5456–5463 *Nucleic Acids Research, 1998, Vol. 26, No. 23*

Analysis of complete genomes suggests that many prokaryotes do not rely on hairpin formation in transcription termination

Takanori Washio[1,2], Junji Sasayama[2,3] and Masaru Tomita[2,3,*]

[1]Laboratory for Bioinformatics, [2]Graduate School of Media and Governance and [3]Department of Environmental Information, Keio University, 5322 Endo, Fujisawa 252, Japan

Received July 20, 1998; Revised and Accepted October 21, 1998

ABSTRACT

Free energy values of mRNA tertiary structures around stop codons were systematically calculated to surmise the hairpin-forming potential for all genes in each of the 16 complete prokaryote genomes. Instead of trying to detect each individual hairpin, we averaged the free energy values around the stop codons over the entire genome to predict how extensively the organism relies on hairpin formation in the process of transcription termination. The free energy values of *Escherichia coli* K-12 shows a sharp drop, as expected, at 30 bp downstream of the stop codon, presumably due to hairpin-forming sequences. Similar drops are observed for *Haemophilus influenzae* Rd, *Bacillus subtilis* and *Chlamydia trachomatis*, suggesting that these organisms also form hairpins at their transcription termination sites. On the other hand, 12 other prokaryotes—*Mycoplasma genitalium*, *Mycoplasma pneumoniae*, *Synechocystis* PCC6803, *Helicobacter pylori*, *Borrelia burgdorferi*, *Methanococcus jannaschii*, *Archaeoglobus fulgidus*, *Methanobacterium thermoautotrophicum*, *Aquifex aeolicus*, *Pyrococcus horikoshii*, *Mycobacterium tuberculosis* and *Treponema pallidum*—show no apparent decrease in free energy value at the corresponding regions. This result suggests that these prokaryotes, or at least some of them, may never form hairpins at their transcription termination sites.

INTRODUCTION

Hairpin loop structures, formed in mRNA of *Escherichia coli*, are believed to be involved with transcription termination. There are two different mechanisms for transcription termination in *E.coli*: rho-dependent and rho-independent termination (1–9). In rho-independent termination, the elongation process of RNA polymerase is believed to be terminated by a hairpin loop in the tertiary structure of mRNA, while in the rho factor-dependent process, transcription is terminated by the interaction of the rho factor protein with RNA polymerase. Many researchers think that this is the general mechanism of prokaryotic transcription termination.

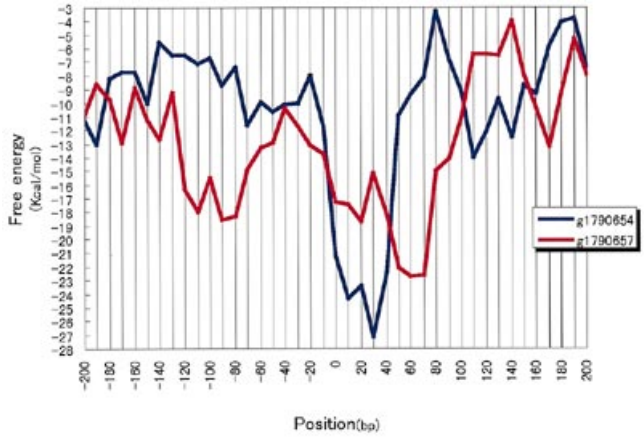


Figure 1. Free energy values around stop codons of individual genes in *E.coli* K-12.

In the present study, we systematically computed the free energy values of mRNA tertiary structures around stop codons in order to surmise the hairpin-forming potential for all genes in each of the 16 complete prokaryote genomes. It has been shown by d'Aubenton Carafa *et al*. (10) that prediction of a hairpin structure is not feasible using the free energy alone. Instead of trying to detect each individual hairpin, however, we classified all genes according to their direction and averaged the free energy values around the stop codons over the entire genome to predict how extensively the organism relies on hairpin formation in the process of transcription termination.

MATERIALS AND METHODS

The entire genome sequences of the following organisms were downloaded from National Center for Biotechnology Information (NCBI) (ftp://ncbi.nlm.nih.gov/): *Haemophilus influenzae* Rd (11), *Mycoplasma genitalium* (12), *Methanococcus jannaschii* (13), *Helicobacter pylori* (14), *Archaeoglobus fulgidus* (15),

*To whom correspondence should be addressed. Tel: +81 466 47 5111; Fax: +81 3 3440 7281; Email: mt@sfc.keio.ac.jp

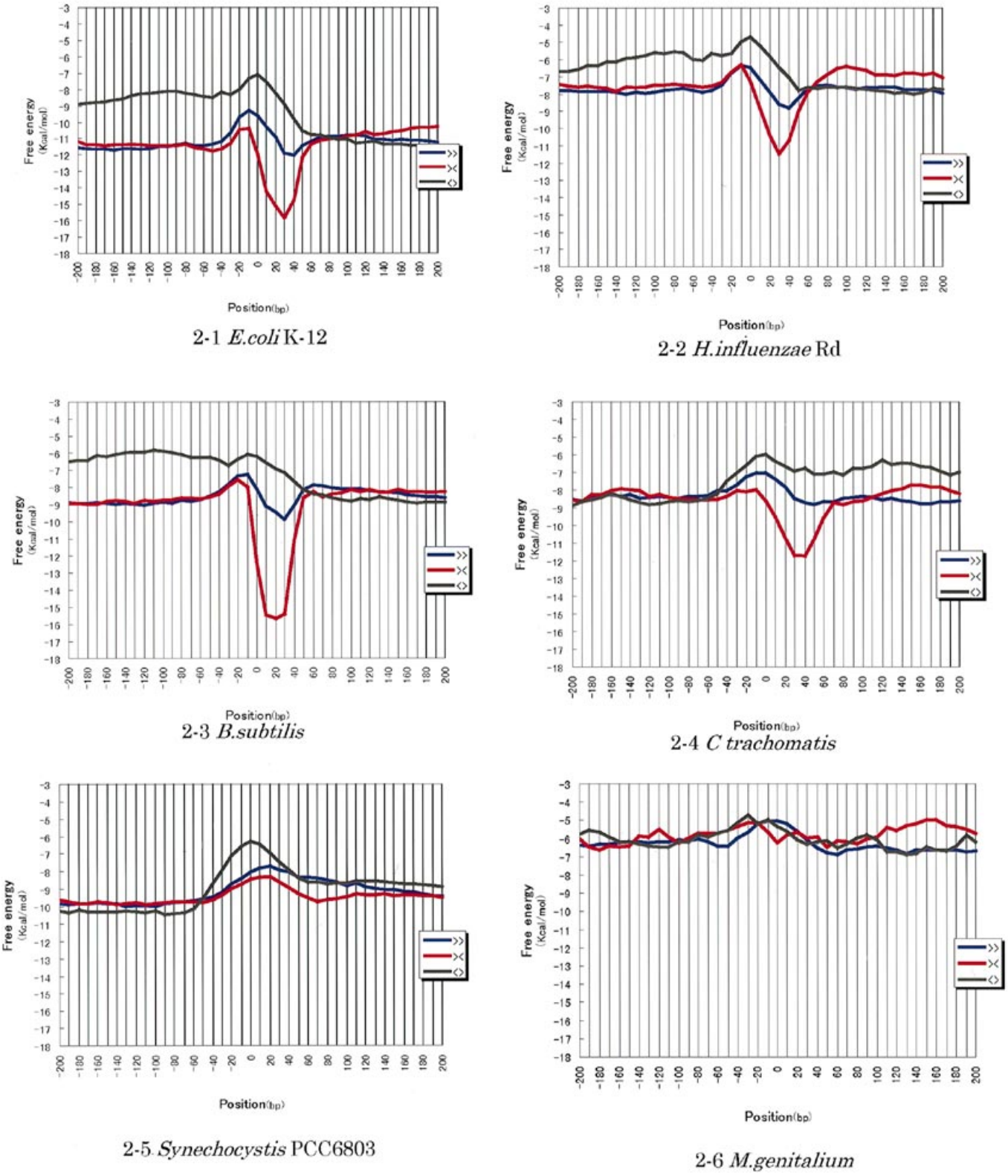
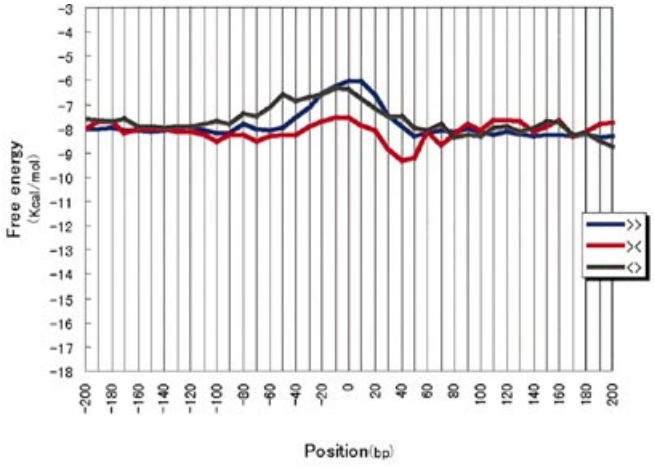
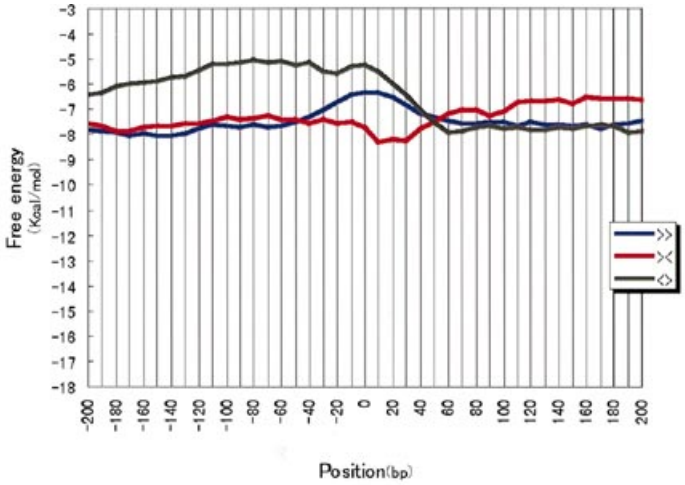


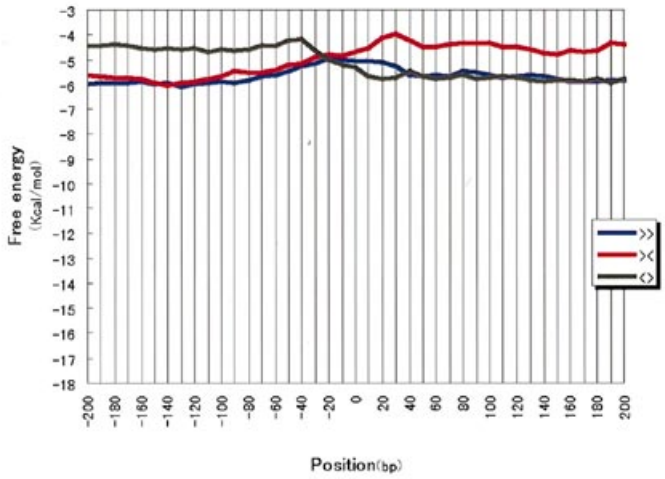
Figure 2. Free energy values around stop codons for each of the three orientations of adjacent genes.



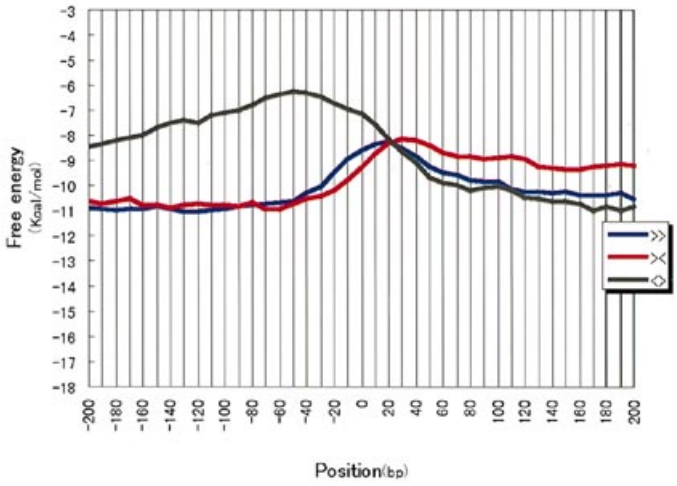
2-7 *M.pneumoniae*



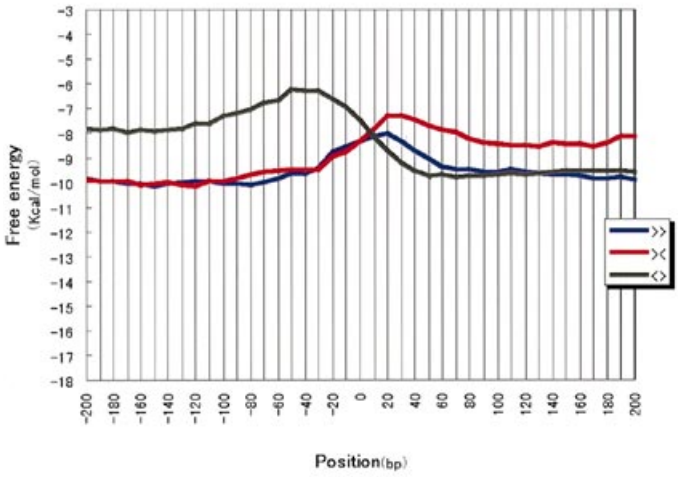
2-8 *H.pylori*



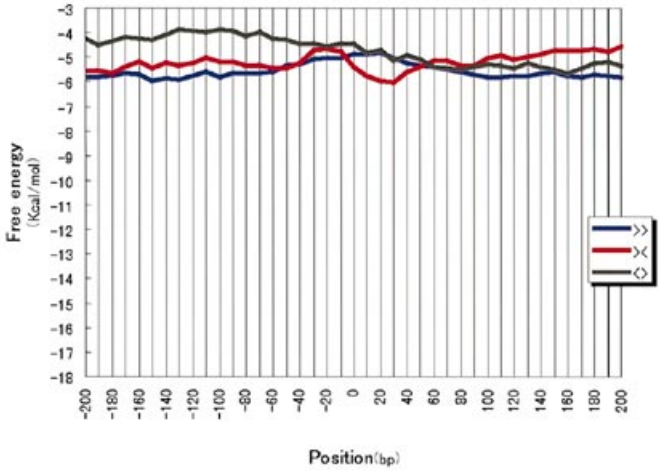
2-9 *M.jannaschii*



2-10 *M.thermoautotrophicum*

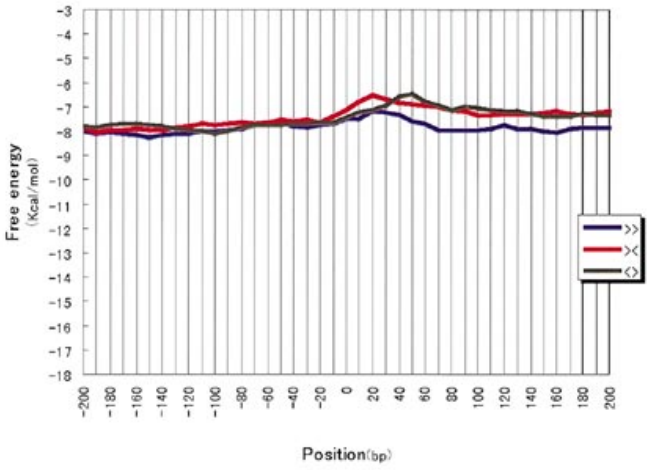


2-11 *A.fulgidus*

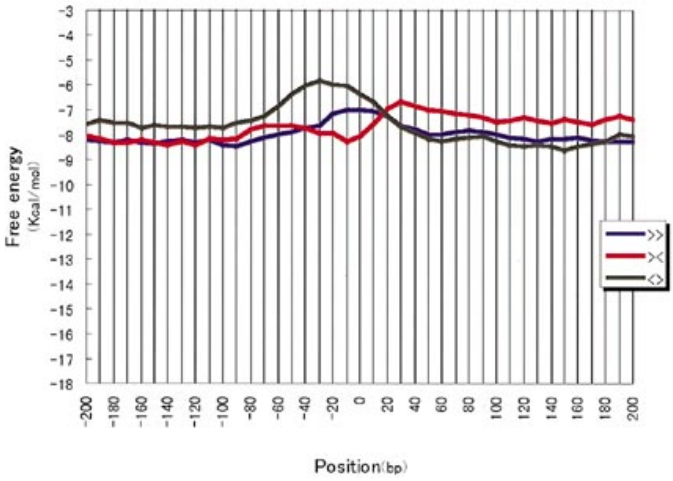


2-12 *B. burgdorferi*

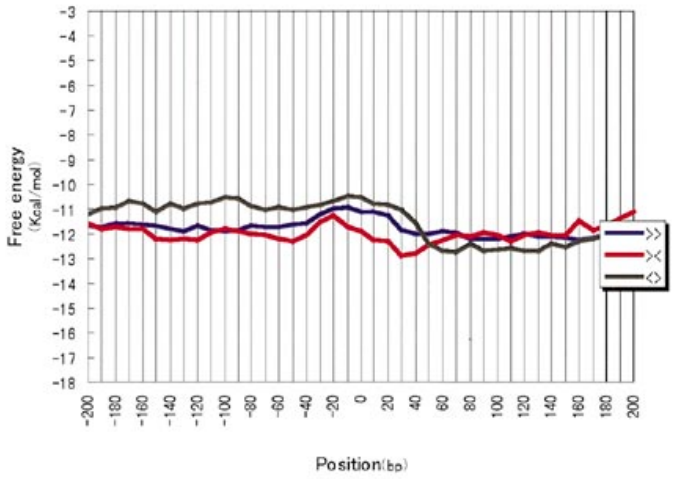
Figure 2. *Continued*



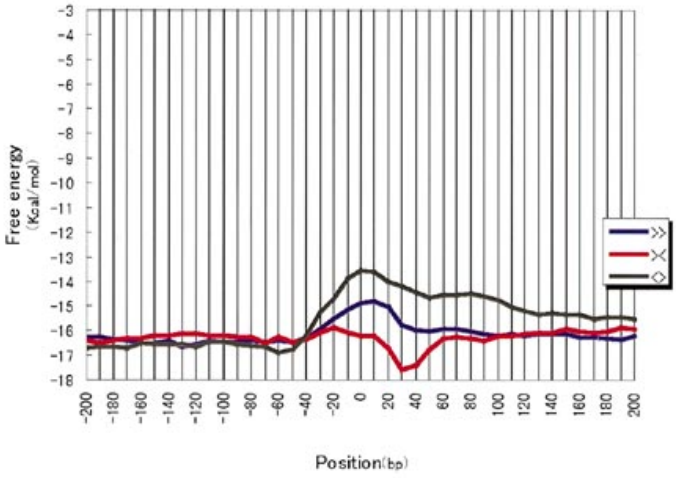
2-13 *P. horikoshii*



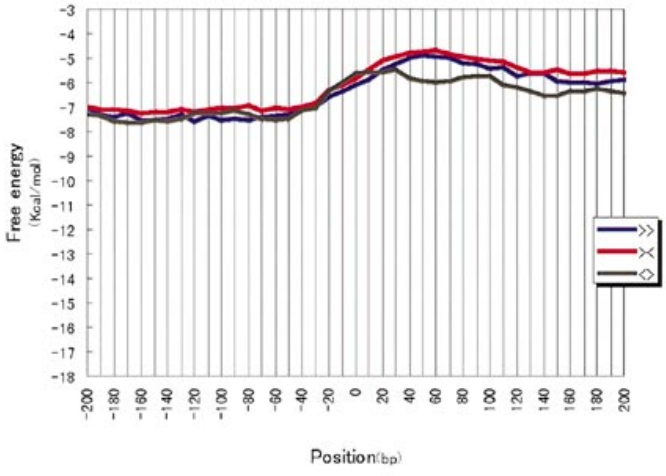
2-14 *A. aeolicus*



2-15 *T. pallidum*



2-16 *M. tuberculosis*



2-17 *S. cerevisiae* Chr VI

Figure 2. *Continued*

Borrelia burgdorferi (16), *Synechocystis* PCC6803 (17), *Mycoplasma pneumoniae* (18), *E.coli* K-12 (19), *Methanobacterium thermoautotrophicum* (20), *Bacillus subtilis* (21), *Chlamydia trachomatis* (22), *Aquifex aeolicus* (23), *Pyrococcus horikoshii* (24), *Mycobacterium tuberculosis* (25), *Treponema pallidum* (26) and *Saccharomyces cerevisiae* chromosome VI (27).

The putative genes including a terminator sequence in *E.coli* were obtained from the *E.coli* database collection (ECDC) (28), which has links to the EMBL Database (29). For reduction of data bias from these sequences, coding sequences without a description of 'terminator sequences' and coding sequences with an ambiguous description such as 'terminator-like sequences' were eliminated from our analysis. The total number of these genes used in this analysis is 111.

The free energy values were calculated using the RNA tertiary structure prediction program, 'RNAfold', developed by Hofacker *et al*. (30), whose algorithm is based on that of 'mfold' (31). We set the window size at 60 bp, and the window is shifted by 10 bp. A nucleotide sequence length of 60 bp is, thus, fed to the RNAfold program each time the window is shifted. In this way, putative free energy values are computed from 200 bp upstream to 200 bp downstream of each stop codon. Free energy values at each base position around the stop codon over all ORFs were averaged and plotted.

We also plotted free energy values for each individual ORF (Fig. 1). However, since the result generally contains too much fluctuation to predict hairpin structures, we shall not discuss each individual hairpin prediction any further in this paper.

RESULTS AND DISCUSSION

In prokaryote genomes, several genes in a specific pathway are often organized into an operon, which is transcribed into a polycistronic mRNA. Those genes within an operon, except the last gene, are not involved in transcription termination. Since reliable operon prediction from genome sequence is not possible, we are unable to reliably distinguish genes with termination sites from those without. However, we know that regions between two 'end-on' genes with opposite directions must contain transcription termination sites for these genes. We classified all intergenic regions into the following three categories: the direction of two adjacent genes can be the same direction (–> –>), 'end-on' (–> <–) or 'head-on' (<– –>). While transcription termination sites must always exist between 'end-on' genes, there may or may not exist a termination site between genes in the same direction, since they may be a part of an operon structure. Termination sites usually do not exist between 'head-on' genes.

Figure 2 shows the free energy values around stop codons for each of the three orientations of adjacent genes for the 16 prokaryotes and *S.cerevisiae*. In *E.coli* K-12, *H.influenzae* Rd, *B.subtilis* and *C.trachomatis*, we observed a sharp drop in free energy values between 'end-on' genes, but much less between genes in the same direction, and no steep drop is found between 'head-on' genes. This is consistent with the report that *E.coli* has palindromic sequences ~30 bp downstream of the stop codon for transcription termination (10). However, in the other 12 prokaryotes, as well as in *S.cerevisiae*, no such characteristic decrease in free energy values is found. These results lead us to suspect that those 12 prokaryote species do not form hairpin loops for transcription termination.

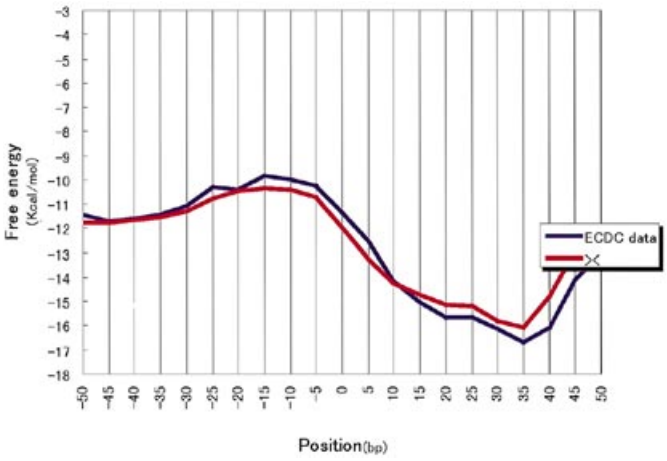


Figure 3. Free energy values of the known terminator sequences in *E.coli* K-12.

To verify that the free energy drop is indeed due to hairpin-forming sequences, we did the same analysis of only *E.coli* transcription termination sequences, which have been confirmed by laboratory experiments to form hairpin loops (28,29). The results are shown in Figure 3. (The window is shifted by 5 bp instead of 10 bp.) The steeper drop appears at the same location as the free energy drop between 'end-on' genes, confirming that those drops are due to hairpin-forming sequences.

We also analyzed *E.coli* and *Synechocystis* PCC6803 genomes with several different window sizes: 30, 40, 50 and 60 bp. Figure 4 shows the correlation between the window size and free energy values between 'end-on' genes in *E.coli* and *Synechocystis* PCC6803. In *E.coli*, we observed similar drops in all of the window sizes, while in *Synechocystis* PCC6803, no drop appears in any of them, confirming that the results of our analyses do not depend on the window size of 60 bp.

We observed a significant peak in free energy values a few bases before the stop codon in all species except *S.cerevisiae*. In general, low GC content yields high free energy values, and coding regions right before a stop codon are reported to have low GC content (32). We thus studied the relationship between free energy values and GC content. In each species examined, we observed an apparent drop in GC content at the same location as the observed peak in free energy values (Fig. 5). Therefore we suspected that the increased free energy values around these regions are merely due to the decreased GC content. In *E.coli* K-12, *H.influenzae* Rd, *B.subtilis* and *C.trachomatis*, however, the sharp drop in free energy values at 30 bp downstream of the stop codon does not appear to be correlated to the GC content (Fig. 6). More precisely, the correlation values between free energy values and GC content for 16 prokaryotes and *S.cerevisiae* are computed and listed in Table 1. The correlation values in *E.coli* K-12, *H.influenzae* Rd, *B.subtilis* and *C.trachomatis* are close to 0, indicating that free energy values of these species are independent from their GC content.

Summary of our analyses is shown in Table 1. A possible explanation for the lack of free energy drops in *M.genitalium*, *M.pneumoniae*, *H.pylori*, *B.burgdorferi*, *Synechocystis* PCC6803, *M.jannaschii*, *A.fulgidus*, *M.thermoautotrophicum*

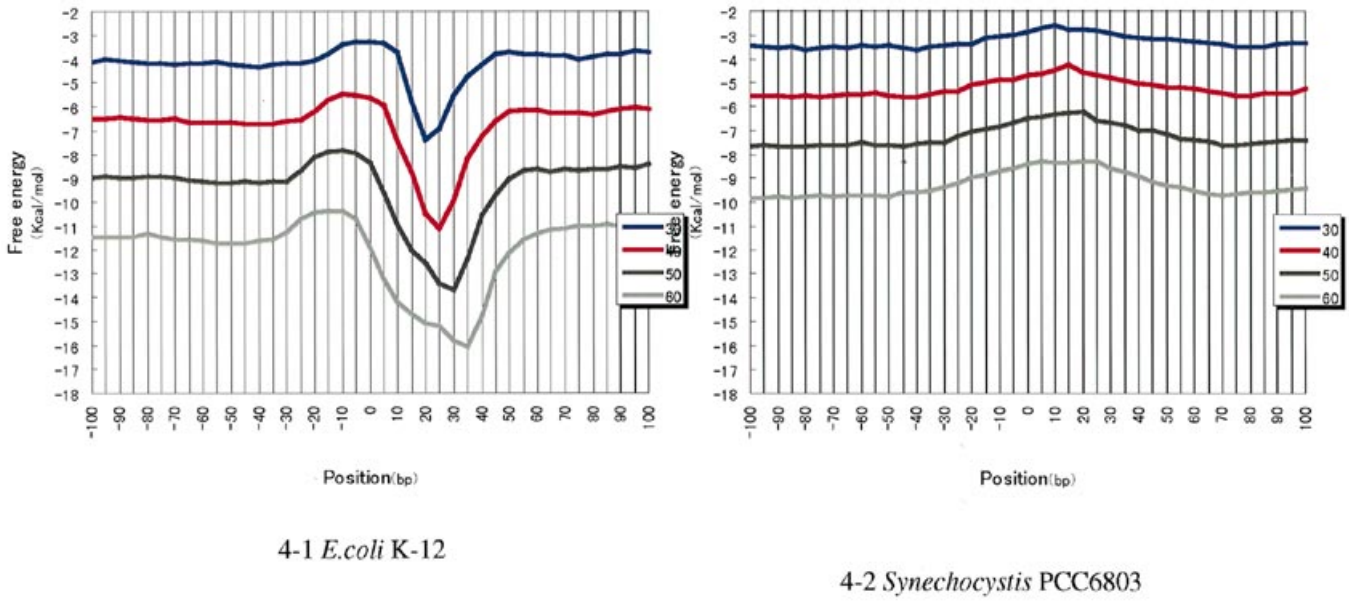


Figure 4. Correlation between the window size and free energy values in *E.coli* K-12 and *Synechocystis* PCC6803.

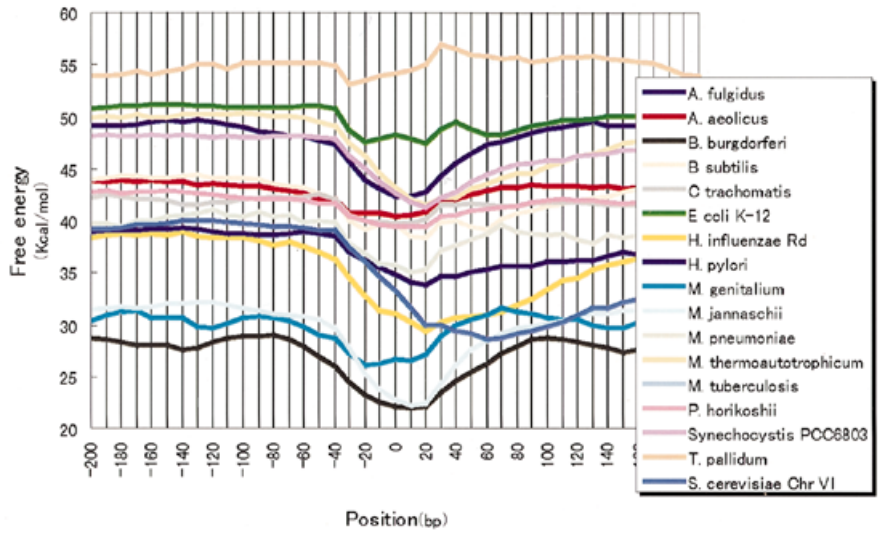


Figure 5. GC content around stop codons in 17 species.

and *T.pallidum* is that their transcription termination processes do not involve hairpin formation.

RNA polymerase II in *S.cerevisiae* is believed to terminate transcription without hairpin formation. RNA polymerase in archaebacteria is known to have structural features more similar to eukaryotic RNA polymerase II than to that of eubacteria. It is therefore plausible that the four archaebacteria (*M.jannaschii*, *A.fulgidus*, *M.thermoautotrophicum* and *P.horikoshii*) do not form hairpin structures for transcription termination.

With respect to the three eubacteria *M.genitalium*, *M.pneumoniae* and *Synechocystis* PCC6803, no gene homologous to the rho factor gene has been identified in their genomes. It is therefore tempting to speculate that they have a qualitatively different and uncharacterized mechanism for transcription termination. *Helicobacter pylori*, *B.burgdorferi*, *A.aeolicus* and *M.tuberculosis* are the only species whose genome has a rho-factor homologous gene but no apparent drop in free energy for unknown reasons. Of the 16 species analyzed, the rho-factor gene or a gene homologous to it has been identified in *E.coli* K-12, *H.influenzae* Rd, *B.subtilis*, *H.pylori*, *B.burgdorferi*, *C.trachomatis*, *M.tuberculosis* and *T.pallidum* (19,11,21,14,16,22,25,26 respectively).

Apart from its role in transcription termination, the hairpin structure is reported to be resistant to exonuclease, and it may play a role in mRNA structure stability (33–35). As seen in Table 1, two of the hairpin-forming species, *E.coli* K-12 and *H.influenzae* Rd, have genes for the exonucleases PNPase and RNaseII (or

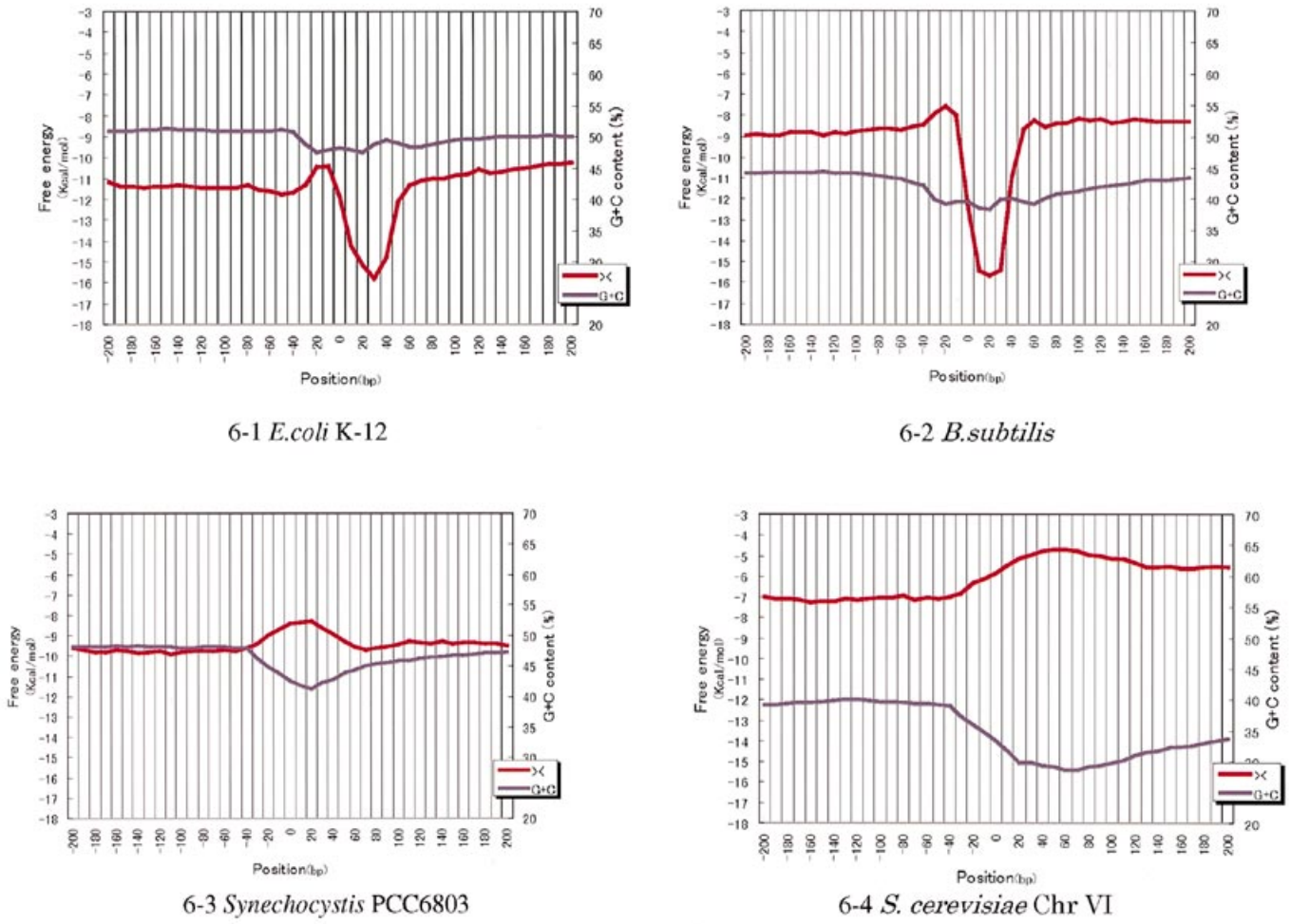


Figure 6. Relation between free energy values and local GC content in *E.coli* K-12, *B.subtilis*, *Synechocystis* PCC6803 and *S.cerevisiae* Chr VI.

Table 1. Summary of the analysis

	>>	><	Domain	Free energy drop	Rho factor	PNPase	RNase II	Correlation
A. fulgidus	1697	710	A	NO	NO	NO	NO	−0.52
A. aeolicus	1046	476	B	QUESTIONABLE	YES	NO	NO	−0.18
B. burgdorferi	620	230	B	QUESTIONABLE	YES	NO	NO	0.23
B subtilis	3012	1085	B	EVIDENT	YES	YES	NO	0.46
C trachomatis	642	252	B	EVIDENT	YES	NO	NO	0.20
E coli K−12	3025	1264	B	EVIDENT	YES	YES	YES	0.28
H. influenzae Rd	1232	446	B	EVIDENT	YES	YES	YES	0.38
H. pylori	1216	350	B	QUESTIONABLE	YES	YES	NO	−0.03
M. genitalium	391	76	B	QUESTIONABLE	NO	NO	NO	−0.33
M. jannaschii	1209	506	A	NO	NO	NO	NO	−0.57
M. pneumoniae	547	130	B	QUESTIONABLE	NO	NO	NO	−0.23
M. thermoautotrophicum	1397	472	A	NO	NO	NO	NO	−0.90
M. tuberculosis	2620	1298	B	QUESTIONABLE	YES	YES	NO	−0.20
P. horikoshii	1019	960	A	NO	NO	NO	NO	−0.72
Synechocystis PCC6803	2095	1074	B	NO	NO	NO	YES	−0.90
T. pallidum	766	266	B	QUESTIONABLE	YES	NO	YES	−0.75
S. cerevisiae Chr VI	381	409	E	NO	NO	NO	NO	−0.99

B, eubacteria; A, archaebacteria; E, eucaryotes. YES, indicates that the gene or its homolog has been identified in the genome; NO, that it has not been identified. Columns with headings >> and >< show the number of uni-directional and 'end-on' gene pairs. The column headed Correlation shows the correlation values between free energy and GC content.

their homologs), but the genomes of *B.subtilis*, *H.pylori* and *M.tuberculosis* contains only PNPase (or their homologs). The genomes of *Synechocystis* PCC6803 and *T.pallidum*, on the other hand, have only RNaseII, but does not show an apparent free energy drop. No clear correlation between possession of exonuclease genes and hairpin formation is seen.

Finally, it should be noted that the lack of apparent free energy drop does not necessarily indicate the complete absence of hairpins. Some of the species may have only small or weak hairpin formation at transcription termination site. Alternatively, if the distance between a hairpin structure and a stop codon varies uniformly, the average free energy values would not show a sharp drop. In either case, however, their transcription termination mechanism is probably different qualitatively from that of the species with an apparent drop in free energy values. We therefore conclude that the widely accepted model of transcription termination of *E.coli* involving hairpin formation is not a universal mechanism of prokaryotes.

ACKNOWLEDGEMENTS

We are deeply grateful to Hiroji Aiba and Hiroyuki Abe (Nagoya University, Nagoya, Japan), Hiroyuki Mori, Takeshi Ito and Naotake Ogasawara (Nara Graduate School of Advanced Science and Technology, Nara, Japan), Takakazu Kaneko (Kazusa DNA

Institute, Kisarazu, Japan) and Hidemi Watanabe (National Center of Biotechnology Information, Maryland, USA) for valuable discussions on the bacterial transcriptional termination systems. We also thank Rintaro Saito, Ryo Matsushima and Takuya Sakaguchi (Keio University, Kanagawa, Japan) for their help in computer programming, and Tom Shimizu for his comments on the early draft of this manuscript. This work is supported in part by a grant-in-aid for Scientific Research on Priority Areas 'Genome Science' from The Ministry of Education, Science, Sports and Culture in Japan.

REFERENCES

1 Roberts,J.W. (1969) *Nature*, **224**, 1168–1174.
2 Adhya,S. and Gottesman,M. (1978) *Annu. Rev. Biochem.*, **47**, 967–996.
3 Platt,T. (1981) *Cell*, **24**, 10–23.
4 Platt,T. (1986) *Annu. Rev. Biochem.*, **55**, 339–372.
5 Friedman,D.I., Imperiale,M.J. and Adhya,S.L. (1987) *Annu. Rev. Genet.*, **21**, 453–488.
6 Das,A. (1993) *Annu. Rev. Biochem.*, **62**, 893–930.
7 Reynolds,R. and Chamberlin,M.J. (1992) *J. Mol. Biol.*, **224**, 53–63.
8 Reynolds,R., Bermudez-Cruz,R.M. and Chamberlin,M.J. (1992) *J. Mol. Biol.*, **224**, 31–51.
9 Opperman,T. and Richardson,J.P. (1994) *J. Bacteriol.*, **176**, 5033–5043.
10 d'Aubenton Carafa,Y., Brody,E. and Thermes,C. (1990) *J. Mol. Biol.*, **216**, 835–858.
11 Fleischmann,R.D., Adams,M.D., White,O., Clayton,R.A., Kirkness,E.F., Kerlavage,A.R., Bult,C.J., Tomb,J.F., Dougherty,B.A., Merrick,J.M., *et al.* (1995) *Science*, **269**, 496–512.
12 Fraser,C.M., Gocayne,J.D., White,O., Adams,M.D., Clayton,R.A., Fleischmann,R.D., Bult,C.J., Kerlavage,A.R., Sutton,G., Kelley,J.M., *et al.* (1995) *Science*, **270**, 397–403.
13 Bult,C.J., White,O., Olsen,G.J., Zhou,L., Fleischmann,R.D., Sutton,G.G., Blake,J.A., FitzGerald,L.M., Clayton,R.A., Gocayne,J.D., Kerlavage,A.R., Dougherty,B.A., Tomb,J.F., Adams,M.D., Reich,C.I., Overbeek,R., Kirkness,E.F., Weinstock,K.G., Merrick,J.M., Glodek,A., Scott,J.L., Geoghagen,N.S.M., Weidman,J.F., Fuhrmann,J.L., Venter,J.C., *et al.* (1996) *Science*, **273**, 1058–1073.
14 Tomb,J.F., White,O., Kerlavage,A.R., Clayton,R.A., Sutton,G.G., Fleischmann,R.D., Ketchum,K.A., Klenk,H.P., Gill,S., Dougherty,B.A., Nelson,K., Quackenbush,J., Zhou,L., Kirkness,E.F., Peterson,S., Loftus,B., Richardson,D., Dodson,R., Khalak,H.G., Glodek,A., McKenney,K., Fitzegerald,L.M., Lee,N., Adams,M.D., Venter,J.C., *et al.* (1997) *Nature*, **388**, 539–547.
15 Klenk,H.P., Clayton,R.A., Tomb,J.F., White,O., Nelson,K.E., Ketchum,K.A., Dodson,R.J., Gwinn,M., Hickey,E.K., Peterson,J.D., Richardson,D.L., Kerlavage,A.R., Graham,D.E., Kyrpides,N.C., Fleischmann,R.D., Quackenbush,J., Lee,N.H., Sutton,G.G., Gill,S., Kirkness,E.F., Dougherty,B.A., McKenney,K., Adams,M.D., Loftus,B., Venter,J.C., *et al.* (1997) *Nature*, **390**, 364–370.
16 Fraser,C.M., Casjens,S., Huang,W.M., Sutton,G.G., Clayton,R., Lathigra,R., White,O., Ketchum,K.A., Dodson,R., Hickey,E.K., Gwinn,M., Dougherty,B., Tomb,J.F., Fleischmann,R.D., Richardson,D., Peterson,J., Kerlavage,A.R., Quackenbush,J., Salzberg,S., Hanson,M., van Vugt,R., Palmer,N., Adams,M.D., Gocayne,J., Venter,J.C., *et al.* (1997) *Nature*, **390**, 580–586.
17 Kaneko,T., Sato,S., Kotani,H., Tanaka,A., Asamizu,E., Nakamura,Y., Miyajima,N., Hirosawa,M., Sugiura,M., Sasamoto,S., Kimura,T., Hosouchi,T., Matsuno,A., Muraki,A., Nakazaki,N., Naruo,K., Okumura,S., Shimpo,S., Takeuchi,C., Wada,T., Watanabe,A., Yamada,M., Yasuda,M. and Tabata,S. (1996) *DNA Res.*, **3**, 109–136.
18 Himmelreich,R., Hilbert,H., Plagens,H., Pirkl,E., Li,B.C. and Herrmann,R. (1996) *Nucleic Acids Res.*, **24**, 4420–4449.
19 Blattner,F.R., Plunkett,G.,III, Bloch,C.A., Perna,N.T., Burland,V., Riley,M., Collado-Vides,J., Glasner,J.D., Rode,C.K., Mayhew,G.F., Gregor,J., Davis,N.W., Kirkpatrick,H.A., Goeden,M.A., Rose,D.J., Mau,B. and Shao,Y. (1997) *Science*, **277**, 1453–1474.
20 Smith,D.R., Doucette-Stamm,L.A., Deloughery,C., Lee,H., Dubois,J., Aldredge,T., Bashirzadeh,R., Blakely,D., Cook,R., Gilbert,K., Harrison,D., Hoang,L., Keagle,P., Lumm,W., Pothier,B., Qiu,D., Spadafora,R., Vicaire,R., Wang,Y., Wierzbowski,J., Gibson,R., Jiwani,N., Caruso,A., Bush,D., Reeve,J.N., *et al.* (1997) *J. Bacteriol.*, **179**, 7135–7155.
21 Kunst,F., Ogasawara,N., Moszer,I., Albertini,A.M., Alloni,G., Azevedo,V., Bertero,M.G., Bessieres,P., Bolotin,A., Borchert,S., Borriss,R., Boursier,L., Brans,A., Braun,M., Brignell,S.C., Bron,S., Brouillet,S., Bruschi,C.V., Caldwell,B., Capuano,V., Carter,N.M., Choi,S.K., Codani,J.J., Connerton,I.F., Danchin,A., *et al.* (1997) *Nature*, **390**, 249–256.
22 Stephens,R.S., Kalman,S., Lammel,C., Fan,J., Marathe,R., Aravind,L., Mitchell,W., Olinger,L., Tatusov,R.L., Zhao,Q., Koonin,E.V. and Davis,R.W. (1998) *Science*, **282**, 754–759.
23 Deckert,G., Warren,P.V., Gaasterland,T., Young,W.G., Lenox,A.L., Graham,D.E., Overbeek,R., Snead,M.A., Keller,M., Aujay,M., Huber,R., Feldman,R.A., Short,J.M., Olsen,G.J. and Swanson,R.V. (1998) *Nature*, **392**, 353–358.
24 Kawarabayasi,Y., Sawada,M., Horikawa,H., Haikawa,Y., Hino,Y., Yamamoto,S., Sekine,M., Baba,S., Kosugi,H., Hosoyama,A., Nagai,Y., Sakai,M., Ogura,K., Otsuka,R., Nakazawa,H., Takamiya,M., Ohfuku,Y., Funahashi,T., Tanaka,T., Kudoh,Y., Yamazaki,J., Kushida,N., Oguchi,A., Aoki,K. and Kikuchi,H. (1998) *DNA Res.*, **5**, 55–76.
25 Cole,S.T., Brosch,R., Parkhill,J., Garnier,T., Churcher,C., Harris,D., Gordon,S.V., Eiglmeier,K., Gas,S., Barry,C.E.,III, Tekaia,F., Badcock,K., Basham,D., Brown,D., Chillingworth,T., Connor,R., Davies,R., Devlin,K., Feltwell,T., Gentles,S., Hamlin,N., Holroyd,S., Hornsby,T., Jagels,K., Barrell,B.G., *et al.* (1998) *Nature*, **393**, 537–544.
26 Fraser,C.M., Norris,S.J., Weinstock,G.M., White,O., Sutton,G.G., Dodson,R., Gwinn,M., Hickey,E.K., Clayton,R., Ketchum,K.A., Sodergren,E., Hardham,J.M., McLeod,M.P., Salzberg,S., Peterson,J., Khalak,H., Richardson,D., Howell,J.K., Chidambaram,M., Utterback,T., McDonald,L., Artiach,P., Bowman,C., Cotton,M.D., Venter,J.C., *et al.* (1998) *Science*, **281**, 375–388.
27 Goffeau,A., Aert,A., Agostini-Carbone,M.L., Ahmed,A., Aigle,M., Alberghina,L., Albermann,K., Albers,M., Aldea,M., Alexandraki,D., *et al.* (1997) *Nature*, **387** (Suppl), 5–105.
28 Rodriguez-Tome,P., Stoehr,P.J., Cameron,G.N. and Flores,T.P. (1996) *Nucleic Acids Res.*, **24**, 6–12.
29 Kroger,M. and Wahl,R. (1997) *Nucleic Acids Res.*, **25**, 39–42.
30 Hofacker,I.L., Fontana,W., Stadler,P.F., Bonherffer,S., Tacker,M. and Schuster,P. (1994) *Monatshefte f. Chemie*, **125**, 167–188.
31 Zuker,M. and Stiegler,P. (1981) *Nucleic Acids Res.*, **9**, 133–148.
32 Watanabe,H., Gojobori,T. and Miura,K. (1997) *Gene*, **205**, 7–18.
33 Abe,H. and Aiba,H. (1996) *Biochimie*, **78**, 1035–1042.
34 Aiba,H., Hanamura,A. and Yamano,H. (1991) *J. Biol. Chem.*, **266**, 1721–1727.
35 Emory,S.A., Bouvet,P. and Belasco,J.G. (1992) *Genes Dev.*, **6**, 135–148.