NUCLEUS
2017, VOL. 8, NO. 6, 600–604
https://doi.org/10.1080/19491034.2017.1380139

Taylor & Francis
Taylor & Francis Group

EXTRA VIEW

OPEN ACCESS Check for updates

Identification of Fkh1 and Fkh2 binding site variants associated with dynamically bound DNA elements including replication origins

A. Zachary Ostrow and Oscar M. Aparicio

Molecular and Computational Biology Program, University of Southern California, Los Angeles, CA, USA

ABSTRACT

Forkhead Box (Fox) DNA binding proteins control multiple genome activities, including transcription, replication, and repair. These activities are organized spatially and temporally in the nucleus, and Fox proteins Fkh1 and Fkh2 have emerged as regulators of long-range chromosomal interactions involved with these activities, such as the clustering of replication origins programmed for early initiation. Fkh1 and Fkh2 bind a subset of replication origins and are thought to dimerize to mediate long-range chromosomal contacts between these origins. The binding of Fkh1 and/or Fkh2 (Fkh1/2) to replication origins and the recombination enhancer (*RE*), which is involved in DNA repair required for mating-type switching, is cell cycle-regulated and thus appears to be more dynamic than Fkh1/2 binding at regulated target genes. Here we report the identification of Fkh1/2 binding sequence variants at replication origins and the *RE* compared with Fkh1/2 binding sequences found at target genes of the *CLB2* group. These different binding sequences have previously been characterized as weak and strong, respectively, suggesting that the presence of weak sites contributes to more dynamic interactions at replication origins and *RE*, possibly facilitated by Fkh1/2 dimerization and cooperative interactions with accessory proteins. We discuss the wealth of regulatory potential imbued in these features of the DNA and its binding proteins.

ARTICLE HISTORY
Received 31 July 2017
Revised 29 August 2017
Accepted 12 September 2017

KEYWORDS
Fox proteins; DNA motif; genome architecture; chromatin binding; DNA replication foci

Introduction

Genome architecture is increasingly recognized as an important contributor to the regulation of fundamental genomic tasks, including DNA replication, repair and transcription.[1] We have recently identified the Forkhead Box (Fox) transcription factors, Fkh1 and Fkh2, as key regulators of replication origin initiation timing in the *S. cerevisiae* genome.[2] Fkh1 and/or Fkh2 (Fkh1/2) bind a subset of replication origins referred to as Fkh-activated origins to stimulate their early initiation in S phase.[3] Deletion of *FKH1* and *FKH2* or their binding sites proximal to Fkh-activated origins results in delayed activation of these origins; as a consequence, other origins, referred to as Fkh-repressed origins are increased in activity in the absence of *FKH1* and *FKH2*, likely due to reduced competition from Fkh-activated origins for dose-limiting replication initiation factors.[2] While it remains unclear exactly how Fkh1/2 stimulate origin initiation (or "firing"), Fkh1/2 are required for the spatial clustering of early replication origins in G1 phase, suggesting a role in establishing chromosomal architecture with functional consequence for origin firing.[2] We have proposed that origin clustering enables cooperativity between origins in the recruitment of limiting initiation factors, thus stimulating early firing.[4]

We have recently discovered the presence of a domain-swapping motif in Fkh1 and Fkh2 that allows for homo-dimerization of Fkh1 and thereby provides a plausible mechanism to establish physical contacts between origins for clustering in anticipation of replication.[5] Mutation of specific residues to create "domain-swap minus" (dsm) alleles of *FKH1* and *FKH2* that cannot dimerize results in loss of Fkh-activated origin clustering and early firing, supporting a causal relationship between origin architecture and activation. In contrast, the dsm mutations appear to have a limited or minimal effect on the functions of

CONTACT A. Zachary Ostrow aostrow@usc.edu Molecular and Computational Biology Program, University of Southern California, Illumina, Inc., San Diego, CA 92122, USA.

FKH1 and *FKH2* in gene regulation.[5] For example, unlike *fkh1Δ fkh2Δ* cells, *fkh1-dsm fkh2Δ* and *fkh1Δ fkh2-dsm* cells do not exhibit pseudohyphal growth that is attributable to mis-regulation of Fkh1/2 target genes. Because of the very high similarity of Fkh1 and Fkh2 amino acid sequences in the Forkhead domain, and because the dsm mutations affects the function of both Fkh1 and Fkh2, we infer that both Fkh1 and Fkh2 are capable of homo-dimerization, though our previous analysis centered on Fkh1 as the major player in replication origin control. We also cannot rule out potential Fkh1-Fkh2 hetero-dimerization, however, the finding that deletion of *FKH2* alone has no effect on replication origin function indicates that Fkh1-Fkh2 hetero-dimers, should they occur, are not required for origin regulation. Thus, we will limit discussion to the potential function of Fkh1 and/or Fkh2 homo-dimers.

Fkh1/2 binding sequence variants are associated with replication origins

Attempting to explain the differential requirement for domain-swapping in replication versus transcription, we considered the following: Fkh1/2 binding to replication origins is cell cycle-regulated, peaking during G1 and S phases,[3] suggesting a dependence on origin-licensing, which occurs in G1 phase and persists until the origin replicates in S phase.[6] In contrast, Fkh1/2 binding to target genes occurs throughout the cell cycle.[3] To gain further insight, we examined the DNA sequences of putative Fkh1/2 binding sites proximal to replication origins and to Fkh1/2-regulated genes, in particular the "*CLB2* cluster" genes that are expressed in S and G2 phases to regulate mitotic functions.[7]

We identified potential Fkh1/2 binding site by searching for sequences matching the core consensus RYMAAYA within 500 bp regions surrounding origins or within 500 bp regions upstream of *CLB2* cluster genes. We identified two main variants of the consensus sequence differing at the first position: GYMAAYA and AYMAAYA, which appeared to differ in their relative abundance at origins versus genes (Fig. 1A). Previous analysis of Fkh1/2 binding sites in the *SWI5* promoter identified GTAAACA as a "strong" Fkh1/2 binding site and ATAAACA as a "weak" Fkh1/2 binding site.[8] Accordingly, *CLB2* cluster genes are significantly enriched for strong binding sites relative to a simulated random distribution whereas Fkh-activated origins are not (Fig. 1A). In comparison, Fkh-repressed origins are significantly depleted of weak sites relative to a random distribution (Fig. 1A). These findings suggest that differences in the density of Fkh1/2 binding sequence variants play a role in controlling the aforementioned binding characteristics of Fkh1/2 to these genetic elements. Indeed, the combination of weak binding sites and proximal licensing proteins at origins may cooperate to regulate loading of Fkh1/2 onto origins only during G1 and S phases. It will be interesting to determine whether altering weak to strong sites will alter Fkh1/2-origin binding dynamics or origin activity.

We extended the analysis of Fkh1/2 binding sequence variants to the recombination enhancer (*RE*) for donor preference in mating-type switching, which depends on *FKH1*[9]. The minimal 700bp *RE* contains numerous Fkh1/2 consensus binding sequences and we find that the AYMAAYA motif is greatly enriched and the GYMAAYA motif is mildly enriched (Fig. 1A). As with replication origins, there is evidence that Fkh1 binding to the *RE* is cell cycle-regulated,[3,10] consistent with the idea that the AYMAAYA motifs may facilitate more dynamic binding. Again suggesting a link between the AYMAAYA motif and domain-swapping, we have found that *fkh1-dsm* strains are defective in donor preference (S.K. Villwock et al., *in review*).

Hierarchies of Fkh1 and Fkh2 chromatin binding?

Fkh1 and Fkh2 are relatively abundant proteins, however, their consensus DNA binding sequences are in great excess relative to protein.[3,11,12] Thus, Fkh1/2 are likely to select binding sites based on sequence variation and context. For example, it is well known that Mcm1 binds DNA cooperatively with Fkh2 (but not Fkh1),[13] distinguishing the *in vivo* distributions of Fkh1 and Fkh2, and replication origin proteins physically interact with and are required for Fkh1 binding to Fkh-activated origins.[2,14] Moreover, DNA bases immediately flanking the seven base-pair consensus sequence we have used for analysis also contribute to binding specificity of Forkhead DNA binding domains.[15,16] We further propose that a hierarchy of Fkh1/2 binding affinities for different sequences indirectly influence binding site selection due to spatial constraints established by robustly bound and dimerized "primary" sites that place some weaker "secondary" sites in favorable locations for binding, such as

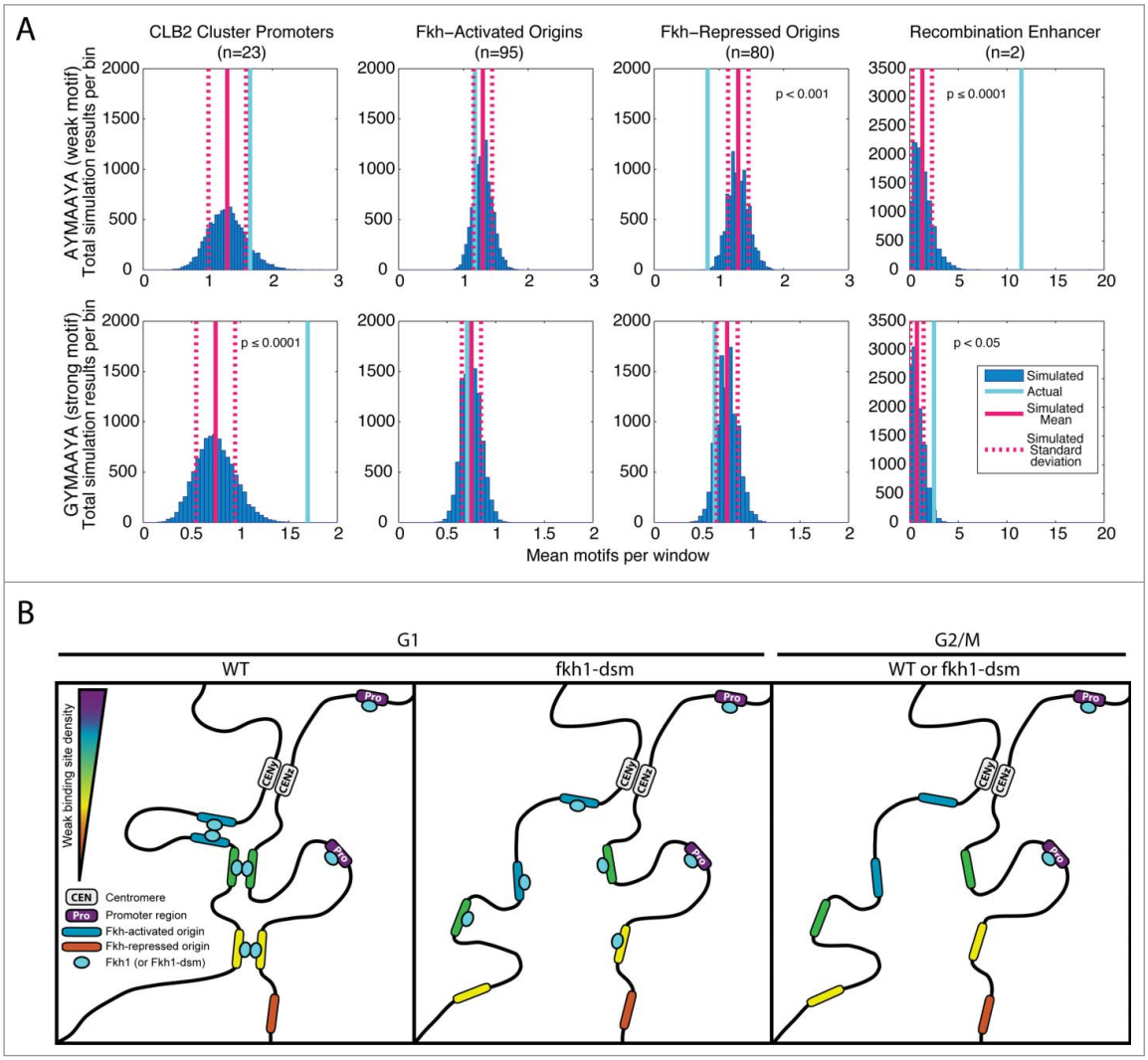

Figure 1. Density of Fkh1/2 binding motif variants correlates with differential regulation of Fkh1/2-chromatin binding. (A) For actual results, genomic sequences of 500 bp centered on the indicated features were queried for the presence of AYMAAYA or GYMAAYA motifs. Significance was calculated by comparison with simulation results, entailing searches (10,000 iterations per feature-binding motif combination) of n random 500 bp genomic regions; n is indicated for each feature class. For the Recombination Enhancer, n = 2 because two non-overlapping, 500 bp regions on each side of the center point were used to calculate mean motifs per window. For each feature class, search results for both motifs were recorded from the same simulation. Fkh-activated and –repressed origins were defined in [2], and *CLB2* cluster genes were defined in[7]. (B) Model depicting the possible effect of Fkh1/2 dimerization and presence of DNA motif variants on binding site hierarchies. Fkh1-dsm binds *CLB2* cluster promoters and Fkh-activated origins similarly to Fkh1, but fails to cluster early origins in G1, potentially leaving Fkh-activated origins with a lower density of Fkh1/2 binding motifs unbound. In G2/M phase (right panel), origins are unbound while *CLB2* cluster promoters (enriched for GYMAAYA motifs) remain bound by either Fkh1 or Fkh1-dsm. Intra-chromosomal origin clustering is likely to be much more frequent than inter-chromosomal origin cluster, which is depicted for emphasis. Cell cycle phase and Fkh1 genotype are depicted above.

loci near concentrations of Fkh1/2, and others in less favored locations (Fig. 1B). Such an effect is consistent with and may be responsible for the linear grouping along chromosomes of Fkh-activated origins separately from Fkh-repressed origins.[2] Consistent with the notion of spatial constraints established by dimerized, primary sites, Fkh1-dsm binds certain loci more avidly than Fkh1,[5] possibly due to an altered chromosomal architecture re-localizing these sites to more favorable positions.

Much remains to be learned about the regulation of Fkh1/2 binding to replication origins in addition to numerous other genomic elements, including centromeres and tRNA genes, which both exhibit spatial clustering.[17,18] A very intriguing possibility is that dimerization allosterically regulates binding specificity

and/or avidity. In this case, dimerization itself may be regulated, for example, by cell cycle kinases, to restrict dimerization and consequently DNA binding spatially and temporally. A related possibility is that the DNA sequence influences dimerization such that only a subset of Fkh1/2 binding sites are bound to dimerized Fkh1 or Fkh2. For example, DNA sequence can influence DNA shape that may in turn influence the structure of the bound protein through an induced-fit mechanism.[19,20] Moreover, the Forkhead DNA binding domain induces a bend in DNA[21,22] such that a change in Fkh1/2 conformation resulting from dimerization might enhance DNA bending to facilitate unwinding, and hence, initiation rate. Clearly, these non-exclusive, hypothetical mechanisms have great potential to regulate genome structure and function. Indeed, dynamic regulation by Fkh1/2 may allow coordination of cell cycle regulation with stress responses to protect against genome instability and replicative aging.[23] Along these lines, a recent study has reported that Fkh1 is degraded during mitosis under stress conditions.[24] Though we have no information on how Fkh1 domain-swapping influences stress responses, one possibility is that dimerization might protect Fkh1 from degradation, thereby protecting replication initiation potential upon cell cycle entry following recovery from stress. On the other hand, degradation of Fkh1 without cell cycle arrest would potentially enable deregulated replication as in *fkh1Δ* cells, and while this might be expected to lengthen S phase and provoke a replication stress response, the length of S phase is not substantially altered and we have not detected replication stress resulting from *FKH1* deletion or overexpression (unpublished and).[25]

Discovery of the domain-swap motif in Fkh1 and Fkh2 in *S. cerevisiae* was based on its prior identification exclusively in the human FoxP family from amongst the 18 Fox protein subfamilies.[26] The other well conserved Fox protein in yeast, Hcm1, shares the α–helix-breaking proline (i.e., domain-swap minus) found in virtually all other human Fox subfamilies, leading to the remarkable realization that exclusively monomeric as well as dimer-capable Fox proteins are conserved from yeast to humans.[5] Thus, a deeper understanding of Fkh1/2 function in yeast should continue to reveal important insights into the wide array of molecular functions controlled by Fox proteins, particularly the still obscure area of genome architecture.

Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank Dr. Matt Dean (University of Southern California) for helpful discussion.

References

[1] Cavalli G, Misteli T. Functional implications of genome topology. Nat Struct Mol Biol. 2013;20(3):290-9. doi:10.1038/nsmb.2474. PMID:23463314

[2] Knott SR, Peace JM, Ostrow AZ, Gan Y, Rex AE, Viggiani CJ, Tavaré S, Aparicio OM. Forkhead transcription factors establish origin timing and long-range clustering in S. cerevisiae. Cell. 2012;148(1-2):99-111. doi:10.1016/j.cell.2011.12.012. PMID:22265405

[3] Ostrow AZ, Nellimoottil T, Knott SR, Fox CA, Tavare S, Aparicio OM. Fkh1 and Fkh2 bind multiple chromosomal elements in the S. cerevisiae genome with distinct specificities and cell cycle dynamics. PLoS One. 2014;9(2):e87647. doi:10.1371/journal.pone.0087647. PMID:24504085.

[4] Aparicio OM. Location, location, location: it's all in the timing for replication origins. Genes Dev. 2013;27(2):117-28. doi:10.1101/gad.209999.112. PMID:23348837

[5] Ostrow AZ, Kalhor R, Gan Y, Villwock SK, Linke C, Barberis M, Chen L, Aparicio OM. Conserved forkhead dimerization motif controls DNA replication timing and spatial organization of chromosomes in S. cerevisiae. Proc Natl Acad Sci U S A. 2017;114:E2411–E2419.

[6] DePamphilis ML, Blow JJ, Ghosh S, Saha T, Noguchi K, Vassilev A. Regulating the licensing of DNA replication origins in metazoa. Curr Opin Cell Biol. 2006;18(3):231-9. doi:10.1016/j.ceb.2006.04.001. PMID:16650748

[7] Zhu G, Spellman PT, Volpe T, Brown PO, Botstein D, Davis TN, Futcher B. Two yeast forkhead genes regulate the cell cycle and pseudohyphal growth. Nature. 2000;406(6791):90-4. doi:10.1038/35017581. PMID:10894548

[8] Hollenhorst PC, Pietz G, Fox CA. Mechanisms controlling differential promoter-occupancy by the yeast forkhead proteins Fkh1p and Fkh2p: implications for regulating the cell cycle and differentiation. Genes Dev. 2001;15(18):2445-56. doi:10.1101/gad.906201. PMID:11562353

[9] Haber JE. Mating-type genes and MAT switching in Saccharomyces cerevisiae. Genetics. 2012;191(1):33-64. doi:10.1534/genetics.111.134577. PMID:22555442

[10] Coic E, Richard GF, Haber JE. Saccharomyces cerevisiae donor preference during mating-type switching is dependent on chromosome architecture and organization. Genetics. 2006;173(3):1197-206. doi:10.1534/genetics.106.055392. PMID:16624909

[11] Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK, Weissman JS. Global

analysis of protein expression in yeast. Nature. 2003;425 (6959):737-41. doi:10.1038/nature02046. PMID:14562106

[12] Kulak NA, Pichler G, Paron I, Nagaraj N, Mann M. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. Nature methods. 2014;11(3):319-24. doi:10.1038/nmeth.2834. PMID:24487582

[13] Boros J, Lim FL, Darieva Z, Pic-Taylor A, Harman R, Morgan BA, Sharrocks AD. Molecular determinants of the cell-cycle regulated Mcm1p-Fkh2p transcription factor complex. Nucleic Acids Res. 2003;31(9):2279-88. doi:10.1093/nar/gkg347. PMID:12711672

[14] Reinapae A, Jalakas K, Avvakumov N, Looke M, Kristjuhan K, Kristjuhan A. Recruitment of Fkh1 to replication origins requires precisely positioned Fkh1/2 binding sites and concurrent assembly of the pre-replicative complex. PLoS Genet. 2017;13(1):e1006588. doi:10.1371/journal.pgen.1006588. PMID:28141805

[15] MacIsaac KD, Wang T, Gordon DB, Gifford DK, Stormo GD, Fraenkel E. An improved map of conserved regulatory sites for Saccharomyces cerevisiae. BMC Bioinformatics. 2006;7:113. doi:10.1186/1471-2105-7-113. PMID:16522208

[16] Zhu C, Byers KJ, McCord RP, Shi Z, Berger MF, Newburger DE, Saulrieta K, Smith Z, Shah MV, Radhakrishnan M, et al. High-resolution DNA-binding specificity analysis of yeast transcription factors. Genome Res. 2009;19 (4):556-66. doi:10.1101/gr.090233.108. PMID:19158363

[17] Haeusler RA, Engelke DR. Genome organization in three dimensions: thinking outside the line. Cell Cycle. 2004;3 (3):273-5. doi:10.4161/cc.3.3.732. PMID:14726665

[18] Jin Q, Trelles-Sticken E, Scherthan H, Loidl J. Yeast nuclei display prominent centromere clustering that is reduced in nondividing cells and in meiotic prophase. J Cell Biol. 1998;141(1):21-9. doi:10.1083/jcb.141.1.21. PMID:9531545

[19] Duguet M. Proteins that shape DNA. Biochimie. 1981;63(8-9):649-69. doi:10.1016/S0300-9084(81)80214-3. PMID:6272878

[20] Rohs R, Jin X, West SM, Joshi R, Honig B, Mann RS. Origins of specificity in protein-DNA recognition. Annu Rev Biochem. 2010;79:233-69. doi:10.1146/annurev-biochem-060408-091030. PMID:20334529

[21] Bravieri R, Shiyanova T, Chen TH, Overdier D, Liao X. Different DNA contact schemes are used by two winged helix proteins to recognize a DNA binding sequence. Nucleic Acids Res. 1997;25(14):2888-96. doi:10.1093/nar/25.14.2888. PMID:9207039

[22] Cirillo LA, Zaret KS. Specific interactions of the wing domains of FOXA1 transcription factor with DNA. J Mol Biol. 2007;366(3):720-4. doi:10.1016/j.jmb.2006.11.087. PMID:17189638

[23] Postnikoff SD, Harkness TA. Mechanistic insights into aging, cell-cycle progression, and stress response. Front Physiol. 2012;3:183. doi:10.3389/fphys.2012.00183. PMID:22675309

[24] Malo ME, Postnikoff SD, Arnason TG, Harkness TA. Mitotic degradation of yeast Fkh1 by the Anaphase Promoting Complex is required for normal longevity, genomic stability and stress resistance. Aging (Albany NY). 2016;8 (4):810-30. doi:10.18632/aging.100949. PMID:27099939

[25] Peace JM, Villwock SK, Zeytounian JL, Gan Y, Aparicio OM. Quantitative BrdU immunoprecipitation method demonstrates that Fkh1 and Fkh2 are rate-limiting activators of replication origins that reprogram replication timing in G1 phase. Genome Res. 2016;26(3):365-75. doi:10.1101/gr.196857.115. PMID:26728715

[26] Stroud JC, Wu Y, Bates DL, Han A, Nowick K, Paabo S, Tong H, Chen L. Structure of the forkhead domain of FOXP2 bound to DNA. Structure. 2006;14(1):159-66. doi:10.1016/j.str.2005.10.005. PMID:16407075