Volume 241, number 1,2, 99–104 FEB 06544 December 1988

A new photosystem II reaction center component (4.8 kDa protein) encoded by chloroplast genome

Masahiko Ikeuchi and Yorinao Inoue

Solar Energy Research Group, The Institute of Physical and Chemical Research (RIKEN), Wako, Saitama 351-01, Japan

Received 19 September 1988

The photosystem II reaction center complex, so-called D1-D2-cytochrome *b*-559 complex, isolated from higher plants contains a new component of about 4.8 kDa [(1988) Plant Cell Physiol. 29, 1233–1239]. The partial amino acid sequence of this component from spinach was determined after release of N-terminal blockage. The determined sequence matched an open reading frame (ORF36) of the chloroplast genome from tobacco and liverwort, which is located downstream from the *psb*K gene and forms an operon with *psb*K. The predicted product consists of 36 amino acid residues and has a single membrane-spanning segment. High homology between the tobacco and liverwort genes, and its presence in the reaction center complex suggest an important role for this component in the photosystem II complex. Since this gene corresponds to a part of the formerly designated *psb*I gene, we propose to revise the definition of *psb*I as the gene encoding the 4.8 kDa reaction center component.

Chloroplast genome; Photosystem II; Reaction center component; *psb*I Gene

1. INTRODUCTION

The PS II reaction center complex isolated by Nanba and Satoh [1] is so far the minimum unit complex to mediate PS II primary photochemical reactions, consisting of D1 protein, D2 protein and the large and small subunits of cytochrome *b*-559. These reaction center components are encoded by *psb*A, *psb*D, *psb*E and *psb*F, respectively, of the chloroplast genome [2].

By using a new SDS-PAGE method affording high resolution of low-molecular-mass proteins, we recently found a fifth component of about 4.8 kDa in the PS II reaction center complex from wheat and spinach [3]. Here, we report its partial amino acid sequence and the corresponding gene identified in the chloroplast genome from tobacco and liverwort.

Correspondence address: M. Ikeuchi, Solar Energy Research Group, The Institute of Physical and Chemical Research (RIKEN), Wako, Saitama 351-01, Japan

Abbreviations: ORF, open reading frame; PS II, photosystem II; SDS-PAGE, SDS-polyacrylamide gel electrophoresis

2. MATERIALS AND METHODS

The reaction center complex was prepared from spinach PS II membrane fragments basically according to the method described in [1] with some modifications. The PS II membrane fragments were solubilized with Triton X-100 (Triton X-100/chlorophyll = 50 (w/w)) and loaded onto a DEAE-Toyopearl column (Toso, Japan). The column was extensively washed with 30 mM NaCl, 0.05% (w/v) Triton X-100 and 50 mM Tris-HCl (pH 7.2) for 20 h. Then the reaction center complex was eluted with 100 mM NaCl, 0.05% (w/v) Triton X-100 and 50 mM Tris-HCl (pH 7.2) and precipitated by centrifugation at 276000 × *g* for 3 h after five times dilution with 50 mM Tris-HCl (pH 7.2). The thus obtained reaction center complex is depleted of detergent and preserves photochemical activities for many hours even at room temperature.

SDS-PAGE with 7.5 M urea and a 16–22% (w/v) polyacrylamide gradient was done according to the method described in [3]. Proteins were stained with Coomassie brilliant blue R-250 (Biorad). Western blotting analysis with the antiserum raised against spinach 4.8 kDa protein was done using a nylon membrane (Zeta probe, Biorad) and a conventional tank-type blot cell (Biorad) according to [3].

For protein sequencing, proteins were transferred onto a polyvinylidene difluoride membrane (Immobilon, Millipore) using a semidry-type electroblotter (Atto, Japan) according to the method described in [4]. The transferred proteins were stained with 0.1% (w/v) Amido black 10B (Biorad) in 50% (v/v) methanol and 10% (v/v) acetic acid for 1 min and destain-

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/88/$3.50

ed with distilled water. The stained band corresponding to the 4.8 kDa protein was cut out, treated with 0.6 N HCl for 24 h at 25°C and then sequenced on an Applied Biosystems model 477A protein sequencer equipped with a model 120A PTH analyzer [5].

3. RESULTS

As already shown in [3], a 4.8 kDa component is nearly quantitatively associated with the PS II reaction center complex. Western blotting analysis of various PS II membrane fragments using an antiserum raised against the spinach 4.8 kDa component showed that the antiserum immunoreacts with pea and wheat components almost the same as with spinach component (fig.1). This suggests that the 4.8 kDa component is highly conserved and ubiquitous in PS II among at least higher plants.

When reaction center proteins were treated with fluorescamine to label free amino moieties of the N-terminus, only the large subunit of cytochrome

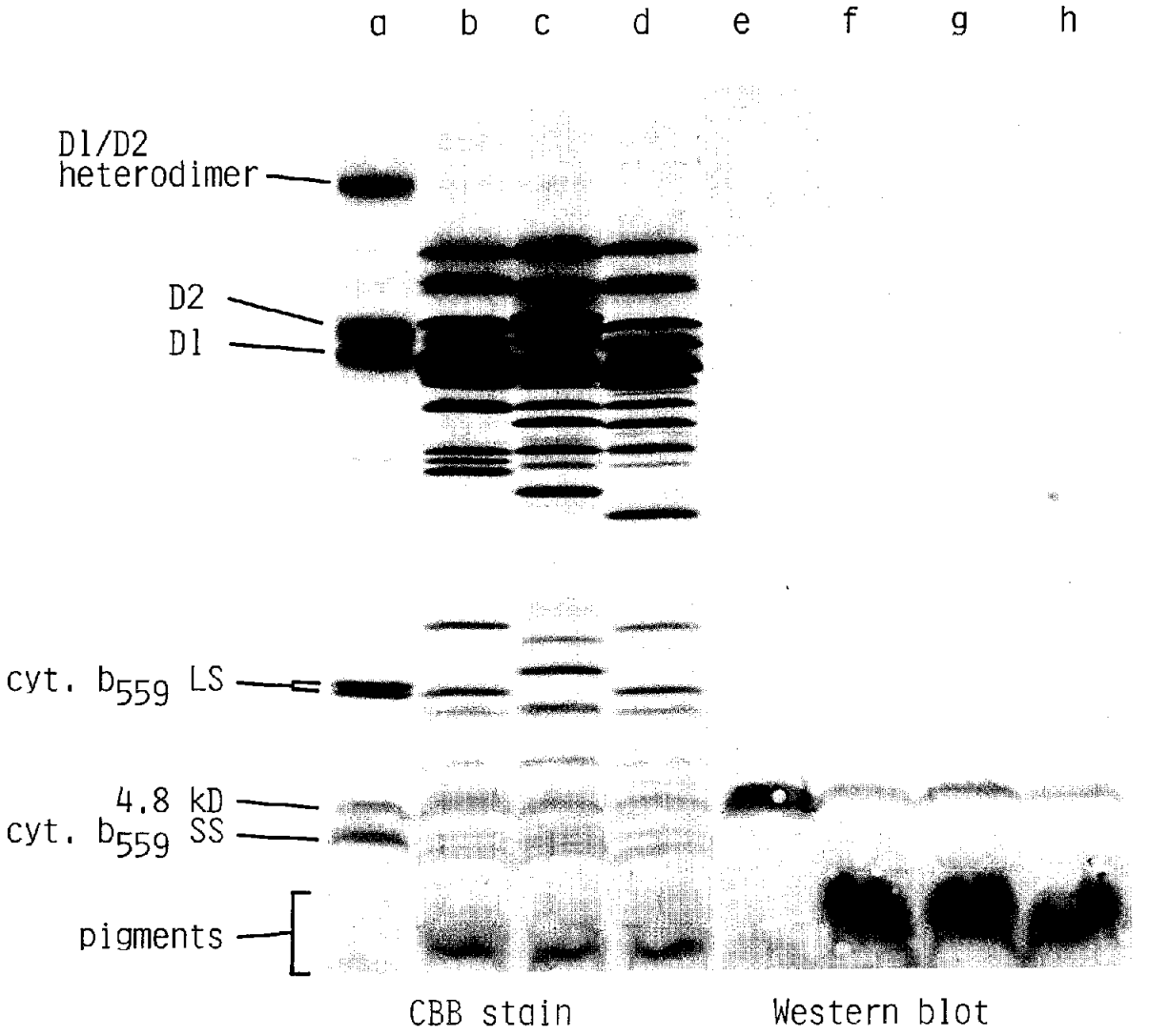

Fig.1. Western blotting analysis using antiserum raised against the spinach 4.8 kDa reaction center component. The reaction center complex from spinach (lanes a,e) was pretreated with fluorescamine. PS II membrane fragments from spinach (b,f), wheat (c,g) and pea (d,h). Note that the doublet band of cytochrome *b*-559 large subunit in lane a is due to partial labeling with fluorescamine. The labeled fluorescent band corresponds to the upper part of the stained doublet (not shown).

b-559 was fluorescently labeled (not shown). This fluorescent band migrated a little bit more slowly than the nonlabeled band, resulting in a doublet band in Coomassie-stained profile (fig.1, lane a). The lack of labeling in the 4.8 kDa component suggested N-terminal blocking. In accordance with this, protein sequencing of the 4.8 kDa component without acid treatment resulted in no amino acid signals like the small subunit of cytochrome *b*-559 [6]. After acid treatment which is presumed to hydrolyze formyl amino bond, the N-terminus of the protein was disclosed and up to 20 amino acid residues except for the third residue were successfully determined as Met-Leu-(Thr?)-Leu-Lys-Leu-Phe-Val-Tyr-Thr-Val-Val-Ile-Phe-Phe-Val-Ser-Leu-Phe-Ile.

A library of all the possible reading frames consisting of longer than 10 amino acid frames sandwiched between a pair of termination codons was constructed by computation out of the chloroplast DNA sequence of tobacco [7] and liverwort [8], in which not only the identified genes and ORFs but also all the possible split genes were listed up. Computer-assisted homology search of the partial amino acid sequence of spinach 4.8 kDa component in the library of tobacco and liverwort DNA revealed the existence of a unique reading frame corresponding to the 4.8 kDa sequence in both chloroplast genomes. Since a putative ribosome-binding site, GGAG, was found at position −12 upstream of the ATG initiation codon corresponding to the first Met residue of the component, we conclude that the 4.8 kDa component is encoded by an ORF which starts with this ATG initiation codon and ends with the TAA termination codon at position 109 (boxed as ORF36 in fig.2). This gene encodes 36 amino acid residues and is almost identical between tobacco and liverwort. The gene is located 300–400 bp downstream from the *psb*K gene [9] and surrounded by genes for $tRNA^{Ser}$ and $tRNA^{Gln}$ on the complementary strand in both plants.

The determined partial sequence of spinach 4.8 kDa component completely matched the

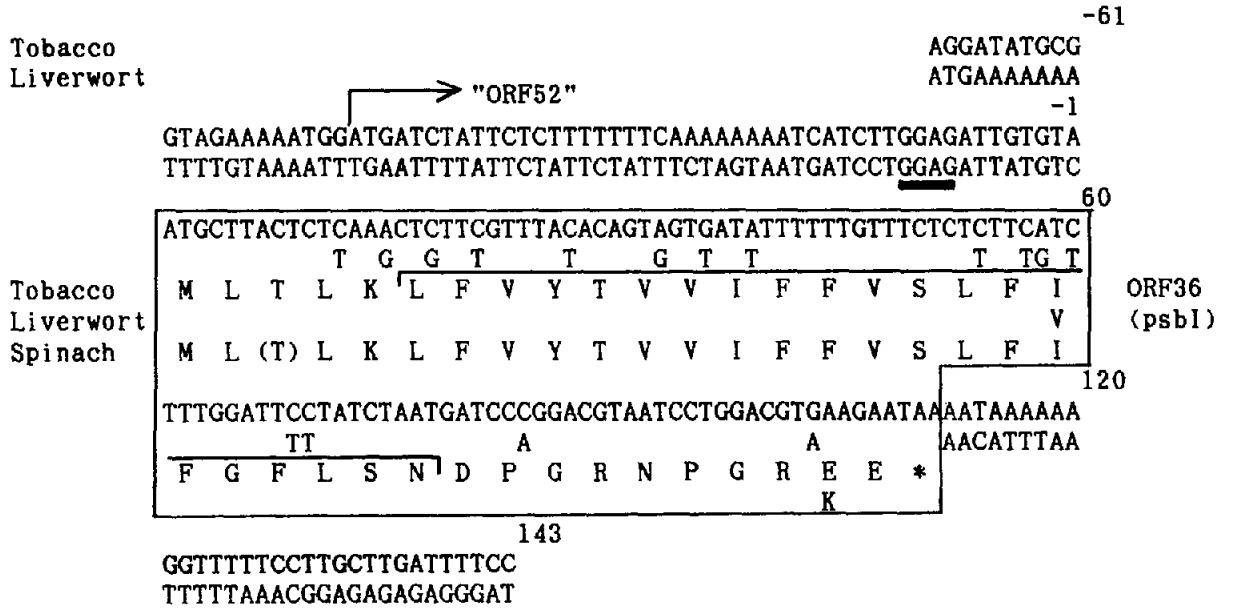

Fig.2. The nucleotide sequences of ORF36 (*psb*I) encoding the 4.8 kDa reaction center component and its flanking region in chloroplast DNA from tobacco (*Nicotiana tabacum* var. Bright Yellow 4) [7] and liverwort (*Marchantia polymorpha*) [8]. Numbering starts at the ATG codon at which translation is probably initiated. Corresponding amino acids deduced from the tobacco sequence and the N-terminal sequence (20 residues) determined for spinach 4.8 kDa component are indicated below respective codons. Nucleotide differences in coding region between tobacco and liverwort and any resulting amino acid differences are shown below the tobacco sequence. The identified gene is boxed. Possible ribosome-binding site is underlined. Hydrophobic region which may span the membrane is indicated by a horizontal bracket. The initiation codon proposed for ORF52, formerly designated as *psb*I of tobacco [13,14], is also shown.

deduced amino acid sequence of the tobacco gene (fig.2). In liverwort, two amino acid residues, 20th Ile and 35th Glu, were replaced by Val and Lys, respectively. In other words, the gene shows 94% homology at amino acid level and 86% homology at nucleotide level between tobacco and liverwort. Amino acid composition of the 4.8 kDa component deduced from the tobacco and liverwort genes revealed the absence of Ala, Cys, Trp, His and Gln. Net charge of the protein estimated as the sum of acidic and basic amino acids is 0 for tobacco and +2 for liverwort. Calculated molecular mass of the gene product is 4168 Da for tobacco and 4153 Da for liverwort. These values are slightly lower than the apparent molecular mass of 4.8 kDa estimated from the mobility in our SDS-PAGE system [3]. It is also of note that the 4.8 kDa component consisting of 36 amino acid residues migrates slower than the small subunit of cytochrome *b*-559 consisting of 39 amino acid residues (fig.1). This may be due to differences in hydrophobicity or total net charge of the protein. No protein having homology with the 4.8 kDa component was found in the Protein Sequence Database (PIR release 17.0).

Fig.3 shows the hydropathy plot of the 4.8 kDa component deduced from the tobacco gene. Profile for the liverwort gene was almost identical (not shown). Obviously, the predicted product contains a single hydrophobic segment consisting of 21 amino acid residues which may span the thylakoid membrane.

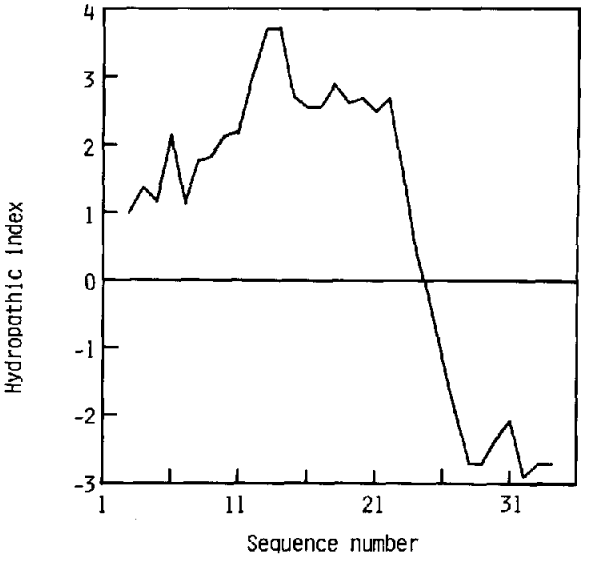

Fig.3. Hydropathy plot of the 4.8 kDa reaction center component deduced from the tobacco gene. Hydropathy was calculated with a 5-point window according to Kyte and Doolittle [15].

4. DISCUSSION

The present study revealed that the 4.8 kDa component of the PS II reaction center complex is encoded by a gene in the chloroplast genome of tobacco and liverwort. Therefore, all the components of the PS II reaction center complex (D1 protein, D2 protein, the large and small subunits of cytochrome *b*-559 and the 4.8 kDa protein) are encoded by the chloroplast genome. Likewise, all of these are hydrophobic proteins, having at least one membrane-spanning segment. The coding region of the chloroplast-encoded PS II genes are compared between tobacco and liverwort at the nucleotide and deduced amino acid level (table 1). *psb*A and *psb*D are highly conserved at both levels in agreement with the important role of their products (D1 and D2 proteins) in the reaction center complex [10]. The gene of the 4.8 kDa component as well as *psb*B, *psb*C, *psb*E and *psb*F are also highly conserved, whereas *psb*G, *psb*H and *psb*K are relatively less conserved. Although the present model of the PS II reaction center consists of D1 and D2 proteins by analogy with the bacterial reaction center [10], *psb*B, *psb*C, *psb*E and *psb*F are also indispensable for PS II functioning and assembly as suggested by genetic inactivation experiments [11,12]. Although we have no conclusive data at present to indicate the presence of a homologous 4.8 kDa component in cyanobacteria, a similar low-molecular-mass protein exists in the cyanobacterial reaction center complex (Koike, H., Ikeuchi, M. and Inoue, Y., unpublished). The cyanobacterial gene, if present, may be a candidate for genetic inactivation to provide information about the role of this component.

In our previous paper [3], the subunit stoichiometry of cytochrome *b*-559 large subunit : small subunit : the 4.8 kDa component in spinach and wheat reaction center complex was determined to be 1 : 1 : 0.75–0.79, based on the apparent molecular mass of 4.8 kDa estimated from the mobility on SDS gels. If we adopt the molecular mass calculated from the gene sequence,

Table 1

Homology of PS II genes between tobacco and liverwort

	Amino acid level (%)	Nucleotide level (%)
*psb*A (D1 protein)	96.9	87.6
*psb*B (CP47 apoprotein)	91.5	84.4
*psb*C (CP43 apoprotein)	96.0	84.8
*psb*D (D2 protein)	96.9	85.0
*psb*E (cyt. *b*-559 LS)	88.0	83.3
*psb*F (cyt. *b*-559 SS)	97.4	85.8
*psb*G	63.0	65.6
*psb*H (10 kDa phosphoprotein)	66.7	73.9
*psb*I (4.8 kDa protein)	94.4	85.6
*psb*K[a]	81.6	82.5

[a] Mature protein region is compared

the stoichiometry should be 1:1:0.87–0.91, indicative of higher retention of the 4.8 kDa component in the reaction center complex. In view of this fact and the high homology mentioned above, we propose to include the 4.8 kDa protein as an integral component in the model of PS II complex of higher plants, although its function is not clarified yet.

In the tobacco chloroplast genome, an ORF52, which starts 48 bp upstream of the initiation codon of ORF36 and shares the common termination codon (fig.2), has been described as a hypothetical gene based on the presence of a ribosome-binding site [13]. Expression of ORF52 and involvement of its product in the PS II were proved by Western blotting analysis using the antibody raised against a synthetic peptide corresponding to the C-terminal region of the putative product. Based on these, ORF52 has been named as *psb*I (Katoh, K., Sayre, R.T. and Bogorad, L., personal communication, and [14]). However, we favor that the 4.8 kDa reaction center component is encoded by a gene corresponding to ORF36 rather than ORF52 for the following reasons: (i) the first ATG codon of ORF36 is preceded by a putative ribosome-binding site, GGAG, in both tobacco and liverwort; (ii) comparison of the DNA sequence between tobacco and liverwort shows very high homology throughout the entire coding region of ORF36 and the GGAG region, but low homology in upstream region from this GGAG (see fig.2); (iii) the presumed initiation codon of tobacco ORF52 is absent in liverwort DNA; (iv) SDS-PAGE analysis of the PS II reaction center complex showed a single band in 4.8 kDa region corresponding to the size of ORF36 but no band around the region of 6 to 7 kDa deduced from the size of ORF52 (fig.1). These eliminate the possibility of processing the ORF52 product. We propose to revise the definition of *psb*I from ORF52 to ORF36 as a gene encoding the 4.8 kDa reaction center component.

As already shown for the tobacco chloroplast genome [9], *psb*I is located about 380 bp downstream of *psb*K. Similar tandem arrangement of *psb*I and *psb*K genes (ORF55 in [8]) is found in liverwort chloroplast genome. Northern hybridization of tobacco transcripts with a *psb*K probe revealed that its mRNA has a size of 1.3 kbp [9]. A mRNA with the same size was also detected by the use of the *psb*I probe (Sugiura, M., personal communication). In addition, a procaryote promoter-like base sequence was found upstream of *psb*K (ORF1 in [13]) but not between *psb*K and *psb*I (ORF2 in [13]). Therefore, it is very likely that *psb*I and *psb*K genes are co-transcribed as an operon.

Acknowledgements: We thank Dr M. Sugiura (Nagoya University) for valuable information and discussion, Dr T. Yasunaga and Mr K. Mamada (RIKEN) for computer calculation and Drs K. Narahashi and K. Takio (RIKEN) for helpful instructions for protein sequencing. This work was supported by a Grant on Solar Energy Conversion by Means of Photosynthesis awarded to The Institute of Physical and Chemical Research (RIKEN) by the Science and Technology Agency of Japan and partly by a Grant on Frontier Research Program at RIKEN awarded by Science and Technology Agency of Japan.

REFERENCES

[1] Nanba, O. and Satoh, K. (1987) Proc. Natl. Acad. Sci. USA 84, 109–112.
[2] Sugiura, M. (1987) Bot. Mag. Tokyo 100, 407–436.
[3] Ikeuchi, M. and Inoue, Y. (1988) Plant Cell Physiol. 29, 1233–1239.
[4] Ikeuchi, M. and Inoue, Y., submitted.
[5] Matsudaira, P. (1987) J. Biol. Chem. 262, 10035–10038.
[6] Widger, W.R., Cramer, W.A., Hermodson, M. and Herrmann, R.G. (1985) FEBS Lett. 191, 186–190.
[7] Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B.Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H. and Sugiura, M. (1986) EMBO J. 5, 2043–2049.

[8] Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H. and Ozeki, H. (1986) Nature 322, 572–574.
[9] Murata, N., Miyao, M., Hayashida, N., Hidaka, T. and Sugiura, M. (1988) FEBS Lett. 235, 283–288.
[10] Trebst, A. (1986) Z. Naturforsch. 41c, 240–245.
[11] Vermaas, W.F.J., Ikeuchi, M. and Inoue, Y. (1988) Photosynth. Res. 17, 97–113.
[12] Pakrasi, H.B., Williams, J.G.K. and Arntzen, C.J. (1988) EMBO J. 7, 325–332.
[13] Deno, H. and Sugiura, M. (1983) Nucleic Acids Res. 11, 8407–8414.
[14] Sugiura, M., Shinozaki, K., Tanaka, M., Hayashida, N., Wakasugi, T., Matsubayashi, T., Ohto, C., Torazawa, K., Meng, B.Y., Hidaka, T. and Zaita, N. (1987) in: Plant Molecular Biology (Von Wettstein, D. and Chua, N.-H. eds) pp.65–76, Plenum, New York.
[15] Kyte, J. and Doolittle, R.F. (1982) J. Mol. Biol. 157, 105–132.