Science Signaling
AAAS

www.sciencesignaling.org/cgi/content/full/7/345/ra91/DC1

Supplementary Materials for

TGF-β–induced epithelial-to-mesenchymal transition proceeds through stepwise activation of multiple feedback loops

Jingyu Zhang, Xiao-Jun Tian, Hang Zhang, Yue Teng, Ruoyan Li, Fan Bai, Subbiah Elankumaran,* Jianhua Xing*

*Corresponding author. E-mail: jxing@vt.edu (J.X.); kumarans@vt.edu (S.E.)

Published 30 September 2014, *Sci. Signal.* **7**, ra91 (2014)
DOI: 10.1126/scisignal.2005304

This PDF file includes:

Other Supplementary Material for this manuscript includes the following:
(available at www.sciencesignaling.org/cgi/content/full/7/345/ra91/DC1)

Text S1. General model for the miRNA-mediated regulation of mRNA

There are two mechanisms for silencing of mRNA by miRNA: translational repression or target degradation. Both mechanisms need formation of mRNA-miRNA complex by basepairing with complementary sequences. Binding of the miRNA and mRNA occurs with a rate constant k_{on}, forming a complex R1 that can either dissociate to miRNA and mRNA with a rate constant k_{off} or be degraded with a rate constant kd_{R1} (fig. S5). During degradation of the complex R1, miRNA can be recycled with a ratio λ (0< λ<1). The translation rate constants of free mRNA and the complex R1 are k_{s0} and k_{s1}, respectively. Thus, the rate equations (described by ordinary differential equations, or ODEs) for the system are:

$$\frac{d[miRNA]_t}{dt} = \mathrm{k_{miR}} - \mathrm{kd_{miR}} * [\mathrm{miRNA}] - \mathrm{kd_{R1}} * (1-\lambda) * [R1],$$

$$\frac{d[mRNA]_t}{dt} = \mathrm{k_{mR}} - \mathrm{kd_{mR}} * [mRNA] - \mathrm{kd_{R1}} * [R1],$$

$$\frac{d[R1]}{dt} = k_{on} * ([miRNA]_t - [R1]) * ([mRNA]_t - [mRNA]) - \mathrm{k_{off}} * [R1] - \mathrm{kd_{R1}} * [R1],$$

$$\frac{d[Protein]}{dt} = k_{s0} * [mRNA] + k_{s1} * [R1] - kd_{Protein} * [Protein],$$

Sometimes, an mRNA contains more than one binding site for the miRNA. Suppose that the number of binding sites is *N*. The number of all possible forms of the miRNA-mRNA complexes and the number of binding and unbinding reactions grow with *N*. Figure S6 shows an example of the possible reactions when $N = 4$. Here, we followed the procedure used by Lu *et al.* and assumed that the rates of binding and unbinding between miRNAs and mRNAs are much faster than the rates of other process, such as transcription, translation, and degradation; therefore, the binding and unbinding processes can be approximated as quasi-equilibrium. In addition, we do not distinguish the difference among different forms of miRNA-mRNA complexes with same number of miRNA molecules bound. We use $[\mathrm{R_i}]$ to represent the abundance of any miRNA-mRNA complex with *i* miRNA bound and assume that the binding affinity for each binding site is the same. Thus, the total abundance of the miRNA-mRNA complex with *i* miRNA bound is $C_N^i R_{\{i\}}$. Under this quasi-equilibrium approximation, the relationship between the $\mathrm{R_i}$ abundance and $\mathrm{R_{i-1}}$ abundance is

$$\mathrm{k_{on}} * [\mathrm{miRNA}] * [\mathrm{R_{i-1}}] = \mathrm{k_{off}} * [\mathrm{R_i}],$$

where [miRNA] is the abundance of the free miRNA in the cell.

We also have conservation equations of miRNA and mRNA:

$$\sum_{i=1}^{N} C_N^i R_i + [mRNA] = [mRNA]_t,$$

$$\sum_{i=1}^{N} i C_N^i R_i + [miRNA] = [miRNA]_t,$$

The equations of total abundance of miRNA, $\mathrm{mRNA_t}$ and protein are

$$\frac{d[miRNA]_t}{dt} = \mathrm{k_{miR}} - \mathrm{kd_{miR}} * [\mathrm{miRNA} - \sum_{i=1}^{N} \mathrm{k_{dRi}} * (1-\lambda_i) * i * R_i,$$

$$\frac{d[mRNA]_t}{dt} = \mathrm{k_{mR}} - \mathrm{kd_{mR}} * [\mathrm{mRNA}] - \sum_{i=1}^{N} \mathrm{k_{dRi}} * R_i,$$

$$\frac{d[Protein]}{dt} = ks_0 * [mRNA] + \sum_{i=1}^{N} ks_i * R_i - kd_{Protein} * [Protein],$$

where $[\mathrm{miRNA}]_t$, $[\mathrm{mRNA}]_t$ are the total abundance of miRNA and mRNA respectively, and λ_i is the recycle ratio for the miRNA following the degradation of R_i.

Text S2. The revised CBS model for EMT system

Tthe core regulatory network for TGF-β-induced EMT that we considered (Fig. 1B) which contains a few interactions not included in the original CBS model. The corresponding mathematical model is similar to the previous model except with some modification(*22*).

SNAIL1/miR-34 module

Exogenous TGF-β promotes expression of *SNAIL1* (*67*), which is translated into SNAIL1 protein. Experimental studies reveal double-negative feedback loops between SNAIL1 and miR-34 (*17, 18*), in which SNAIL1 inhibits the transcription of *miR-34*, and miR-34 represses the translation of *SNAIL1*. Two conserved SNAIL1-binding sites are found in *miR-34* promoter (*17*) and one conserved miR-34 binding site is found in the sequence of *SNAIL1* (*17*). Different from our previous model (*68*), we included the negative autoregulation of SNAIL1 here. We expect that the impact of the auto-regulation of SNAIL1 is modest because there is only one E-box in the *SNAIL1* promoter 600 basepairs (bp) from the transcription start site (*68*).

$$\frac{\mathrm{d[snail1]}_t}{dt} =$$

$$\mathrm{k0_{snail}} + \mathrm{k_{snail}} * \frac{([\mathrm{TGF}]\mathrm{t}/\mathrm{J_{snail0}})^2}{1+([\mathrm{TGF}]\mathrm{t}/\mathrm{J_{snail0}})^2} * \frac{1}{1+[\mathrm{SNAIL1}]/\mathrm{J_{snail1}}} - \mathrm{kd_{snail}} * [\mathrm{snail1}] - \mathrm{kd_{SR}} * [\mathrm{SR}],$$

$$\frac{d[\mathrm{miR34}]_t}{dt} = \mathrm{k0_{34}} + \frac{\mathrm{k_{34}}}{1 + ([\mathrm{SNAIL1}]/\mathrm{J1_{34}})^2 + ([\mathrm{ZEB}]/\mathrm{J2_{34}})^2} - \mathrm{kd_{34}} * [\mathrm{miR34}] - (1-\lambda\mathrm{s}) * \mathrm{kd_{SR1}} * [\mathrm{SR}],$$

$$\frac{d[\mathrm{SNAIL1}]}{dt} = \mathrm{k_{SNAIL}} * [\mathrm{snail1}] - \mathrm{kd_{SNAIL}} * [\mathrm{SNAIL1}],$$

$$[\mathrm{miR34}] = [\mathrm{miR34}]_t - [\mathrm{SR}],$$

$$[\mathrm{snail1}] = [\mathrm{snail1}]_t - [\mathrm{SR}],$$

$$[\mathrm{SR}] = \mathrm{K_s} * [\mathrm{snail1}] * [\mathrm{miR34}],$$

$$[\mathrm{TGF}]_\mathrm{t} = [\mathrm{TGF}] + [\mathrm{TGF0}],$$

where [TGF0] is the abundance of the exogenous TGF-β as the input of the system and [TGF] is the abundance of the autocrine TGF, which is inhibited by miR-200 (see the TGF-β/miR-200 module section). [SR] indicates the abundance of *SNAIL1*/miR-34 complex.

ZEB1/miR-200 module

SNAIL1 transmits the signal to the ZEB1/miR-200 module by promoting *ZEB1* transcription (*34, 69*), and by repressing *miR-200* transcription (*70*). A double-negative feedback loop is also found between ZEB1 and miR-200 (*19-21*), in which ZEB1 represses the transcription of *miR-200* and miR-200 inhibits the translation of *ZEB1*. Additionally, ZEB1 inhibits *miR-34* expression by binding to two sites in its promoter region (*17*) (*71*). Two ZEB1-binding sites are found in three of the *miR-200* family (*miR-200a*, *miR-200c*, and *miR-429*) and three binding sites are found in two other members of the *miR-200* family (*miR-200*c and *miR-141*) (*41*). Five conserved binding sites for miR-200b, miR-200c, and miR-429 and three for miR-141 and miR-200a are found in the sequence of *ZEB1* (*41*).

$$\frac{\mathrm{d[zeb]_t}}{dt} = \mathrm{k0_{zeb}} + \mathrm{k_{zeb}} * \frac{([\mathrm{SNAIL1}]/\mathrm{J_{zeb}})^2}{1 + ([\mathrm{SNAIL1}]/\mathrm{J_{zeb}})^2} - \mathrm{kd_{zeb}} * [\mathrm{zeb}] - \sum_{i=1}^{5} \mathrm{kd_{ZR_i}} * \mathrm{C_5^i} * [\mathrm{ZR}_i],$$

$$\frac{\mathrm{d[ZEB]}}{dt} = \mathrm{k_{ZEB}} * [\mathrm{zeb}] - \mathrm{kd_{ZEB}} * [\mathrm{ZEB}],$$

$$\frac{\mathrm{d[miR200]_t}}{dt} = \mathrm{k0_{200}} + \mathrm{k_{200}} \frac{1}{1 + ([\mathrm{SNAIL}]/\mathrm{J1_{200}})^3 + ([\mathrm{ZEB}]/\mathrm{J2_{200}})^2} - \mathrm{kd_{200}} * [\mathrm{miR200}] - \sum_{i=1}^{5}(1 - \lambda_i) * \mathrm{kd_{ZR_i}} * \mathrm{C_5^i} * \mathrm{i} * [\mathrm{ZR}_i] - (1 - \lambda_{\mathrm{TR}}) * \mathrm{kd_{TR}} * [\mathrm{TR}],$$

$$[\mathrm{ZR}_i] = \mathrm{K}_i * [\mathrm{miR200}] * [\mathrm{ZR}_{i-1}] \ (i = 1 \ldots 5),$$

$$[\mathrm{zeb}] = [\mathrm{zeb}]_t - \sum_{i=1}^{5} \mathrm{C_5^i} * [ZR_i],$$

$$[\mathrm{miR200}] = [\mathrm{miR200}]_\mathrm{t} - \sum_{i=1}^{5} \mathrm{i} * \mathrm{C_5^i} * [ZR_i],$$

where [$\mathrm{ZR_i}$] is the abundance of *ZEB1*/miR-200 complex with *i* copies of miR200 bound on the sequence of *ZEB1*.

TGF-β/miR-200 module

Furthermore, miR-200 may also inhibit the autocrine expression of transcripts encoding TGF-β because one conserved miR-200 binding site is present on the TGF-β2-encoding transcript (*13, 21*). This constitutes another feedback loop in the core regulatory network.

$$\frac{\mathrm{d[tgf]}_t}{dt} = \mathrm{k_{tgf}} - \mathrm{kd_{tgf}} * [\mathrm{tgf}] - \mathrm{kd_{TR}} * [\mathrm{TR}],$$

$$\frac{\mathrm{d[TGF]}}{dt} = \mathrm{k_{TGF}} * [\mathrm{tgf}] - \mathrm{kd_{TGF}} * [\mathrm{TGF}],$$

$$[\mathrm{tgf}] = [\mathrm{tgf}]_t - [\mathrm{TR}],$$

$$[\mathrm{TR}] = \mathrm{K_{TGF}} * [\mathrm{miR200}] * [\mathrm{tgf}],$$

where [TR] is the abundance of tgf/*miR-200* complex.

Output of the system

The gene encoding E-cadherin, the marker of epithelial cells, is inhibited by SNAIL1 and ZEB1, whereas genes encoding N-cadherin and vimentin, markers of mesenchymal cells, are stimulated by SNAIL1 and ZEB1 (*1*).

$$\frac{d[\text{E\_marker}]}{dt} = \mathrm{k_{e0}} + \mathrm{k_{e1}}\frac{1}{([\text{SNAIL1}]/\mathrm{J_{e1}})^2 + 1} + \mathrm{k_{e2}}\frac{1}{([\text{ZEB}]/\mathrm{J_{e2}})^2 + 1} - \mathrm{kd_e} * [\mathrm{E_{marker}}],$$

$$\frac{d[\text{M\_marker}]}{dt} = \mathrm{k_{m0}} + \mathrm{k_{m1}}\frac{([\text{SNAIL1}]/\mathrm{J_{m1}})^2}{([\text{SNAIL1}]/\mathrm{J_{m1}})^2 + 1} + \mathrm{k_{m2}}\frac{(\text{ZEB}/\mathrm{J_{m2}})^2}{(\text{ZEB}/\mathrm{J_{m2}})^2 + 1} - \mathrm{kd_m} * [\mathrm{M_{marker}}],$$

In all of the equations, a Hill function is used for the transcription factor-dependent activation or inhibition, with the Hill coefficient as the binding site number, consistent with the treatments in both the CBS and TCS models (*22, 23*).

Text S3. Qualitatively, but not quantitatively, different predictions can distinguish the CBS and TBS models for EMT

Lu *et al.* (*23*) propose that the CBS and TCS models can be distinguished by measuring the abundance of miR-200 and ZEB1 of cells in the P state. If we refer to the expression patterns as low, medium, and high with values of 0, ½, 1, then the E state (high miR-200, low ZEB) and the M state (low miR-200, high ZEB1) would be represented as (1, 0) and (0, 1). Lu *et al*. argued that the expression pattern of the P state predicted by the CBS model is (1, 0), and is (½, ½) when by the TCS model. Our measurements with MCF10A cells showed that the P state has the pattern (1, 0) as does the E state, which is consistent with the CBS model.

We recommend caution when distinguishing competing models on the basis of quantitative differences in the model predictions. Even if a (½, ½) pattern were observed, one could not rule out the CBS model. The reason is that, although the CBS model predicts that miR-200 (ZEB1) remains in the high (low) branch of the TGF-β- miR-200 (ZEB1) bifurcation diagram as in Fig. 1C during the transition from E to P, the abundance does not necessarily remain unchanged. With different choices of parameters, the CBS model can still predict a decrease (increase) of miR-200 (ZEB1) abundance from the E to the P state. Given that there are typically large uncertainties and heterogeneities in cell biology measurements, practically it is difficult to differentiate some subtle quantitatively different predictions. To avoid this issue, we evaluated the two models on the basis of qualitatively different predictions.

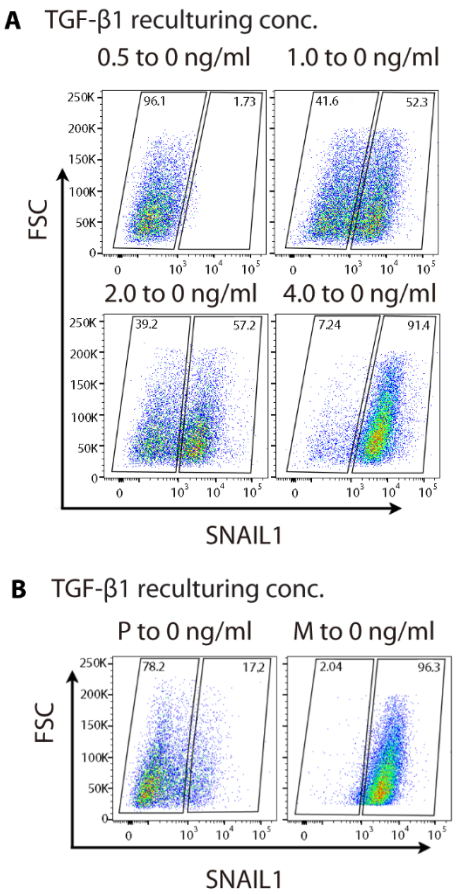

Fig. S1. Flow cytometric analysis of TGF-β1–induced cells recultured in the absence of TGF-β1 confirms that the SNAIL1/miR-34 module functions as a bistable switch. **(A)** After 7 days with the indicated concentrations of hTGF-β1, cells were recultured in the absence of hTGF-β1 for 10 days before flow cytometric analysis of the SNAIL1 abundance. (**B**) Cells cultured in the presence of hTGF-β1 to induce EMT were sorted and cells in P and M states were isolated and recultured without hTGF-β1 for 10 days before flow cytometric analysis of the SNAIL1 abundance.

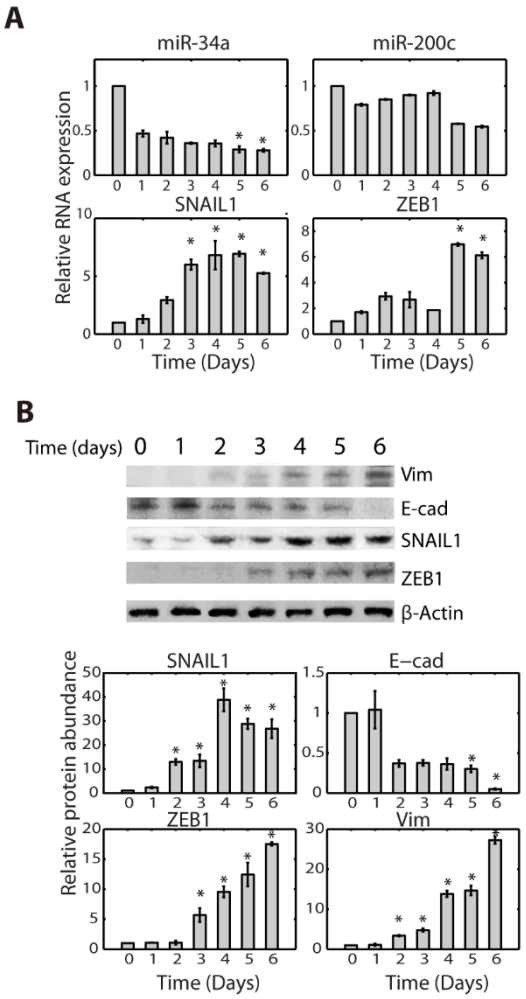

Fig. S2. Temporal dynamics of TGF-β1-induced EMT in MCF10A cells. Cells were exposed to 4 ng/ml TGF-β1 for 0-6 days. (**A**) The abundance of RNAs was measured by qRT-PCR. (**B**) The abundance of proteins was measured by Western blot. Plots were obtained from the average concentrations of targeted proteins and normalized with the average concentrations of actin. Data in all graphs represent the mean ± SD, n = 3. Asterisks indicate $P < 0.05$ (Mann-Whitney test) and at least 3-fold changes from the control group.

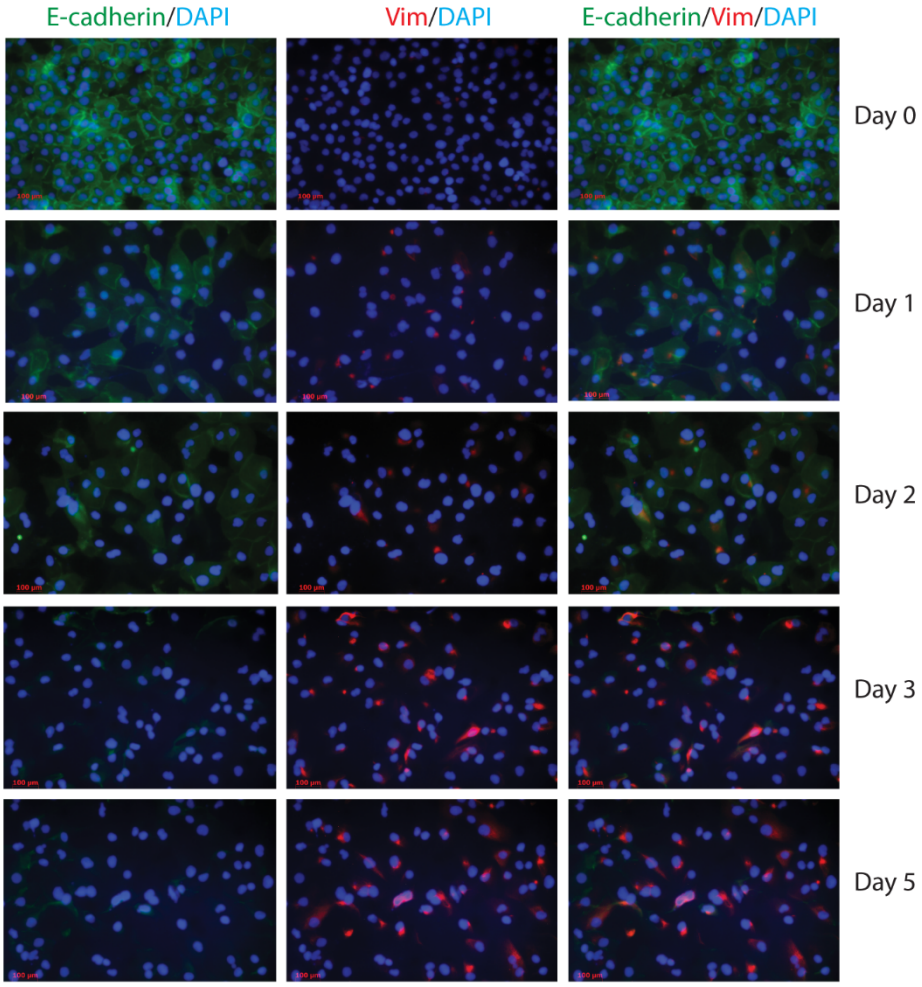

Fig. S3. Immunofluorescence of E-cadherin and vimentin at different time points in response to 4 ng/ml TGF-β1. Nuclei were stained with DAPI (blue). E-cadherin (E-cad) is green and vimentin (Vim) is red.

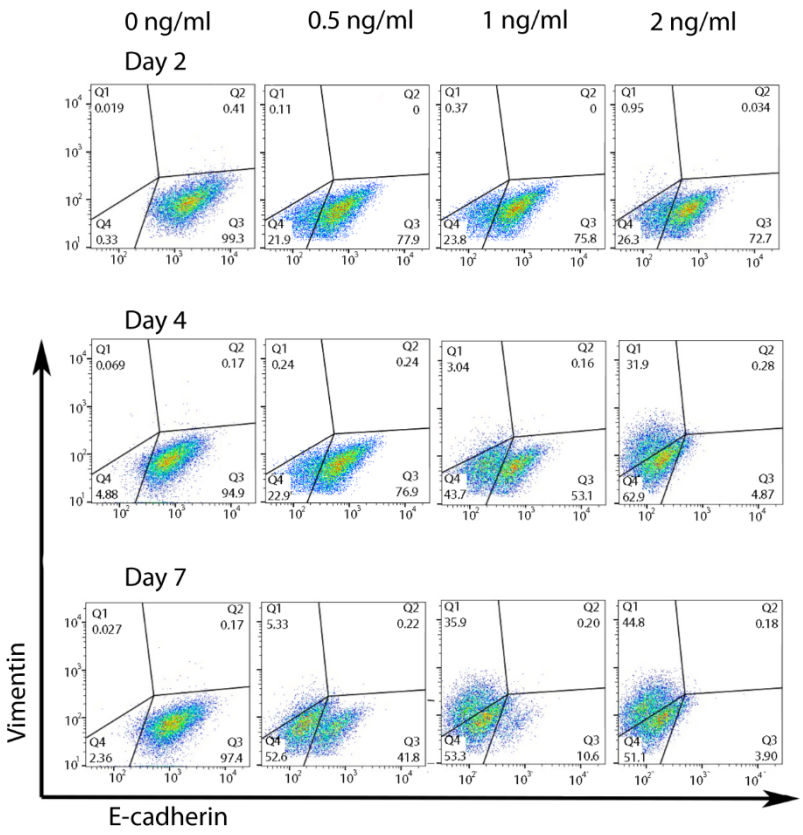

Fig. S4. Analyses of the three subgroups at the indicated time points by flow cytometry in response to different concentrations of TGF-β1.

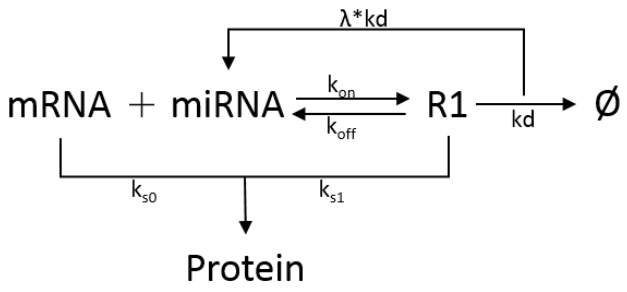

Fig.S5. Generic model for the regulation of mRNA by miRNA. ∅ represents degradation.

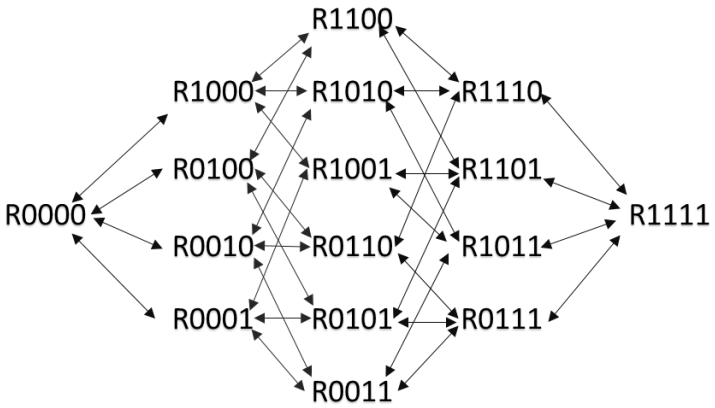

Fig. S6. Full reaction diagram of miRNA regulation of mRNA with four binding sites when considering all the possible miRNA-mRNA complexes. There are 15 possible miRNA-mRNA complexes and 64 reactions. The miRNA-mRNA complex is represented by Rxxxx with an miRNA bound on the binding site where x = 1. For example, R0101 indicates the miRNA-mRNA complex with miRNA bound on the 2nd and 4th binding sites, R0000 indicates the free mRNA because no miRNA is bound and R1111 is an mRNA with all four sites bound by miRNA.

Table S1. Comparison between the original CBS and TCS models. H, High; M, Medium; L, Low; -, Not specified.

	Original CBS Model	**TCS Model**
The functional role of SNAIL1/miR-34 and ZEB1/miR-200 modules		
SNAIL1/miR-34	Bistable switch controlling E-P transition	Monostable, noise-buffering Integrator
ZEB1/miR-200	Bistable switch controlling P-M transition	Ternary switch for E to P then to M
Other regulations		
SNAIL1 self-inhibition	Not included because it was assumed not essential for the key mechanism	Included as the mechanism of noise-buffering
ZEB1 self-activation	Not included because there is no direct evidence	Included as the requirement of ternary switch
TGF-β autocrine loop	Included as the mechanism of the irreversibility of EMT	Not included because irreversibility is not discussed
The regulator abundance in each state		
E	SNAIL1, miR-34, ZEB1, miR-200 = [L H L H]	[L H L H]
P	[H L L H]	[- - M M]
M	[H L H L]	[- - H L]

Table S2. Primers used for quantitative RT-PCR.

Targeted genes	**Primers**
SNAIL1 Forward	GGCCCACCTCCAGACCCACT
SNAIL1 Reverse	GCGGGGACATCCTGAGCAGC
ZEB1 Forward	AGTGGTCATGATGAAAATGGAAC
ZEB1 Reverse	AGGTGTAACTGCACAGGGAGC
18S ribosomal RNA Forward	GTAACCCGTTGAACCCCATT
18S ribosomal RNA Reverse	CCATCCAATCGGTAGTAGCG
miR-34a Forward	CACGCATGGCAGTGTCTTAGC
miR-34a loop	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACACAACC
miR-200c Forward	CACGCATAATACTGCCGGGTAATGAT
miR-200c loop	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTCCATCT CCATC

Table S3. Parameters used in the present model studies. Most of the values of parameters are from our previous model (*1*) and other parameters are fitted with the experimental data (Fig. 4B). The fitted parameters are in shade.

Parameter	Value	Parameter	Value
SNAIL1/miR-34 module			
$k0_{snail}$	0.0006 μM/hr	$k0_{34}$	0.0012 μM/hr
k_{snail}	0.05 μM/hr	k_{34}	0.012 μM/hr
J_{snail0}	0 .62 μM	$J1_{34}$	0.15 μM
J_{snail1}	0.67 μM	$J2_{34}$	0.36 μM
kd_{snail}	0.09 /hr	kd_{34}	0.035 /hr
kd_{SR}	0.9 /hr	K_s	100 /μM
k_{SNAIL}	17 μM/hr	λs	0.5
kd_{SNAIL}	1.66 /hr		
ZEB1/miR-200 module			
$k0_{zeb}$	0 .003 μM/hr	$k0_{200}$	0.0002 μM/hr
k_{zeb}	0.06 μM/hr	k_{200}	0.02 μM/hr
J_{zeb}	3.5 μM	$J1_{200}$	3.25 μM
kd_{zeb}	0.09 /hr	$J2_{200}$	0.2 μM
kd_{ZR_i}	0.9 /hr	kd_{200}	0.035 /hr
k_{ZEB}	17 μM/hr	K_i	10 /μM
kd_{ZEB}	1.66 /hr	λ_i	0.5
TGF-β module			
k_{tgf}	0.05 μM/hr	k_{TGF}	1.6 μM/hr
kd_{tgf}	0.09 /hr	kd_{TGF}	1 /hr
kd_{TR}	0.9 /hr	K_{TGF}	20 /μM
λ_{TR}	0.8		
Markers module			
k_{e0}	0.01 μM/hr	k_{m0}	0.01 μM/hr
k_{e1}	0.15 μM/hr	k_{m1}	0.1 μM/hr
k_{e2}	0.05 μM/hr	k_{m2}	0.06 μM/hr
J_{e1}	0.2 μM	J_{m1}	0.2 μM
J_{e2}	0.5 μM	J_{m2}	0.5 μM
kd_e	0.05 /hr	kd_m	0.05 /hr

Table S4. Initial conditions of the variables in the model.

Variable	Description	Initial condition(μ M)
$[\mathrm{snail1}]_t$	Total concentration of *SNAIL1* transcripts	0
[SNAIL1]	Concentration of SNAIL1 protein	0
$[\mathrm{miR34}]_t$	Total concentration of miR-34	0.37
[SR]	Concentration of SNAIL1/miR-34 complex.	0
$[\mathrm{zeb}]_\mathrm{t}$	Total concentration of *ZEB1* transcripts	0
[ZEB]	Concentration of ZEB1 protein	0
$[\mathrm{miR200}]_\mathrm{t}$	Total concentration of miR-200	0.23
$[\mathrm{ZR}_i]$	Concentration of ZEB1/miR-200 complex with i copies of miR200 bound	0
$[\mathrm{tgf}]_t$	Total concentration of *TGF-β* transcripts	0.07
[TGF]	Concentration of endogenous TGF-β	0.03
[TR]	Concentration of TGF/miR-200 complex	0.05
[E_marker]	Concentration of the marker of epithelial cells	4.2
[M_marker]	Concentration of the markers of mesenchymal cells	0.2

Model S1. Computer code of the theoretical model. This is an ".ode" script file that produces the modeling results (Fig.4B and Fig.5CD) in this work. It can be used with XPP/XPPAUT or Oscill8.