nature structural & molecular biology

Heme is involved in microRNA processing

Michael Faller[1], Michio Matsunaga[1], Sheng Yin[2], Joseph A Loo[1,2] & Feng Guo[1]

MicroRNAs (miRNAs) regulate the expression of a large number of protein-coding genes. Their primary transcripts (pri-miRNAs) have to undergo multiple processing steps to reach the functional form. Little is known about how the processing of miRNAs is modulated. Here we show that the RNA-binding protein DiGeorge critical region-8 (DGCR8), which is essential for the first processing step, is a heme-binding protein. The association with heme promotes dimerization of DGCR8. The heme-bound DGCR8 dimer seems to trimerize upon binding pri-miRNAs and is active in triggering pri-miRNA cleavage, whereas the heme-free monomer is much less active. A heme-binding region of DGCR8 inhibits the pri-miRNA–processing activity of the monomer. This putative autoinhibition is overcome by heme. Our finding that heme is involved in pri-miRNA processing suggests that the gene-regulation network of miRNAs and signal-transduction pathways involving heme might be connected.

miRNAs are a class of noncoding RNAs that are about 22 nucleotides (nt) in length[1]. They are involved in a variety of important biological processes, such as developmental timing and patterning, apoptosis, hematopoietic differentiation, cell proliferation, organ development and tumorigenesis[2–6]. They specifically regulate the expression of many protein-coding genes by targeting their messenger RNAs through pairing interactions that direct cleavage or translational repression in a combinatorial fashion[7]. miRNAs are encoded in the genomes of multicellular organisms[8–10] as well as some large DNA viruses[11]. Several hundred human miRNAs have been cloned, representing 2%–3% of total human genes[12]. It is estimated that 10%–30% of all protein-coding genes are regulated by miRNAs[13].

In animals, pri-miRNAs can be several thousand nucleotides in length[14] and undergo sequential cleavage steps in the nucleus and cytoplasm, separated by nuclear export mediated by exportin-5 (ref. 15–17). In the first step, pri-miRNAs are cleaved into ~65-nt hairpin-like intermediates (termed precursor miRNAs or pre-miRNAs) by the RNase Drosha along with an RNA-binding protein, DGCR8 (its fly and worm homologs are called Pasha)[18–22]. These two proteins (known as the Microprocessor complex, because they cofractionate in cell extracts) are necessary and sufficient for pri-miRNA processing. Pri-miRNA processing activity can be reconstituted *in vitro* using recombinant Drosha and DGCR8 (refs. 20,21). It is thought that DGCR8 is important for recognition of pri-miRNAs and functions as an RNA-binding partner of Drosha[19,23]. In the cytoplasm, pre-miRNAs are further cleaved by another RNase III family member, Dicer, to give rise to a miRNA duplex[24–28]. One or both strands of this duplex are incorporated into the RNA-induced silencing complex (RISC)[29,30].

To elucidate the roles of the human DGCR8 protein in pri-miRNA processing, we characterized this protein using biochemical methods. We found that DGCR8 is a heme-binding protein. The association with a heme cofactor (a complex of iron with protoporphyrin IX) promotes dimerization of DGCR8. The monomer and dimer of DGCR8 were found to be functionally distinct in pri-miRNA–binding and pri-miRNA–processing assays. Our results indicate a possible mechanism of molecular regulation during the maturation of miRNAs. The implications for the biological functions of miRNAs and heme are discussed.

RESULTS

DGCR8 is a heme-binding protein

It has been demonstrated that the N-terminal 275 residues of DGCR8 are dispensable for pri-miRNA processing activity *in vitro*[21]. To investigate the functional roles of DGCR8 in pri-miRNA processing, we overexpressed a DGCR8 construct containing residues 276–751 (called NC1) in *Escherichia coli*. NC1 contains a putative WW motif (a protein module with two highly conserved tryptophan residues) and two double-stranded RNA–binding domains (dsRBDs) (**Fig. 1a**).

The highly purified recombinant NC1 protein had a yellow color (**Fig. 1a**), suggesting that it binds a cofactor. Three lines of evidence established that the cofactor is heme. First, the UV-visible absorption spectrum of NC1 had peaks at 367, 450 and 556 nm (**Fig. 1b**). Two types of common cofactors absorb at these wavelengths, heme and flavin derivatives (such as FAD). Whereas flavin derivatives fluoresce strongly at UV and visible wavelengths, heme does not. The fluorescence excitation and emission spectra of NC1 showed little fluorescence signal between 300 and 600 nm (**Supplementary Fig. 1** online). Second, the cofactor of NC1 was extracted using an organic solvent and characterized using HPLC, following a protocol used to characterize the heme bound to a nitric oxide synthase[31]. The retention time of a compound from a reverse-phase column depends on its chemical properties and thus contains information about its identity. The cofactor extracted from NC1 was eluted from the reverse-phase column at the same retention time as the heme from horse heart

[1]Department of Biological Chemistry, David Geffen School of Medicine, [2]Department of Chemistry and Biochemistry, Molecular Biology Institute, University of California at Los Angeles (UCLA), Los Angeles, California 90095, USA. Correspondence should be addressed to F.G. (fguo@mbi.ucla.edu).

Received 29 August; accepted 17 November; published online 10 December 2006; doi:10.1038/nsmb1182

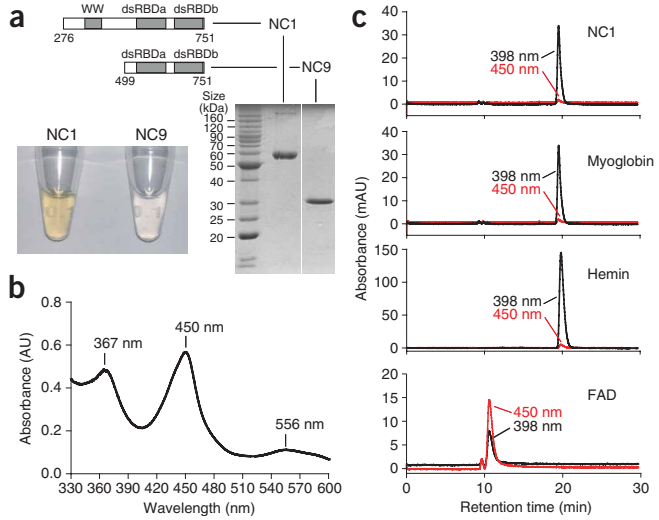

Figure 1 DGCR8 contains a heme cofactor. (**a**) A truncated DGCR8 (NC1, 4.0 mg ml^{-1}) has a yellow color, whereas NC9 (15.5 mg ml^{-1}) does not. Both proteins were overexpressed in *E. coli* and highly purified, as indicated by SDS-PAGE. The domain structures and ranges of amino acid residues of these proteins are indicated. Full-length human DGCR8 contains 773 residues. (**b**) UV-visible absorption spectrum of NC1. AU, absorbance units. (**c**) Extraction and HPLC characterization of the DGCR8-bound cofactor, using horse heart myoglobin, pure hemin and FAD as controls.

myoglobin and pure hemin, but at a later time than FAD (**Fig. 1c**). After extraction, the absorbance of the cofactor at 450 nm (A_{450}; bound to protein) was greatly reduced and a large absorption signal appeared at 398 nm, as expected for a free heme. The NC1 protein was eluted from the column much earlier than heme, suggesting that it is not covalently linked to the protoporphyrin of heme. Third, ESI-MS of NC1 under denaturing conditions showed an abundant mass/charge ratio (m/z) peak at 616, the molecular weight of free heme (**Supplementary Fig. 2** online).

Heme serves as a prosthetic group of hemoproteins, such as hemoglobins, cytochromes and guanylate cyclases. It is involved in important transportation, catalytic, electron transfer and signaling activities[32]. The finding that DGCR8 is a heme-binding protein led to the hypothesis that DGCR8 may be involved in the regulation of pri-miRNA processing, in addition to acting as an RNA-binding partner of Drosha.

Heme promotes dimerization of DGCR8

The initial evidence of a function of the heme bound to DGCR8 came from examination of the protein's oligomerization state using size-exclusion chromatography (SEC). In an SEC experiment, the elution volume of a protein can be used to estimate its molecular weight, thus defining its oligomerization state (see **Supplementary Fig. 3** online for standard curves). As the last step of the NC1 purification, the SEC gave two major protein peaks (**Fig. 2a**). Both contained essentially pure NC1, as indicated by denaturing gels. These peaks corresponded to a monomer and dimer of NC1, respectively. Only the NC1 dimer contained heme, as indicated by the A_{450}, whereas the monomer bound no heme.

Does heme mediate the dimerization of DGCR8, or does the dimerization cause an association with heme that does not affect the monomer-dimer equilibrium? To distinguish these possibilities, we added δ-aminolevulinic acid (δ-ALA) to the *E. coli* culture overexpressing NC1. δ-ALA is a heme biosynthesis intermediate that the bacterium can use to produce heme; addition of δ-ALA has been shown to increase the heme content in bacterial cells and to increase the yield and activity of overexpressed heme-binding proteins[33]. The addition of δ-ALA tripled the yield of NC1 in the overexpression and purification experiments and, moreover, shifted NC1 nearly exclusively to its dimeric state (**Fig. 2b**). This result indicates that the association with heme promotes the dimerization of DGCR8.

Two independent approaches were used to determine the protein/heme molar ratio in the NC1 dimer complex. First, integration of the absorption peaks in the HPLC experiments (**Fig. 1c**) provided an estimate of 2.4:1 in the pure NC1 dimer sample (**Supplementary Fig. 4** online). Second, ESI-MS, which has been used to examine noncovalently bound macromolecular complexes[34], showed a difference of ~618 Da between the heme-bound and heme-free NC1 dimers, close to the molecular weight (616 Da) of a heme cofactor (**Fig. 2c**). Thus, each NC1 dimer binds one heme.

DGCR8 interacts with heme through Cys352

The absorption peak of DGCR8 at 450 nm is characteristic of a ferrous heme with a cysteine as an axial ligand, best known from the cytochrome P450 and nitric oxide synthase enzymes[31,35]. A region (residues 276–498) of DGCR8 that includes the putative WW motif (residues 303–332) is required for the association with heme, because a further truncated DGCR8 protein (NC9, residues 499–751) is colorless (**Fig. 1a**) and does not have any major peak in its absorption spectrum between 350 and 600 nm. There are two cysteine residues in

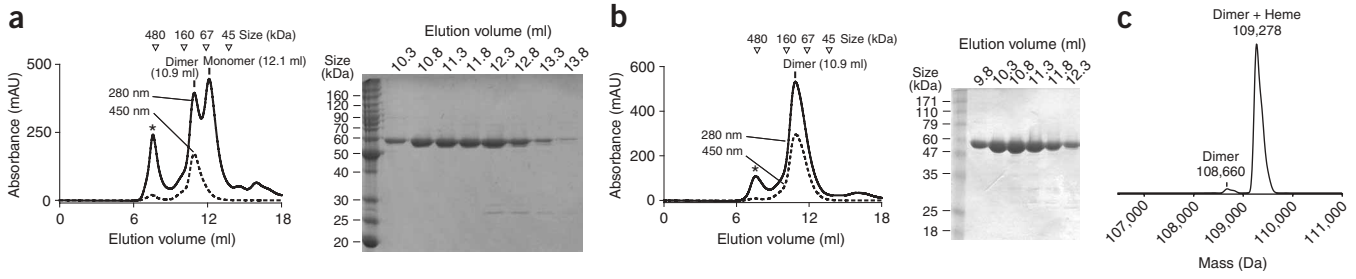

Figure 2 Heme promotes the dimerization of DGCR8. (**a**,**b**) Size-exclusion chromatograms in the last step of the purification of the NC1 protein are shown. NC1 was overexpressed in *E. coli* without (**a**) and with (**b**) 1 mM δ-ALA. Before SEC, the NC1 protein was already over 90% pure, as judged by denaturing PAGE (data not shown). Assignment of the monomer and dimer species of NC1 was based on their elution volumes (indicated in parentheses) and a standard curve (**Supplementary Fig. 3**). Open triangles mark elution volumes of standard proteins. Asterisks indicate an impurity (mostly nucleic acids) from the bacterial extract. AU, absorbance units. Fractions around the peaks were examined using SDS-PAGE (right). Starting volume of each 0.5-ml fraction is indicated above gel. (**c**) ESI mass spectrum of NC1 dimer in **b** confirmed the protein/heme stoichiometry of 2:1. The spectrum was acquired at pH 6.8 and was deconvoluted to the mass domain.

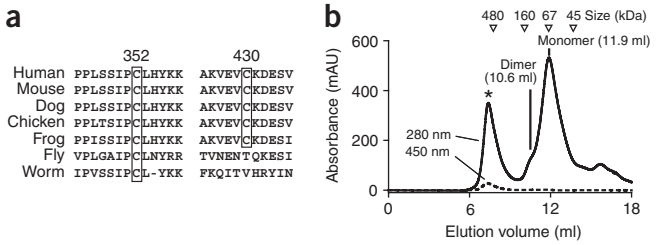

Figure 3 A conserved cysteine probably serves as an axial ligand of heme. (**a**) Sequence alignments of DGCR8 homologs. (**b**) The C352A mutant does not bind heme and is mostly a monomer, as indicated by the SEC in the last step of the purification (shown as in **Fig. 2a**,**b**).

this region: Cys352 is conserved in all DGCR8 homologs; Cys430 is conserved only in organisms that are evolutionarily close to human (**Fig. 3a**). We mutated Cys352 to alanine and histidine in the context of NC1. These mutants did not associate with heme when over-expressed and purified using the same protocol as for NC1 (**Fig. 3b** and **Supplementary Fig. 5** online). In contrast, the C430A mutant was capable of binding heme and the complex had absorption peaks (at 366, 450 and 556 nm) similar to those of the wild-type NC1 (**Supplementary Fig. 5**). These results suggest that Cys352 probably serves as an axial ligand to the iron in the heme cofactor.

The heme-free C352A and C352H mutants were mostly monomeric (**Fig. 3b** and **Supplementary Fig. 5**), further supporting the model that heme is involved in regulating the oligomerization state of DGCR8. On the other hand, small fractions of C352A and C352H were present in a dimeric state, suggesting that heme is not absolutely required for the dimerization of DGCR8. This observation led us to pose the question, are the monomeric and dimeric forms of DGCR8 functionally distinct?

Distinct forms of DGCR8–pri-miRNA complexes

A filter-binding assay system was used to examine the interaction between pri-miRNA and DGCR8. A fragment of pri-miR-30a containing the miRNA hairpin and its immediate flanking regions was used in all the RNA-binding assays and most processing assays reported in this study. The binding curves (fraction RNA bound to protein versus DGCR8 concentration) of the heme-bound dimer and the monomer were highly cooperative and could not be fit with a simple model in which one protein molecule associates with one RNA molecule (**Fig. 4a**). Hill plots of these data indicate straight lines in the range of the binding transition, with distinct slopes of 2.9 ± 0.3 for the heme-bound dimer and 1.8 ± 0.2 for the heme-free monomer (mean ± s.e.m. from three independent experiments; **Fig. 4b**). The slope of a curve on a Hill plot (the Hill constant, n) represents the lower limit on the oligomerization of the protein in a binding reaction to another molecule. When cooperativity is high, few intermediate oligomers are populated and the Hill constant should be close to the number of protein subunits oligomerized. This situation has been observed in the tetrameric binding of the Arc repressor to its operator DNA[36]. The fact that Hill plots of the NC1 proteins binding pri-miRNA are straight lines suggests that the cooperativity is high. Indeed, the binding curves are best fit by simple models in which three (for the heme-bound dimer) or two (for the heme-free monomer) NC1 molecules bind the pri-miRNA highly cooperatively (**Fig. 4a**). In these models, the heme-bound dimer binds a pri-miRNA as a trimer of dimers (a hexamer overall), and the heme-free monomer binds as a dimer. The x-intercept of the Hill plot was used to calculate the K-value (the protein concentration at which 50% RNA is bound) as 26 ± 2 nM (dimer concentration) for the heme-bound dimer and 52 ± 8 nM for the heme-free monomer.

To further characterize the NC1–pri-miR-30a RNA complexes, we mixed NC1 heme-bound dimer and heme-free monomer with pri-miR-30a RNA at around micromolar concentrations, and SEC was performed in a low-salt buffer, measuring the A_{260}, A_{280} and A_{450} simultaneously. Despite the fact that the protein/RNA molar ratios in both samples were the same (6:1), the chromatograms indicate that distinct complex species eluted at very different volumes (**Fig. 4c**).

The SEC results (**Fig. 4c**) support the models described above in four respects. First, the NC1 dimer and monomer proteins complexed with the pri-miR-30a RNA were eluted from the sizing column in nearly single peaks, indicating that a major complex species is formed in each binding reaction. This observation is consistent with the assumption of high cooperativity used to derive the model. Second, the elution volumes of the complex peaks were used to estimate their molecular weights on the basis of a standard curve (**Supplementary Fig. 3**): 310 kDa for the NC1 dimer–RNA complex and 170 kDa for the monomer in complex with the pri-miR-30a RNA. These values were consistent with expectations (372 kDa for an NC1 hexamer with one pri-miR-30a RNA; 156 kDa for an NC1 dimer with one pri-miR-30a RNA), within experimental error.

Third, the A_{260}/A_{280} ratio (R) was used to estimate the stoichiometry of the protein and RNA in their complexes. The absorption ratios of the complexes ($R_{complex}$, typically between 1.7 and 2.1) determined from the chromatograms were distinct from those of the RNA (R_{RNA}, 2.36 for pri-miR-30a) and the NC1 proteins ($R_{protein}$, 0.639 for NC1 dimer and 0.579 for NC1 monomer) individually and

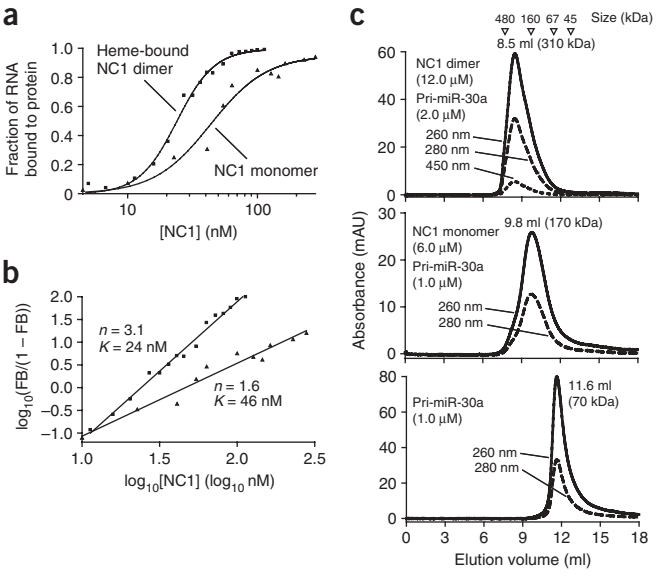

Figure 4 Dimer and monomer of NC1 form different higher-order structures upon binding pri-miRNA. (**a**) Filter-binding assays demonstrated that both forms of NC1 associate with the pri-miR-30a RNA cooperatively. The data were best fit using a cooperative trimer model for the heme-bound NC1 dimer (squares) and a cooperative dimer model for the heme-free NC1 monomer (triangles). The x-axis represents the concentration of either NC1 dimer (squares) or monomer (triangles). (**b**) Hill plots of the data in **a** around the binding transition. The result shown is one of three experiments. K is defined as $10^{(x\text{-intercept})}$. (**c**) SEC (shown as in **Fig. 2a**,**b**) was used to estimate the molecular weights of the NC1–pri-miR-30a complexes.

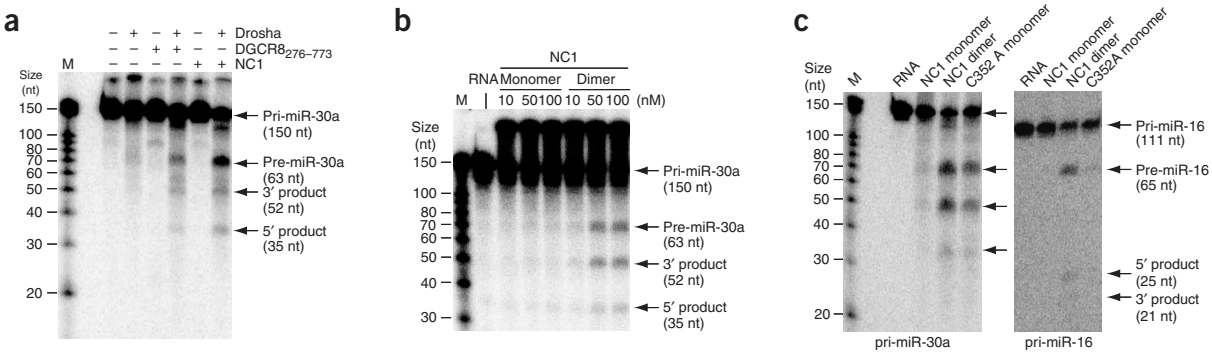

Figure 5 The DGCR8 dimer is active in pri-miRNAs processing *in vitro*. (**a**) Recombinant Drosha and DGCR8 were used to reconstitute processing activity using the pri-miR-30a RNA as a substrate. (**b**) Cleavage assays of pri-miR-30a by Drosha in the presence of NC1 monomer and dimer proteins, respectively. (**c**) Cleavage assays of pri-miR-30a (left) and pri-miR-16 (right) by Drosha in the presence of 100 nM NC1 monomer or dimer proteins or C352A mutant monomer. M, size markers.

thus were used to calculate the molar ratio of the protein and RNA in their complexes, [protein]/[RNA], using the following equation (see **Supplementary Methods** online):

$$\frac{[\text{protein}]}{[\text{RNA}]} = \frac{R_{\text{RNA}} - R_{\text{complex}}}{R_{\text{RNA}}(R_{\text{complex}} - R_{\text{protein}})} \cdot \frac{\varepsilon_{\text{RNA},260}}{\varepsilon_{\text{protein},280}}$$

where $\varepsilon_{\text{RNA},260}$ is the extinction coefficient of the RNA, which is 1.20×10^6 M^{-1} cm^{-1} for the pri-miR-30a RNA, and $\varepsilon_{\text{protein},280}$ is the extinction coefficient of the protein, which is 4.72×10^4 M^{-1} cm^{-1} for an NC1 protomer (a monomer or half a dimer). The protein/RNA ratios were 5.1 ± 0.9 for the NC1 dimer–pri-miR-30a complex and 2.4 ± 0.4 for the NC1 monomer–pri-miR-30a complex.

Fourth, the A_{450} in the chromatograms of the NC1 dimer–RNA complex (**Fig. 4c**) provided an independent estimate of the stoichiometry of the protein and RNA components. The ratios of the absorptions of the NC1 dimer protein (typically, A_{260}/A_{450} is around 1.22 and A_{280}/A_{450} is about 1.94) were used to estimate the contribution of the protein to the total A_{260} and A_{280}, assuming that these ratios were not changed upon association with the pri-miRNA. The A_{260} contributed by the RNA was obtained by subtracting the value contributed by protein from the total. The protein/RNA molar ratio was calculated from their estimated contributions to A_{280} and A_{260}, respectively, as 5.4 ± 0.3. Given the stoichiometry of DGCR8 and pri-miRNA determined experimentally, it is unlikely that each complex contains more than one RNA molecule, because the molecular weight would be much larger than the values estimated from the elution volumes (**Fig. 4c**).

In conclusion, we found that the heme-bound NC1 dimer and the heme-free monomer form distinct complexes with a pri-miRNA. Our data are consistent with a model in which the dimer trimerizes to form a hexamer on the pri-miRNA and the monomer forms mainly a dimer upon association with the pri-miRNA. However, because the estimated molecular weights of the DGCR8 monomer in complex with the pri-miRNA, based on the SEC experiments, are slightly higher than the calculated ones, we cannot rule out the possibility that the monomeric DGCR8 can also trimerize in complex with the pri-miRNA.

DGCR8 contains two putative dsRBDs and is thought to be the RNA-binding partner of Drosha[19–22]. A UV cross-linking study has recently demonstrated that DGCR8 associates directly and specifically with pri-miRNAs[23]. We have shown here that both dimeric and monomeric DGCR8 proteins bind pri-miRNAs with affinities between 20 and 50 nM. Moreover, different forms of DGCR8 form distinct higher-order structures upon binding pri-miRNAs and could therefore function differently in pri-miRNA processing.

Heme-bound DGCR8 is active in pri-miRNA processing

To evaluate the function of heme in miRNA maturation, we overexpressed Drosha in insect cells and reconstituted pri-miRNA processing assays *in vitro* using purified Drosha and NC1 (**Fig. 5a**). The cleavage activity depends on the presence of both Drosha and NC1. NC1 does not contain the C-terminal 22 residues of DGCR8 but is at least as active as the minimal DGCR8 construct reported in the literature[21], which contains residues 276–773 (**Fig. 5a**).

The DGCR8 constructs used in the initial studies (**Fig. 5a**) were a mixture of dimer and monomer. We subsequently examined the purified NC1 dimer (heme-bound) and monomer (heme-free) using *in vitro* cleavage assays. The heme-bound NC1 dimer was active in triggering pri-miRNA cleavage by Drosha at concentrations above the RNA-binding transition; the heme-free monomeric NC1 was much less active at all the concentrations tested (**Fig. 5b**). The monomeric C352A mutant was found to be active at a lower level than the wild-type dimer (**Fig. 5c**). It is possible that the C352A mutation affects both heme binding and an autoinhibition of DGCR8 that will be discussed below. Different forms of DGCR8 were also tested using pri-miR-16, which was used previously in reconstituted processing assays[21]. The results were similar to those obtained using pri-miR-30a as the substrate (**Fig. 5c**).

The dsRBDs are the effector domains

To evaluate the function of the conserved region required for association with heme (residues 276–498), we constructed the NC9 protein, which contains the two dsRBDs but not the WW motif (**Fig. 1a**). We found that NC9 is monomeric in the absence of RNA (**Fig. 6a**). NC9 bound pri-miR-30a cooperatively, with a Hill constant (n) of 2.8 ± 0.1, and with essentially the same affinity (K = 25 ± 10 nM; **Fig. 6b,c**) as the heme-bound NC1 dimer. The binding curve was best fit with a model in which three NC9 molecules bind cooperatively to a pri-miRNA (**Fig. 6b**). SEC of the NC9–pri-miR-30a complex mixed at a 3:1 molar ratio (**Fig. 6d**) indicated that the complex contained a major oligomeric species, suggesting that the Hill constant is likely to be close to the oligomerization state of NC9 in the complex. The elution volume of this peak provided an estimated molecular weight of 111 kDa, close to those of complexes containing one RNA molecule and two or three NC9 subunits (106 kDa or 135 kDa, respectively).

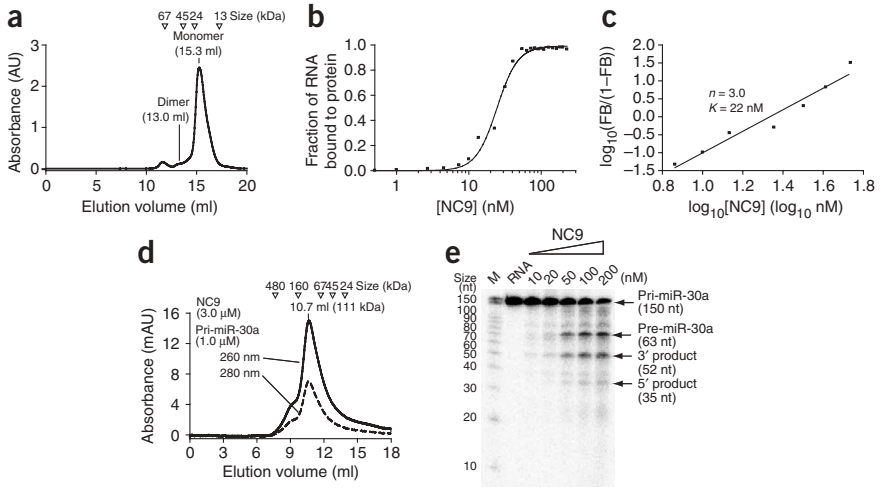

Figure 6 The dsRBD domains (NC9) contain the core pri-miRNA–binding and pri-miRNA–processing activities. (**a**) The NC9 protein is mostly monomeric, as indicated by the SEC in the last step of the purification (shown as in **Fig. 2a,b**). The purity of NC9 in the major peak was shown by SDS-PAGE (**Fig. 1a**). (**b**) Filter-binding assays showed that NC9 binds the pri-miR-30a RNA cooperatively, and the data were best fit using a cooperative trimer model. (**c**) The Hill plot confirmed the cooperativity. The result shown is one of three experiments. (**d**) SEC was used to estimate the molecular weight of the NC9–pri-miR-30a complex. (**e**) NC9 was active in triggering the cleavage of pri-miRNA by Drosha when present at concentrations higher than the binding transition. M, size markers.

Because NC9 is relatively small (29 kDa), the NC9–pri-miR-30a complex is predicted to be elongated. Thus, the error of the estimated molecular weight is expected to be relatively large and it is difficult to determine the true number of NC9 protomers present in the complex based on the elution volume. The A_{260}/A_{280} ratio for the complex peak (**Fig. 6d**) was 2.11. Using the formula described earlier, the molar ratio of NC9 to pri-miR-30a was calculated to be 3.4, also consistent with a trimeric NC9 bound to a pri-miRNA. Furthermore, NC9 was found to be active in pri-miRNA processing assays when present at concentrations higher than the binding transition indicated by the filter-binding assays (**Fig. 6e**). Therefore, we conclude that the two dsRBDs of DGCR8 are the effector domains for pri-miRNA processing, the pri-miRNA–binding and pri-miRNA–processing activities of which seem to be similar to those of the heme-bound DGCR8 dimer.

The fact that the NC1 monomer is much less active than the heme-bound dimer in pri-miRNA processing assays suggests that the region between residues 276 and 498 has an autoinhibitory function. The association with heme promotes the dimerization of DGCR8 and activates its pri-miRNA–processing activity. The presence of this autoinhibitory mechanism strongly suggests that the DGCR8 monomer is physiologically relevant. The biological activities of many proteins, including the Src kinase[37], are regulated through their autoinhibitory domains[38]. DGCR8 seems to be a previously uncharacterized type of autoinhibitory protein, in which autoinhibition modulates an interaction with target RNA.

DISCUSSION

We showed that the human DGCR8 protein interacts with heme through a conserved residue, Cys352, which probably serves as an axial ligand of the iron. In the absence of heme, DGCR8 is mostly monomeric. The binding of heme promotes the dimerization of DGCR8, forming a complex containing one heme molecule per dimer. Both forms of DGCR8 contact pri-miRNA with rather similar affinities, but they form distinct higher-order structures. The heme-bound dimer seems to bind as a cooperative trimer (of dimers) and is fully active in the pri-miRNA cleavage assay. The heme-free monomer probably binds as a cooperative dimer and is much less active than the heme-bound dimer. The region important for association with heme also has an autoinhibitory function in pri-miRNA processing, probably preventing the trimerization of DGCR8 upon binding to pri-miRNAs. Thus, we have elucidated a possible molecular mechanism through which heme regulates the processing of primary miRNAs.

The experimental evidence presented here suggests that trimerization of DGCR8 upon association with pri-miRNAs might be important for triggering cleavage by Drosha. The heme-bound dimer seems to act as a functional unit, equivalent to an NC9 monomer. Few nucleic acid–binding proteins function in trimeric forms. The yeast heat-shock transcription factor forms a constitutive trimer and binds DNA elements containing arrays of a 5-base-pair recognition unit[39]. Recently, it has been shown that the Drosha–DGCR8 complex determines the cleavage site by measuring ~11 nt from the base of the pri-miRNA hairpin. DGCR8 binds directly to pri-miRNAs, suggesting that DGCR8 may function as the molecular anchor[23]. The trimerization of DGCR8 we demonstrated might be the molecular basis for this measuring mechanism.

Biogenesis of miRNAs is likely to be regulated at both transcription and processing steps. miRNAs are synthesized by RNA polymerase II[40,41], which also transcribes most mRNAs and is known to be highly regulated. Thus, it is conceivable that the transcription of miRNAs is regulated through mechanisms similar to those that have been elucidated for protein-coding genes. It has been shown that a large portion of pri-miRNAs are cotranscribed with other RNA[42]. Regulation at the processing steps would provide a way to generate miRNA expression patterns distinct from those of the host genes.

It is also possible for cells to reduce the global expression of miRNAs by downregulating the processing steps. In fact, a very recent study has demonstrated that during development and in primary tumors, many miRNA primary transcripts are not efficient processed by Drosha[43]. It is noteworthy that the expression levels of Drosha and DGCR8 are not induced concordantly with the production of mature miRNAs. Thus, the post-transcriptional regulation of microRNAs observed in ref. 43 may be mediated by the interaction between heme and the DGCR8 protein.

Our findings suggest that miRNAs might be involved in the signaling network of heme. Heme has been shown to regulate its own biosynthesis and to stimulate the differentiation of erythroid, adipose and neuronal cells in mammals[32]. Furthermore, heme biosynthesis and the circadian clock are reciprocally regulated[44]. Indeed, a search of the miRanda server[45] indicates that many genes involved in heme biosynthesis and the circadian clock are predicted to contain considerably larger numbers of conserved miRNA target sites in the 3′ untranslated region than are found in a typical gene. Our experimental evidence is consistent with the model that DGCR8 is a sensor for heme. Another possible role for the heme bound to DGCR8 is to mediate a response to diatomic gas signaling molecules such as carbon

monoxide, nitric oxide or oxygen, similarly to a class of hemoprotein sensors found in a variety of organisms[46]. We have started to test the latter possibility using the assays described above and have not found that carbon monoxide alters DGCR8 activity (data not shown).

Our study suggests that heme is required for the activation of DGCR8. Under conditions of heme deficiency, the pri-miRNA cleavage is predicted to be slowed and the abundances of mature miRNAs should be decreased globally. In humans, defects in heme biosynthesis cause diseases such as porphyrias. The neurological symptoms of porphyrias include peripheral and autonomic neuropathy[47]. In addition, heme deficiency has also been shown to decrease the abundance of mitochondrial complex IV, to activate nitric oxide synthase, to alter amyloid precursor protein and to disrupt iron and zinc homeostasis in brain cells—phenotypes similar to those of dysfunctional neurons in patients with Alzheimer disease[48,49]. The underlying mechanisms for these observations are not clear. As miRNAs have been shown to be involved in neuron development[50], it is possible that heme deficiency causes some of these phenotypes through a decrease in miRNA processing.

METHODS

Protein expression and purification. The complementary DNA clone of DGCR8 was obtained from Open Biosystems. Sequence encoding residues 275–751 (NC1) was subcloned into the pET-24a vector (Novagen) between the NdeI and EcoRI sites. NC1 was overexpressed in *E. coli* BL21(DE3) CodonPlus cells (Stratagene) by addition of 1 mM IPTG. Where indicated, δ-ALA was added to the culture at a concentration of 1 mM together with IPTG. The NC1 protein was extracted from the cells by sonication in 20 mM Tris (pH 8.0), 100 mM NaCl, 10 mM β-mercaptoethanol (β-ME) and 0.25 mM phenylmethanesulfonyl fluoride. NC1 was purified by two rounds of ion-exchange chromatography using a 5-ml HiTrap SP column (Amersham), followed by an SEC using a Superdex 200 column (Amersham). The running buffer contained 20 mM Tris (pH 8.0), 400 mM NaCl and 10 mM β-ME. The monomeric NC1 was purified further by concentration of the corresponding peak fractions from the SEC using an Amicon Ultra-15 concentrator (Millipore), followed by an additional SEC step. The NC9 protein was overexpressed and purified using a similar procedure, except that only one ion-exchange purification step was necessary and the protein solution was exchanged to a buffer containing 20 mM sodium acetate (pH 5.1), 300 mM NaCl and 1 mM DTT before the SEC in the same buffer. The $DGCR8_{276–773}$ construct was expressed using similar protocol to that of NC1. It contains a C-terminal His_6 tag and was purified using affinity purification followed by an SEC.

The Drosha sequence (see Acknowledgments) was subcloned into the pFastBac-HTb vector (Invitrogen), which was used in the overexpression experiment in Sf9 insect cells. The Drosha protein was extracted from the cells by sonication in 20 mM Tris (pH 8.0) and 400 mM NaCl and was purified by SEC using a Superdex 200 column in the same buffer. The early fractions, containing partially purified Drosha, were pooled and dialyzed against a storage buffer containing 20 mM Tris (pH 8.0), 100 mM NaCl, 0.2 mM EDTA and 10% (v/v) glycerol. The protein was stored at –80 °C until use.

Heme extraction and quantification using high-performance liquid chromatography. Protein solutions were mixed with four volumes of solvent A, which contained acetonitrile, water and trifluoroacetic acid at the volume ratio of 60:40:0.1. The mix was injected onto an Ultrasphere Semi-prep C18 reverse-phase column (Beckman). Elution was performed isocratically with solvent A at a flow rate of 1 ml min^{-1} and monitored at wavelengths of 398 and 450 nm simultaneously. The integrated peak area was obtained using UNICORN (Amersham).

Size-exclusion chromatography. The SEC characterization of DGCR8 was performed at room temperature under the same conditions as described in the purification procedure. The running buffer used in SEC of the DGCR8–RNA complexes (**Fig. 4c** and **Fig. 6d**) contained 20 mM Tris (pH 8.0) and 80 mM NaCl.

Electrospray ionization mass spectrometry. A nanoESI-QqTOF analyzer (QSTAR Pulsar XL, m/z range 40,000; Applied Biosystems/MDS Sciex) was used for the ESI-MS experiments[34]. Borosilicate glass nanospray emitters coated with gold and palladium to allow for spray operation at 10–50 nl min^{-1} were obtained from Proxeon. Before ESI-MS analyses, samples were desalted and concentrated using Millipore Microcon centrifugal filter devices in 20 mM ammonium acetate (pH 6.8) and 0.2 mM DTT. ESI-MS measurements of noncovalently bound complexes are sensitive to the energy used in the atmosphere-vacuum interface of the ESI mass spectrometer to desolvate the hydrated complexes as they transition from the liquid state to the gaseous state. The spectrum shown in **Figure 2c** represents a low interface potential (+140 V) that minimizes protein-ligand dissociation. Using higher interface potentials allows release of the heme molecule from the protein dimer, as shown by an abundant peak at $m/z = 616$ (**Supplementary Fig. 2**).

Pri-miRNA binding and *in vitro* processing assays. The sequence of the pri-miR-30a was cloned from human genomic DNA into the pUC19 vector. The plasmid was linearized and was used as the template in transcription reactions using T7 RNA polymerase. The transcribed RNA covers 33- and 53-nt sequences flanking the pre-miRNA region at its 5′ and 3′ termini, respectively. Two guanosine residues were added at the 5′ end to improve the yield of the transcription. The RNA was uniformly labeled with ^{32}P.

Each binding reaction (22 µl total volume) contained 2.2 µl of DGCR8 protein solution, 1 µl annealed RNA and a buffer containing 20 mM Tris (pH 7.5), 50 mM NaCl and 1 mM DTT. Series of DGCR8 dilutions were prepared in 20 mM Tris (pH 8.0), 400 mM NaCl and 10 mM β-ME. The uniformly labeled RNA (described above) was annealed in 10 mM sodium HEPES buffer (pH 7.0) and 50 mM NaCl by heating to 65 °C for 3 min and cooling at room temperature for 15 min. The binding reactions were filtered through a nitrocellulose membrane and a positively charged nylon membrane sequentially, using a spot-blot apparatus. The membranes were air-dried and exposed to an image plate. The intensity of the spots was integrated using Quantity One (Bio-Rad). The data were fit to curves using PRISM (GraphPad). The equation used in the fitting with cooperative binding models is $FB = FB_0[DGCR8]^n/(K_d + [DGCR8]^n)$, where FB is the fraction of RNA bound to DGCR8, FB_0 is the fitted value of the maximum FB, [DGCR8] is the molar concentration of DGCR8 and n is 2 for a dimer model or 3 for a trimer model and is fixed in a fitting process.

The processing assays were performed essentially as reported[21]. The reactions contained purified Drosha (14 nM) and DGCR8 at concentrations indicated.

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

A clone containing human Drosha sequence was kindly provided by V.N. Kim (Seoul National University). We thank T. Cech, D. Eisenberg, J. Fukuto, C.W. Goulding and S. Kurtistani for comments of the manuscript, R. Landgraf, V. Kickhoefer and L. Simpson for assistance and advice on insect-cell protein expression, and D. Black, S. Clarke, A. Fire, A. Fulco, O. Hankinson, S. Merchant and J.S. Valentine for discussion. This work was supported in part by seed grants from the Stein Oppenheimer Endowment Fund and the Jonsson Comprehensive Cancer Center at UCLA to F.G. The UCLA Functional Proteomics Center was established and equipped by a grant from the W.M. Keck Foundation. J.A.L. acknowledges support from the UCLA–US Department of Energy Institute for Genomics and Proteomics, and the US National Institutes of Health (RR 20004).

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/nsmb/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Bartel, D.P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281–297 (2004).
2. Lee, R.C., Feinbaum, R.L. & Ambros, V. The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75**, 843–854 (1993).
3. Brennecke, J., Hipfner, D.R., Stark, A., Russell, R.B. & Cohen, S.M. *bantam* encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in *Drosophila*. *Cell* **113**, 25–36 (2003).

4. Chen, C.Z., Li, L., Lodish, H.F. & Bartel, D.P. MicroRNAs modulate hematopoietic lineage differentiation. *Science* **303**, 83–86 (2004).
5. Cimmino, A. *et al.* miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proc. Natl. Acad. Sci. USA* **102**, 13944–13949 (2005).
6. Johnson, S.M. *et al.* RAS is regulated by the let-7 microRNA family. *Cell* **120**, 635–647 (2005).
7. Meister, G. & Tuschl, T. Mechanisms of gene silencing by double-stranded RNA. *Nature* **431**, 343–349 (2004).
8. Lagos-Quintana, M., Rauhut, R., Lendeckel, W. & Tuschl, T. Identification of novel genes coding for small expressed RNAs. *Science* **294**, 853–858 (2001).
9. Lau, N.C., Lim, L.P., Weinstein, E.G. & Bartel, D.P. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* **294**, 858–862 (2001).
10. Lee, R.C. & Ambros, V. An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* **294**, 862–864 (2001).
11. Pfeffer, S. *et al.* Identification of microRNAs of the herpesvirus family. *Nat. Methods* **2**, 269–276 (2005).
12. Griffiths-Jones, S., Grocock, R.J., van Dongen, S., Bateman, A. & Enright, A.J. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.* **34**, D140–D144 (2006).
13. Lewis, B.P., Burge, C.B. & Bartel, D.P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**, 15–20 (2005).
14. Lee, Y., Jeon, K., Lee, J.T., Kim, S. & Kim, V.N. MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J.* **21**, 4663–4670 (2002).
15. Kim, V.N. MicroRNA biogenesis: coordinated cropping and dicing. *Nat. Rev. Mol. Cell Biol.* **6**, 376–385 (2005).
16. Lund, E., Guttinger, S., Calado, A., Dahlberg, J.E. & Kutay, U. Nuclear export of microRNA precursors. *Science* **303**, 95–98 (2004).
17. Yi, R., Qin, Y., Macara, I.G. & Cullen, B.R. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev.* **17**, 3011–3016 (2003).
18. Lee, Y. *et al.* The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415–419 (2003).
19. Denli, A.M., Tops, B.B., Plasterk, R.H., Ketting, R.F. & Hannon, G.J. Processing of primary microRNAs by the Microprocessor complex. *Nature* **432**, 231–235 (2004).
20. Gregory, R.I. *et al.* The Microprocessor complex mediates the genesis of microRNAs. *Nature* **432**, 235–240 (2004).
21. Han, J. *et al.* The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev.* **18**, 3016–3027 (2004).
22. Landthaler, M., Yalcin, A. & Tuschl, T. The human DiGeorge syndrome critical region gene 8 and its *D. melanogaster* homolog are required for miRNA biogenesis. *Curr. Biol.* **14**, 2162–2167 (2004).
23. Han, J. *et al.* Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* **125**, 887–901 (2006).
24. Bernstein, E., Caudy, A.A., Hammond, S.M. & Hannon, G.J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**, 363–366 (2001).
25. Grishok, A. *et al.* Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* **106**, 23–34 (2001).
26. Hutvagner, G. *et al.* A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* **293**, 834–838 (2001).
27. Ketting, R.F. *et al.* Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev.* **15**, 2654–2659 (2001).
28. Knight, S.W. & Bass, B.L. A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*. *Science* **293**, 2269–2271 (2001).
29. Hammond, S.M., Bernstein, E., Beach, D. & Hannon, G.J. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* **404**, 293–296 (2000).
30. Gregory, R.I., Chendrimada, T.P., Cooch, N. & Shiekhattar, R. Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* **123**, 631–640 (2005).
31. Mayer, B.J. *et al.* Purification of brain nitric oxide synthase from baculovirus overexpression system and determination of cofactors. in *Nitric Oxide Synthase: Characterization and Functional Analysis* (ed. Maines, M.D.) 130–139 (Academic Press, San Diego, 1996).
32. Ponka, P. Cell biology of heme. *Am. J. Med. Sci.* **318**, 241–256 (1999).
33. Kery, V., Elleder, D. & Kraus, J.P. Delta-aminolevulinate increases heme saturation and yield of human cystathionine beta-synthase expressed in *Escherichia coli*. *Arch. Biochem. Biophys.* **316**, 24–29 (1995).
34. Loo, J.A. *et al.* Electrospray ionization mass spectrometry and ion mobility analysis of the 20S proteasome complex. *J. Am. Soc. Mass Spectrom.* **16**, 998–1008 (2005).
35. Dawson, J.H. & Sono, M. Cytochrome P-450 and chloroperoxidase: thiolate-ligated heme enzymes. Spectroscopic determination of their active-site structures and mechanistic implications of thiolate ligation. *Chem. Rev.* **87**, 1255–1276 (1987).
36. Brown, B.M., Bowie, J.U. & Sauer, R.T. Arc repressor is tetrameric when bound to operator DNA. *Biochemistry* **29**, 11189–11195 (1990).
37. Xu, W., Doshi, A., Lei, M., Eck, M.J. & Harrison, S.C. Crystal structures of c-Src reveal features of its autoinhibitory mechanism. *Mol. Cell* **3**, 629–638 (1999).
38. Pufall, M.A. & Graves, B.J. Autoinhibitory domains: modular effectors of cellular regulation. *Annu. Rev. Cell Dev. Biol.* **18**, 421–462 (2002).
39. Sorger, P.K. & Nelson, H.C. Trimerization of a yeast transcriptional activator via a coiled-coil motif. *Cell* **59**, 807–813 (1989).
40. Lee, Y. *et al.* MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* **23**, 4051–4060 (2004).
41. Cai, X., Hagedorn, C.H. & Cullen, B.R. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA* **10**, 1957–1966 (2004).
42. Rodriguez, A., Griffiths-Jones, S., Ashurst, J.L. & Bradley, A. Identification of mammalian microRNA host genes and transcription units. *Genome Res.* **14**, 1902–1910 (2004).
43. Thomson, J.M. *et al.* Extensive post-transcriptional regulation of microRNAs and its implications for cancer. *Genes Dev.* **20**, 2202–2207 (2006).
44. Kaasik, K. & Lee, C.C. Reciprocal regulation of haem biosynthesis and the circadian clock in mammals. *Nature* **430**, 467–471 (2004).
45. John, B. *et al.* Human microRNA targets. *PLoS Biol.* **2**, e363 (2004).
46. Gilles-Gonzalez, M.A. & Gonzalez, G. Heme-based sensors: defining characteristics, recent developments, and regulatory hypotheses. *J. Inorg. Biochem.* **99**, 1–22 (2005).
47. Meyer, U.A., Schuurmans, M.M. & Lindberg, R.L. Acute porphyrias: pathogenesis of neurological manifestations. *Semin. Liver Dis.* **18**, 43–52 (1998).
48. Atamna, H., Killilea, D.W., Killilea, A.N. & Ames, B.N. Heme deficiency may be a factor in the mitochondrial and neuronal decay of aging. *Proc. Natl. Acad. Sci. USA* **99**, 14807–14812 (2002).
49. Atamna, H. & Frey, W.H., II. A role for heme in Alzheimer's disease: heme binds amyloid beta and has altered metabolism. *Proc. Natl. Acad. Sci. USA* **101**, 11153–11158 (2004).
50. Chang, S., Johnston, R.J., Jr., Frokjaer-Jensen, C., Lockery, S. & Hobert, O. MicroRNAs act sequentially and asymmetrically to control chemosensory laterality in the nematode. *Nature* **430**, 785–789 (2004).