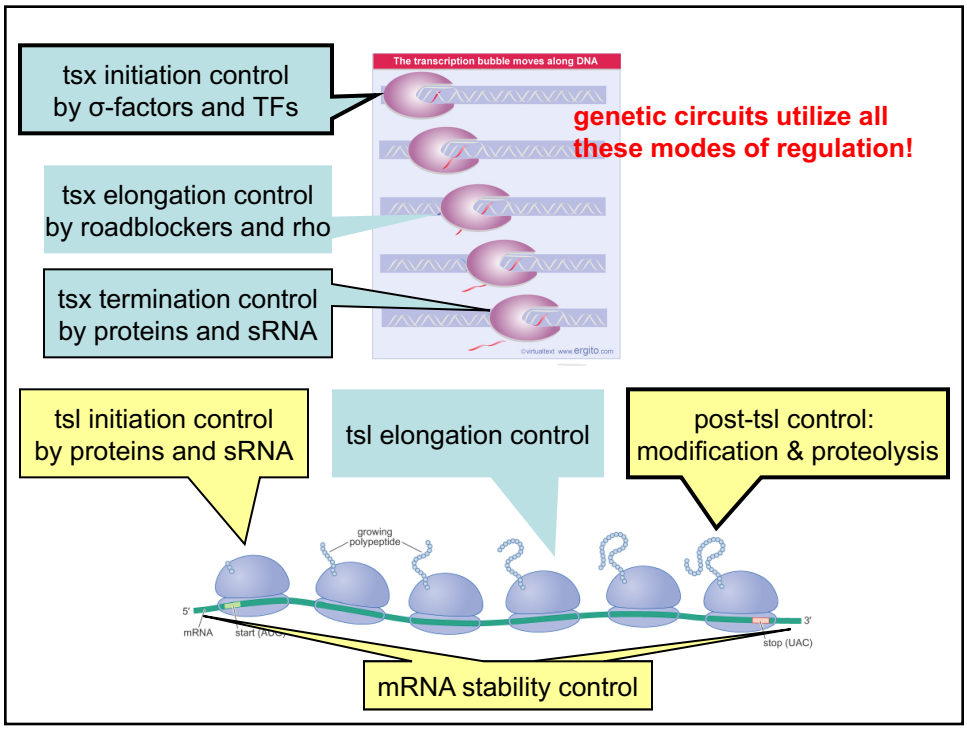


1

Topic 3: Post-transcriptional control

A. Transcriptional elongation and termination
   1. Basic models of tsx elongation and termination
   2. mechanisms of elongation (intrinsic vs rho-dependent)

B. Control of termination (=anti-termination or AT)
   1. AT at a single termination site (various mechanisms)
   2. processive AT (Q, N, Nus)

C. Translational mechanisms (initiation, elongation, termination)

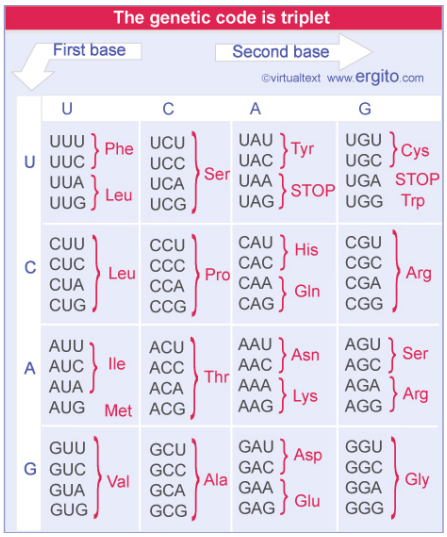
D. Translational control
   1. RNA-binding protein
   2. riboswitch
   3. small regulatory RNA

E. Protein degradation and post-translational control
   1. proteolytic machinery
   2. protein unfolding
   3. substrate selection
   4. effect on gene regulation

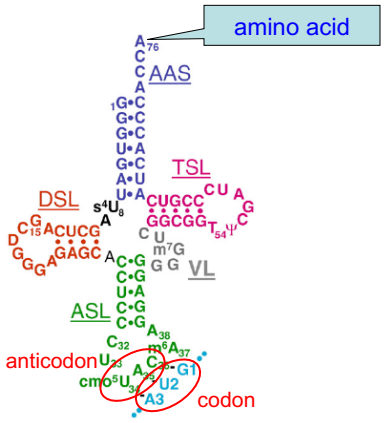
2

C. Translation

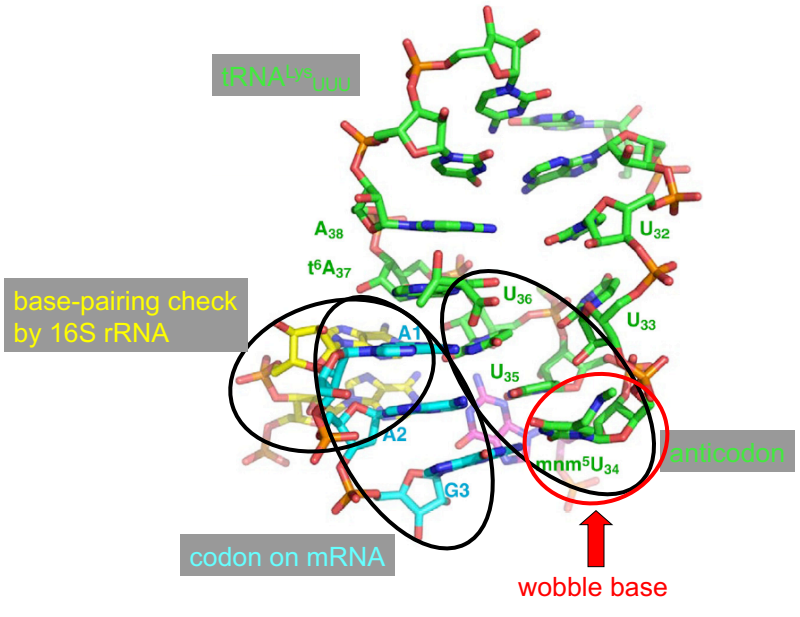
1. tRNA and the Genetic code



secondary structure of $tRNA^{Val}_{UAC}$

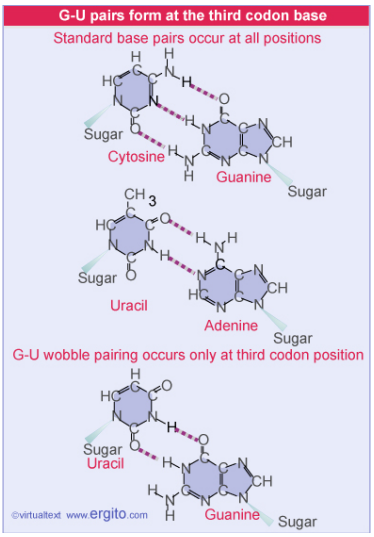


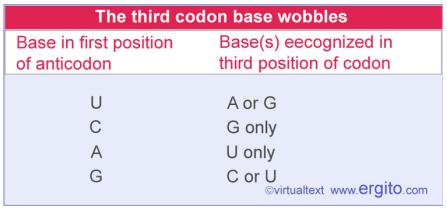
3

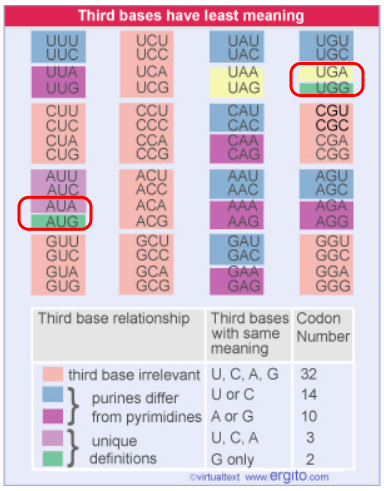


4

wobble pairing at 3rd codon position

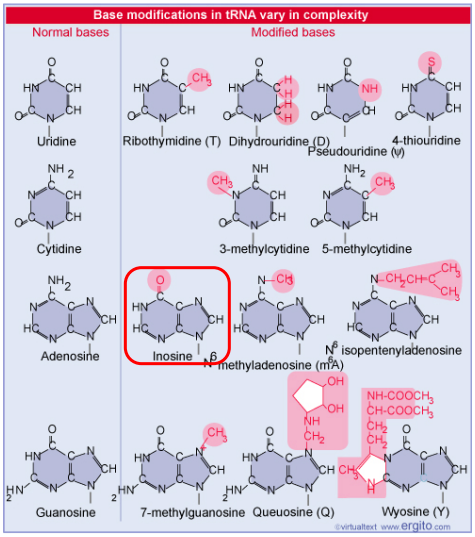


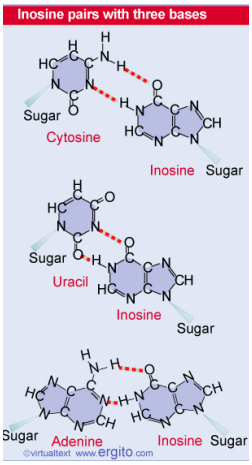




5

base modification & alternative pairing

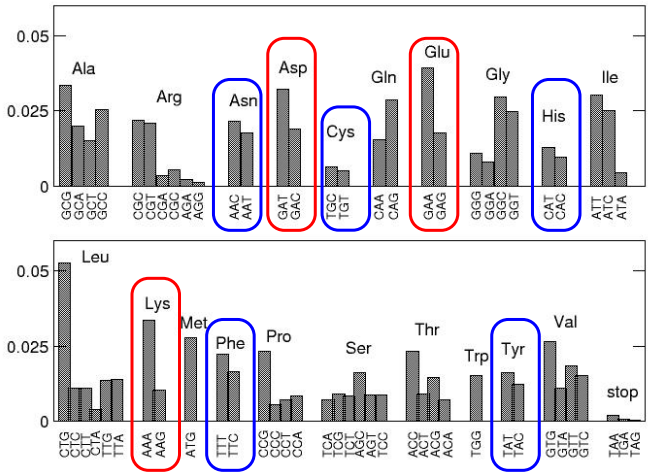




Wobble Hypothesis[1] nucleoside		Modified Wobble Hypothesis[2a] nucleoside	
Anticodon N_{34}	Codon N3	Anticodon N_{34}	Codon N3
G	C,U	G	C,U
C	G	C	G
I	U,C,A	I	U,C,A
U	A,G	xm^5U^a s^2U^b xo^5U^c	G A,G A,G,U (C)

6

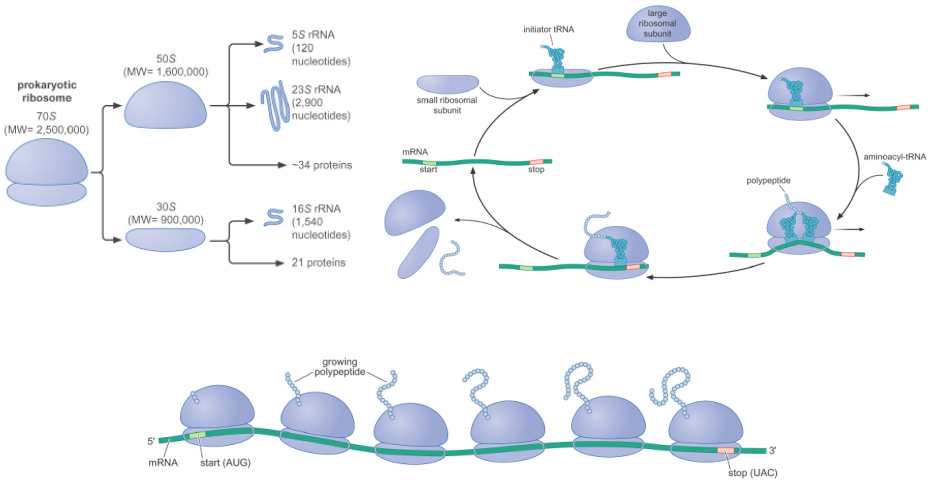
Codon frequency in the E. coli genome



strong codon bias exists even for those with single tRNA species??
(related to AT/GC bias?)

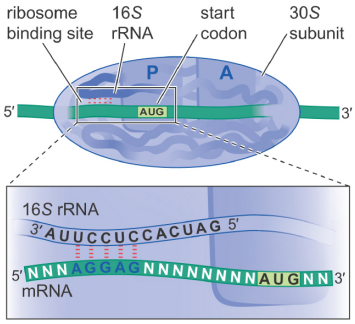
7

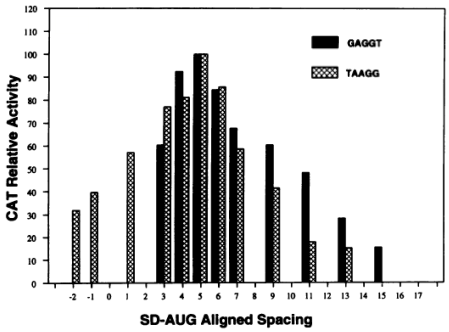
2. Translational mechanisms



8

a. translational initiation





• broad range of allowed spacing

• dependence on distance between RBS and ATG

The complete SD sequence
A: SD-ATG spacing series D_n

5′...**TAAGGAGGT**.....ATG
SD

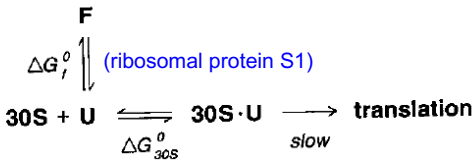
D_3	TCGAT**GAGGT**CATATG
D_4	TCGAT**GAGGT**CCATATG
D_5	TCGAT**GAGGT**CACATATG
D_6	TCGAT**GAGGT**CTTCATATG
D_7	TCGAT**GAGGT**CATTCATATG
D_9	TCGAT**GAGGT**CATTATCATATG
D_{11}	TCGAT**GAGGT**CTTATTATCATATG
D_{13}	TCGAT**GAGGT**CTATTATTATCATATG
D_{15}	TCGAT**GAGGT**CATTATTATTATCATATG

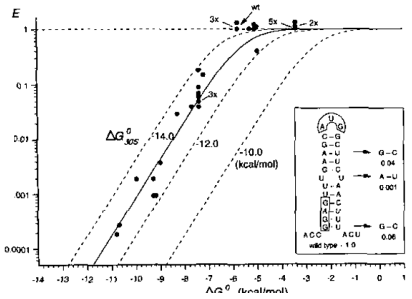
SD

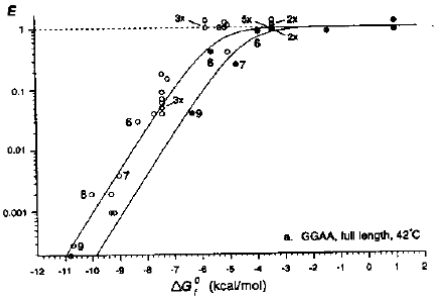
d_2	TCGATT**AAGG**CTATG
d_3	TCGATT**AAGG**CATATG
d_5	TCGATT**AAGG**CACATATG
d_7	TCGATT**AAGG**CATTCATATG
d_8	TCGATT**AAGG**CATTACATATG
d_9	TCGATT**AAGG**CATTATCATATG
d_{10}	TCGATT**AAGG**CTATTATCATATG
d_{11}	TCGATT**AAGG**CTTATTATCATATG
d_{13}	TCGATT**AAGG**CTATTATTATCATATG
d_{15}	TCGATT**AAGG**CATTATTATTATCATATG
d_{17}	TCGATT**AAGG**CATATTATTATTATCATATG

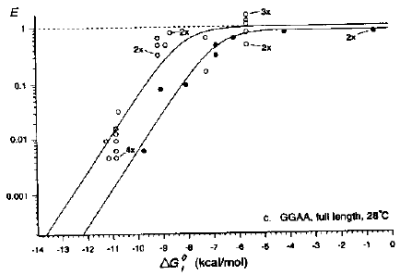
9

• dependence on 5’-UTR sequence







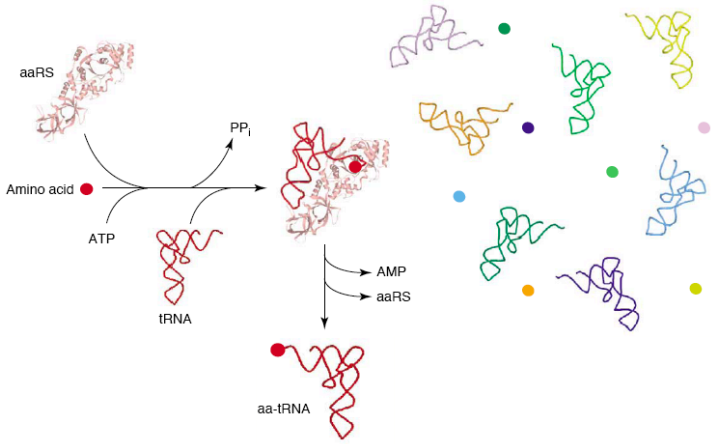


➔ quantitatively accounted for by thermodynamics RNA pairing
➔ translational efficiency is highly tunable; often explored in synthetic biology
➔ but endogenous genes mostly have high translational efficiency (later)

10

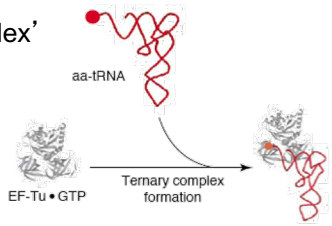
b. translational elongation

- tRNA charging
  - associates the correct a.a. to the tRNA
  - uses a dedicated tRNA synthetase for each a.a. (and all isoacceptors)
  - consumes ATP
  - aa-tRNA recognition not necessarily dependent on anticodon

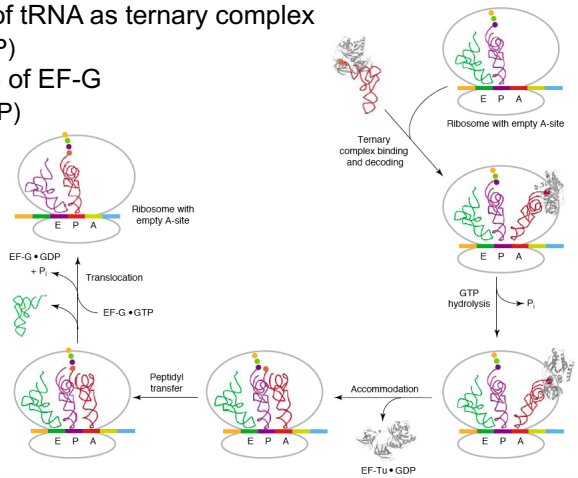


11

- formation of tRNA-aa•EF-TU•GFP 'ternary complex'
  - tRNA-aa unstable otherwise
  - almost all tRNA-aa present in ternary complex
  - large demand for EF-TU (~40kD)
  - most abundant protein in fast growing cells (~5x no. ribosomes; sets the total tRNA amount)



- ribosomal incorporation of tRNA as ternary complex
  note: spends energy (GTP)
- translocation via the help of EF-G
  again spends energy (GTP)
- total energy:
  4ATP/peptide bond



12

TABLE 4 Stoichiometric content of transcription-translation proteins in *E. coli*

Protein	Mol wt (10^3)	α_i^a (t = 40 min) (%)	Molecules (τ = 40 min)	
			Per OD_{460} (10^{12})	Per ribosome
r-Protein	850	13.5	10.2	1.00
L7/L12	12	0.81	40.8	4.00
EF-Tu	42	5.55	55.1	5.40
EF-G	84	1.66	8.2	0.80
EF-Ts	31	0.13	1.8	0.18
IF1	8	0.04	2.5	0.25
IF2	115	0.52	3.1	0.30
IF3	20	0.07	2.0	0.20
Leu S	100	0.12	0.5	0.05
Phe S-β	94	0.21	1.0	0.10
Lys S	58	0.11	0.8	0.08
Arg S	58	0.08	0.6	0.06
Gly S	77	0.17	0.9	0.09
Val S	106	0.14	0.6	0.06
Glu S-β	48	0.10	0.9	0.09
Ile S	107	0.24	1.0	0.10
Phe S-α	36	0.11	1.2	0.12
Gln S	61	0.11	0.8	0.08
Thr S	65	0.09	0.6	0.06
RNA polymerase β	150	0.52	1.4	0.14
RNA polymerase α	39	0.37	3.8	0.37
RNA polymerase, core	375	1.30	1.9	0.19

13

- translational accuracy?
  - translational error rate = 10^{-3} to 10^{-4}
  - but thermo probab of base mismatch much larger

➔kinetic proof reading (Hopfield, Ninio)

spend energy to enhance specificity

[gromadski & rodina, 04]

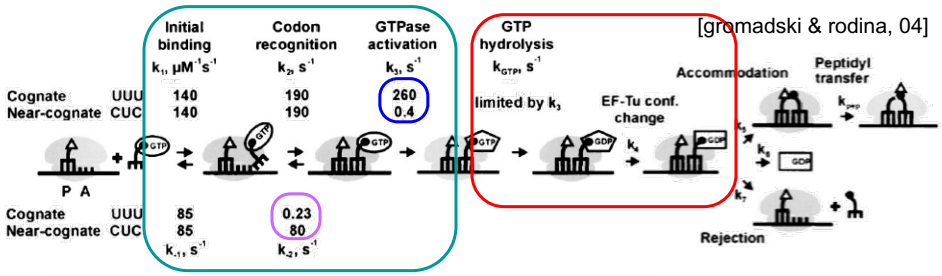


Table 1. Parameters of aa-tRNA Discrimination

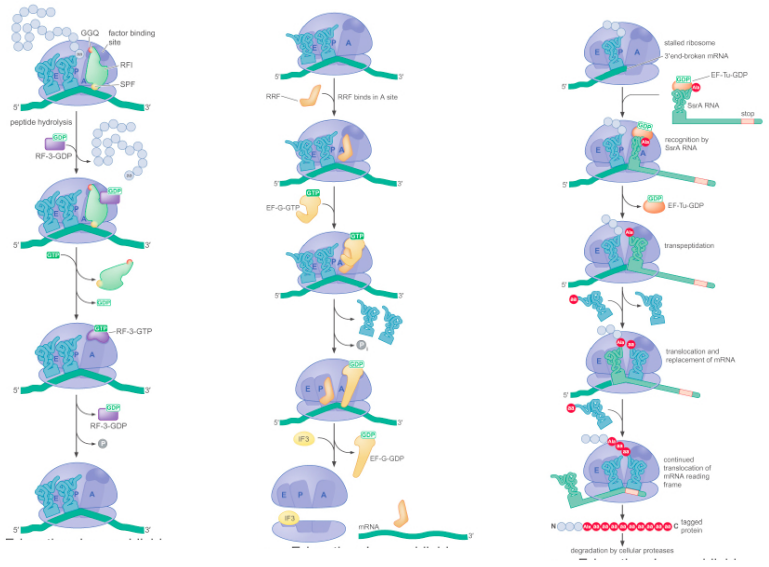
	Initial Selection			Proofreading
Codon	k_{cat}	K_M	k_{cat}/K_M	$k_5/(k_5+k_7)$
UUU	190 ± 20	2.0 ± 0.6	100 ± 20	1.0 ± 0.1
CUC	0.4 ± 0.1	0.25 ± 0.1	1.6 ± 0.5	0.06 ± 0.02
			$F_{initial\ selection}$ = 60 ± 20	$F_{proofreading}$ = 15 ± 5

~1000x discrimination against near-cognate!

14

c. translational termination

release of peptide | recycling of ribosome | rescue of broken mRNA

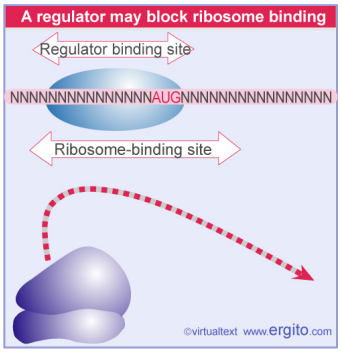


15

D. Translational control

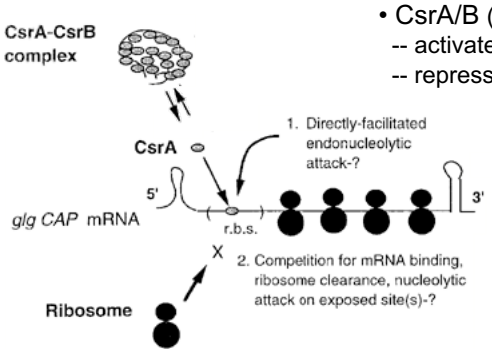
➔ access to ribosomal binding site (RBS) may be blocked by either proteins or RNA

➔ untranslated mRNA rapidly degraded (avoids problems with truncated polypeptides)

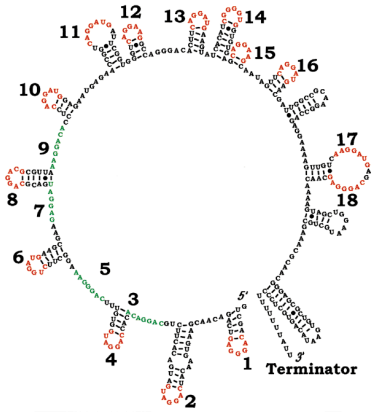


16

1. via RNA-binding proteins



- CsrA/B (global regulation of carbon utilization)
  -- activates glycolysis
  -- represses gluconeogenesis

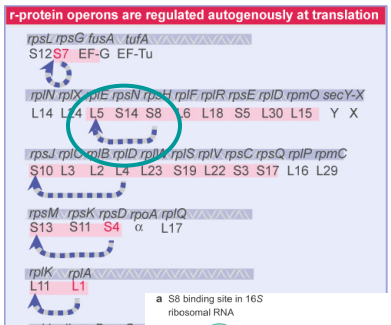


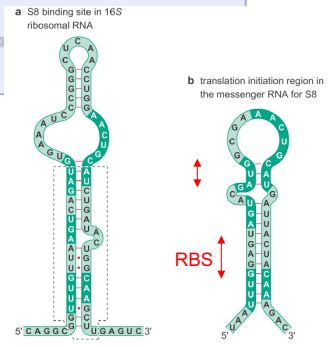
Repeated Elements		
1. CAGGGAG	7. GAGGAUG	13. CAGGAUG
2. CAGGAUG	8. CAGGACG	14. CGGGGUG
3. CAGGACA	9. AAGGACA	15. CAGGAAG
4. CAGGAUG	10. CAGGAUG	16. CAGGAUG
5. CAGGGAA	11. CAGGAUG	17. AAGGAUG
6. CUGGAUG	12. CAGGAAG	18. CAGGGAG

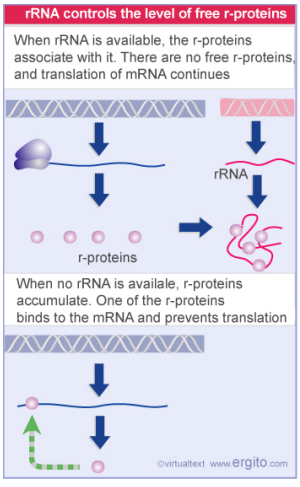
➔ binding of CsrA in 5' -UTR inhibits translation and accelerates mRNA degradation (τ ~ 1min)
➔ CsrA is sequestered by CsrB (regulatory RNA with 18 binding motifs resembling those of target mRNA); csrB tsx regulated by BarA/UvrY
➔ CsrA stimulates csrB tsx via UvrY (-ve feedback)
➔ 2nd antagonizer of CsrA (CsrC recently found)

17

- ribosomal proteins



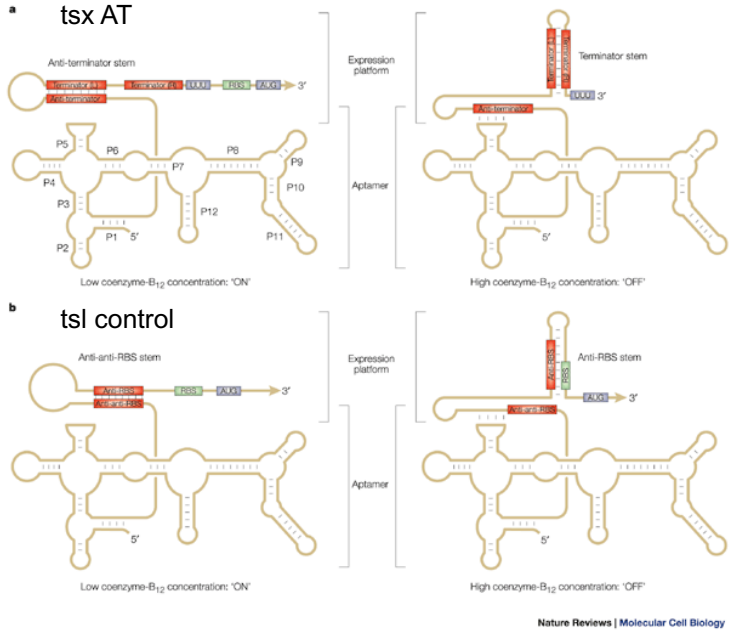




– weak 2nd structure stabilized by ribosomal protein (-ve feedback)
– rRNA itself regulated by (p)ppGpp (stringent response)

18

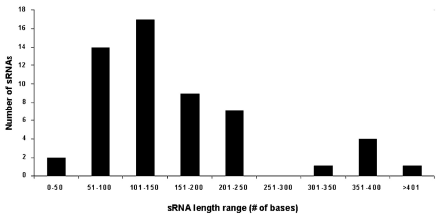
2. via riboswitch (block RBS by alternative 2nd structure)

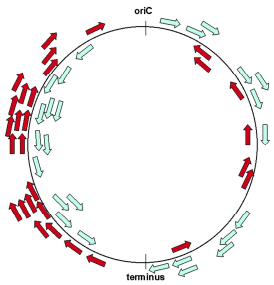


19

3. via sRNA control

length and location of sRNA



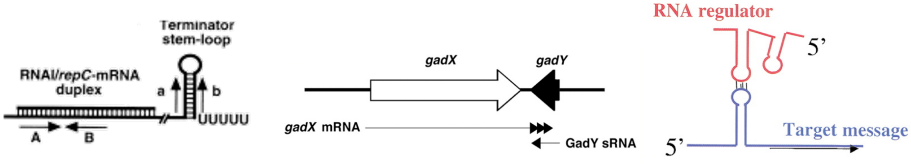


sRNA-mediated control found in

-- transcriptional termination (plasmic copy # control)

-- mRNA stability control

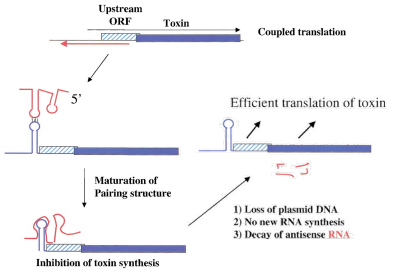
➔ translational inhibition/activation & mRNA degradation



20

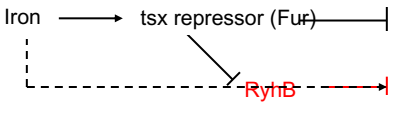
❖ cis-acting

e.g., hok/sok (toxin-antidote system)
-- inhibits translation
-- unique target, spatially localized
-- rapid sRNA decay



❖ trans-acting

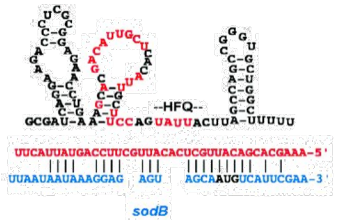
• Regulation of iron metabolism by the repressor Fur



Iron acquisition genes

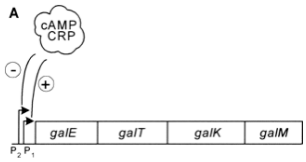
Iron storage genes,
Iron consuming genes,
oxidative stress relief, …

-- RyhB binds to translational initiation regime of mRNA (sodB, sdh, …)



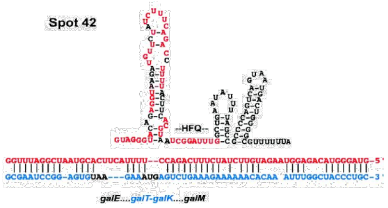
21

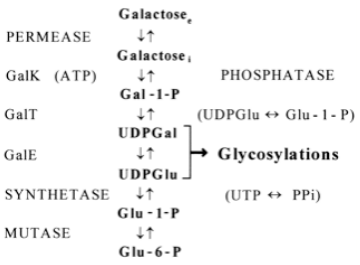
• discoordinate expression of the *gal* operon
-- cAMP high (glucose shortage): operon activated
-- cAMP low (glucose-rich): GalE remain high, GalK:GalE ~ 1/4



• physiology:
-- GalK needed **only** for galactose metabolism
-- GalE needed in galactose metabolism **AND** UDP-galactose synthesis (building block for cell wall and capsule)

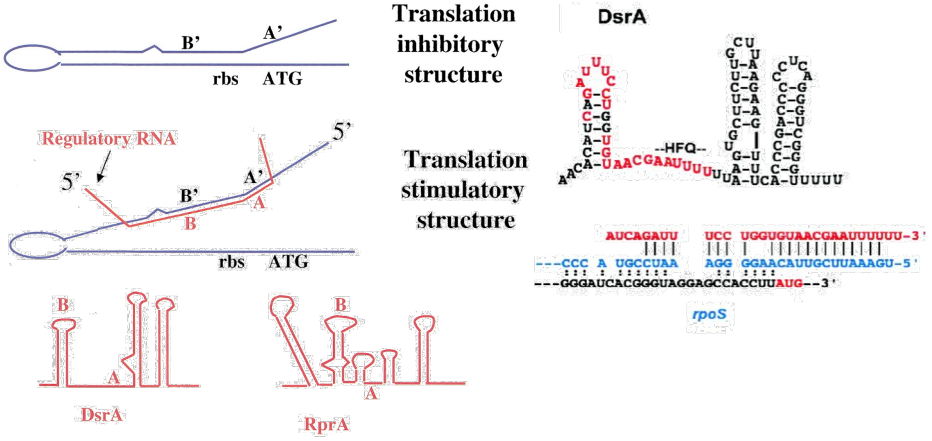
• GalK translationally repressed by Spot42 (repressed by cAMP-CRP)





22

• translational control of rpoS (encodes σ^S)
-- default translation "off" (by 2nd structure of 5' -UTR)

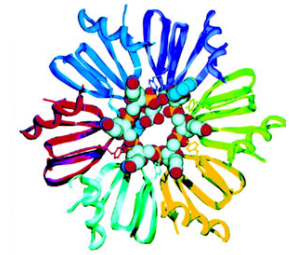


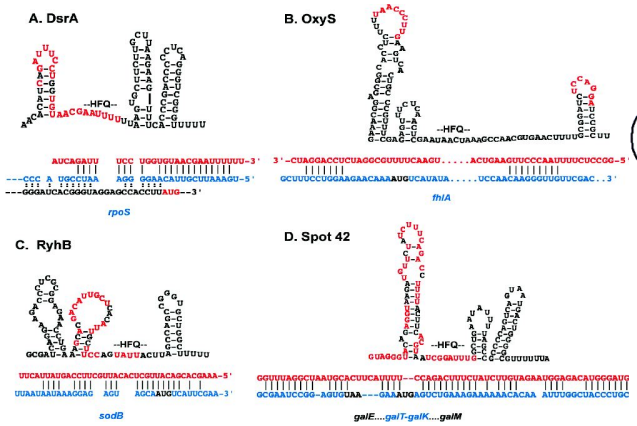
-- activated by sRNAs DsrA (low temperature) and RprA (cell surface stress)
-- negatively regulated by another sRNA OxyS (oxidative stress, ~40 targets)

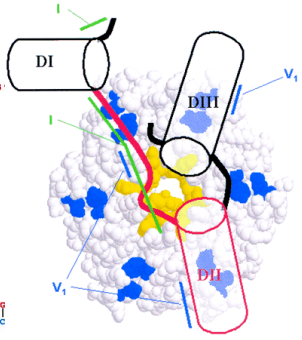
23

❖ Mechanism(s) of sRNA-mediated regulation

• many require Hfq (e.g., DsrA, OxyS, Spot42, RyhB)
-- hexameric protein forming a ring
-- homologous to eukaryotic Sm-like proteins that function in RNA splicing
-- binds A/U-rich single-stranded RNA next to stem-loop region

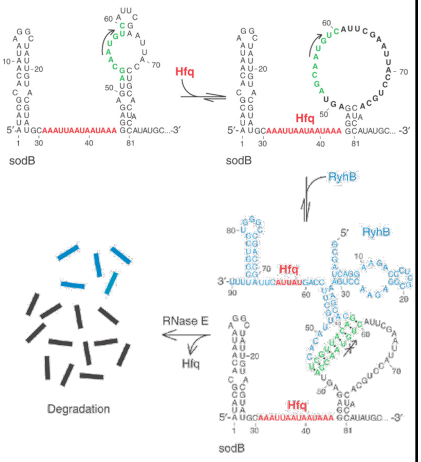






24

- effect of Hfq
  - can stabilize sRNA (half-life > 30 min)
  - can also bind to target mRNA
    (and induce alternative 2nd structure)
  - can stimulate sRNA/target mRNA pairing
    (e.g., acting as RNA chaperone)



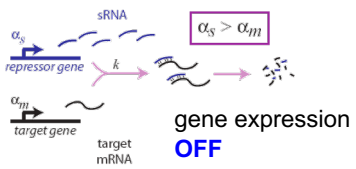
- effect of sRNA/mRNA pairing
  - can change ribosome accessability
  - can lead to rapid mRNA degradation
  - can lead to rapid sRNA degradation
    (stoichiometric rather than catalytic)

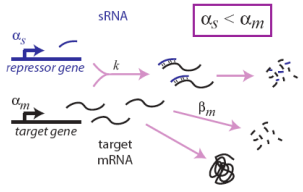
- effect of RNase E
  - required in mRNA and sRNA degradation
  - recognition motif similarly to Hfq
    (dissociation of Hfq upon pairing?)
  - mRNA degradation must be blocked in some cases; how?
    (e.g., Spot42 regulation of *galK* and DsrA stimluation of *rpoS*)

25

sRNA-mediated gene silencing

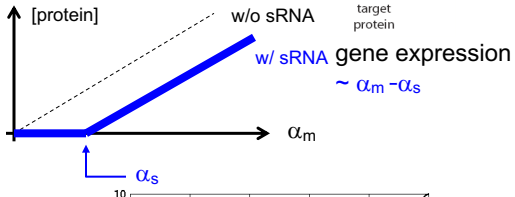
[E. Levine et al, PLoS Biol. 2007]





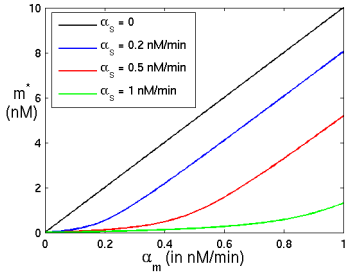
- qualitative expectation:

  threshold-linear response
  - tight repression for $\alpha_m \lesssim \alpha_s$
  - weak repression for $\alpha_m \gg \alpha_s$



- quantitative prediction:

$$\begin{cases} \frac{dm}{dt} = \alpha_m - \beta_m m - k \cdot m \cdot s \\ \frac{ds}{dt} = \alpha_s - \beta_s s - k \cdot m \cdot s \end{cases}$$



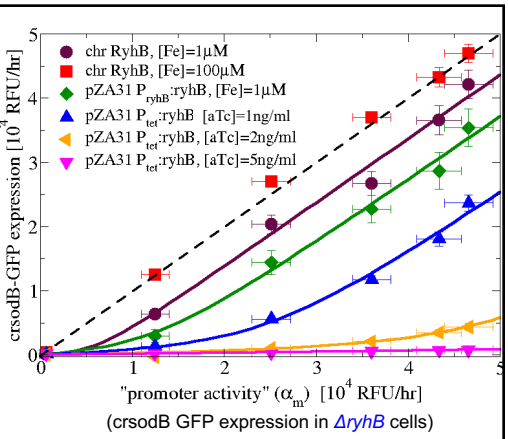
parameters characterized for RyhB/sodB,
with $\beta_m^{-1} \approx 6$ min; $\beta_s^{-1} \approx 30$ min
and $k^{-1} \sim 50$ nM-min (diffusion-limited)

27

Expt' l characterization

plasmid-encoded inducible GFP reporter
translationally fused with truncated *sodB*

various sources of *ryhB*



(crsodB GFP expression in *ΔryhB* cells)

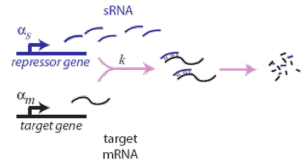
➔single-parameter fit to the expected threshold-linear form

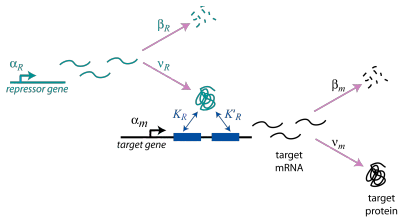
further verification by

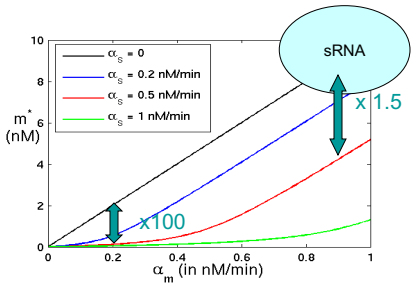
- direct mRNA measurement: similar to GFP
- mutant crsodB: lose nonlinear response
- compensating mutation in RyhB: restores threshold-linear response

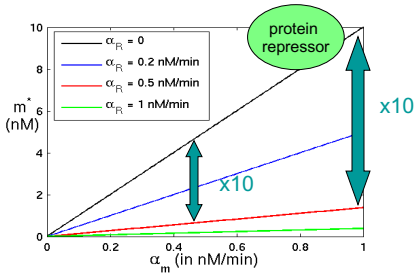
28

Compare to protein regulators







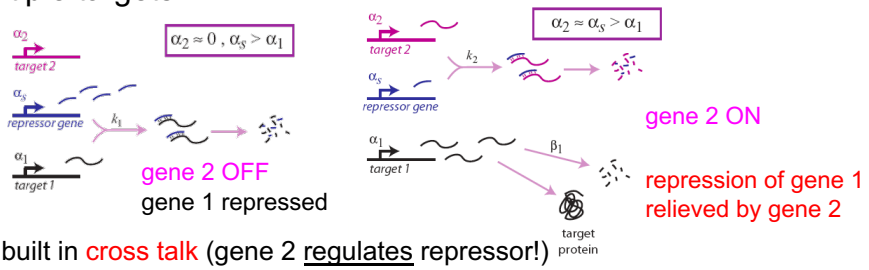


sRNA-mediated tsl regulation
➔ fold-repression depends sensitively on α_m

protein-mediated tsx regulation
➔ fold repression independent of promoter activity (α_m)

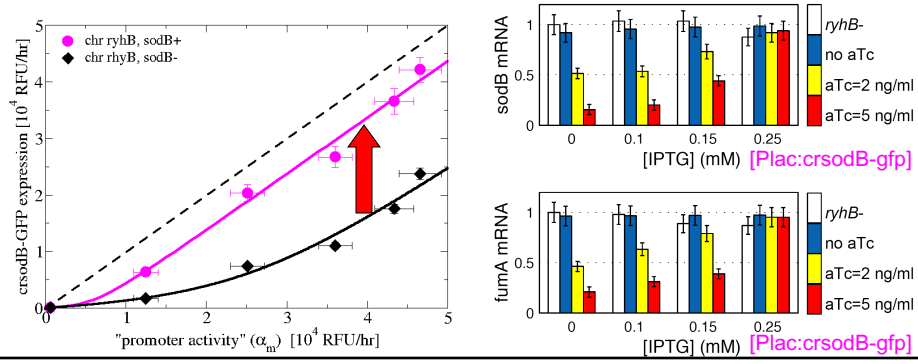
29

multiple targets



➔ built in cross talk (gene 2 regulates repressor!)

➔ nontrivial circuits from apparently simple feedforward motifs



30

Sequence dependence on sRNA-mediated gene silencing

[Hao et al, PNAS 2011]

Approach:

- focus on RyhB-sodB interaction in *E. coli* as a model system
- construct mutants of *ryhB* and/or *sodB* in selected regions
- express the mutants using titratable promoters
- quantify degree of repression on reporters and native targets
- analyze results using mathematical models

Goal: construct biophysical model of gene regulation by sRNA

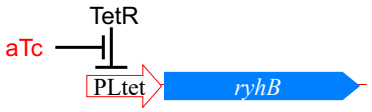
(c.f. von Hippel's work on protein-DNA interaction)

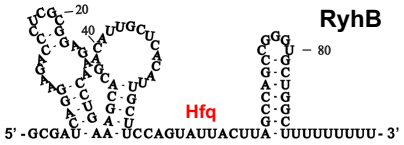
- inform bioinformatic algorithms to improve sRNA/target predictions
- provide framework to characterize sRNA/target interaction
- design of sRNA regulators for synthetic biology applications

31

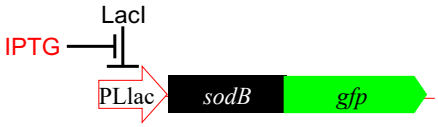
Experimental system

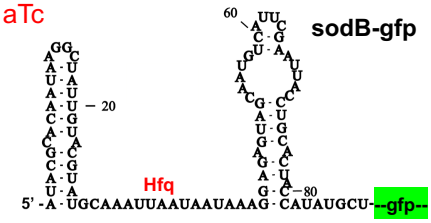
strain: *E. coli* K12 MG1655 with *ΔryhB*, *lacIq*, *tetR*





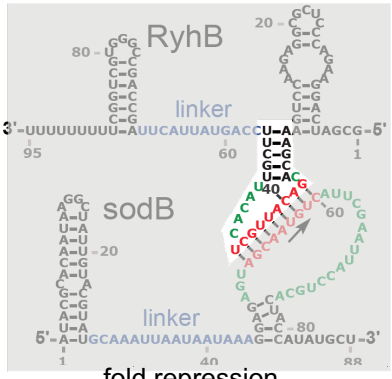
- wild-type or mutant *ryhB* encoded on pZA plasmid (p15A ori, 10~15 copies) or chromosomally integrated
- titratable expresion controlled by aTc



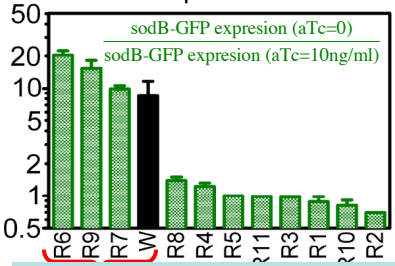


- translational fusion of truncated (wt or mutant) sodB with GFP reporter
- encoded on pZE plasmid (colE1 ori, ~30 copies)
- titratable expresion of *sodB-gfp* controlled by IPTG

32



fold repression

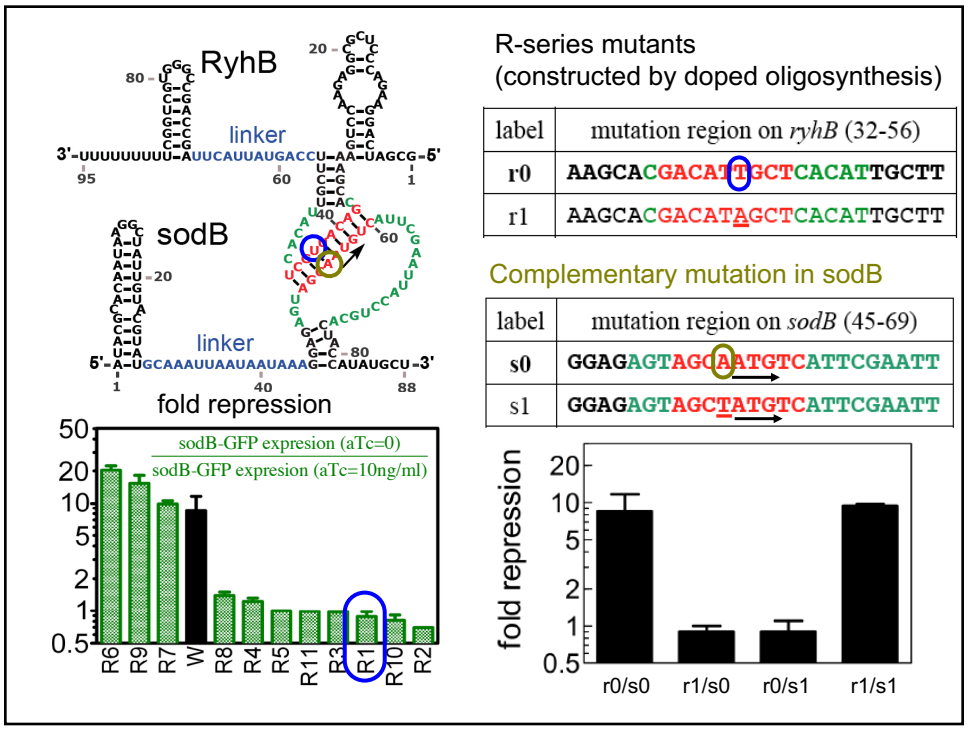


R-series mutants
(constructed by doped oligosynthesis)

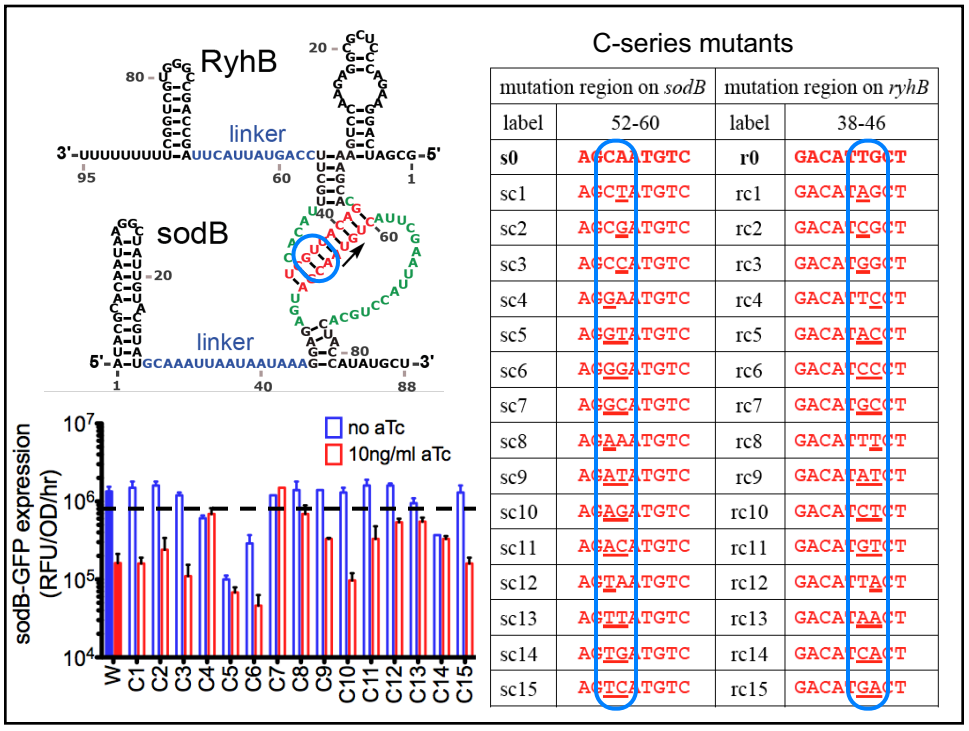
label	mutation region on *ryhB* (32-56)
r0	AAGCACGACATTGCTCACATTGCTT
r1	AAGCACGACATAGCTCACATTGCTT
r2	AAGCACGACATTGCTTCATTGCTT
r3	AAGCACGAATTGCTAACATTGCTT
r4	AAGCAAACACTGCTCATTATTGCTT
r5	AAGCACGACATTGATCACATTGCTT
r6	AAGCACGACATTGCTGACATTGCTT
r7	AAGCACGACATTGCTCACACTGCTT
r8	AAGCACGACGTTGCGCACATTGCTT
r9	AAGCACGACATTGCTCACATTTCTT
r10	AAGCATGCCAATGCTCACATTGCTT
r11	AAGCACTACATGCTAACATTGCTT

➔ Mutation in core interaction region eliminates repression

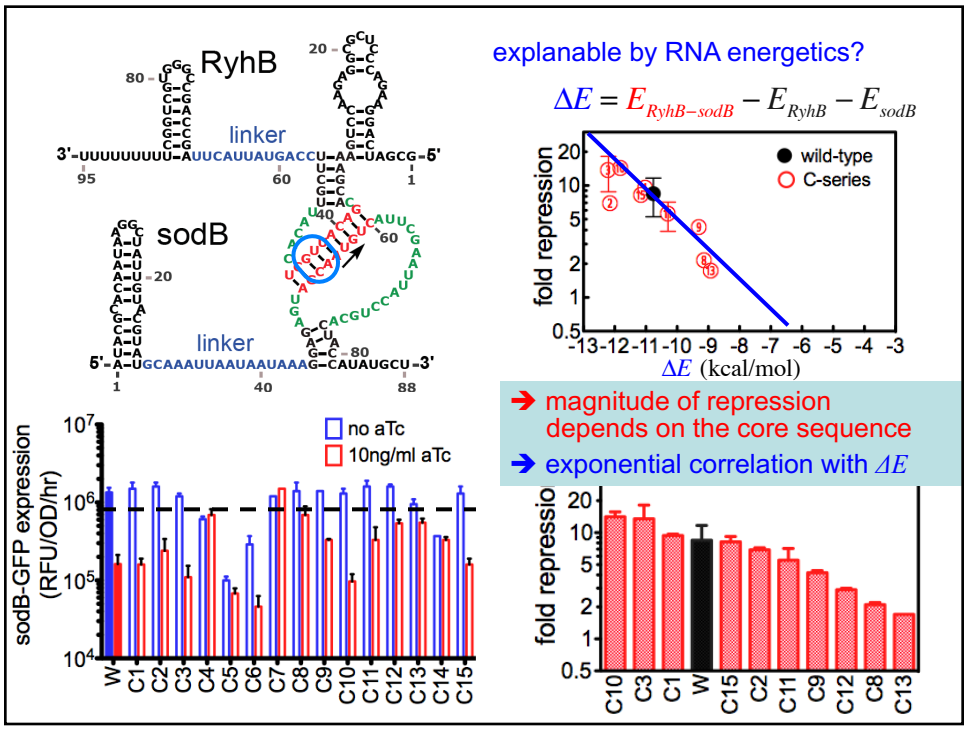
33


RyhB
linker
3'-UUUUUUUUU-AUUCAUUAUGACCU-AA-UAGCG-5'
sodB
linker
5'-A-UGCAAAUUAAUAAUAAAG-CAUAUGCU-3'
fold repression
sodB-GFP expresion (aTc=0)
sodB-GFP expresion (aTc=10ng/ml)
R6 R9 R7 W R8 R4 R5 R11 R3 R1 R10 R2
R-series mutants
(constructed by doped oligosynthesis)
label
mutation region on ryhB (32-56)
r0 AAGCACGACATTGCTCACATTGCTT
r1 AAGCACGACATAGCTCACATTGCTT
Complementary mutation in sodB
label
mutation region on sodB (45-69)
s0 GGAGAGTAGCAATGTCATTCGAATT
s1 GGAGAGTAGCTATGTCATTCGAATT
fold repression
r0/s0 r1/s0 r0/s1 r1/s1

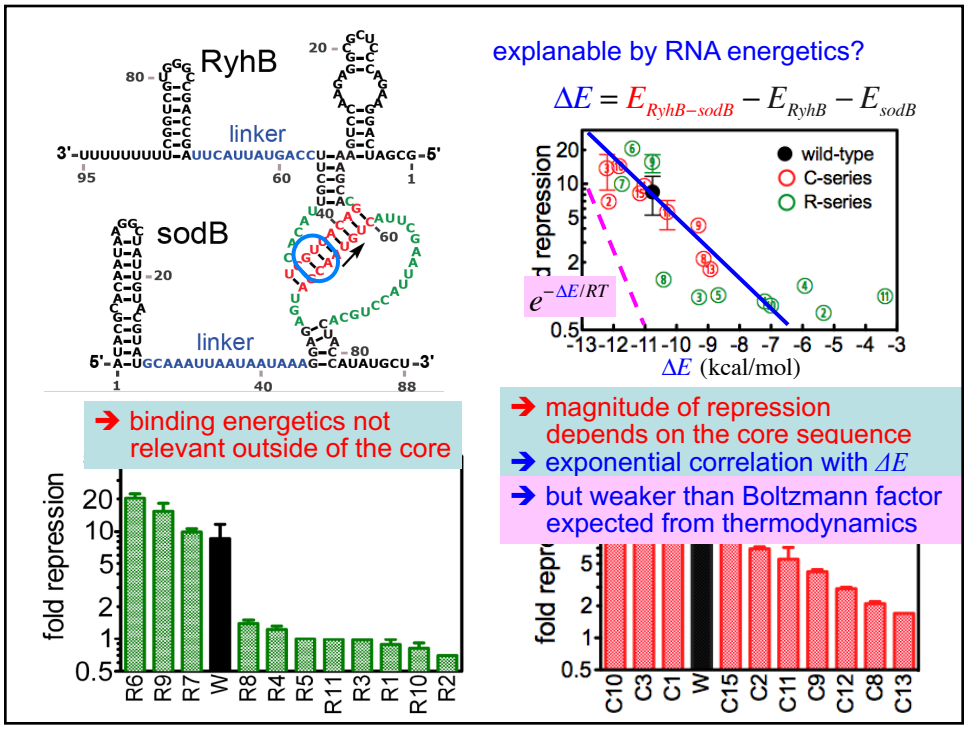
34


RyhB
linker
3'-UUUUUUUUU-AUUCAUUAUGACCU-AA-UAGCG-5'
sodB
linker
5'-A-UGCAAAUUAAUAAUAAAG-CAUAUGCU-3'
sodB-GFP expression (RFU/OD/hr)
no aTc
10ng/ml aTc
W C1 C2 C3 C4 C5 C6 C7 C8 C9 C10 C11 C12 C13 C14 C15
C-series mutants
mutation region on sodB
mutation region on ryhB
label 52-60 label 38-46
s0 AGCAATGTC r0 GACATTGCT
sc1 AGCTATGTC rc1 GACATAGCT
sc2 AGCGATGTC rc2 GACATCGCT
sc3 AGCCATGTC rc3 GACATGGCT
sc4 AGGAATGTC rc4 GACATTCCT
sc5 AGGTATGTC rc5 GACATACCT
sc6 AGGGATGTC rc6 GACATCCCT
sc7 AGGCATGTC rc7 GACATGCCT
sc8 AGAAATGTC rc8 GACATTTCT
sc9 AGATATGTC rc9 GACATATCT
sc10 AGAGATGTC rc10 GACATCTCT
sc11 AGACATGTC rc11 GACATGTCT
sc12 AGTAATGTC rc12 GACATTACT
sc13 AGTTATGTC rc13 GACATAACT
sc14 AGTGATGTC rc14 GACATCACT
sc15 AGTCATGTC rc15 GACATGACT

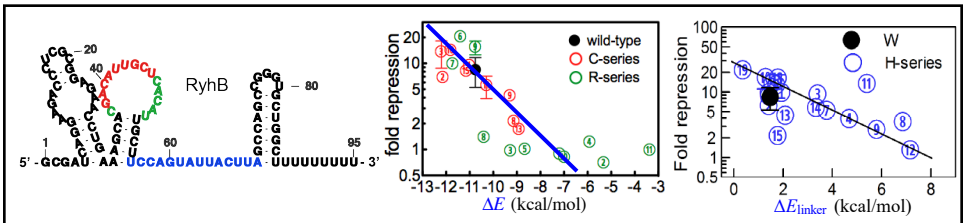
35


RyhB
linker
3'-UUUUUUUUU-AUUCAUUAUGACCU-AA-UAGCG-5'
sodB
linker
5'-A-UGCAAAUUAAUAAUAAAG-CAUAUGCU-3'
explainable by RNA energetics?
$\Delta E = E_{RyhB-sodB} - E_{RyhB} - E_{sodB}$
wild-type
C-series
fold repression
ΔE (kcal/mol)
magnitude of repression depends on the core sequence
exponential correlation with ΔE
sodB-GFP expression (RFU/OD/hr)
no aTc
10ng/ml aTc
W C1 C2 C3 C4 C5 C6 C7 C8 C9 C10 C11 C12 C13 C14 C15
C10 C3 C1 W C15 C2 C11 C9 C12 C8 C13

36


RyhB
linker
3'-UUUUUUUUU-AUUCAUUAUGACCU-AA-UAGCG-5'
sodB
linker
5'-A-UGCAAAUUAAUAAUAAAG-CAUAUGCU-3'
explainable by RNA energetics?
$\Delta E = E_{RyhB-sodB} - E_{RyhB} - E_{sodB}$
wild-type
C-series
R-series
$e^{-\Delta E/RT}$
ΔE (kcal/mol)
binding energetics not relevant outside of the core
magnitude of repression depends on the core sequence
exponential correlation with ΔE
but weaker than Boltzmann factor expected from thermodynamics
fold repression
R6 R9 R7 W R8 R4 R5 R11 R3 R1 R10 R2
C10 C3 C1 W C15 C2 C11 C9 C12 C8 C13

37



Summary

- Core base-pairing sequence:
  - -- complementarity necessary, but not sufficient
  - -- activity depends exponentially on the energy of hybridization (ΔE)
  - ➔ allows for fine-tuning of activity
  - ➔ strong binding may be a useful filter for bioinformatic searches
- Hfq-binding sequence:
  - -- activity depends exponentially on the energy opening the linker (ΔE_{linker})
  - -- not necessary for function
  - -- RyhBt does not require Hfq for function
  - ➔ Why does RyhB make itself dependent on Hfq ?
  - ➔ allows fine tuning of activity

46