the remaining 53 genes were associated with at least eight cell processes, including three biomodules associated with phototrophy: the purple membrane (bR) biogenesis and the carotenoid synthesis biomodules were induced, and the arginine fermentation biomodule was repressed (Table 1). Consistent with the purple membrane overexpression phenotype, the four bR regulon genes were induced in *bat*<sup>+</sup> with respect to the wild type (6), suggesting that *bat* also controls the expression of key genes in the bR biomodule (increased *bop*, *brp*) and the carotenoid biomodule (increased *crtB1*). In contrast, the genes controlling the arginine fermenta-

tion biomodules were down-regulated (*arcA*, *arcB*, and *arcC*) in the *bat*<sup>+</sup> strain, as were individual genes in the glutamate and pyrimidine biomodules (Table 1).

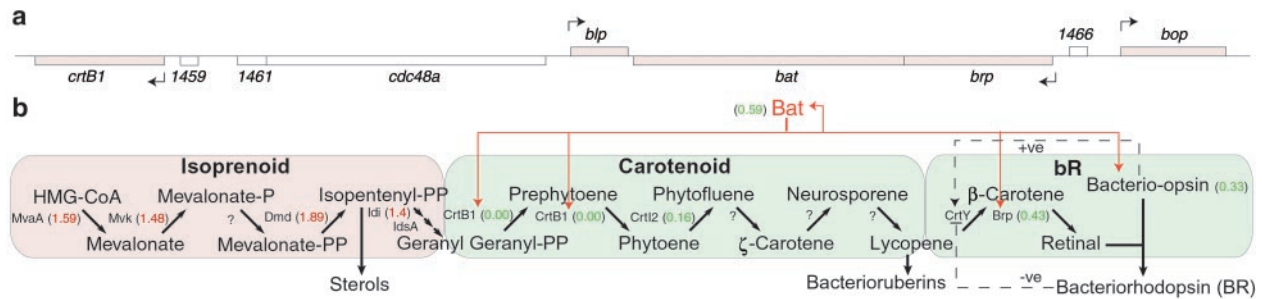
**Comparisons of *bat*<sup>+</sup> and *bat*<sup>-</sup> Strains.** Insertional inactivation of the *bat* gene resulted in induction of 62 genes and repression of 95 genes relative to the parent *bat*<sup>+</sup> strain (Fig. 1b; Table 2). Of these 157 genes, 75 were of unknown function. Thirty-one of the 157 genes were also differentially regulated in *bat*<sup>+</sup> vs. wild type. The biomodules associated with phototrophy in the *bat*<sup>-</sup> strain behaved in an inverse manner to their *bat*<sup>+</sup> counterparts: the purple membrane biosynthesis and the carotenoid synthesis biomodules were suppressed, and the arginine fermentation biomodule was induced (Table 1). The repressed levels of all four purple membrane regulon genes in the *bat*<sup>-</sup> strain were consistent with depletion of bR resulting from inactivation of *bat* (6, 10). A gene involved in arginine synthesis (*argG*) was repressed, and the gene encoding an arginyl tRNA synthetase (*argS*) was induced in the *bat*<sup>-</sup> strain. Several enzymes in pyrimidine metabolism were induced, as were several TCA enzymes (Table 1). Both Northern blot and RT-PCR analyses of several of these genes confirmed their differential expression in the *bat*<sup>-</sup> vs. *bat*<sup>+</sup> strains (Fig. 1d).

**Comparisons of *bop*<sup>-</sup> and *bat*<sup>+</sup> Strains.** In stark contrast, only 15 genes were induced and 8 repressed in the *bop*<sup>-</sup> strain relative to the parent *bat*<sup>+</sup> strain (Fig. 1c; Table 2). Of all the genes analyzed in the various biomodules related to phototrophy, only one (*bop*) changed. Of these 23 genes, 8 also changed in *bat*<sup>+</sup>/*bat*<sup>-</sup> comparison and only three in *bat*<sup>+</sup>/wild type comparison. With respect to purple membrane biogenesis, only expression of the *bop* gene was repressed in the *bop*<sup>-</sup> strain, suggesting that Bop does not participate in transcriptional regulation of the carotenoid biosynthesis enzymes (Table 1). Similarly, *bop* inactivation did not affect the transcriptional regulation of genes encoding the arginine synthesis and fermentation biomodules, implying that the regulation of these genes was also tied to the function of Bat and not bR. Clearly, perturbing the transcriptional regulator Bat had significantly more effect on the global transcription network for phototrophy than perturbing the structural protein Bop, implying that Bat directly or indirectly controlled a variety of interrelated phototrophy biomodules.

**Proteome Analysis.** We used the ICAT approach to analyze the patterns of the protein expression in the *bat*<sup>-</sup> and *bat*<sup>+</sup> strains, because we were interested in comparing the patterns of mRNA and protein expression (11). Total proteins were isolated from the *bat*<sup>+</sup> and *bat*<sup>-</sup> strains and labeled, respectively, with the heavy and light ICAT reagent and then analyzed by mass spectrometry. The ICAT procedure is capable of detecting as little as a 30% change in protein levels (14). A total of 1,120 tryptic peptides, containing cysteine-attached light and heavy ICAT isotopes, were analyzed from 272 different proteins. In the wild-type strain, 1,576 of 2,413 proteins have cysteine-containing tryptic peptide(s) in the mass range of 700–4,000 Da. Hence, we have analyzed 11% of the proteome (272/2,413). At least 50 of these proteins were differentially expressed. Of these 50 proteins, only 7 encoded unknown functions, and only 17 changed in a corresponding manner at the mRNA level (Fig. 2; Table 3, which is published as supporting information on the PNAS web site).

In the *bat*<sup>-</sup> mutant, the induced proteins related directly or indirectly to phototrophy, including those involved with isoprenoid synthesis, arginine fermentation, glutamate metabolism, and pyrimidine metabolism biomodules; proteins in the carotenoid synthesis and arginine synthesis biomodules and one protein encoded within the bR regulon were suppressed. Several



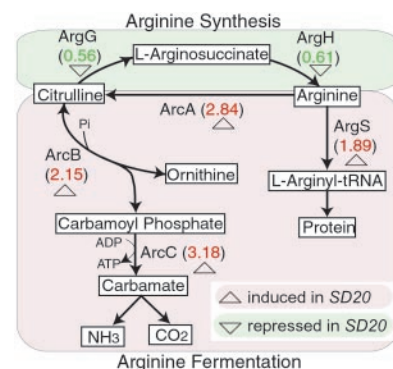


**Fig. 4.** The bR regulon and the isoprenoid, carotenoid, and bR biomodules in *Halobacterium* sp. (6, 8, 21). (a) The bR regulon. Genes shaded in red are transcriptionally regulated by Bat. (b) The isoprenoid, carotenoid, and bR biomodules. Pink shading indicates up-regulation, and green denotes down-regulation in the *bat*<sup>-</sup> strain relative to the *bat*<sup>+</sup> strain. The intermediates and enzymes catalyzing the various biochemical conversions are indicated with ratios of mRNA or protein levels in the *bat*<sup>-</sup> strain relative to the *bat*<sup>+</sup> strain in parentheses. A question mark (?) indicates that the enzyme catalyzing the reaction has yet to be identified. The multiple arrows from IPP to GGPP indicate this conversion requires several steps (22). The red lines and arrows indicate enzymatic steps regulated transcriptionally by Bat. The dashed lines within the bR biomodule indicate positive (+ve) or negative (-ve) feedback loops controlling conversion of lycopene to  $\beta$ -carotene (7).

**An Analysis of Transcriptional and Posttranscriptional Levels of Control.** The comparison of the *bat*<sup>+</sup> and *bat*<sup>-</sup> strains for changes in global expression of both mRNA and proteins permits the identification of genes and their proteins whose expression patterns change in a similar (coordinate) manner (up or down), consistent with transcriptional control, and those whose patterns are not coordinated, implying some type of posttranscriptional control. About 50 of 272 proteins analyzed changed in the comparison of the *bat*<sup>+</sup> and *bat*<sup>-</sup> strains. If the remainder of the 2,413 wild-type proteins behaved in a similar manner, 18% of the proteome would be changed. Of these 50 proteins, 17 demonstrated corresponding changes in their mRNAs, and 33 did not (Fig. 2). Most of these proteins on average had adequate levels of mRNA so that changes should have been detected. Hence, almost two-thirds of the protein changes could be posttranscriptionally regulated. Nearly 25% of the differentially expressed proteins were associated with biomodules directly or indirectly related to phototrophy: isoprenoid synthesis, carotenoid synthesis, arginine fermentation, arginine synthesis, glutamate metabolism, and pyrimidine metabolism. Clearly, components of the TCA cycle change in these perturbations, but the interface with phototrophy is complex, because both the *bat*<sup>+</sup> and *bat*<sup>-</sup> mutants induce TCA-cycle proteins.

**The *bat* Transcription Factor Has Far-Reaching Effects on the Expression of Biomodules in the *Halobacterium* sp.** As noted earlier, the *bat*<sup>+</sup> mutant overrides the control by redox potential of the *bat* gene, and operationally its increased expression level is associated with a permanent phototrophy state. Approximately 7% of the *Halobacterium* sp. genes change their expression as a result of this perturbation in *bat* function (vs. wild type). The *bat* knockout (vs. *bat*<sup>+</sup>) also changed 7% of its gene expression patterns as contrasted with a 1% mRNA expression change for the *bop* knockout (vs. *bat*<sup>+</sup>). Moreover, 18% (50/272) of the proteins sampled in the *bat*<sup>-</sup> vs. *bat*<sup>+</sup> strain comparisons were perturbed. Bat controls the coordinated expression of itself and the three other genes in the bR regulon through the UAS motif (6). Thus, the *bat* transcription factor presumably regulates only a few genes directly, genes that play key roles in the bR (*bop*, *brp*) and the carotenoid (*crtB1*) biomodules. The changes in gene expression for the remainder of the genes arise from indirect (downstream effects) in either a positive or negative sense. These indirect changes are reflected in a variety of other biomodules, some obviously related to phototrophy and others to general metabolism (TCA cycle) (Table 1). These data emphasize the importance of integrating data from both genes and their regulatory networks.

**The Mutations of Genes of the bR Regulon Provide Systems-Level Hypotheses Concerning the Interrelationships of the Biomodules Participating in Phototrophy and Establish Their Links to Other Aspects of Metabolism.** The integration of quantitative changes in the mRNA and protein levels in the *Halobacterium* wild type and the three mutants reveals a coherent picture of the regulation of the biomodules mediating phototrophy and their interactions with other biomodules of metabolism (Table 1). Phototrophy involves the induction of the bR regulon. The bR regulon encodes one structural protein, Bop; two enzymes, Brp and CrtB1; and a transcription factor, Bat, that regulates each of these genes and itself (Fig. 4). Under anaerobic conditions, *bat* is up-regulated, and phototrophy is induced. The *bat*<sup>+</sup> mutant has the *bat* gene permanently up-regulated, so it is possible to study phototrophy under both aerobic and anaerobic conditions. Two biomodules participate directly in bR synthesis, the carotenoid biomodules provide the precursor lycopene, and in the bR biomodule, lycopene is converted to retinal, which complexes with Bop to produce bR. The other members of the regulon control key reactions in these biomodules, *crtB1* in the carotenoid biomodule, and *bop* and *brp* in the bR biomodule (Fig. 4). As indicated in Table 1, the overexpression of *bat* in the *bat*<sup>+</sup> strain induces all four regulon genes and, correspondingly, presumably increases the activity of the carotenoid and bR modules to generate the purple membrane. Two genes with related activities are also induced: *hop*, encoding halorhodopsin, a chloride pump that works in the opposite direction from bR



**Fig. 5.** The arginine synthesis and fermentation biomodule in *Halobacterium* sp. The enzymes are indicated, with message or protein expression ratios in the *bat*<sup>-</sup> to *bat*<sup>+</sup> strains in parentheses. Green shading indicates down-regulation and pink denotes up-regulation in the *bat*<sup>-</sup> to *bat*<sup>+</sup> strain comparisons.

proton pump to maintain intracellular pH homeostasis and prevent hyperpolarization of the membrane; and *crtI2*, another key gene in the carotenoid biomodule. It is striking to note that the genes of the arginine fermentation biomodule, *arcA*, *arcB*, and *arcC*, are down-regulated in keeping with the fact the bR biomodule is generating significant ATP, and presumably a balance in ATP production is maintained by down-regulating the second major source of ATP under anaerobic conditions, arginine fermentation. Several genes in the TCA cycle, *korA* and *korB*, encoding the subunits of 2-oxoglutarate ferredoxin reductase and *fer2*, encoding ferredoxin, are induced presumably providing a complex interface between these two energy-producing biomodules.

The *bat* knockout strain, *bat*<sup>-</sup>, in contrast, suppresses the bR regulon genes and the arginine synthesis, the carotenoid synthesis, and the bR biomodules (Table 1). It up-regulates the genes of arginine fermentation presumably to maintain ATP production with the loss of bR-mediated ATP production. It also up-regulates the arg-tRNA synthetase gene *argS*, presumably to capture sufficient arginine to maintain protein synthesis (Fig. 5), and an amino acid transporter *yhdG*, which may transport more arginine into the cell. Seven genes of the TCA cycle (*acc*, *icd*, *korA*, *fer2*, *sdh*, *sdhB*, and *sdhC*) are also up-regulated, presumably forming an impedance match between arginine fermentation and the TCA cycle. However, the up-regulation of the *korA*, *korB*, and *fer2* genes, but the absence of differential regulation of other genes involved with the TCA cycle in comparing the *bat*<sup>+</sup> strain to the wild type, suggests that the TCA cycle has complex interfaces with phototrophy and arginine fermentation. On the other hand, the up-regulation of genes involved in glutamate metabolism (*gdhB*, *glnA*, and *carB*) and pyrimidine metabolism (*carB*, *pyrB*, *pyrE2*, *nrdb2*, and *txa2*) appears to be related to an increase in arginine fermentation. The up-regulation of the carbamoyl phosphate synthetase subunit (CarB) may compensate for the breakdown of carbamoyl phosphate by carbamate kinase (ArcC) for ATP production during arginine fermentation (Fig. 5). The carbamoyl phosphate is then presumably captured by aspartate carbamoyl-transferase catalytic subunit (PyrB) for pyrimidine metabolism. Hence, an increase in arginine fermentation in the *bat*<sup>-</sup> mutant appears to be responsible for induction of genes in at least two biomodules: pyrimidine synthesis and glutamate metabolism. Thus, a variety of systems-level hypotheses have been generated that can be tested through additional hypothesis-driven perturbations and global analyses.

**Posttranscriptional Control of the Isoprenoid Biomodule.** It is striking that the levels of four genes in the isoprenoid biomodule, *mvaA*, *mvk*, *dmd*, and *idi*, do not change in their expression patterns in the mRNA level in the *bat*<sup>+</sup> and *bat*<sup>-</sup> strain comparisons but are up-regulated at the protein levels in the *bat*<sup>-</sup> strain (Fig. 4b). The

induction of the isoprenoid synthesis biomodule is perhaps a consequence of the repression of the carotenoid synthesis and the bR synthesis biomodules in the *bat*<sup>-</sup> strain. The mechanism of this posttranscriptional regulation is clearly complex and not fully understood. Two mechanisms of posttranscriptional regulation (positive and negative) have already been proposed to regulate the conversion of lycopene to  $\beta$  carotene (7). Together with the Bat-mediated transcriptional regulation of the bR regulon genes, these posttranscriptional regulatory mechanisms might optimally coordinate the isoprenoid, carotenoid, and bR synthesis biomodules on induction of phototrophy in the *Halobacterium* sp. (Fig. 4b).

**The Regulatory Network for Phototrophy Includes at Least Two Transcription Factors.** We identified a potential transcription factor-binding motif associated with five genes, two of which (*arcB* and *arcC*) are associated with arginine fermentation. The activity of this transcription regulator is presumably indirectly suppressed by Bat, because the genes encoding arginine fermentation are repressed on overproduction of Bat in the *bat*<sup>+</sup> strain and induced on depletion of Bat in the *bat*<sup>-</sup> strain. The close proximity of *arcR*, which encodes a transcription regulator, to the *arcBCA* gene cluster on the pNRC200 minichromosome, suggests it may be a candidate member of this regulatory network. Unlike *arcA*, *arcB*, and *arcC*, *arcR* is not transcriptionally regulated and has similar mRNA levels in both aerobically and anaerobically grown cultures (2). Perhaps it is activated post-transcriptionally.

The systems-level approaches reported here have revealed fundamental new insights and hypotheses about phototrophy in *Halobacterium* sp. The biomodules for ATP production, bR, and arginine fermentation are coordinately and inversely regulated to maintain a balance of ATP production under anaerobic conditions. Furthermore, the redox and light-sensitive transcription factor, Bat, regulated key enzymes in the carotenoid and bR biomodules. A second unknown transcription factor may regulate the arginine fermentation biomodule in a similar manner. Thus, transcription factors may play a key role in balancing anaerobic ATP production. Survival in high salinity can be energetically taxing, because it requires the buildup and maintenance of steep ion concentration gradients across the cell membrane (20). Therefore, the presence of such networks for regulating energy transduction in *Halobacterium* sp. might be crucial for its survival in its natural, extreme, and dynamic habitats. Thus, the systems-level insights reported here highlight the importance of integrated global genomic and proteomic analyses.

We thank Michael J. Danson and Mark P. Krebs for helpful suggestions. This work was supported by funds from Merck & Inc. to the Institute for Systems Biology.

- Ng, W. V., Kennedy, S. P., Mahairas, G. G., Berquist, B., Pan, M., Shukla, H. D., Lasky, S. R., Baliga, N. S., Thorsson, V., Sbrogna, J., et al. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 12176–12181.
- Ruepp, A. & Soppa, J. (1996) *J. Bacteriol.* **178**, 4942–4947.
- Oesterheld, D. & Stoekenius, W. (1973) *Proc. Natl. Acad. Sci. USA* **70**, 2853–2857.
- Hoff, W. D., Jung, K. H. & Spudich, J. L. (1997) *Annu. Rev. Biophys. Biomol. Struct.* **26**, 223–258.
- Yang, C. F. & DasSarma, S. (1990) *J. Bacteriol.* **172**, 4118–4121.
- Baliga, N. S., Kennedy, S. P., Ng, W. V., Hood, L. & DasSarma, S. (2001) *Proc. Natl. Acad. Sci. USA* **98**, 2521–2525.
- Deshpande, A. & Sonar, S. (1999) *J. Biol. Chem.* **274**, 23535–23440.
- Peck, R. F., Echavarri-Erasun, C., Johnson, E. A., Ng, W. V., Kennedy, S. P., Hood, L., DasSarma, S. & Krebs, M. P. (2001) *J. Biol. Chem.* **276**, 5739–5744.
- Gropp, F. & Betlach, M. C. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 5475–5479.
- Yang, C. F., Kim, J. M., Molinari, E. & DasSarma, S. (1996) *J. Bacteriol.* **178**, 840–845.
- Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H. & Aebersold, R. (1999) *Nat. Biotechnol.* **17**, 994–999.
- Ideker, T., Thorsson, V., Siegel, A. F. & Hood, L. E. (2000) *J. Comput. Biol.* **7**, 805–817.
- DasSarma, S. & Fleischmann, E. M. (1995) *Halophiles* (Cold Spring Harbor Lab. Press, Plainview, NY).
- Han, D. K., Eng, J., Zhou, H. & Aebersold, R. (2001) *Nat. Biotechnol.* **19**, 946–951.
- Bailey, T. L. & Elkan, C. (1994) in *Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology* (American Association for Artificial Intelligence, Menlo Park, CA), pp. 28–36.
- Schneider, T. D. & Stephens, R. M. (1990) *Nucleic Acids Res.* **18**, 6097–6100.
- Schneider, T. D. (1996) *Methods Enzymol.* **274**, 445–455.
- Gropp, F., Gropp, R. & Betlach, M. C. (1995) *Mol. Microbiol.* **16**, 357–364.
- Delcher, A. L., Harmon, D., Kasif, S., White, O. & Salzberg, S. L. (1999) *Nucleic Acids Res.* **27**, 4636–4641.
- Oren, A. (1999) *Microbiol. Mol. Biol. Rev.* **63**, 334–348.
- Peck, R. F., Johnson, E. A. & Krebs, M. P. (2002) *J. Bacteriol.* **184**, 2889–2897.
- Sumper, M., Reitmeier, H. & Oesterheld, D. (1976) *Angew. Chem. Int. Ed. Engl.* **15**, 187–194.