Learning about genomics and disease from the anucleate human red blood cell

Edward J. Benz Jr.

Commentary

During the differentiation of an erythrocyte, the developing erythroblast shuts down expression of most of its genes but preserves high levels of expression of certain key genes, such as those encoding hemoglobin and critical membrane proteins. In this issue of the *JCI*, Gallagher et al. show that a specialized type of DNA sequence element known as an insulator protects the expression of ankyrin, a key membrane protein. In several kindreds, mutations in the insulator led to impaired ankyrin expression and congenital hemolytic anemia. This work provides important insights into ways in which epigenetic changes can alter gene expression and thereby lead to human disease.

Find the latest version:

https://jci.me/45433/pdf

Learning about genomics and disease from the anucleate human red blood cell

Edward J. Benz Jr.

Dana Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts, USA.

During the differentiation of an erythrocyte, the developing erythroblast shuts down expression of most of its genes but preserves high levels of expression of certain key genes, such as those encoding hemoglobin and critical membrane proteins. In this issue of the *JCI*, Gallagher et al. show that a specialized type of DNA sequence element known as an insulator protects the expression of ankyrin, a key membrane protein. In several kindreds, mutations in the insulator led to impaired ankyrin expression and congenital hemolytic anemia. This work provides important insights into ways in which epigenetic changes can alter gene expression and thereby lead to human disease.

During erythropoiesis, erythrocytes (red blood cells) develop from proerythroblasts in the bone marrow. This process is marked by progressive condensation of nuclear chromatin into its inactive form, heterochromatin, and increasing accumulation of only those proteins needed for the structure and function of the erythrocyte during its long life in the circulation. As maturation proceeds, erythroblasts become smaller, the nucleus becomes more pyknotic (condensed), and the cytoplasm redder, reflecting the increasing accumulation of hemoglobin and the decline in basophilic staining due to reductions in the amounts of RNA and ribosomes. These visible changes reflect the fact that expression of much of the genome is repressed as the cells mature. Indeed, the maturing erythroblast ultimately loses its nucleus and all of its organelles. Yet a few hundred genes whose protein products are critical to the oxygen transport function and viability of the mature erythrocyte continue to be expressed robustly almost until the last stages of terminal maturation. In this issue of the *JCI*, Gallagher et al. (1) shed light on how one of these key genes — that encoding the crucial cytoskeletal protein ankyrin — maintains its activity during terminal erythroid differentiation, by studying patients with hereditary spherocytosis (HS), a congenital anemia caused by mutations in the ankyrin-encoding gene. To place this article

in context, it is important first to consider the importance of the red cell cytoskeleton to erythrocyte viability, and the specific role played by ankyrin.

Adaptation of erythrocytes to life in the circulation

Mature erythrocytes circulate for 120 days despite having almost no capacity for either repair or regeneration of damaged or worn-out components. This longevity is remarkable considering the mechanical, metabolic, and geometric challenges the circulation presents. The average diameter of a human erythrocyte is 7.5 μm, while the diameter of many capillaries is 2–3 μm. During its lifespan, an erythrocyte will squeeze through these narrow channels hundreds of thousands of times. Erythrocytes also pass through regions of varying osmolality, swelling and shrinking as they do, and encounter turbulence and shear stresses wherever laminar flow is disrupted.

The erythrocyte membrane is a remarkably evolved structure perfectly adapted to deal with these challenges (2) (Figure 1). The erythrocyte possesses more membrane surface area than is needed to encase the volume of its cytoplasm. This allows for a biconcave disc geometry. This shape can be stretched, twisted, distended, and compressed without permanent damage. The fragile lipid bilayer is buttressed by a protein cytoskeleton consisting primarily of the highly helical proteins α- and β-spectrin. Arrayed in a hexagonal lattice, spectrin functions like a coiled spring able to stretch and snap back as the mature erythrocyte squeezes through capillaries, swells

and shrinks, and is distorted by shear stresses. Spectrin attaches to the membrane via two "molecular swivels," protein 4.1R and ankyrin, which bind simultaneously to spectrin and to the cytoplasmic domain of transmembrane proteins such as glycophorin and the anion channel, band 3. This confers sufficient tensile strength to withstand mechanical stresses and adequate flexibility to change shape as needed. Defects in key cytoskeletal proteins cause congenital hemolytic anemias such as HS, which is caused by mutations in the ankyrin-encoding gene (3).

Control of specialized cytoskeletal protein expression during erythropoiesis

The assembly of the extraordinary erythrocyte membrane structure during erythropoiesis is a critical process not yet well understood. The production and distribution to appropriate sites of all of the relevant proteins must be exquisitely timed. It is known that there is a clear transition in maturing erythroblasts from the synthesis of widely expressed forms of cytoskeletal proteins to production of erythroid-specific subtypes. However, the mechanisms governing the precise timing and amounts of erythroid cytoskeletal protein production remain obscure.

In their study, Gallagher et al. (1) identified kindreds with HS in whom the causative mutations were located near the promoter governing erythroid-specific ankyrin expression, but not at the positions affecting promoter function, for example, transcription initiation and factor binding. In transgenic mice, these mutations ablated a curious property of the ankyrin promoter, its capacity to confer "position independence," that is, its ability to direct erythroid tissue–specific expression at high levels, regardless of where the transgene is integrated into the mouse host genome. In other words, the ankyrin promoter seemed to be able to protect ("insulate") the ankyrin gene from the general shutdown of gene expression during late erythropoiesis.
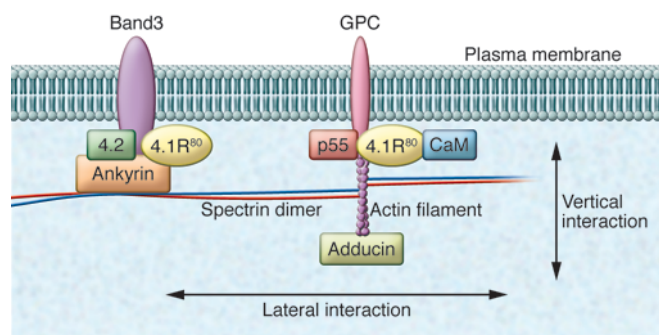
Figure 1
Basic structure of the erythrocyte membrane and cytoskeleton. The diagram shows that the membrane is composed of a lipid bilayer punctuated by transmembrane proteins including Band3 and glycophorin C (GPC). The erythrocyte gains its strength and flexibility from the underlying cytoskeleton, which consists primarily of α- and β-spectrin. Spectrin (α- and β-spectrin) is attached flexibly to the membrane via ankyrin, which binds to Band 3, and via protein 4.1, which binds GPC. The other proteins shown enhance the strength of the binding. "Vertical interactions" provide strong attachments of the cytoskeleton to the lipid bilayer. "Horizontal interactions" allow the cell to stretch reversibly under shear stress without fragmentation. CaM, calmodulin; 4.1R[80], 80-kDa erythrocyte isoform of protein 4.1R. Figure adapted from *Frontiers in Bioscience* (2).

A role for insulator sequences in erythropoiesis

Gallagher and colleagues hypothesized that the ankyrin promoter region contained an additional element — a "barrier insulator" (1). Insulator elements have been defined largely by their ability to "mark" or "protect" regions of chromatin in an "open" configuration that is accessible to components needed to activate gene expression (4–9). Barrier insulators protect active legions of chromatin from being inactivated by the "invasion" of heterochromatin. They exhibit a distinctive chromatin configuration, associate with certain insulator binding proteins, and confer position independence.

Gallagher et al. (1) showed clearly that the normal ankyrin promoter region contains an element exhibiting the properties of a barrier insulator sequence, and localized the element precisely. The two mutations leading to HS studied by Gallagher and colleagues fell within this putative barrier element. Moreover, introduction of these mutations into gene constructs expressed in transgenic mice disrupted the position independence function, chromatin configuration, and protein binding properties of the barrier element in the appropriate and predicted manner. These results provide persuasive evidence that the ankyrin promoter is tightly linked to a barrier insulator and that the mutations causing HS in the kindreds studied by these researchers acted by disrupting insulator function.

Insulators and disease

The results of Gallagher et al. (1) define a novel mechanism leading to HS and may have additional implications. First, this work identifies an important element governing ankyrin gene expression during erythropoiesis. Insulator elements also surround genes for globin and erythroid spectrin (10–12). It would be interesting to learn whether similar insulators are critical to the expression of other genes whose expression must be maintained during late erythroid maturation, even as most genes are engulfed by heterochromatin and silenced. It is attractive to speculate that common *trans* elements might interact with these insulators in a concerted fashion to preserve the activities of these genes. To date, there has been no other explanation for the coordination of the gene expression program in late erythroblasts. This article (1) also provides an important example of epigenetic phenomena causing disease. Finally, the demonstration that insulator mutations can lead to disease might partially explain why so many disease-associated SNPs are located in extragenic DNA. It has been hard to envision how a genetic variation located in DNA that does not appear to be expressed could influence a disease. Many insulators are found close to gene clusters, but others are located at greater distances, for example, the locus control regions of globin genes (13). Thus, at least some SNPs in noncoding DNA could reside in insulator sequences or other

DNA elements governing chromatin configuration and function.

The work of Gallagher et al. (1) could not resolve some questions that vex studies of insulator sequences. First, many insulators are adjacent to or intermingled with sequence elements that appear to be promoters (13). It is technically almost impossible to separate cleanly mutations impacting promoter integrity from those impairing insulator functions (13–18). Thus, an open question remains: Do insulators hold chromatin domains open so that promoters can function, or do promoters open chromatin by their activity? A second question is how and when erythroid insulators are recognized by regulatory proteins in the nucleus and activated. A third is whether the "opening" of chromatin by insulators and actual expression of the "insulated" genes occur simultaneously or asynchronously.

One approach to answering these questions in erythroid cells might be to revisit studies done almost 40 years ago suggesting that genes destined for expression in erythroblasts might become "poised" or "programmed" at earlier stages of erythroid progenitor differentiation (19, 20). Using more modern erythroid cell culture systems and far more sensitive and direct assays for open chromatin, such as histone acetylation or DNAse hypersensitive site mapping, one might be able to compare the time courses of chromatin remodeling and gene transcription during early and late erythropoiesis. If it proves to be true that there is a temporal separation, then one could dissect the insulator and promoter functions of these elements independently and search directly for the events that activate specific chromatin regions.

Genome physiology and human disease

The article by Gallagher et al. (1) provides a powerful illustration of the importance of regarding the genome as a dynamic physiologic entity. Understanding genome anatomy alone is unlikely to reveal the behavior and disease-modifying impact of epigenetic changes. These will be understood only if the physiologic functions of chromatin are linked appropriately to our rapidly increasing understanding of the pathologic anatomy of the genome. Changes correlating with altered phenotype or disease, whether they be SNPs, mutations, rearrangements, copy number changes, or epigenetic alterations, must be assessed

in the best functional assays available in appropriate cell culture or animal models, in order to assess their net effects on the tissue specificity, timing, and amount of expression of the relevant genes. Given the overwhelming number of changes that are being correlated with phenotypic traits and diseases, it is important to know which changes are good bets to be meaningful before these complex and expensive studies are undertaken. Identification and thorough, incisive analysis of informative patients, such as done by Gallagher and colleagues (1), remains the best way to know that the change being studied matters in the pathogenesis of a disease.

Address correspondence to: Edward J. Benz Jr., President and CEO, Dana Farber Cancer Institute, Room 1628D, 44 Binney Street, Boston, Massachusetts 02115, USA. Phone: 617.632.2159; Fax: 617.632.2161; E-mail: ebenz@comcast.net.

1. Gallagher PG, et al. Mutation of a barrier insulator in the human ankyrin-1 gene is associated with hereditary spherocytosis. *J Clin Invest*. 2010; 120(12):4453–4465.
2. Nunomura W, Takakuwa Y. Regulation of protein 4.1R interactions with membrane proteins by Ca^{2+} and calmodulin. *Front Biosci*. 2006;11:1522–1539.
3. Eber S, Lux SE. Hereditary spherocytosis — defects in proteins that connect the membrane skeleton to the lipid bilayer. *Semin Hematol*. 2004;41(2):118–141.
4. Valenzuela L, Kamakaka RT. Chromatin insulators. *Annu Rev Genet*. 2006;40:107–138.
5. Gaszner M, Felsenfeld G. Insulators: exploiting transcriptional and epigenetic mechanisms. *Nat Rev Genet*. 2006;7(9):703–713.
6. Kuhn EJ, Geyer PK. Genomic insulators: connecting properties to mechanism. *Curr Opin Cell Biol*. 2003;15(3):259–265.
7. Singh V, Srivastava M. Enhancer blocking activity of the insulator at H19-ICR is independent of chromatin barrier establishment. *Mol Cell Biol*. 2008; 28(11):3767–3775.
8. Capelson M, Corces VG. Boundary elements and nuclear organization. *Biol Cell*. 2004;96(8):617–629.
9. Felsenfeld G, et al. Chromatin boundaries and chromatin domains. *Cold Spring Harb Symp Quant Biol*. 2004;69:245–250.
10. Gallagher PG, Nilson DG, Steiner LA, Maksimova YD, Lin JY, Bodine DM. An insulator with barrier-element activity promotes alpha-spectrin gene expression in erythroid cells. *Blood*. 2009; 113(7):1547–1554.
11. Steiner LA, et al. Chromatin architecture and transcription factor binding regulate expression of erythrocyte membrane protein genes. *Mol Cell Biol*. 2009; 29(20):5399–5412.
12. Chung JH, Whiteley M, Felsenfeld G. A 5′ element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in Drosophila. *Cell*. 1993; 74(3):505–514.
13. Raab JR, Kamakaka RT. Insulators and promoters: closer than we think. *Nat Rev Genet*. 2010; 11(6):439–446.
14. Lunyak VV, et al. Developmentally regulated activation of a SINE B2 repeat as a domain boundary in organogenesis. *Science*. 2007;317(5835):248–251.
15. Noma K, Cam HP, Maraia RJ, Grewal SI. A role for TFIIIC transcription factor complex in genome organization. *Cell*. 2006;125(5):859–872.
16. Scott KC, Merrett SL, Willard HF. A heterochromatin barrier partitions the fission yeast centromere into discrete chromatin domains. *Curr Biol*. 2006; 16(2):119–129.
17. Valenzuela L, Dhillon N, Kamakaka RT. Transcription independent insulation at TFIIIC-dependent insulators. *Genetics*. 2009;183(1):131–148.
18. Simms TA, et al. TFIIIC binding sites function as both heterochromatin barriers and chromatin insulators in Saccharomyces cerevisiae. *Eukaryot Cell*. 2008;7(12):2078–2086.
19. Young NS, Benz EJ, Kantor JA, Kretschmer P, Nienhuis AW. Hemoglobin switching in sheep: only the γ gene is in the active conformation in fetal liver but all the β and γ genes are in the active conformation in bone marrow. *Proc Natl Acad Sci*. 1978; 75(12):5884–5888.
20. Benz EJ, Barker JE, Pierce JE, Turner PA, Nienhuis AW. Hemoglobin switching in sheep: commitment of erythroid stem cells to expression of the $β^C$–globin gene and accumulation of $β^C$-globin mRNA. *Cell*. 1978;14(3):733–740.