



Preparation of *Escherichia coli* cell extract for highly productive cell-free protein expression

Takanori Kigawa¹, Takashi Yabuki¹, Natsuko Matsuda¹, Takayoshi Matsuda¹, Rie Nakajima¹, Akiko Tanaka¹ & Shigeyuki Yokoyama^{1–3,*}

¹Protein Research Group, RIKEN Genomic Sciences Center, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan; ²Cellular Signaling Laboratory and Structurome Research Group, RIKEN Harima Institute at SPring-8, 1-1-1 Kouto, Mikazuki-cho, Sayo, Hyogo 679-5148, Japan; ³Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-0033, Japan; *Author for correspondence (e-mail: yokoyama@biochem.s.u-tokyo.ac.jp; fax: +81-45-503-9195)

Received 10 December 2003; accepted in revised form 20 January 2004

Key words: cell-free protein synthesis, high-throughput, structural genomics, structural proteomics

Abstract

As structural genomics and proteomics research has become popular, the importance of cell-free protein synthesis systems has been realized for high-throughput expression. Our group has established a high-throughput pipeline for protein sample preparation for structural genomics and proteomics by using cell-free protein synthesis. Among the many procedures for cell-free protein synthesis, the preparation of the cell extract is a crucial step to establish a highly efficient and reproducible workflow. In this article, we describe a detailed protocol for *E. coli* cell extract preparation for cell-free protein synthesis, which we have developed and routinely use. The cell extract prepared according to this protocol is used for many of our cell-free synthesis applications, including high-throughput protein expression using PCR-amplified templates and large-scale protein production for structure determinations.

Abbreviations: 2-mEt – 2-mercaptoethanol; AcCoA – acetyl coenzyme A; BL21 CP – BL21 codon-plus RIL; CAT – chloramphenicol acetyl transferase; CK – creatine kinase; Cm – chloramphenicol; DEPC – diethylpyrocarbonate; CP – creatine phosphate; DTNB – 5,5'-dithiobis-2-nitrobenzoic acid; DTT – dithiothreitol; Folic acid – L(-)-5-formyl-5,6,7,8-tetrahydrofolic acid; KGlu – polyethyleneglycol; PEG – potassium glutamate; PEP – phospho-*enol*pyruvate; PK – pyruvate kinase.

Introduction

Cell-free protein synthesis has become a powerful protein expression method due to recent advancements. The low productivity problem, which was considered to be the most serious obstacle of cell-free synthesis, has been solved by optimization of the reaction conditions (Kigawa *et al.*, 1995, 1999; Kim *et al.*, 1996; Yabuki *et al.*, 1998), modification of the conditions, such as the energy-regenerating system (Kim and Swartz, 1999, 2000, 2001), and improvement of the cell extract preparation (Madin *et al.*, 2000; Shimizu *et al.*, 2001). The development of

novel reaction methods, including the dialysis method (Kigawa *et al.*, 1999; Kigawa and Yokoyama, 1991; Kim and Choi, 1996; Spirin *et al.*, 1988) and the bilayer method (Sawasaki *et al.*, 2002a), also improved the productivity. The packaged kits incorporating various aspects of these improved cell-free protein synthesis systems are now available from several companies (Roche Applied Science, Promega, Invitrogen, Qiagen, and Toyobo).

We have been studying and developing the cell-free protein synthesis system that uses the *E. coli* cell extract originally developed by Zubay (Zubay, 1973), in order to establish it as one of the most powerful

protein expression methods. A high-throughput protein expression method has already been established by combining cell-free protein synthesis with PCR, which is performed on multi-well plates and is thus adapted for robotics. In this method, the PCR-amplified linear DNA fragment is utilized as a template for protein synthesis, without any cloning procedures. This method has been applied to protein sample screening for structural and functional analyses (Yokoyama *et al.*, 2000), and, so far, more than 100,000 expression constructs have been tested for the productivity and solubility of the products. We have also established a large-scale sample production method using the dialysis cell-free method for structure analyses (Kigawa *et al.*, 1999, 2002), which could produce our target proteins in milligram quantities. This is extremely useful for producing uniformly [^{13}C , ^{15}N]-labeled samples for NMR spectroscopy (Kigawa *et al.*, 1999) and seleno-methionine-substituted samples for X-ray crystallography (Kigawa *et al.*, 2002). We have already produced more than 100 uniformly [^{13}C , ^{15}N]-labeled protein samples using the cell-free system, and then determined their structures.

In recent years, the research area of structural genomics and proteomics, which aims to identify all protein structures and/or functions and then to compile an encyclopedia, has rapidly emerged, based on the fruits of genome sequencing projects. As extremely large numbers of protein samples need to be prepared efficiently and rapidly, protein expression methods adapted to a high-throughput format are required to facilitate this effort. Cell-free protein synthesis could produce proteins directly from PCR-amplified linear DNA fragments, without any cloning procedures. All of the reactions can be carried out in a multi-well plate format, which could be easily adapted to automated procedures. Thus, it has been widely noticed as a promising and suitable protein expression method for these purposes (Busso *et al.*, 2003; Sawasaki *et al.*, 2002b; Yokoyama *et al.*, 2000), mainly owing to our efforts toward its development and applications.

Much effort has been expended while developing the protocol for the extract preparation from *E. coli* cells. Although this protocol is fundamentally based on that of Pratt (Pratt, 1984), our modifications dramatically improved the reproducibility of the extract preparation and thus the productivity of the cell-free synthesis. Within a year, we prepared approximately 5 l of the cell extract, which corresponds to 16 l of

the cell-free reaction mixture, and almost all of the extract has reproducible and sufficient activity for the cell-free synthesis. The detailed protocol that we have established thus far is described in the present article.

Materials and methods

Apparatus and reagents

All glassware, centrifuge tubes, and cylinders should be treated with RNase AWAY (Molecular BioProducts), AbSolve (DuPont/NEN), or equivalent reagents to avoid RNase contamination. Creatine phosphate (CP), creatine kinase (CK), phospho-*enol*pyruvate (PEP), and an *E. coli* total tRNA mixture prepared from MRE600 strain were purchased from Roche. ATP, GTP, UTP, and CTP were from Seikagaku Corporation. Pyruvate kinase (PK), L(-)-5-formyl-5,6,7,8-tetrahydrofolic acid (folinic acid), cAMP, potassium glutamate, polyethylene glycol (PEG) (average MW 8,000), chloramphenicol (Cm), and acetyl coenzyme A (AcCoA) were obtained from Sigma Aldrich. Amino acids, dithiothreitol (DTT), and diethylpyrocarbonate (DTNB) were obtained from Nacalai Tesque. Other reagents used in this study were obtained from Wako Pure Chemicals. T7 RNA polymerase was prepared according to reported procedures (Davanloo *et al.*, 1984; Grodberg and Dunn, 1988; Zawadzki and Gross, 1991). For CAT protein synthesis, plasmid pK7-CAT (Kim *et al.*, 1996), which has the T7 promoter and the gene for the CAT protein, was used as the DNA template.

Usually, the water used for cell-free protein synthesis is treated with diethylpyrocarbonate (DEPC) for RNase inactivation. However, in our experience, ultra-pure water prepared by the Milli-Q Synthesis system (Millipore) has sufficient quality for the cell-free synthesis. Thus, this ultra-pure water is routinely used as the RNase-free water for stock solution preparation, and is simply denoted as 'Milli-Q water' in this article.

Preparation of cell extract

We usually use the *E. coli* BL21 codon-plus RIL strain (BL21 CP, Stratagene) for the cell extract preparation. Some parts of this protocol have been optimized for the BL21 CP strain, and should therefore be modified if another strain is used.

The BL21 CP strain was grown at 37 °C in six 2-l baffled flasks, each with 500 ml 2 × YT medium containing 34 µg/ml Cm, by circular shaking at 160 rpm. Cells were harvested in the mid-log phase ($OD_{600} \cong 3$, usually 3–4 h), and then were washed three times with the S30 buffer (A) [10 mM Tris-acetate buffer (pH 8.2) containing 14 mM Mg(OAc)₂, 60 mM KOAc, 1 mM DTT, and 0.5 ml/l 2-mercaptoethanol]. At this stage, we used a Polytron cell homogenizer (Kinematica AG, PT-MR3100) to facilitate resuspension of the cell pellets. The collected cells were immediately and completely frozen by being immersed in liquid nitrogen for at least 2 min, and then stored at –80 °C for 1–3 days. From our experience, storage periods longer than 3 days result in lower extract activity. In a typical case, about 16 g of loosely packed cells are obtained from a 3 l culture.

The frozen cells were thawed in 50 ml of the S30 buffer (A), and then collected by centrifugation at $16,000 \times g$ for 30 min at 4 °C. Cells (7 g) were suspended in 8.9 ml of the S30 buffer (B) (the same as the S30 buffer (A) except for the absence of 2-mercaptoethanol), and disrupted with 22.7 g of glass beads (B. Braun Melsungen AG 854–150/7, $\phi = 0.17$ – 0.18 mm) by the Multi-beads Shocker Type MB301 (Yasui Kikai, <http://www.yasuikikai.co.jp>) with the program: 30 sec on, 30 sec off, 30 sec on, 30 sec off, 30 sec on. The preparation was centrifuged twice at $30,000 \times g$ for 30 min at 4 °C, to remove the glass beads. The resultant supernatant (4.9 mL) was incubated at 37 °C for 80 min with 0.3 volume of the preincubation buffer [293 mM Tris-OAc buffer (pH 8.2) containing 9.2 mM Mg(OAc)₂, 13.2 mM ATP (pH 7.0), 84 mM PEP (pH 7.0), 4.4 mM DTT, 40 µM of each of the 20 amino acids, and 6.7 U/ml PK (Sigma P7768)]. The incubated sample was dialyzed 4 times against 50 volumes of the S30 buffer (B) for 45 min each at 4 °C, and then centrifuged at $4,000 \times g$ for 10 min at 4 °C to obtain the supernatant (the S30 extract). The S30 extract was dispensed into 0.2–2.0 ml aliquots, immediately and completely frozen by being immersed in liquid nitrogen for at least 2 min, and then stored in liquid nitrogen or below –80 °C. The S30 extract remains active for at least 1 year if stored in liquid nitrogen.

The cell-free protein synthesis reaction

To validate the activity of the prepared S30 extract, the optimum magnesium concentration of CAT synthesis was investigated using the batch mode of cell-

free protein synthesis. The batch-mode of the cell-free reaction consisted of (per 30 µl) 55 mM Hepes-KOH buffer (pH 7.5) containing 1.7 mM DTT, 1.2 mM ATP (pH 7.0), 0.8 mM each of CTP (pH 7.0), GTP (pH 7.0), and UTP (pH 7.0), 80 mM CP, 250 µg/ml CK, 4.0 % PEG 8000, 0.64 mM 3',5'-cyclic AMP, 68 µM L(–)-5-formyl-5,6,7,8-tetrahydrofolic acid, 175 µg/ml *E. coli* total tRNA, 210 mM potassium glutamate, 27.5 mM ammonium acetate, 10.7 mM magnesium acetate, 1.0 mM of each of the 20 amino acids, 6.7 µg/ml of the pK7-CAT plasmid, 93 µg/ml T7 RNA polymerase, and 9.0 µl S30 extract. The reaction mixture was incubated at 37 °C for 1 h.

The dialysis mode of the cell-free reaction (Kigawa *et al.*, 1999) was carried out as follows. The internal solution (3 ml) consisted of all of the components used for the batch mode reaction mixture and 0.05% sodium azide. The external solution (30 ml) contained the components of the internal solution except for CK, the plasmid vector, the T7 RNA polymerase *E. coli* total tRNA, and the S30 extract. The internal solution was dialyzed in a dialysis tube (Spectra/Por 7 MWCo:15,000, Spectrum) against the external solution at 30 °C for 8 h with shaking.

The CAT productivity was calculated from CAT activity of the reaction as follows. Three µl of diluted (1000 times in a typical case) reaction was mixed with 400 µl of color reagent (0.4 mg/ml DTNB, 0.1 mM AcCoA, and 0.1 mM Cm), and then incubated at 37 °C for 30 min. The CAT protein content was calculated from the A_{412} value of the reaction using the following formulas (Shaw, 1975):

$$\text{CAT [units/ml]} = (A_{412} - A_{412\text{blk}}) \times 1000 \\ \times (1000/T/\epsilon_{412})$$

$$\text{CAT [\mu g/ml (assay mixture)]} = \text{CAT [units/ml]} / A$$

$$\text{CAT [\mu g/ml (cell-free reaction mixture)]} \\ = \text{CAT [\mu g/ml (assay mixture)]} / (V/S) \times D$$

where: ϵ_{412} [M/cm]: the molar extinction coefficient for reduced DTNB at 412 nm, 13600/M/cm; A [units/µg]: specific activity for CAT, 150 units/mg; T [min]: color development time, 30 min; V [µl]: sample volume used for assay, 3 µl; S [µl]: assay scale, 400 µl; D : sample dilution factor, 1000.

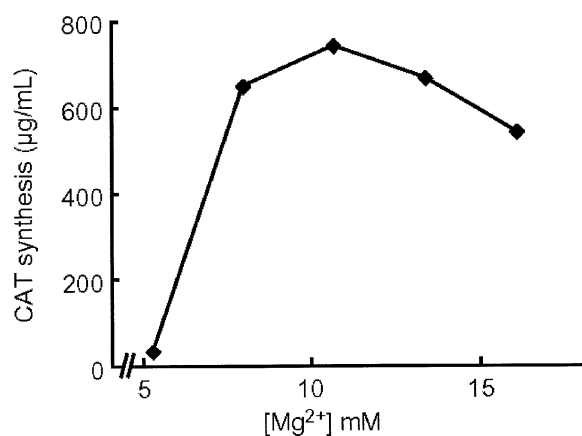


Figure 1. Typical magnesium concentration dependency profile of CAT production by the batch mode of cell-free protein synthesis.

Results and discussion

Extract validation

In our experience, the optimum magnesium concentration of the cell-free system that uses the S30 extract prepared according to this protocol is in the range of 10–14 mM, and the CAT productivity at the optimum magnesium concentration is in the range of 600–800 µg/ml reaction mixture. Figure 1 illustrates a typical magnesium concentration dependency profile of the protein synthesis. As the kinetics of the cell-free translation reaction is strongly affected by the magnesium concentration, the optimum concentration should be examined for each batch to obtain the best result. It is also important to monitor the protein products by SDS-PAGE analysis (Figure 2). We usually examine (i) the optimum magnesium concentration by the batch system using the plasmid template, and (ii) the protein productivity of the dialysis systems as well as the batch system, using both the plasmid and linear templates.

E. coli strains

E. coli strains that lack a major RNase, such as the MRE600 and A19 strains, were previously favored as the source of the cell extract (Pratt, 1984; Zubay, 1973). However, we usually use the BL21 codon-plus RIL strain, a BL21-derivative strain containing extra copies of the *argU*, *ileY*, and *leuW* tRNA genes, which encode tRNA species that recognize the argin-

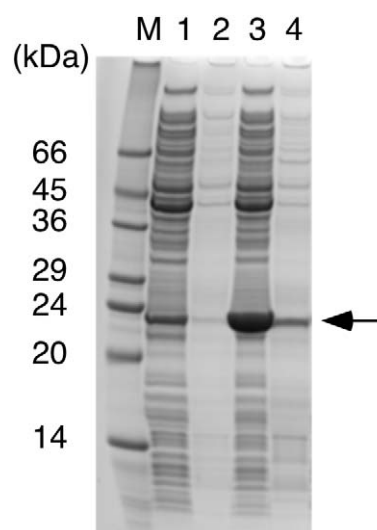


Figure 2. SDS-PAGE analysis of the CAT protein expressed using the plasmid template by the batch mode (lanes 1 and 2) and the dialysis mode (lanes 3 and 4). The expression reaction was divided into supernatant (lanes 1 and 3) and precipitate (lanes 2 and 4) fractions. On each lane, the sample equivalent to 2.5 µl of the reaction mixture was loaded. The gel was stained with Coomassie Brilliant Blue.

ine minor codons AGA and AGG, the isoleucine minor codon AUA, and the leucine minor codon CUA, respectively. From our examinations, cell extracts prepared from strains based on the BL21 strain and containing extra copies of tRNA genes, such as BL21 codon-plus strains (Stratagene) or Rosetta strains (Novagen), are 50% more active for cell-free protein synthesis. There might be three reasons for the higher activity of the extracts from these strains. First, the mRNA productivity has been enhanced by the use of an exogenous RNA polymerase (Kigawa *et al.*, 1995), and thus mRNA degradation by the intrinsic RNases has less of an effect on the protein expression. Second, the protein expression is also strongly influenced by the existence of minor codons in the coding sequence. Third, proteolytic degradation during protein synthesis by intrinsic proteases also has an effect on the protein expression, and this problem is overcome with the BL21-derivative strain, which belongs to the *E. coli* B strain family lacking OmpT endoproteinase and lon protease. We thus assume that if a BL21-derivative strain lacking a major RNase exists, then it would be one of the best strains for the extract source.

Cell disruption method

Pratt's protocol (Pratt, 1984), on which our protocol is based, uses a French press (Aminco) to disrupt the *E. coli* cells. However, in our experience, it is difficult to obtain reproducible cell disruption results by using a French press in a laboratory-scale preparation, mainly because it is quite difficult to disrupt *E. coli* cells uniformly. We now accomplish the disruption with small glass beads, which are usually used to disrupt cells with rigid cell walls, such as yeast and small algae, and thus always obtain highly active cell extracts. We had previously used the MSK cell homogenizer (B. Braun) for this purpose, and then switched to the Multi-beads Shocker (Yasui Kikai) because of better sample cooling during disruption and higher capacity. In our tests, sonication, which is another popular cell disruption method, especially for *E. coli* cells, is not suitable for extract preparation, due to sample heating and difficulty of management.

Purification of the template DNA

The quality of the template DNA is generally important for cell-free protein synthesis, and this is also true for the *E. coli* cell-free synthesis. Usually, plasmid DNA purified by the commercially available kits (Promega, Qiagen, etc.) has acceptable quality, with little RNase contamination if extra extractions with phenol and then phenol–chloroform are performed before ethanol precipitation. We also think it is better to use kits (Qiagen EndoFree Plasmid kit, etc.) or procedures that remove endotoxin contamination, because, from our experience, plasmid DNA purified by these kits or procedures gives at least 20% more productivity.

T7 RNA polymerase

We usually purify the T7 RNA polymerase used for the cell-free synthesis according to reported procedures (Davanloo *et al.*, 1984; Grodberg and Dunn, 1988; Zawadzki and Gross, 1991), mainly because the enzymes from commercial companies are not sufficiently concentrated. However, based on our tests, highly concentrated lots from the following companies have almost the same activity as ours for the cell-free synthesis: Ambion, Takara, Toyobo (not suitable for the linear template in our assays), and Promega.

Stable-isotope labeling for NMR analysis

We have reported that the cell-free system is extremely useful for three different kinds of stable-isotope labeling of proteins: amino acid-selective (Kigawa *et al.*, 1995; Yabuki *et al.*, 1998), uniform (Kigawa *et al.*, 1999), and site-directed (Hirao *et al.*, 2002; Kiga *et al.*, 2002; Kigawa *et al.*, 1999) labeling. Amino acid-selective or uniform labeling can be achieved by simply replacing the amino acid(s) of interest in the cell-free reaction mixture with labeled one(s). A relatively inexpensive stable-isotope labeled algal amino acid mixture can be used instead of expensive purified labeled amino acids (Kigawa *et al.*, 1999). In this case, labeled cysteine, tryptophan, asparagine, and glutamine are still required for uniform labeling, because the commercially available algal amino acid mixture lacks these four amino acids. The optimum concentration (from our experience, 2–4 mg/ml) of the mixture should be examined because the productivity of the cell-free synthesis becomes lower at higher concentration of the mixture. We think that this is because some unidentified contaminant(s) in the algal amino acid mixture inhibit(s) the cell-free protein synthesis. From our experience, this inhibitory effect differs between each manufacturer and also between the preparations from one manufacturer. We strongly recommend that as many available preparations as possible should be examined to obtain the best result.

Cell-free protein synthesis is an extremely powerful protein expression method, especially for high-throughput applications. The extract preparation is one of the most important matters for cell-free protein synthesis. The protocol for cell extract preparation, which we routinely use, is described in this article, and will yield a highly active cell extract if followed precisely.

Acknowledgements

This work was supported by the RIKEN Structural Genomics/Proteomics Initiative (RSGI), the National Project on Protein Structural and Functional Analyses, Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

- Busso, D., Kim, R. and Kim, S.H. (2003) *J. Biochem. Biophys. Meth.* **55**, 233–240.
- Davanloo, P., Rosenberg, A.H., Dunn, J.J. and Studier, F.W. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 2035–2039.
- Grodberg, J. and Dunn, J.J. (1988) *J. Bacteriol.* **170**, 1245–1253.
- Hirao, I., Ohtsuki, T., Fujiwara, T., Mitsui, T., Yokogawa, T., Okuni, T., Nakayama, H., Takio, K., Yabuki, T., Kigawa, T., Kodama, K., Nishikawa, K. and Yokoyama, S. (2002) *Nat. Biotechnol.* **20**, 177–182.
- Kiga, D., Sakamoto, K., Kodama, K., Kigawa, T., Matsuda, T., Yabuki, T., Shirouzu, M., Harada, Y., Nakayama, H., Takio, K., Hasegawa, Y., Endo, Y., Hirao, I. and Yokoyama, S. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 9715–9720.
- Kigawa, T., Muto, Y. and Yokoyama, S. (1995) *J. Biomol. NMR* **6**, 129–134.
- Kigawa, T., Yabuki, T., Yoshida, Y., Tsutsui, M., Ito, Y., Shibata, T. and Yokoyama, S. (1999) *FEBS Lett.* **442**, 15–19.
- Kigawa, T., Yamaguchi-Nunokawa, E., Kodama, K., Matsuda, T., Yabuki, T., Matsuda, N., Ishitani, R., Nureki, O. and Yokoyama, S. (2002) *J. Struct. Funct. Genom.* **2**, 29–35.
- Kigawa, T. and Yokoyama, S. (1991) *J. Biochem. (Tokyo)* **110**, 166–168.
- Kim, D.M. and Choi, C.Y. (1996) *Biotechnol. Prog.* **12**, 645–649.
- Kim, D.M., Kigawa, T., Choi, C.Y. and Yokoyama, S. (1996) *Eur. J. Biochem.* **239**, 881–886.
- Kim, D.M. and Swartz, J.R. (1999) *Biotechnol. Bioeng.* **66**, 180–188.
- Kim, D.M. and Swartz, J.R. (2000) *Biotechnol. Prog.* **16**, 385–390.
- Kim, D.M. and Swartz, J.R. (2001) *Biotechnol. Bioeng.* **74**, 309–316.
- Madin, K., Sawasaki, T., Ogasawara, T. and Endo, Y. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 559–564.
- Pratt, J.M. (1984) In *Transcription and Translation* (Eds., Hames, B.D. and Higgins, S.J.), IRL Press, Oxford, UK and Washington, DC, pp. 179–209.
- Sawasaki, T., Hasegawa, Y., Tsuchimochi, M., Kamura, N., Ogasawara, T., Kuroita, T. and Endo, Y. (2002a) *FEBS Lett.* **514**, 102–105.
- Sawasaki, T., Ogasawara, T., Morishita, R. and Endo, Y. (2002b) *Proc. Natl. Acad. Sci. USA* **99**, 14652–14657.
- Shaw, W.V. (1975) *Methods Enzymol.* **43**, 737–755.
- Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K. and Ueda, T. (2001) *Nat. Biotechnol.* **19**, 751–755.
- Spirin, A.S., Baranov, V.I., Ryabova, L.A., Ovodov, S.Y. and Alakhov, Y.B. (1988) *Science* **242**, 1162–1164.
- Yabuki, T., Kigawa, T., Dohmae, N., Takio, K., Terada, T., Ito, Y., Laue, E.D., Cooper, J.A., Kainosho, M. and Yokoyama, S. (1998) *J. Biomol. NMR* **11**, 295–306.
- Yokoyama, S., Hirota, H., Kigawa, T., Yabuki, T., Shirouzu, M., Terada, T., Ito, Y., Matsuo, Y., Kuroda, Y., Nishimura, Y., Kyogoku, Y., Miki, K., Masui, R. and Kuramitsu, S. (2000) *Nat. Struct. Biol.* **7 Suppl.**, 943–945.
- Zawadzki, V. and Gross, H.J. (1991) *Nucleic Acids Res.* **19**, 1948.
- Zubay, G. (1973) *Ann. Rev. Genet.* **7**, 267–287.