Molecular BioSystems

RSC Publishing

PAPER

View Article Online
View Journal | View Issue

The proteomic response to mutants of the *Escherichia coli* RNA degradosome†

Cite this: *Mol. BioSyst.*, 2013, **9**, 750

Li Zhou,‡[ab] Ang B. Zhang,‡[c] Rong Wang,[ad] Edward M. Marcotte[a] and Christine Vogel*[ac]

The *Escherichia coli* RNA degradosome recognizes and degrades RNA through the coordination of four main protein components, the endonuclease RNase E, the exonuclease PNPase, the RhlB helicase and the metabolic enzyme enolase. To help our understanding of the functions of the RNA degradosome, we quantified expression changes of >2300 proteins using mass spectrometry based shotgun proteomics in *E. coli* strains deficient in *rhlB*, *eno*, *pnp* (which displays temperature sensitive growth), or *rne*(1-602) which encodes a C-terminal truncation mutant of RNase E and is deficient in degradosome assembly. Global protein expression changes are most similar between the *pnp* and *rhlB* mutants, confirming the functional relationship between the genes. We observe down-regulation of protein chaperones including GroEL and DnaK (which associate with the degradosome), a decrease in translation related proteins in Δ*pnp*, Δ*rhlB and rne*(1-602) cells, and a significant increase in the abundance of aminoacyl-tRNA synthetases. Analysis of the observed proteomic changes points to a shared motif, CGCTGG, that may be associated with RNA degradosome targets. Further, our data provide information on the expression modulation of known degradosome-associated proteins, such as DeaD and RNase G, as well as other RNA helicases and RNases – suggesting or confirming functional complementarity in some cases. Taken together, our results emphasize the role of the RNA degradosome in the modulation of the bacterial proteome and provide the first large-scale proteomic description of the response to perturbation of this major pathway of RNA degradation.

Received 10th November 2012,
Accepted 23rd January 2013

DOI: 10.1039/c3mb25513a

www.rsc.org/molecularbiosystems

Introduction

The RNA degradosome is a vital component of mRNA degradation in *Escherichia coli*[1] and a main contributor to post-transcriptional gene regulation. The catalytic core enzyme of the RNA degradosome, the endonuclease RNase E (RNase E), is conserved amongst proteobacteria, but homologs have also been identified in archaea and plants.[1] It consists of an N-terminal catalytic region and a C-terminal non-catalytic region which provides a scaffold for binding of other proteins, RNA or other cellular components (Fig. 1). The canonical RNA degradosome in *E. coli* is a 500–700 kDa complex formed by the assembly of polynucleotide phosphorylase (PNPase, the product of the *pnp* gene) which functions as an exonuclease, the DEAD-box helicase RhlB, and the glycolytic enzyme enolase (Eno) on the C-terminal end of the RNase E.[1] Cells lacking RNase E are not viable, most likely due to defects in tRNA transcript processing.[2] However, truncation mutants lacking the C-terminal scaffolding domain and thus incapable of assembling the degradosome complex are viable although they exhibit slow growth and temperature sensitivity.[2] RNase III (Rnc) and RNase G (Rng) can complement *rne* mutants whereas RNase R (Rnr) and RNase II (Rnb) can substitute for PNPase.[1]

Several other proteins have been suggested to associate with the degradosome or play a role in RNA degradation. For example, the poly(A) polymerase PcnB interacts with the DEAD-box helicases SrmB, CsdA and RhlE.[3] RNase E co-precipitates with Ppk, DnaK and GroEL in sub-stoichiometric amounts relative to the major degradosome components.[1] Pull-down experiments have further identified a number of proteins bound to degradosome components (*e.g.* see Fig. 3), but the role of these interactions is unclear.[4]

[a] *University of Texas at Austin, Center for Systems and Synthetic Biology, Institute for Cellular and Molecular Biology, Austin, TX, USA. E-mail: cvogel@nyu.edu*
[b] *Department of Molecular Biology, Department of Chemistry and The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA 92037, USA*
[c] *New York University, Department of Biology, Center for Genomics and Systems Biology, New York, NY, USA*
[d] *National Heart Lung and Blood Institute, NIH, Bethesda, Maryland, USA*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c3mb25513a

‡ These authors contributed equally.

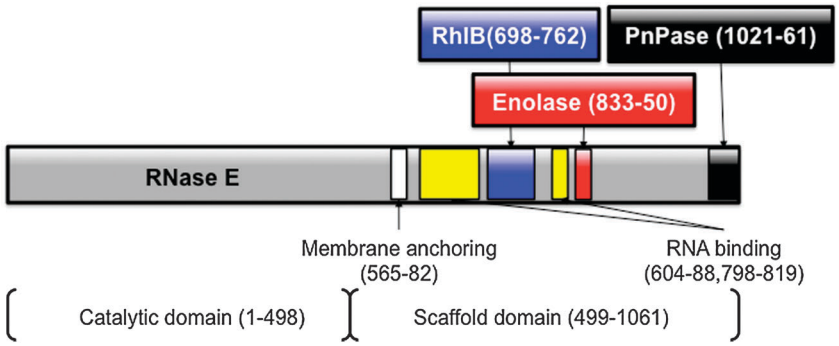

Fig. 1 The *Escherichia coli* RNA degradosome complex. The figure shows the four main RNA degradosome components, RNase E, Enolase, RhlB, and PNPase. RNase E comprises a catalytic N-terminal region and a non-catalytic C-terminal region to which the other three proteins, RhlB, Eno and Pnp, bind. In addition to interacting with RNase E, RhlB and Pnp also physically interact with each other.[7] The RNase E truncation mutant contains only the catalytic N-terminal domain; the other three mutants are gene deletions.

While the physiological significance of the *E. coli* degradosome is well-established, there are still many unresolved questions regarding its function.[1] For example, the exact role of the RNA degradosome in modulating mRNA decay and thus protein abundance *in vivo* is unclear.[5] The significance of the enolase binding to the C-terminal half of RNase E has not been defined apart from its role in modulating the mRNA stability of the glucose transporter ptsG during metabolic stress.[6] Enolase does not bind RNA, and only 5–10% of the total enolase in cell extracts is complexed with the degradosome.[5] RhlB and PNPase can also form a complex that degrades double-stranded RNA independent of RNase E.[7] Furthermore, it is not known how the individual component enzymes cooperate in degrading RNA or how changes in the degradosome composition affect the decay of different transcripts.[8]

To answer some of these questions at a systems level, several studies have analyzed changes in transcript abundance and mRNA half-lives in *E. coli* strains deficient in degradosome components, namely in cells carrying null mutations in *pnp*, *eno* or *rhlB* or an RNase E C-terminal truncation that abrogates degradosome assembly (*rne*(1-602)).[9–12] Interference with degradosome function may lead to accumulation of un-degraded RNAs in the cell and an increase in their RNA half-lives. However, transcript studies observed a wide range in responses, with a substantial number of transcripts that decreased in half-lives or did not change in concentration,[11] suggesting that the regulation of RNA degradation is complex.

The ultimate effect of mutations that influence mRNA half-life is a change in the steady state proteome. Here we present a large-scale quantitative analysis of changes in the proteome of strains lacking the canonical degradosome components and compare our data with the published information on the mRNA half-lives in these mutants. Deep protein abundance analysis was carried out using high-resolution mass spectrometry and a label-free quantitation method developed in our lab[13,14] to estimate the concentration changes of >2300 proteins. These data provide novel information on how the function of the degradosome shapes the *E. coli* proteome.

Methods

Data collection

E. coli strains carrying *pnp*, *rhlB*, *eno* mutations or *rne*(1-602) were kindly provided by Dr S. N. Cohen (Stanford University)[10] (Table 1). Cells were grown at 30 °C in M9 media supplemented with 0.2% tryptone, 0.2% glycerol, 1 mM $MgSO_4$ and 0.0001% (w/v) thiamine. 40 mM succinate was added to strains K10 and DF261 as described in ref. 10. Cell pellets were harvested when the culture absorbance reached A_{600} = 0.6, the cells were resuspended in lysis buffer (25 mM Tris HCl, pH 7.5, 1 mM EDTA, 2.5 mM DTT, and 1× Roche (Nutley, NJ) Protease inhibitor) and then lysed by passing three times through a French Press at 20 000 psi. Soluble fractions were collected by centrifugation (10 000×*g*, 10 min, at 4 °C) and diluted to 4 mg ml^{-1}. The protein concentration was determined by Nanodrop (Thermo Scientific). Protein samples were heat denatured at 95 °C for 15 min. After cooling down to room temperature, trypsin (Sigma) was added

Table 1 Strains carrying mutations in the major degradosome components. All experiments were conducted as described in ref. 10. The protein expression levels were quantified in mutant and wild-type strains separately, and the *Z*-score calculation from ref. 14 used to estimate significance and direction of the expression changes. See ESI for assessment of data quality

Name of experiment	Mutant strain	Genotype	No. of proteins quantified in expression change (5% FDR)	Wild-type	Genotype
Δeno	DF261	K10 except for *eno-2*	1576	K10	garB10, fhuA22, ompF627(T_2^R), fadL701(T_2^R), relA1, pit-10, spoT1, rrnB-2, mcrB1, creC510
Δpnp	YHC012	N3433 except for *Tn5::pnp*	1651	N3433	lacZ, relA, spot1, thi1
ΔrhlB	SU02	N3433 except for Δ*rhlB*	1540	N3433	lacZ, relA, spot1, thi1
Truncated rne	BZ453	SH3208 except for *RNase E* truncation	1545	SH3208	his ΔtrpE5(λ)

at 1 : 50 (w/w). Digestions were carried out at 37 °C for 24 hours. The sample was again lyophilized to 10 μl, resuspended in 120 μl buffer (95% H_2O, 5% acetonitrile, 0.1% formic acid) and filtered through a Microcon-10 filter at 12 000*g*. The samples were stored at −80 °C until LC-MS/MS analysis.

Samples (10 μl) were injected into an LTQ-Orbitrap Classic (Thermo Electron) mass spectrometer and peptides were resolved using a 5 to 90% acetonitrile gradient over five hours *via* reverse phase chromatography on a Zorbax BioBasic-18 column 150 mm × 0.10 mm ID (Thermo Fisher). Each of the runs was analyzed independently with Bioworks (Thermo Fisher), searching a database of *E. coli* protein sequences imported from the NCBI genomic database. The results were combined for analysis by PeptideProphet,[15] ProteinProphet[16] and post-processed in the APEX pipeline[13,14] to estimate absolute and differential protein expression based on adjusted spectral counts. We accepted proteins as confidently identified if the ProteinProphet probability was above a cutoff corresponding to <5% global false discovery rate. Relative protein expression changes were calculated with respect to measurements in the respective wild-type strain (Table 1), using a *Z*-score calculation described by Lu *et al.*[14] The *Z*-score adjusts the error model to the absolute concentration (spectral counts). A *Z*-score of $|Z| > 1.96$ corresponds to a *P*-value < 0.05. Note that *Z*-scores can be calculated even if one of the samples (mutant or wild-type) has zero counts measured, but a fold-change in expression cannot. We collected at least five technical replicates per sample, *i.e.* repeat mass spectrometry measurements which were pooled for the quantitative analysis. These replicates also included repeat sample preparations. Replicate concentration measurements correlated well (R^2 = 0.74 to 0.94, not shown) and were pooled for further analysis. The raw data are available at Tranche (ProteomeCommons), data hash: zO+qZFSpP8Z5XzXrSSTm3KYpF2fE2J8F3pDUZpSL5WtS0dIdQ+lF/eGXJHUoXMsRtcdkHzfFmStufDON4S9BaOSjQq0AAAAAAAADIA==. More information on parameters of the mass spectometry experiment and a discussion of data quality is provided in Section 1 (ESI†).

Computational analysis

The mRNA microarray expression data and estimated half-lives were obtained from Bernstein *et al.*[10] and consist of relative mRNA expression values for 1825 genes with missing values estimated using a *K* nearest neighbor method. Tiling-array data for *rne*(1-602) cells were from Stead *et al.*[11] and consist of 2720 genes for which significant up- or down-regulation is reported.

All data were analyzed with a combination of Perl scripts, R, and the software Perseus.[17] The combined *Z*-score for a protein's expression change (Fig. 3 and 4) was calculated as the sum of all *Z*-scores across all four mutants divided by the square root of the number of *Z*-scores available for the particular protein: $Z_{comb} = \sum [Z \text{ of all contributing mutants}] / \sqrt{[\text{number of contributing mutants}]}$. The combined *Z*-score is a measure for the general trend in the data: we assume that a protein (mRNA) which is consistently up-regulated across the four RNA degradosome mutants is more likely a direct target of the degradosome than proteins (mRNAs) with low combined *Z*-scores.

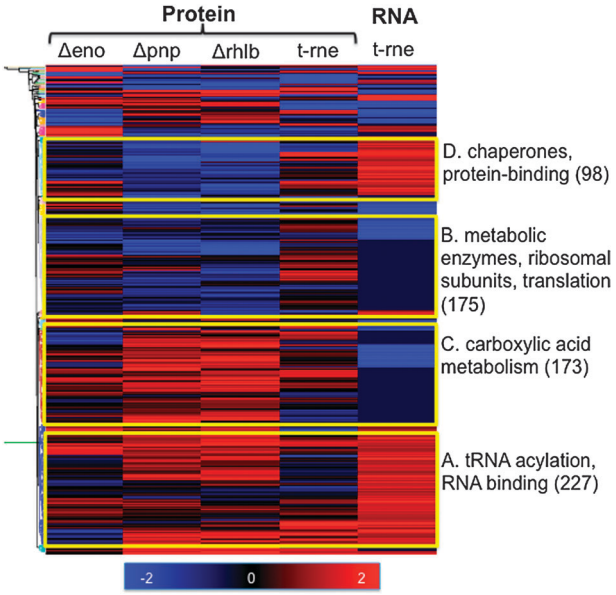

Fig. 2 Protein and RNA expression response of degradosome mutants. The heatmap shows the *Z*-scores of the protein expression data (first four columns) of the four degradosome mutants mapped to tiling-array based mRNA expression data of the RNase E truncation mutant[11] (rightmost column). Significant function enrichments (adjusted *P*-value < 0.01, compared to entire *E. coli* genome) are indicated for the four largest clusters (yellow boxes), with cluster sizes in brackets. Only genes without missing data are shown (*N* = 904). Colors saturate at values of |2| to illustrate statistically significant expression changes (*P*-value < 0.05) – actual *Z*-scores range from −15 to 8. Fold-changes of expression range from −6.0 to 5.7 on a log scale (base 2). t-rne denotes the truncation mutant *rne*(1-602).

For clustering of the data in Fig. 2, up to two missing values were accepted for the proteomics data, and all missing values in the tiling array data were substituted by 0. Data were re-normalized to a mean of 0 and standard deviation 1. Clustering was based on the Euclidean distance measure using an average linkage algorithm (Perseus[17]), with a 1.8 Euclidian distance cutoff to determine genes clusters with similar expression profiles. Gene function analysis was performed for the largest clusters using FuncAssociate.[18] Sequence motifs were analyzed using MEME[19] and fuzznuc.[20]

Results

Effects of Δ*pnp*, Δ*eno*, Δ*rhlB* or *rne*(1-602) on the *E. coli* proteome – data quality and general results

We grew *E. coli* strains carrying Δ*eno*, Δ*rhlB*, Δ*pnp*, or *rne*(1-602) alleles to the mid-exponential phase and used mass spectrometry-based proteomics of cell lysates to quantify soluble protein expression changes relative to the respective parental (wild-type) strains, in a manner similar to the published mRNA analysis.[12] We provide statistical scores of expression changes for a total of 2390 proteins, quantifying >1500 proteins per mutant (Table 1). Fig. 2 shows the direction and statistical significance (*Z*-scores) of the protein expression changes

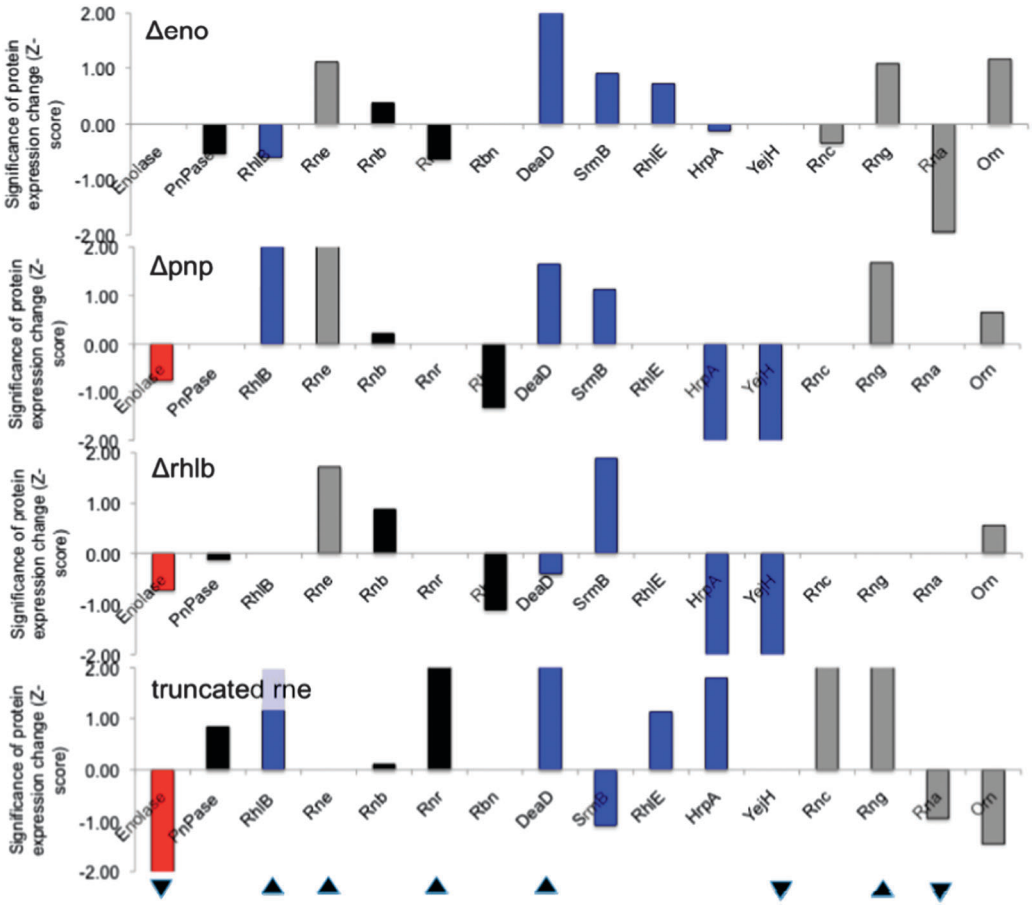

Fig. 3 Expression changes of RNA degradosome proteins and related helicases and RNases. The figure shows the significance (*Z*-score) and direction of the protein expression change of the four main mutants as determined by mass spectrometry experiments for the degradosome components and other known RNases and RNA helicases that can potentially buffer for the loss of gene function. Color codes follow the designation in Fig. 1. A *Z*-score of $|Z| > 1.96$ is significant within a *P*-value of 0.05. The figure shows a range of *Z* between {−2, 2} for clarity reasons; actual *Z*-scores range from −15 to 8 (ESI†). Corresponding fold-changes of protein expression can be found in Fig. S4 (ESI†). The value for the gene in the respective mutant is left out (see Fig. S1, ESI†). Triangles indicate the significance of the combined *Z*-score at the 5% level (see Methods).

between the mutant strains and their respective wild type cells for the subset of 904 proteins that are completely characterized (no missing values, capped at $|Z| = 2$ corresponding to a *P*-value = 0.05).

Label-free quantitative proteomics reports protein expression changes with a precision of about 2- to 3-fold.[14] The proteomics experiment conducted here correctly detects expression changes of the four mutant proteins with highly significant *Z*-scores: as expected, the proteins encoded by *eno*, *pnp*, and *rhlB* are absent in the respective null strains whereas in *rne*(1-602) cells the truncated RNase E(1-602) protein is over-expressed, consistent with previous observations[12] (Fig. S1, ESI†). For further quality control, we conducted western blot experiments on two soluble proteins, GlmS and CspE (Fig. S2, ESI†). For both proteins, the proteomics and western blot data are in strong agreement for the PNPase mutant, and less so for the RhlB mutant. GlmS concentration drastically increases in the PNPase mutant compared to wild-type both in the proteomics and western blot data. This increase contrasts the concentration decrease at the mRNA level which can be explained by strong regulation of GlmS translation and mRNA degradation by two small non-coding RNAs (GlmY and GlmZ) which in turn interact with the RNA degradosome.[21] This result is also consistent with our recent study in which we searched for mutants which can up-regulate *rraB* transcription. In our study, we isolated a *glmS* mutant in which a Tn5 insertion disrupts the coding region of the GlmS protein.[22] RraB disrupts PNPase associated with RNA degradosome both *in vivo*[23] and *in vitro* (L. Zhou, G. Georgiou, unpublished). The proteomics data are consistent with the notion that there is negative feedback between PNPase and GlmS.

In general, the fold-change in expression values between mutant and wild-type for a protein appears to be of similar, if not larger, magnitude to that of the corresponding mRNA (Fig. S4 and S6, ESI†). Examples of fold-changes of protein expression for some proteins are provided in Fig. S5 (ESI†). About 25% of the proteins detected an increase in relative abundance in all four mutant strains (Δ*eno*, Δ*rhlB*, Δ*pnp*, *rne*(1-602)) (Fig. 2). When grouping protein expression profiles according to the Euclidean distance, the Δ*rhlB* and Δ*pnp* cells proved to be more similar to each other than any other mutant

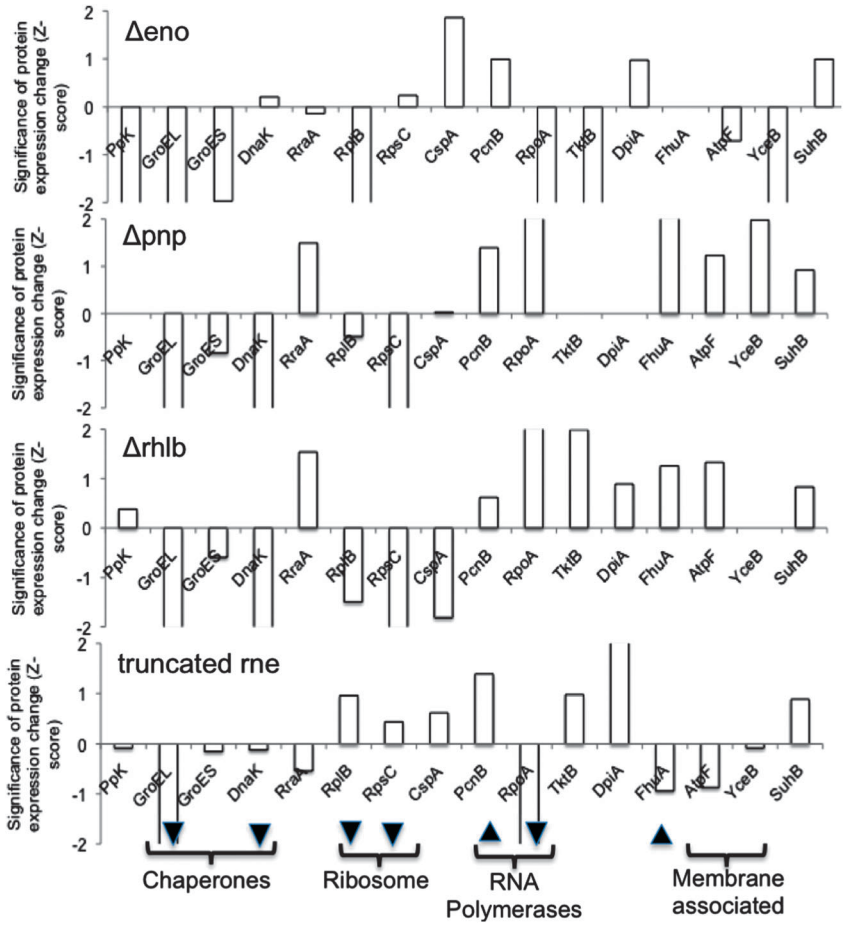

Fig. 4 Expression changes of proteins associated with the RNA degradosome. The figure shows the significance (*Z*-score) and direction of the protein expression change of several genes with known association with the RNA degradosome or RNA degradation. A *Z*-score of $|Z| > 1.96$ is significant within a *P*-value of 0.05. The figure shows a range of *Z* between {−2, 2}; actual *Z*-scores range from −15 to 8 (ESI†). Triangles indicate the significance of the combined *Z*-score at the 5% level (see Methods).

pairs (Fig. 2; Fig. S4, ESI†), supporting the known physical and functional interactions between the two enzymes.[7]

Functional analysis of the four large clusters in Fig. 2 indicates a global down-regulation of translation-related proteins (chaperones, ribosomal subunits), primarily in the Δ*pnp* and Δ*rhlB* mutants (cluster B). Such down-regulation is consistent with regulatory coupling between mRNA availability and translation. However, the proteomics data also demonstrate an up-regulation of 14 of the 20 *E. coli* aminoacyl-tRNA synthetases in all four mutants (cluster A, Fig. 2; Fig. S3, ESI,† *P*-value < 0.001). To the best of our knowledge, a possible link between the RNA degradosome and aminoacylation of tRNAs has not yet been described. We hypothesize that this up-regulation may occur as the cell's effort to counteract the imbalanced concentrations of mRNAs and mature tRNAs due to impaired degradosome function, or possibly as part of a general stress response that produces alarmones.[24]

Next, we examined likely degradosome targets for enrichment of sequence motifs which may be recognized by components of the RNA degradation machinery. We derived these likely targets by combining the *Z*-scores describing the expression change in the four individual mutants into one combined *Z*-score. Across the protein and RNA expression data, as well as mRNA half-life data,[25] we consistently observe enrichment in the CGCTGG motif in which the center CTG triplet is the most conserved part (datafile 3, ESI†). This motif occurs more frequently in mRNAs of short half-life than mRNAs with long half-life, and it is enriched in sequences that increase in either mRNA or protein expression levels when RNA degradosome function is impaired. Thus, the CGCTGG can be loosely associated with direct targets of the RNA degradosome. The ESI† discusses this observation and analysis of other sequence features.

The expression response in degradosome components and protein family members

Fig. 3 indicates the protein expression changes for the four main degradosome components as well as for other *E. coli* RNases and RNA helicases detected by the proteomics analysis. Note that not all RNases and RNA helicases are associated with the degradosome, and that the figure caps expression changes at a *Z*-score of >2 which corresponds to significant expression changes at a *P*-value of <0.05.

RNase R (Rnr) is a 3′ to 5′ exoribonuclease closely related to RNase II (Rnb) which is important for decay of mRNA with extensive secondary structures.[26] Since the coordination between RNA degradosome components organized by the C-terminal region of RNase E is required for decay of structured RNA substrates, it is tempting to assume that the observed increase in the cellular concentration of RNase R in cells expressing the truncated RNase E(1-602) (Fig. 3) is due to a defect in structured RNA observed in this strain. Similarly, RNase BN (Rbn) is implicated in the processing of tRNA precursors.[27] RNase BN abundance decreases in both the *rhlB* and *pnp* mutants (Fig. 3), but is unchanged in *rne*(1-602) (or in the *eno* mutant). The RNase BN's (Rbn) inverse expression pattern with RNase II (Rnb) is consistent with their overlapping functions *in vivo*:[28] down-regulation of one enzyme may be compensated for by up-regulation of the other enzyme and *vice versa*.

E. coli has five known DeaD-box RNA helicases (ribonucleases): RhlB, RhlE, SmrB, DeaD (also called CspA), and DbpA.[1] Null mutants of these helicases result in viable cells, but can affect the growth rate.[29] DeaD, SrmB and RhlE have been reported to associate with RNase E *in vitro*.[5,30,31] DeaD and RhlE have evidence for their role in RNA degradation;[1] they can complement the RNA helicase function in the absence of RhlB,[32] whereas SrmB cannot.[33] The binding sites for DeaD, SrmB and RhlE within the C-terminal end of RNase E are different from those of RhlB, and overlap with the enolase binding site (residues 833–850, Fig. 1). We observe an overall increase in the protein expression levels of DeaD, SrmB, and RhlE across the four mutants which is largely consistent with the functional complementarity between the enzymes and degradosome components (Fig. 3).

HrpA is also an RNA helicase (although not from the DEAD-box family) that functions in mRNA processing.[34] Its expression level is drastically down-regulated in both the *rhlB* and *pnp* mutants, similar to the effect observed for another RNA helicase, YejH (Fig. 3). The expression patterns of these two helicases contrast those of the DeaD-box helicases.

RNase III (Rnc) is involved in ribosomal RNA processing,[35] an activity which overlaps with the essential role of RNase E in *E. coli*. Similar to RNase R, we observe significant up-regulation of RNase III in the *rne*(1-602) mutant consistent with their complementary roles in rRNA processing and in ribosome biogenesis (Fig. 3). Although RNase G (Rng) is an RNase E parolog that cannot fully complement RNase E function,[36] RNase G concentration also increases in three of the four mutants (Fig. 3). Both the endonuclease RNase G and the DeaD helicase are up-regulated in the *eno*, *pnp* and *rne* mutants, a finding consistent with a recent genetic analysis by Tamura *et al.* showing that mutation in DeaD can suppress RNase E in a temperature-sensitive *rne* mutant.[12]

RNase A (*rna*) is involved in tRNA and rRNA processing,[37] and our data show slight down-regulation in the RNase E mutant (Fig. 3). The oligonucleotide-ribonuclease encoded by *orn* acts downstream of the RNA degradosome where it regenerates single nucleotides from oligonucleotides.[38] Its expression levels changes are similar to RNase E and RNase G except for the *rne* mutant (Fig. 3).

The expression response of other degradosome associated proteins

We also assembled a number of genes which have some known association with the RNA degradosome based on pull-down interaction data and other evidence,[1] and for which our proteomics data provided information (Fig. 4). Since these proteins are neither RNases nor RNA helicases, which are discussed above, their role in RNA degradation is presumably indirect.

In particular, three chaperones (GroEL, GroES, and DnaK) and two ribosomal proteins (RplB, RpsC) are strongly down-regulated in the four mutants (somewhat less in the RNase E truncation) (Fig. 4) – corresponding to cluster B in Fig. 2. The RNA polymerase RpoA is significantly up-regulated in the *rhlB* and *pnp* mutants, but down-regulated in the *eno* and *rne*(1-602) mutant strains (P-value < 0.05) (Fig. 4), suggesting feedback between transcription and degradation regulation. The poly(A) polymerase PcnB is up-regulated in all four mutants in our data as might be expected since these strains display varying degrees of defects in RNA processing and PcnB may counteract this defect. The addition of poly(A) tails (by PcnB) generally destabilizes RNA in bacteria,[39] and it facilitates degradation of REP-stabilizers, *i.e.* structural elements preventing degradation, with the help of RhlB.[1]

The protein expression response to perturbation of the RNA degradosome differs from the RNA response

We compared the protein expression changes to published steady state transcriptome and tiling array data[10,11] to investigate the relationship between transcript and protein expression. Since partially degraded RNA can still hybridize to arrays, and translation efficiency of partially degraded RNAs is likely affected in the mutants as well, we did not expect much agreement between the protein and mRNA data. Indeed, the protein expression data do not correlate with either mRNA expression changes or mRNA half-life changes measured by microarrays (Fig. S4, ESI†).[10]

mRNAs that are not processed in the *rne*(1-602) mutant (which is unable to form the degradosome complex) may still be translated into proteins. Inspection of the published tiling array data[11] in Fig. 2 (rightmost column) indicates that about one third of the transcripts for which protein abundance data could be obtained are possibly processed in a degradosome-dependent manner (Fig. 2, clusters A, D, (227 + 98)/904 = 36%): their expression levels increase in tiling array in the *rne*(1-602) mutant. Of these, a large number of transcripts also show positive protein expression changes in the *rne*(1-602) mutant (Fig. 2, second column from the right), but the agreement between tiling array data and protein expression data for the RNase E mutant is generally low. Only one quarter (227/904, cluster A) of the proteins in the dataset are largely up-regulated in their protein expression patterns across all four mutants and in the tiling array data. The remaining set of proteins may accumulate at the mRNA level, but their functional half-life does not seem to be affected, and protein expression levels are not up-regulated (cluster D). Similarly, a number of genes are

either unaffected in their RNA expression levels by the RNase E truncation or down-regulated (clusters B and C), and these genes may be degraded by pathways other than the canonical degradosome. Again, the protein data match the mRNA data only partially (cluster B).

Conclusions

We present a large-scale proteomics analysis of mutants of the *E. coli* RNA degradosome, complementing existing data on mRNA expression and half-life changes.[10] Since measurements of mRNA half-lives do not inform on translation activity of the transcripts and the final effect of RNA degradation on protein abundance, proteomics analysis provides an important additional angle to our understanding of RNA degradation. The proteome-wide data provided here indicate several trends and novel associations that can be followed up by future experiments. Expression data of >2300 proteins reveal substantial differences between the effects of perturbed RNA degradation on transcript and on protein concentrations: fewer than half of the genes in our dataset appear to follow the canonical model of RNA degradation (Fig. 2). A trivial explanation of this discrepancy lies in technical measurement noise. However, the estimated error in protein expression measurements is smaller than the observed variation between experiment and control here,[14,40] supporting the existence of true biological trends.

The complexity of *E. coli* RNA degradation can be partially explained by the up-regulation of several RNA helicases and RNases that may buffer for loss of enzyme function in the degradosome mutants. We are able to provide proteomic characterization for RNA helicases and RNases known to have such buffering roles (DeaD, Rnr, Rng),[1] but also suggest possible complementary functionality for additional enzymes (HrpA, YejH, Rna, Orn) – but their precise interaction with the degradosome needs to be tested in future experiments. We provide two novel findings that have not yet been discussed in the context of the RNA-degradosome. First, we observe a large number of aminoacyl-tRNA synthetases to be up-regulated in the *pnp*, *rhlB*, and *rne* mutants, suggesting a link to tRNA metabolism. Second, we observe enrichment of the CGCTGG motif in sequences that are likely direct targets of the RNA degradosome – the role of which is entirely unknown.

Our results provide insights into the cellular role and structure of the RNA degradosome itself. Since only a small fraction of the PNPase and enolase molecules participate in the degradosome complex,[5,41] it is unsurprising that their mutant expression profiles are highly discordant with those of the other mutants, and highlight the roles of these proteins outside the degradosome. However, the profiles of the *rhlB* and *pnp* mutants correlate better with each other than the profiles of the other mutants ($R = 0.72$), consistent with their physical and functional interaction independent of the RNase E scaffold.[7] Interestingly, this correlation between the *rhlB* and *pnp* mutant is observed for protein expression (this study) and mRNA half-lives,[10] but not for published mRNA expression data (Fig. S4, ESI†), suggesting a highly dynamic and complex system of RNA degradation.

The lack of correlation amongst the datasets also suggests extensive regulation at the level of translation and protein degradation as well as RNA degradosome functions outside the removal of mRNAs. Indeed, even in unperturbed wild-type bacterial systems, a low correlation between mRNA and protein concentrations has indicated that translation and protein degradation regulation account for substantial amounts of protein expression variation.[14,42] The same is true for eukaryotes; transcription and RNA degradation account for only half of the variation in steady-state protein expression levels.[40,43] Further, mRNA half-lives and transcript concentrations are inversely correlated in *E. coli*,[9] suggesting additional layers of regulatory complexity.

Finally, the discordance between protein and RNA expression levels observed is also consistent with a model that distinguishes between the chemical and functional half-life of mRNAs.[1] The mRNA expression data indicate changes in the chemical half-life, *i.e.* the physical presence of the transcript. The functional half-life may be very different from the chemical half-life, affecting the translation activity associated with the mRNA and influencing the final protein expression levels. An mRNA may be physically present in the cell, but translationally inactive and thus non-functional. Our proteomics data deliver supporting evidence for the distinction between chemical and functional half-lives of mRNAs, but detailed biochemical studies will be required in order to confirm and further characterize this mode of regulation.

Acknowledgements

We thank Dr Stanley Cohen from Stanford University for providing *E. coli* strains. We thank George Georgiou and Manuel Santos for helpful discussions. EMM acknowledges funding from the N.I.H., U.S. Army Research (58343-MA), Office of Naval Research (BAA 11-26), and Welch Foundation (F1515).

References

1 A. J. Carpousis, The RNA degradosome of *Escherichia coli*: an mRNA-degrading machine assembled on RNase E, *Annu. Rev. Microbiol.*, 2007, 71–87.
2 M. C. Ow and S. R. Kushner, Initiation of tRNA maturation by RNase E is essential for cell viability in *E. coli*, *Genes Dev.*, 2002, 1102–1115.
3 L. C. Raynal and A. J. Carpousis, Poly(A) polymerase I of *Escherichia coli*: characterization of the catalytic domain, an RNA binding site and regions for the interaction with proteins involved in mRNA degradation, *Mol. Microbiol.*, 1999, 765–775.
4 M. E. Regonesi, M. Del Favero, F. Basilico, F. Briani, L. Benazzi, P. Tortora, P. Mauri and G. Deho, *Biochimie*, 2006, **88**(2), 151–161.
5 B. Py, C. F. Higgins, H. M. Krisch and A. J. Carpousis, A DEAD-box RNA helicase in the *Escherichia coli* RNA degradosome, *Nature*, 1996, 169–172.

6 T. Morita, H. Kawamoto, T. Mizota, T. Inada and H. Aiba, Enolase in the RNA degradosome plays a crucial role in the rapid decay of glucose transporter mRNA in the response to phosphosugar stress in *Escherichia coli*, *Mol. Microbiol.*, 2004, 1063–1075.
7 G.-G. Liou, H.-Y. Chang, C.-S. Lin and S. Lin-Chao, DEAD box RhlB RNA helicase physically associates with exoribonuclease PNPase to degrade double-stranded RNA independent of the degradosome-assembling region of RNase E, *J. Biol. Chem.*, 2002, 41157–41162.
8 A. J. Carpousis, B. F. Luisi and K. J. McDowall, Endonucleolytic initiation of mRNA decay in *Escherichia coli*, *Prog. Mol. Biol. Transl. Sci.*, 2009, 91–135.
9 J. A. Bernstein, A. B. Khodursky, P. H. Lin, S. Lin-Chao and S. N. Cohen, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**(15), 9697–9702.
10 J. A. Bernstein, P.-H. Lin, S. N. Cohen and S. Lin-Chao, Global analysis of *Escherichia coli* RNA degradosome function using DNA microarrays, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, 2758–2763.
11 M. B. Stead, S. Marshburn, B. K. Mohanty, J. Mitra, L. Pena Castillo, D. Ray, H. van Bakel, T. R. Hughes and S. R. Kushner, Analysis of *Escherichia coli* RNase E and RNase III activity *in vivo* using tiling microarrays, *Nucleic Acids Res.*, 2011, **39**(8), 3188–3203.
12 M. Tamura, J. A. Kers and S. N. Cohen, *J. Bacteriol.*, 2012, **194**(8), 1919–1926.
13 C. Vogel and E. M. Marcotte, *Nat. Protocols*, 2008, **3**(9), 1444–1451.
14 P. Lu, C. Vogel, R. Wang, X. Yao and E. M. Marcotte, *Nat. Biotechnol.*, 2007, **25**(1), 117–124.
15 A. Keller, A. I. Nesvizhskii, E. Kolker and R. Aebersold, *Anal. Chem.*, 2002, **74**(20), 5383–5392.
16 A. I. Nesvizhskii, A. Keller, E. Kolker and R. Aebersold, *Anal. Chem.*, 2003, **75**(17), 4646–4658.
17 J. Cox, N. Neuhauser, A. Michalski, R. A. Scheltema, J. V. Olsen and M. Mann, *J. Proteome Res.*, 2011, **10**(4), 1794–1805.
18 G. F. Berriz, O. D. King, B. Bryant, C. Sander and F. P. Roth, *Bioinformatics*, 2003, **19**(18), 2502–2504.
19 T. L. Bailey, N. Williams, C. Misleh and W. W. Li, *Nucleic Acids Res.*, 2006, **34**(Web Server), W369–W373.
20 P. Rice, I. Longden and A. Bleasby, *Trends Genet.*, 2000, **16**(6), 276–277.
21 J. H. Urban and J. Vogel, *PLoS Biol.*, 2008, **6**(3), e64.
22 L. Zhou, M. Zhao, R. Z. Wolf, D. E. Graham and G. Georgiou, *J. Bacteriol.*, 2009, **191**(21), 6665–6674.
23 J. Gao, K. Lee, M. Zhao, J. Qiu, X. Zhan, A. Saxena, C. J. Moore, S. N. Cohen and G. Georgiou, *Mol. Microbiol.*, 2006, **61**(2), 394–406.
24 B. R. Bochner, P. C. Lee, S. W. Wilson, C. W. Cutler and B. N. Ames, *Cell*, 1984, **37**(1), 225–232.
25 D. W. Selinger, R. M. Saxena, K. J. Cheung, G. M. Church and C. Rosenow, *Genome Res.*, 2003, **13**(2), 216–223.
26 Z. F. Cheng and M. P. Deutscher, *Mol. Cell*, 2005, **17**(2), 313–318.
27 T. Dutta and M. P. Deutscher, *J. Biol. Chem.*, 2010, **285**(30), 22874–22881.
28 K. O. Kelly and M. P. Deutscher, *J. Bacteriol.*, 1992, **174**(20), 6682–6684.
29 K. L. Jagessar and C. Jain, Functional and molecular analysis of *Escherichia coli* strains lacking multiple DEAD-box helicases., *RNA*, 2010, 1386–1392.
30 I. Iost and M. Dreyfus, DEAD-box RNA helicases in *Escherichia coli*, *Nucleic Acids Res.*, 2006, 4189–4197.
31 V. Khemici, I. Toesca, L. Poljak, N. F. Vanzo and A. J. Carpousis, *Mol. Microbiol.*, 2004, **54**(5), 1422–1430.
32 V. Khemici and A. J. Carpousis, *Mol. Microbiol.*, 2004, **51**(3), 777–790.
33 M. Kalman, H. Murphy and M. Cashel, rhlB, a new *Escherichia coli* K-12 gene with an RNA helicase-like protein sequence motif, one of at least five such possible genes in a prokaryote., *New Biol.*, 1991, 886–895.
34 J. T. Koo, J. Choe and S. L. Moseley, *Mol. Microbiol.*, 2004, **52**(6), 1813–1826.
35 P. Gegenheimer and D. Apirion, *Nucleic Acids Res.*, 1980, **8**(8), 1873–1891.
36 D. H. Chung, Z. Min, B. C. Wang and S. R. Kushner, *RNA*, 2010, **16**(7), 1371–1385.
37 R. Kaplan and D. Apirion, *J. Biol. Chem.*, 1974, **249**(1), 149–151.
38 X. Zhang, L. Zhu and M. P. Deutscher, *J. Bacteriol.*, 1998, **180**(10), 2779–2781.
39 M. Dreyfus and P. Regnier, *Cell*, 2002, **111**(5), 611–613.
40 C. Vogel, S. Abreu Rde, D. Ko, S. Y. Le, B. A. Shapiro, S. C. Burns, D. Sandhu, D. R. Boutz, E. M. Marcotte and L. O. Penalva, *Mol. Syst. Biol.*, 2010, **6**, 400.
41 G. G. Liou, W. N. Jane, S. N. Cohen, N. S. Lin and S. Lin-Chao, RNA degradosomes exist *in vivo* in *Escherichia coli* as multicomponent complexes associated with the cytoplasmic membrane *via* the N-terminal region of ribonuclease E., *Proc. Natl. Acad. Sci. U. S. A.*, 2001, 63–68.
42 W. Zhang, M. A. Gritsenko, R. J. Moore, D. E. Culley, L. Nie, K. Petritis, E. F. Strittmatter, D. G. Camp, II, R. D. Smith and F. J. Brockman, *Proteomics*, 2006, **6**(15), 4286–4299.
43 B. Schwanhausser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen and M. Selbach, *Nature*, 2011, **473**(7347), 337–342.