

Manuscript EMBO-2015-41465

Mechanism Governing Heme Synthesis Reveals a GATA Factor-Heme Circuit that Controls Differentiation

Nobuyuki Tanimura, Eli Miller, Kazuhiko Igarashi, David Yang, Judith N. Burstyn, Colin N. Dewey, and Emery H. Bresnick

Corresponding author: Emery H. Bresnick, University of Wisconsin School of Medicine and Public Health

Review timeline:

Submission date: Editorial Decision: Revision received: Editorial Decision: Revision received: Accepted: 30 September 2015 01 October 2015 27 October 2015 18 November 2015 20 November 2015 24 November 2015

Editor: Esther Schnapp

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

This manuscript was peer-reviewed at The EMBO Journal and transferred to EMBO reports post-review together with the original referee reports.

1st Editorial Decision

01 October 2015

Thank you for the transfer of your manuscript to EMBO reports.

We can work with the referee reports at hand, and given the constructive comments, we would like to invite you to revise your manuscript with the understanding that the referee concerns must be fully addressed and their suggestions taken on board. Please address all referee concerns in a complete point-by-point response. Acceptance of the manuscript will depend on a positive outcome of a second round of review. It is EMBO reports policy to allow a single round of revision only and acceptance or rejection of the manuscript will therefore depend on the completeness of your responses included in the next, final version of the manuscript.

Revised manuscripts should be submitted within three months of a request for revision; they will otherwise be treated as new submissions. Please contact us if a 3-months time frame is not sufficient for the revisions so that we can discuss the revisions further. Given your 8 main figures, we will publish the manuscript as a normal article for which there are no length limitations. Please note that the results and discussion sections must be separate and the entire materials and methods included in the main manuscript file. Supplementary figures are now called expanded view (EV) figures and are embedded in the main text online. Please upload EV figures individually and please include the

legends for EV figures at the end of the main manuscript file. You can find more information in our guide to authors.

Please remember to change the reference style to the numbered EMBO reports style.

Regarding data quantification, please note that "n" specifies the number of independently performed experiments and not technical repeats or the sum of both. n must equal three or more in order for statistics to be calculated. For you cases where n=2 please remove the error bars and show all individual data points instead. Please also specify the statistical test used to calculate p values in each figure legend.

I look forward to seeing a revised version of your manuscript when it is ready. Please let me know if you have questions or comments regarding the revision.

REFEREE REPORTS

Referee #1:

In this manuscript, Tanimura et al. set up to study the role of heme in the regulation of GATA1dependent genes in erythroid cells. As a model system they used the well-characterized G1E-ER-GATA1 erythroid cells where GATA1-mediated erythroid differentiation is induced by incubation with estradiol that leads to GATA1 translocation to the nucleus.

First, the authors have used CRISPR/Cas9 to mutate GATA1 binding sites within 2 intronic regions of the Alas2 gene, which codes for an enzyme necessary for heme biosynthesis. As expected, the mutations decrease GATA1 binding to their respective sites upon estradiol induction (shown by ChIP), and lead to a strong decrease in Alas2 transcript levels. The biggest effects are seen in cells with the double mutants. Furthermore, hemoglobinization of the cells is severely compromised in the mutants. Interestingly, hemoglobinization of the double mutant cells can be rescued by incubation with 5-ALA, a heme precursor and the product of the ALAS-2 enzyme. Also interestingly, the authors found that expression of some (e.g. globins) but not all genes are rescued by 5-ALA, suggesting that GATA1 and heme may regulate common as well as specific gene expression programs during erythroid differentiation.

The authors then attempt to delineate the common versus specific roles of GATA1 and heme in the regulation of gene expression upon GATA1-mediated erythroid differentiation. First, they analyze changes in gene expression by RNAseq in 3 conditions: 1) -/+ estradiol (i.e. +/- GATA1); 2) WT and double mutant in the presence of estradiol; and 3) double mutant and double mutant rescued with 5-ALA (in the presence of estradiol). Genes that change in condition 2 above are deemed"Alas2-regulated genes", and genes that change in conditions 2 and 3 are deemed "heme-regulated genes". These comparative analyses identified 94 changing genes in the 3 conditions, and the majority of them (i.e. 66) are "GATA1-activated and heme-activated". This result confirms, not surprisingly, that heme participates in the regulation of some but not all GATA1-regulated genes, but the extent to which this is a direct vs indirect effect or the mechanism(s) through which heme may regulate those genes remain(s) unresolved.

Based on the previously known role of heme in promoting degradation of the transcription factor Bach1, the authors then examined Bach1 levels in WT and mutants cells and, again not surprisingly, they observe that 1) Bach1 levels are elevated in mutant cells that cannot express heme; and 2) Bach1 levels are decreased upon heme production induced by providing exogenous 5-ALA. Then, to determine, among some selected GATA1/heme-regulated genes, those that are also dependent on Bach1, the authors induce the knockdown of Bach1 by shRNA and found that some genes are "Bach1-sensitive" and others are not. These results confirm that some of the heme-mediated effects on gene expression are mediated through its Bach1-degradation function. Furthermore, it suggests that heme may have additional functions, independent of its Bach1-degradation role. Unfortunately, the paper does not identify these potential additional role(s).

Finally, based on their data and previously published data, the authors propose a model for GATA1 and heme regulation of Globin and Heme oxygenase genes, including a "Type I incoherent feed-

forward loop".

Main comment:

Overall, this is an interesting manuscript attempting to decipher the role of heme in the regulation of transcription for hemoglobin production during GATA1-mediated erythroid differentiation. The involvement of heme in regulating gene expression/transcription has been demonstrated previously (i.e. through heme-mediated degradation of the transcription factor Bach1) but more work is necessary to fully understand the extent of heme participation in this process. In that sense, the manuscript is original and addresses an important biological question. The authors also clearly demonstrate that heme is important for hemoglobinization through the regulation of the expression of some GATA1-dependent genes (including the globin genes). The 5-ALA-mediated rescue of hemoglobinization in the double mutant (Fig. 4) is quite striking, and provides the authors with a useful model to uncouple heme-mediated effects from GATA1-mediated effects. Up until this point in the manuscript the data is convincing and the experiments appear of high quality.

Unfortunately, the paper does not provide any information regarding the mechanism(s) (direct or indirect) through which heme participates in the regulation of GATA1-dependent genes (or other genes), other than the previously known heme-mediated Bach1 degradation. For instance, could heme mediate the degradation of other transcription factors? Could it stabilize some proteins? Therefore, in the opinion of this reviewer, there is not enough novel biological information in the manuscript. Indeed, there is little to no mechanistic insight on the role of heme/GATA1 in the regulation of gene expression. The comparative analyses of changes in gene expression by RNAseq (presented on Fig.6) appear quite superficial, mostly correlative and sometimes misleading. The authors have also often over-interpreted these data. Therefore, in the opinion of this reviewer, the data presented do not convincingly support the model proposed on Fig. 8B. This model is also incomplete (and misleading) as it does not incorporate critical players in the regulation of globin gene expression. These points are detailed below.

Major Point 1:

The manuscript is lacking novel mechanistic insight on the role of heme in the regulation of gene expression, other than the previously described induction of Bach1 degradation.

Major Point 2:

The model on Fig.8B and most conclusions on the role of heme/GATA1/Bach1 are based solely on changes in gene expression measured by RNAseq or RT-PCR. This is problematic since some of these genes could be direct or indirect targets of GATA1 and Bach1. To delineate the roles of heme/GATA1/Bach1 in regulating gene expression, the authors need to differentiate changes in gene expression that occur on GATA1 and Bach1 direct target genes vs changes on other genes (i.e. indirect effects). The GATA1 ChIP-seq data is available. It's not clear why this information was not integrated in the comparative analysis. Importantly, the authors also need to identify Bach1 binding genome-wide by ChIP-seq if they are to address the question "whether Bach1 counteracts GATA1-mediated activation, and whether this mechanism is restricted to b-like globin gene regulation or broadly impacts the erythroid cell transcriptome" as stated on p.9.

Major Point 3:

On Fig. 6A, the authors have defined the genes that are differentially expressed in the double mutant cells as being "Alas2-regulated genes". However, this is not necessarily true for all these genes. Indeed, the authors cannot exclude that disruption of the GATA-binding sites in the Alas2 gene introns, has additional (undefined) consequence(s) on the expression of other genes, independently of Alas2 downregulation. For instance these mutations could lead to re-organization of the 3D chromatin structure in the nucleus, disrupting putative interactions of these enhancers with other genes. Indeed, it is now well established that enhancers often co-regulate several genes through chromatin looping, including genes that are located at a large distance in cis. Therefore, it is incorrect to equate the mutations of GATA-binding sites in the Alas-2 gene (i.e. the double mutant cells) to a knockdown of Alas-2. In fact, the RNAseq data showing that the majority of changing genes in the double mutant cells are not rescued by 5-ALA treatment strongly support an Alas-2 decrease-independent function of the intronic enhancers. As such the authors should repeat their experiments (phenotypic and gene expression) after induction of ALAS2 knockdown or knockout instead of the double mutant cells. The 5-ALA rescue should also be repeated in these conditions. Alternatively, the authors could attempt to demonstrate that the intronic enhancers of Alas-2 do not

interact with any promoter other than that of the Alas-2 gene. This would need to be done genomewide.

Major Point 4:

The analyses correlating changes in gene expression presented on Fig. 6B are not clearly explained. It appears that the authors have not distinguished genes that are up- and down- regulated. i.e. a gene that is upregulated 2-fold in double mutant cells and down-regulated 2-fold in the 5-ALA rescued cells would appear "correlated". The authors need to distinguish genes that go up- (negative log2 fold change) from those that go down- (positive log2 fold change) and to provide scatter plots representing changes in expression for all genes, not only the "significant" genes. The "significant" genes could be highlighted in a different color. Also, the authors need to provide correlation coefficients for all these plots. For instance, they conclude on p.12 that "the Alas2 regulatory effect correlated well with the 5-ALA regulatory effect". However, this is a subjective statement, and the scatter plot on Fig. 6B (right panel) would suggest that the correlation is not so strong.

Commenting on the GO analysis presented on Fig. 6D,E, the authors note on page 12 that " No GO term was enriched for GATA1-activated/heme-repressed genes". This is very misleading since only 4 genes belong to the group of GATA1-activated/heme-repressed genes (Fig. 6D), and as such no GO term can possibly be found enriched. It is just not appropriate to perform a GO analysis on a group of 4 genes.

Also, the authors need to provide the full list of enriched GO terms for the group "GATA1-activated Heme-activated" genes. Not only a few selected categories as shown on Fig. 6E.

In general in the manuscript, the RNAseq experiments and bioinformatics analyses are not properly described and justified. For instance, for the RNAseq experiments, how many replicates were performed for each condition? A minimum of 3 would be required unless properly justified. Also, access to the raw data should be provided to the reviewers.

Major Point 5:

On pages 12-13, the authors note that "Bach1 represses Hbb-b1 and Hba-a1 expression in a hemedependent manner (Sun et al. 2004; Tahara et al. 2004)" and incorporate this in their model on Globin gene expression (Fig. 8B). However, this statement is misleading since the data published by Sun et al. shows that the absence of Bach1 in erythroid cells is not sufficient to re-activate Hbb-b1. In fact, Brand et al. (NSMB 11(1): 73-80, 2004) have shown that activation of Hbb-b1 is mediated by an exchange of MafK interacting partners from Bach1 to NFE2p45. MafK and NFE2p45 are therefore an integral part of Bach1 and heme -mediated regulation of Hbb-b1 transcription. It is not clear why the authors have not considered these critical players in their study of heme-mediated gene regulation. As presented, the model on Fig. 8B. is misleading. Along the same lines, it is not clear why, on Fig. 8A, the authors have considered the role of several transcription factors (i.e. FOG1, TAL1, Lmo2 and Mi2b) on Bach1 and heme-mediated gene regulation but have ignored MafK and NFE2p45.

Minor Point 1:

Fig. 2B, right panel: there appear to be vacuoles present in int1/8 Mut1 cells + estradiol.

Minor Point 2:

On page 8, the author mention that [upon the double mutations] "the cells did not exhibit gross proliferative or survival defects". However, no data is shown to sustain this conclusion. The authors should test proliferation (e.g. BrdU incorporation) and survival (e.g. AnnexinV staining to measure apoptosis) upon estradiol induction to sustain this claim.

Referee #2:

Tanimura et al report a previously unrecognised regulatory circuit that controls the maturation of red blood cells, where a lineage-specifying transcription factor is connected by means of a feedback mechanism responding to the level of heme production. The function of Gata1 in red blood cell development has provided many insights of wider relevance for the field of cell type differentiation. A better understanding of red blood cell development also has potential translation implications. The regulatory circuit reported here appears to be much more than just another "small addition" to what we already know about Gata1 function, because it provides a feedback link between transcription factors, metabolites and metal ion concentrations. As such, it has novelty that goes significantly beyond the field of red blood cell development, and may indeed stimulate the discovery of similar regulatory circuits in other mammalian tissues.

Specific Comments:

1) RNA-Seq: This is not described in enough detail. I could not see for example whether the analysis was performed using biological replicates. Also, at some point a fold-change cut-off was mentioned, in another place the authors talked about an FDR cut-off. If this analysis was not done using biological replicates, then it would be advisable to do so. At a minimum, the authors need to compare their results to previous studies with the G1ER cells, because one of their pair-wise comparisons (straightforward +/- Gata1 induction) has been analysed by expression profiling already by multiple groups. A good overlap with previous studies would go some way to provide added confidence in the results reported here.

2) The authors need to provide shRNA sequences used for Fog1, Tal1, Lmo2 and Mi2b, or cite relevant references. They should also show the levels of knock-down achieved for these genes.

3) On page 19, the authors state that Gata2 promotes proliferation, but at least in blood stem cells, Gata2 has been reported to do the opposite (put a break on proliferation)

4) The authors need to provide a GEO or Arrayexpress accession number for their RNA-Seq data before publication can be considered

5) Figure 2D: The difference in the right hand panel (labelled with stars as being statistically significant) is not really discussed in the text. Nevertheless, this seems a rather interesting result, as it suggests that in the absence of the intron 1 Gatal binding, the intron 8 Gatal binding has some function in terms of Alas2 expression. It seems therefore yet another illustration that the whole issue of how we can assign enhancer function is really complicated. I agree with the authors that deletion of enhancers is one important assay, but their results here show that in fact deletion of single enhancers may not be enough, and that complex experiments involving the deletion of multiple elements may be necessary.

6) Figures 3 and 4 could be combined into a single figure if the authors want to reduce the number of display elements.

7) Figure 6D: The authors state that GO analysis only revealed results for some of the gene lists from this figure. This is not surprising since some of the gene lists are too small to perform meaningful analysis. The part of the text corresponding to this figure should therefore be rephrased.

8) Figure 8B: It was not clear to me why some of the arrows were red, and some black (appeared to not be explained in legend). For example, why is the arrow from Gata1 to globin a different colour to the one to Alas2?

1st Revision - authors' response

27 October 2015

Reviewer 1:

Major point 1: The reviewer indicated that the manuscript lacks novel mechanistic insight on the role of heme in the regulation of gene expression.

Response: After discussions with the Editor, we were instructed not to expand the mechanistic component of the manuscript. However, we would like to emphasize that our discovery of a mechanism in which heme amplifies GATA-1-mediated activation and repression illustrates an entirely new concept that could not be predicted from the literature and/or existing knowledge. Furthermore, our discovery that the heme amplification mechanism has two components (Bach1-sensitive and –insensitive) represents a significant advance in fleshing out the molecular details of the heme-dependent mechanism. Finally, our new data incorporated into this revision indicating

that $\sim 25\%$ of Bach1-occupied chromatin sites are GATA-1-occupied illustrates the broad scope of the mechanism that we have discovered in the context of the genome.

Major point 2: The reviewer asked whether GATA-1 and Bach1 directly regulate target genes and requested that we integrate Bach1 and GATA-1 ChIP-seq data to address our contention that the mechanism impacts broadly on the erythroid cell transcriptome.

Response: We analyzed GATA-1 and Bach1 occupancy at five GATA-1/heme-regulated and Bach1-sensitive genes from previous human ChIP-seq datasets in peripheral blood-derived erythroblast (PBDE) and K562 cells (Figure EV2A). GATA-1 and Bach1 occupied similar sites in three genes (*HBB*, *HBA1*, and *SLC7A11*). Bach1 only occupied *FBXO30*, and GATA-1 only occupied *TBCEL*. These results indicate that four out of five genes are directly regulated by Bach1. We also compared GATA-1 occupancy and Bach1 occupancy genome-wide in K562 cells (Figure EV2B and Table EV3). GATA-1 and Bach1 occupied 3132 and 2976 genes, respectively. 778 genes were occupied by both GATA-1 and Bach1, and this overlap was highly significant ($p < 2.2 \times 10^{-16}$; naïve Fisher's exact test). These results illustrate the broad scope of the mechanism that we have discovered in the context of the genome. This new data is presented on page 13, line 13.

Major point 3: The referee indicated that the expression of many genes differentially expressed in *Alas2*-enhancer-lacking mutants was not rescued by 5-ALA treatment and suggested that the deleted *Alas2* enhancers might regulate other genes through chromatin looping.

Response: A major discovery from our study is that heme amplifies the capacity of GATA-1 to regulate genes, and Bach1 opposes GATA-1 function. The finding that certain genes affected in *Alas2* intron1/8 mutant cells were not rescued by 5-ALA could be considered to be a distraction from this major discovery, and we plan to analyze the underlying mechanisms in the future. We revised the manuscript to focus on genes regulated by the *Alas2* enhancers ("*Alas2*-enhancer-regulated genes") that are also rescued by 5-ALA. In aggregate, this data provides strong evidence that these genes are heme-regulated. In addition, we added new text on page 11, line 12 describing models that may explain why certain genes dysregulated by the enhancer deletions were not reactivated by 5-ALA.

Major point 4: The referee indicated that we should include hexbin plots representing all up- and down-regulated genes, including genes not deemed significant and requested that we include correlation coefficients. The referee noted that Gene Ontology (GO) analysis should not be performed on a small cohort of 4 genes and requested a full list of GO terms enriched in GATA-1-activated/heme-activated genes. The referee also requested the accession number of the raw data and asked how many replicates were examined in RNA-seq experiments.

Response: We replaced the hexbin plots showing only the upregulated genes with plots depicting both the upregulated and downregulated genes (Figure 6A).

With regard to including genes that are not statistically significant, one cannot reject the null hypothesis that the true fold changes of these genes are zero and cannot distinguish significant versus insignificant terms with the hexbin plot. Thus, genes that are not statistically significant were not included in the hexbin plots.

We provided coefficient of determination (r^2) values, which indicate a correlation between *Alas2*-enhancer and 5-ALA regulatory effects $(r^2 = 0.57;$ Figure 6A).

We conducted GO analysis on only the GATA-1-activated/heme-activated genes, and the text corresponding to this analysis was modified in the Results (page 12, line 7). Enriched GO terms were clustered, and all GO terms deemed significant (P < 0.05) were listed in Table EV1. We also described this analysis in Material and Methods and in the legend for Table EV1.

The raw data from our RNA-seq analysis was submitted (accession number: GSE74371).

Our RNA-seq analysis was conducted rigorously with 3 biological replicates. This information was added to Material and Methods (page 25, line 11).

Major point 5: The referee indicated that Sun et al. suggested that Bach1 dissociation from b-globin locus control region HS2 did not correlate with activation of *Hbb-b1* expression. The referee requested a discussion about the function of MafK and p45/NF-E2 in heme/Bach1-mediated transcriptional regulation.

Response: This relationship between GATA-1 and Bach1 could not have been predicted from existing knowledge of Bach1 function nor from the prior work on p45/NF-E2 mechanisms in Mouse Erythroleukemia Cells (MEL). We added text on page 18, line 1 to discuss the prior work, some of which our coauthor Dr. Igarashi was involved with. While these prior mechanistic studies implicated Bach1 in repressing b-like globin gene expression in MEL cells, how (and whether) GATA-1 might interface with the respective mechanisms was entirely unclear. Furthermore, since p45/NF-E2 knockout mice have been reported to have normal beta-like globin gene expression and normal erythroid cell development/function, we did not focus on p45/NF-E2-dependent mechanisms in our erythroid cell studies.

Minor point 1: The referee indicated that b-estradiol-treated int1/8 mutant cells might have vacuole-like structures.

Response: We re-evaluated multiple slides of Giemsa-stained cells, and in some cases, there were modest differences in cytoplasmic morphology. We revised text on page 7, line 2 to indicate this apparent morphological difference between b-estradiol-treated double-mutant cells and wild type cells.

Minor point 2: The referee requested that we should confirm that proliferative defects were not observed in the double-mutant cells.

Response: We conducted new experimentation to quantitate cell proliferation of untreated or bestradiol-treated wild type and double-mutant cells. The proliferative capacity of the cells was indistinguishable (Figure 2C). This result was described on page 7, line 3.

Reviewer 2:

Query 1: The referee requested a detailed explanation of biological replicates and the cut-off used in the RNA-seq analysis.

Response: RNA from three biological replicates was analyzed, and this information is stated on page 25, line 11.

The information regarding the cut-off has been incorporated into Figure Legends and Results (page 11, line 1). Because three biological replicates were analyzed, and rigorous statistical analysis was conducted, we did not compare our results to prior studies of GATA-1-regulated genes.

Query 2: The referee indicated we should cite the appropriate references regarding FOG-1, TAL1, Lmo2 and Mi2b and requested information on the knockdown efficiencies.

Response: We cited the references on page 37, line 18, and the knockdown efficiencies are now described in the legend of Figure 8A.

Query 3: The referee indicated that GATA-2 has been reported to inhibit proliferation of hematopoietic stem cells.

Response: We deleted the unnecessary statement regarding GATA-2 function and cell proliferation.

Query 4: The referee requested submission of raw RNA-seq data.

Response. We submitted the raw RNA-seq data to GEO (accession number: GSE74371) and indicated this on page 26, line 1.

Query 5: The referee requested that we describe the importance of *Alas2* intron 8 in the context of the intron 1 mutation.

Response: We incorporated text on page 7, line 9 to indicate "in the context of the double *cis*element mutant, the intron 8 GATA motif contributed to the very low-level *Alas2* expression, suggesting a cooperative function of the enhancers".

Query 6: The referee indicated that Figure 3 could be combined with Figure 4 if the number of figures should be reduced.

Response: As our manuscript falls within the space constraints of a "normal article", we maintained Figure 3 and Figure 4 as separate figures.

Query 7: The referee requested revision of text regarding the GO analyses.

Response: Given the low number of genes comprising all categories except the GATA-1/hemeactivated gene category, we only subjected the GATA-1/heme-activated gene cohort to Gene Ontology (GO) analysis. We modified text on page 12, line 7.

Query 8: The referee indicated that the different colors of the arrows in Fig. 8B require an explanation.

Response: The red arrows indicate the new mechanistic links derived from our study, while the black arrows indicated links that had been reported previously and/or could be readily inferred from existing knowledge. We indicated this in the legend to Figure 8B.

We thank the reviewers/editors for providing constructive comments.

18 November 2015

Thank you for the submission of your revised manuscript to EMBO reports. We have now received the comments from the referees, and I am happy to tell you that both support its publication now. Referee 1 only has a minor suggestion that I would like you to address, if you agree, before we can proceed with the official acceptance of your manuscript.

The abstract of the manuscript needs to be written in present tense, and if you agree, we will make these changes for you.

EMBO press papers are accompanied online by A) a short (1-2 sentences) summary of the findings and their significance, B) 2-3 bullet points highlighting key results and C) a synopsis image that is 550x200-400 pixels large (the height is variable). You can either show a model or key data in the synopsis image. Please note that text needs to be readable at the final size. Please send us this information along with the revised manuscript.

I look forward to receiving a new final version of your manuscript as soon as possible.

REFEREE REPORTS

Referee #1:

The authors have addressed all my previous points. I think the described findings are novel and important, and the manuscript is highly improved.

I just have one last minor comment regarding the Discussion, which is interesting but very long. A

suggestion to reduce it would be to remove the last paragraph at the end of page 20 i.e. the proposed role of GATA/heme/Bach in AML. I feel that this paragraph is a little bit far fetched and it detracts from the main focus of the manuscript.

Referee #2:

I am happy with the authors' responses and revisions relating to my comments on the original submission

2nd Revision - authors' response

20 November 2015

Referee #1

Comment: The referee requested that we delete the last paragraph of page 20. Response: We deleted the paragraph.

3rd Editorial Decision

24 November 2015

I am very pleased to accept your manuscript for publication in the next available issue of EMBO reports. Thank you for your contribution to our journal.