Receptor-Binding Affinities of Human Epidermal Growth Factor Variants Having Unnatural Amino Acid Residues in Position 23†

Hiroshi Koide,[‡,§] Shigeyuki Yokoyama,[*,‡] Yasushi Katayama,[‡,∥] Yutaka Muto,[‡] Takanori Kigawa,[‡] Toshiyuki Kohno,[‡,⊥] Hiromi Takusari,[‡,#] Mitsuko Oishi,[‡,#] Seizo Takahashi,[#] Ken-ichi Tsukumo,[▽] Tetsuyuki Sasaki,[▽] Tetsuo Miyake,[▽] Tohru Fuwa,[▽] Gota Kawai,[∥,○] and Tatsuo Miyazawa[△,∞]

Department of Biophysics and Biochemistry, School of Science, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113, Japan, Faculty of Engineering, Yokohama National University, Tokiwadai, Hodogaya-ku, Yokohama 240, Japan, Protein Engineering Research Institute, Furuedai, Suita, Osaka 565, Japan, Department of Chemistry, Japan Women's University, Mejirodai, Bunkyo-ku, Tokyo 112, Japan, and The Central Research Laboratories, Wakunaga Pharmaceutical Co., Ltd., Koda-cho, Takata-gun, Hiroshima 729-64, Japan

Received January 3, 1994; Revised Manuscript Received April 5, 1994[⊛]

ABSTRACT: Ile-23 of human epidermal growth factor (hEGF) has been indicated, by mutagenesis and NMR studies, to be directly recognized by the receptor. In the present study, an unnatural phenylalanine analog, either 2-azaphenylalanine (2aF), 3-azaphenylalanine (3aF), 4-azaphenylalanine (4aF), or 4-fluorophenylalanine (4fF), was incorporated by an *in vivo* protein synthesis system into position 23 of [Phe[23]]hEGF, which retains appreciable receptor-binding affintiy (about 20% of the wild type). We compared the receptor-binding affinities of the variants with that of [Phe[23]]hEGF and found that substitution of Phe-23 with 2aF or 3aF raised the affinity, while substitution with 4aF or 4fF remarkably reduced the affinity. The tertiary structure of [Phe[23]]hEGF was not significantly affected by the substitution of Phe-23 with 2aF, as shown by the two-dimensional nuclear magnetic resonance analysis. In addition, the substitution of residue 23 with His or Tyr produced an hEGF variant with a slightly higher receptor-binding affinity than that of [Phe[23]]hEGF. Our results suggest that the receptor has an asymmetric hydrophobic pocket for recognition of the side chain in position 23 of hEGF. Furthermore, on the receptor surface, this pocket seems to be adjacent to a less hydrophobic region with a hydrogen-bond acceptor and donor. Thus, the use of unnatural amino acids in addition to the 20 natural ones allows analyses of the structure–function relationship of a protein at a higher resolution than conventional site-directed substitution by only natural amino acid residues.

In protein engineering, site-directed mutagenesis is an indispensable technique by which one amino acid residue is replaced by another. To establish the structure–function relationships of a protein at atomic resolution, it is sometimes necessary to introduce drastic changes, and sometimes fine changes, in the structure of the amino acid residue under study. However, if the substitution is restricted to only the 20 natural amino acids, the changes in structure may be either more drastic or more minute than desired. This limitation in protein engineering may be overcome by preparing proteins bearing unnatural amino acid residues in addition to the 20 natural ones (Koide et al., 1988; Bain et al., 1989; Noren et al., 1989). Such a protein is referred to as an "alloprotein" (Koide et al., 1988).

One method for the preparation of alloproteins is *in vitro* protein synthesis. In fact, *in vitro* systems for site-directed incorporation of unnatural amino acids into proteins have been developed by the use of chemical methods for charging nonsense-suppressor tRNAs with unnatural amino acids (Heckler et al., 1984; Bain et al., 1989, 1992; Noren et al., 1989; Ellman et al., 1991, 1992; Mendel et al., 1992; Chung et al., 1993; Judice et al., 1993). Thus, the *in vitro* systems allow the incorporation of unnatural amino acids that markedly differ in chemical structure from the 20 natural amino acids. Examples include an unnatural amino acid with a photocleavable site (Mendel et al., 1991) and unnatural amino acids with novel backbone structures (Ellman et al., 1992). However, it is difficult, at present, to construct a high-yield *in vitro* system with the chemical aminoacylation method, although an *in vitro* system with a high yield for ordinary protein synthesis has been developed (Spirin et al., 1988; Kigawa & Yokoyama, 1991).

As compared to the *in vitro* protein synthesis systems, *in vivo* systems for protein production exhibit much higher yields in general. When the *in vivo* systems are used for the preparation of alloproteins, unnatural amino acids must be misrecognized and charged to tRNAs by aminoacyl-tRNA synthetases. This does not allow the incorporation of unnatural

† This work was supported in part by Grants-in-Aid for Scientific Research on Priority Areas 01656002, 02238102, 02250106, 03222102, and 03236106 from the Ministry of Education, Science and Culture of Japan and a Grant for "Biodesign Research Program" from the Institute of Physical and Chemical Research (RIKEN) to S.Y.

* To whom correspondence should be addressed: Department of Biophysics and Biochemistry, School of Science, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113, Japan. Telephone number: +81-3-3812-1805. Fax number: +81-3-5684-2394.

‡ The University of Tokyo.

§ Present address: Faculty of Bioscience and Biotechnology, Tokyo Institute of Technology, Midori-ku, Yokohama 227, Japan.

∥ Yokohama National University.

⊥ Present address: Mitsubishi Kasei Institute of Life Science, Minamiooya, Machida-shi, Tokyo 194, Japan.

Japan Women's University.

▽ Wakunaga Pharmaceutical Co., Ltd.

○ Present address: Department of Industrial Chemistry, Faculty of Engineering, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113, Japan.

△ Protein Engineering Research Institute.

∞ Deceased.

⊛ Abstract published in *Advance ACS Abstracts,* May 15, 1994.

0006-2960/94/0433-7470$04.50/0

amino acids that drastically differ from any of the natural ones. However, we are still able to incorporate an unnatural amino acid with a remarkably different chemical structure from those of any of the natural ones, when the enzyme-bound conformation of the unnatural amino acid is similar to that of the intrinsic substrate (Kohno et al., 1990). This may be the reason for the incorporation of a wide variety of unnatural amino acids into proteins *in vivo* (Richmond, 1962; Wilson & Hatfield, 1984). Note that incorporation of unnatural amino acid analogs with fine variations in structure into proteins is useful for studying structure–function relationships at atomic resolution. In principle, such unnatural amino acids are incorporable into proteins *in vivo* to a high yield, sufficient for structural analyses.

As for the *in vivo* synthesis of alloproteins, one of the most serious problems is the toxicity of the unnatural amino acids to the host cells, such as *Escherichia coli* (Richmond, 1962; Fowden et al., 1967). We have already developed a strategy to avoid the cytotoxicity of unnatural amino acids in *in vivo* alloprotein synthesis and have incorporated unnatural amino acids into human epidermal growth factor (hEGF[1]): Met-21 of hEGF has been replaced by norleucine, yielding an oxidation-resistant hEGF that has receptor-binding and mitogenic activities equally potent to those of the wild-type hEGF (Koide et al., 1988). Norleucine has also been substituted for methionine in some other proteins (Gilles et al., 1988; Tsai et al., 1988; Bogosian et al., 1989). Furthermore, since hEGF has no Phe residues, the Tyr codon in position 22 or 29 of hEGF has been replaced by a Phe codon in the hEGF gene, and then a phenylalanine analog, 4-fluorophenylalanine (4fF) has been incorporated into position 22 or 29 (Koide et al., 1990). Therefore, phenylalanine analogs should also be incorporable site-specifically into other positions of hEGF.

The hydrophobic side chain of Ile-23 of hEGF has been shown to be directly recognized by the receptor (Matsunami et al., 1990; Koide et al., 1992a). The [Phe^{23}]hEGF retains appreciable receptor-binding affinity (about 20% of that of the wild type) (Koide et al., 1992a), although the sizes of the Ile and Phe side chains are significantly different. In the present study, a detailed analysis of the receptor binding of [Phe^{23}]hEGF has been undertaken by substituting the unique Phe residue of [Phe^{23}]hEGF with a variety of phenylalanine analogs, 2-azaphenylalanine (2aF), 3-azaphenylalanine (3aF), 4-azaphenylalanine (4aF), and 4-fluorophenylalanine (4fF) (Figure 1). For comparison, changes in the side-chain structure of Phe-23 have also been introduced by replacement with other natural aromatic amino acid residues (His and Tyr).

MATERIALS AND METHODS

Materials. Escherichia coli strain JE8487 (φ_3Phe) was kindly provided by Dr. A. Nishimura. 3-Aza-L-phenylalanine was purchased from Chiba Chemicals (Chiba, Japan). 4-Fluoro-DL-phenylalanine was purchased from Sigma.

Construction of hEGF Variant Genes. Oligonucleotides for site-directed mutagenesis were synthesized as shown below using Cyclone Plus (MilliGen/Biosearch) (the target codons are underlined).

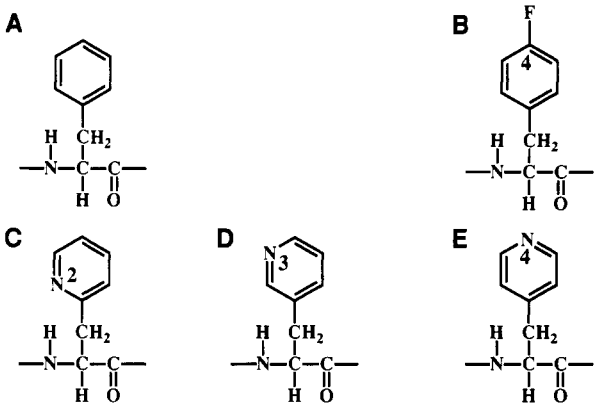

FIGURE 1: Chemical structures of (A) phenylalanine, (B) 4-fluorophenylalanine, (C) 2-azaphenylalanine, (D) 3-azaphenylalanine, and (E) 4-azaphenylalanine.

5′-TTGCATGTATTTTGAAGCTCT-3′ (21-mer)
(for Ile-23 → Phe)

5′-TGCATGTATCACGAAGCTCTG-3′ (21-mer)
(for Ile-23 → His)

5′-TGCATGTATTATGAAGCTCTG-3′ (21-mer)
(for Ile-23 → Tyr)

The fragment carrying the hEGF gene was prepared from plasmid pTA1522 (Oka et al., 1985) by digestion with *Eco*RI and *Sal*I and was transferred into the multiple cloning site of M13mp19 DNA. Mutagenesis was performed using the Muta-Gene kit (Bio-Rad). The codon changes of the hEGF gene were confirmed by sequencing the gene on M13mp19 DNA. The *Eco*RI–*Sal*I fragment of each mutant clone from M13mp19 was inserted into plasmid pTA1522. Thus, plasmids pTA152-23F, pTA152-23H, and pTA152-23Y, which carry the genes encoding [Phe^{23}]hEGF, [His^{23}]hEGF, and [Tyr^{23}]hEGF, respectively, were constructed.

Preparation of [His^{23}]hEGF and [Tyr^{23}]hEGF. E. coli strain YK537 (F$^-$ *leuB6 thi hsdR hsdM lacY rpsL20 galK2 ara-14 xyl-5 mtl-1 supE44 endI$^-$ phoA8 recA1*) harboring plasmid pTA152-23H or pTA152-23Y was cultivated overnight at 37 °C in high-phosphate medium (TG+20) (Oka et al., 1985) containing leucine at a concentration of 50 μg/mL. The cells were collected by centrifugation, suspended in the same volume of low-phosphate medium (TG+1) (Oka et al., 1985) containing leucine at a concentration of 50 μg/mL, and cultivated for 6 h at 37 °C. Then, the cells were harvested and the solution containing the hEGF variant was recovered by osmotic shock and lyophilized (Koide et al., 1988).

Chemical Synthesis. 2-Aza-L-phenylalanine and 4-aza-L-phenylalanine were synthesized from 2- and 4-(chloromethyl)pyridine, respectively, by the procedure described (Albertson & Archer, 1945). Each of 2-, 3-, and 4-azaphenylalanines was converted to a blocked derivative of *N*-acetylazaphenylalanine methylamide by the procedures described (Dewitt & Ingersoll, 1951; Woodward et al., 1961).

pH Titration. In order to determine the pK_a values of the pyridyl groups of the azaphenylalanines, the pH dependence (pH 3–11) of the ultraviolet absorbance of the azaphenylalanines and *N*-acetylazaphenylalanine methylamides was measured. The wavelength of maximum absorbance was chosen as 255 nm for 4-azaphenylalanine and *N*-acetyl-4-

[1] Abbreviations: 2aF, 2-azaphenylalanine; 3aF, 3-azaphenylalanine; 4aF, 4-azaphenylalanine; CAT, chloramphenicol acetyltransferase; 2D HOHAHA, two-dimensional homonuclear Hartmann–Hahn spectroscopy; 4fF, 4-fluorophenylalanine; hEGF, human epidermal growth factor; HPLC, high-performance liquid chromatography; IC_{50}, the concentration of a competitor which inhibits binding of [^{125}I]mEGF to the receptor down to 50%; [^{125}I]mEGF, ^{125}I-labeled mouse epidermal growth factor.

azaphenylalanine methylamide, 260 nm for 2-azaphenylalanine, 3-azaphenylalanine, and *N*-acetyl-3-azaphenylalanine methylamide, and 265 nm for *N*-acetyl-2-azaphenylalanine methylamide.

Cytotoxicities of Phenylalanine Analogs. *E. coli* strain W3110 was cultured in LB medium. The overnight culture (75 μL) was inoculated into 5 mL of TG+20 medium including a phenylalanine analog at a concentration of 400 μg/mL. Growth of W3110 cells was monitored by measuring the optical density at 600 nm.

In Vitro Incorporation of Unnatural Amino Acids. Incorporation of azaphenylalanines into proteins was examined with an *in vitro* system (Kohno et al., 1990). The reaction mixture contained 83 μg/mL pBR325 carrying the gene for chloramphenicol acetyltransferase (CAT), 1.3 μM [^{35}S]-methionine (Amersham, specific activity 3.7 MBq/μg), and natural amino acids (except phenylalanine) each at a concentration of about 200 μM, in addition to an *E. coli* S-30 extract and a supplement solution from a prokaryotic DNA-directed translation kit (Amersham). The reaction mixture was incubated for 1 h, in the absence or the presence of phenylalanine or a phenylalanine analog (5.5 mM). Then, the methionine chase solution was added to the reaction mixture, and the incubation was continued for 5 min. Finally, the reaction mixture was subjected to sodium dodecyl sulfate–polyacrylamide gel electrophoresis.

In Vivo Preparation of hEGFs Containing Phenylalanine Analogs in Position 23. *E. coli* strain JE8487 harboring plasmid pTA152-23F was cultivated overnight at 37 °C in high-phosphate medium (TG+20) containing phenylalanine at a concentration of 50 μg/mL. Then, by the same preparation method as used for [His^{23}]hEGF and [Tyr^{23}]-hEGF, the cells were transferred to low-phosphate medium (TG+1) containing a phenylalanine analog at a concentration of 50 μg/mL, but lacking phenylalanine. After cultivation for 6 h at 37 °C, the cells were harvested and the solution containing the hEGF variant was recovered by the osmotic shock procedure and lyophilized.

Purification of hEGF Variants. The lyophilizate containing the hEGF variant was dissolved in 25 mM ammonium acetate buffer (pH 5.8) and loaded on a column (2.6 × 90 cm) of Sephadex G-50 (particle size 50–150 μm, Pharmacia) equilibrated with the same buffer. Isocratic elution was performed with the same buffer at a flow rate of 1 mL/min. Fractions containing the hEGF variant were pooled and lyophilized. The lyophilizate was dissolved in 20% CH_3CN/0.1% trifluoroacetate (vol/vol) and subjected to reverse-phase HPLC (LC-4A system; Shimadzu, Kyoto) with a YMC-Pack ODS-A column (0.46 × 15 cm) (YMC, Kyoto). Elution was carried out with a linear gradient of 20–40% CH_3CN at a flow rate of 1 mL/min. The peak of the hEGF variant was collected and lyophilized. The lyophilizate was dissolved in 50 mM ammonium acetate buffer (pH 6.0) and subjected to ion-exchange HPLC (LC-6A system; Shimadzu) with a DEAE-5PW column (0.75 × 7.5 cm) (Tosoh, Tokyo). Elution was performed with a linear gradient of 50–350 mM ammonium acetate buffer (pH 6.0) at a rate of 0.5 mL/min. The purified preparations were subjected to amino acid composition analyses with an L-8500 system (Hitachi Ltd., Tokyo) in order to confirm the incorporation of the phenylalanine analogs and, at the same time, to accurately determine the concentrations of the hEGF alloproteins.

Receptor-Binding Assay. The receptor-binding affinities of the hEGF variants were measured by radioreceptor competition assay (Koide et al., 1992a). Human KB cells (Dainippon Pharmaceutical Co., Osaka) were suspended in a binding buffer (Dulbecco's modified Eagle's medium containing 0.1% bovine serum albumin and 20 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid at pH 7.4) at 2×10^6 cells/mL. The binding buffer, containing ^{125}I-labeled mouse EGF ([^{125}I]mEGF) (Amersham, specific activity 3.7 MBq/μg) at 2×10^5 cpm/mL, was added to the same volume of the binding buffer containing either the wild-type or variant hEGF at various concentrations (approximately 0.1–20 nM). Then, this solution was mixed with the same volume of the KB cell suspension. After incubation at 4 °C for 3 h, the unbound [^{125}I]mEGF was washed away with cold phosphate-buffered saline. The radioactivity in the resultant cell pellet was measured with a γ spectrometer, Model ARC-301 (Aloka, Tokyo). The radioactivities in the presence of the hEGF variants at various concentrations were corrected for the effect of nonspecific binding of [^{125}I]mEGF to the KB cells. The correction was obtained from the radioactivity of the cell pellet in the presence of a large excess of hEGF (640 nM); the magnitude of the correction was smaller than 5% of the radioactivity in the absence of hEGF.

NMR Measurements. The purified preparations of [Phe^{23}]hEGF and [2aF^{23}]hEGF were dissolved in 99.85% ^2H_2O, and the pH was adjusted to 2.5. Proton NMR spectra (400 MHz) were recorded on a Bruker AM-400 spectrometer at a probe temperature of 25 °C. Sodium 2,2-dimethyl-2-silapentane-5-sulfonate was used as the internal standard for the proton chemical shifts. Two-dimensional homonuclear Hartmann–Hahn spectroscopy (2D HOHAHA; Bax & Davis, 1985) experiments were performed with a mixing time of 45 ms. Ninety-six free induction decays of 2K data points in the t_2 domain were accumulated for each of the 460 t_1 increments, using time proportional phase incrementation (Marion & Wüthrich, 1983). The spectra of 1K × 2K points were obtained with zero-filling in the t_1 domain and resolution enhancement with Gaussian window in both the t_1 and t_2 domains prior to two-dimensional Fourier transformation.

RESULTS

Protonation of Side Chains of Azaphenylalanines. Azaphenylalanines, the unnatural phenylalanine analogs incorporated into hEGF in the present study, have side-chain pyridyl groups that are protonated in acidic solutions. From the pH dependence of the ultraviolet absorbance (data not shown), the side-chain pK_a values of the 2-, 3-, and 4-azaphenylalanines were determined to be 3.9, 4.5, and 4.8, respectively. However, the protonation of the pyridy groups in these phenylalanine analogs may be affected by the ionization of the free amino and carboxyl groups, particularly in 2-azaphenylalanine. Therefore, the blocked azaphenylalanines, namely, the *N*-acetylazaphenylalanine methylamides, are more useful as models for azaphenylalanine residues in proteins. From the pH dependence of the ultraviolet absorbance (data not shown), the pK_a values of the side chains of the N-acetylated 2-, 3-, and 4-azaphenylalanine methylamides were determined to be 5.1, 4.6, and 5.1, respectively. As expected from the location of the nitrogen atom in the pyridyl groups, the pK_a change on the terminal blocking is much larger for 2-azaphenylalanine (+1.2) than for 3-azaphenylalanine (+0.1) and 4-azaphenylalanine (+0.3). Thus, for the blocked analogs of the three azaphenylalanines, the pK_a values were all found to be nearly equal to 5, which is consequently expected to be the case for azaphenylalanine residues in polypeptide chains.

Cytotoxicity of Phenylalanine Analogs. Unnatural amino acids are usually cytotoxic and disturb cell growth (Richmond,

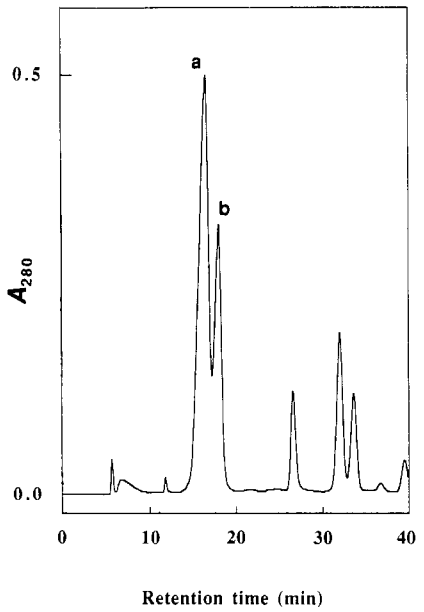

FIGURE 2: Elution profile of ion-exchange HPLC for the separation of [4fF^{23}]hEGF (peak b) from [Phe^{23}]hEGF (peak a). Chromatography was performed by an isocratic elution with 100 mM ammonium acetate buffer (pH 6.0) for the first 2 min and then by a linear-gradient elution with 100 → 400 mM ammonium acetate buffer (pH 6.0) for 40 min.

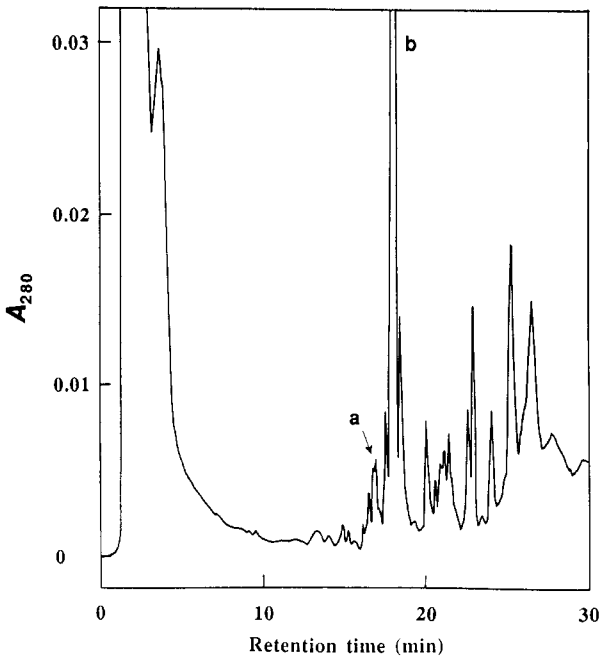

FIGURE 3: Elution profile of reverse-phase HPLC for the purification of [3aF^{23}]hEGF (peak a) from [Phe^{23}]hEGF (peak b). The elution was performed with a linear gradient of 20 → 40% CH_3CN containing 0.1% trifluoroacetate for 20 min.

1962; Fowden et al., 1967). Therefore, we examined the cytotoxic effects of the four phenylalanine analogs on the growth of *E. coli* strain W3110. Although this strain is not a phenylalanine auxotroph, the growth rate in the presence of any phenylalanine analog is reduced as compared to the rate in the absence of the analog (data not shown). Among the four analogs, 4-fluorophenylalanine was the most toxic, while 3-azaphenylalanine was the least toxic. It is likely that these analogs are incorporated into the proteins of the bacterial cells, thus disturbing the cellular physiology.

In Vitro Incorporation of Azaphenylalanines into Proteins. It has been found that 4-fluorophenylalanine is incorporated into proteins *in vivo* (Richmond, 1962; Wheatley, 1978; Koide et al., 1990). Such incorporation into proteins is probably responsible for its cytotoxicity. Therefore, in the present study, the incorporation of azaphenylalanines into a protein was preliminarily examined by an *in vitro* system, using plasmid pBR325 carrying the gene for chloramphenicol acetyltransferase (CAT) as a template. In the autoradiograms of synthesized proteins, a major band, due to CAT, was observed from the reaction mixture containing all 20 natural amino acids, while the CAT band was hardly detectable from the mixture lacking phenylalanine (data not shown). However, when azaphenylalanine, instead of phenylalanine, was added to the reaction mixture, we observed a major band corresponding to CAT (data not shown). The results indicated that the 2-, 3-, and 4-azaphenylalanines were incorporated into CAT in place of Phe and encouraged us to perform *in vivo* incorporation of azaphenylalanines into hEGF.

In Vivo Synthesis of [4fF^{23}]hEGF. As in the previous cases of [4fF^{22}]hEGF and [4fF^{29}]hEGF (Koide et al., 1990), a phenylalanine auxotroph, *E. coli* strain JE8487, was used as the host cell in the present study to enhance the incorporation of 4-fluorophenylalanine. In the purification with reverse-phase HPLC, [4fF^{23}]hEGF eluted together with [Phe^{23}]hEGF (data not shown). However, in the next step (ion-exchange HPLC) (Figure 2), [4fF^{23}]hEGF was eluted (peak b) later than [Phe^{23}]hEGF (peak a). The fractions containing [4fF^{23}]-hEGF were collected and subjected again to the same ion-exchange HPLC. The purity of the final preparation of [4fF^{23}]hEGF was found to be nearly 100% by the same ion-exchange HPLC. The final yield of [4fF^{23}]hEGF was 3 μg/L from the *E. coli* culture.

In Vivo Synthesis of hEGF with an Azaphenylalanine Residue. E. coli strain JE8487 was also used as the host for efficient incorporation of azaphenylalanine into hEGF *in vivo*. For the highest yield of the substituted protein, the concentration of azaphenylalanine in the low-phosphate medium was chosen to be 50 μg/mL, from the results of incorporation tests of azaphenylalaine at concentrations ranging from 10 to 200 μg/mL. The side chain of the azaphenylalanine residue has a pK_a value of about 5 and should therefore be protonated under the reverse-phase HPLC conditions (at about pH 2). In fact, in the reverse-phase HPLC step (Figure 3), the [3aF^{23}]-hEGF (peak a) was clearly separated from the [Phe^{23}]hEGF (peak b). In contrast, in the ion-exchange HPLC step (at pH 6.0), the [3aF^{23}]hEGF was not separated from the [Phe^{23}]-hEGF (data not shown), probably because the side chain of the 3aF residue was unprotonated at pH 6.0, as was the case for Phe. The final yields of [2aF^{23}]hEGF, [3aF^{23}]hEGF, and [4aF^{23}]hEGF were 2, 0.8, and 1 μg/L, respectively, from the *E. coli* culture, each with a purity of nearly 100% as determined by reverse-phase HPLC.

Amino Acid Composition of hEGF Variants with Phenylalanine Analogs. Site-specific incorporation of phenylalanine analogs into hEGF was confirmed by the amino acid composition analysis of the finally purified preparations. On an HPLC elution profile for the analysis of [4fF^{23}]hEGF, a peak due to 4-fluorophenylalanine is observed, with an area corresponding to one residue, while a peak due to phenylalanine is absent (data not shown). Moreover, the composition of amino acid residues other than 4fF^{23} (Table 1) was practically the same as that of [Phe^{23}]hEGF, indicating the substitution of 4-fluorophenylalanine for Phe-23. Thus, the cytotoxic 4-fluorophenylalanine was site-specifically incorporated into position 23, in the present study, as well as into positions 22

Table 1: Amino Acid Compositions[a] of [4fF^{23}]hEGF and [3aF^{23}]hEGF (Residues/Mole[b])

	[4fF^{23}]hEGF		[3aF^{23}]hEGF	
amino acid	predicted	analyzed	predicted	analyzed
Asx	7	6.9	7	7.0
Thr	0	0.0	0	0.1
Ser	3	2.5	3	2.6
Glx	5	5.2	5	5.3
Gly	4	4.0	4	4.1
Ala	2	2.0	2	2.1
Val	3	2.8	3	2.8
Met	1	1.1	1	1.1
Ile	1	1.0	1	1.1
Leu	5	5.0	5	5.0
Tyr	5	5.0	5	4.9
Lys	2	1.9	2	1.9
His	2	2.0	2	2.1
Arg	3	2.9	3	3.0
Pro	1	1.0	1	1.0
Phe	0	0.0	0	0.0
4fF	1	0.9	0	0.0
3aF	0	0.0	1	0.7

[a] The absorption coefficient of Phe was used. The Pro content was analyzed by the absorbance at 440 nm, although the contents of Trp and Cys were not determined. [b] Leu was used as the standard for normalization.

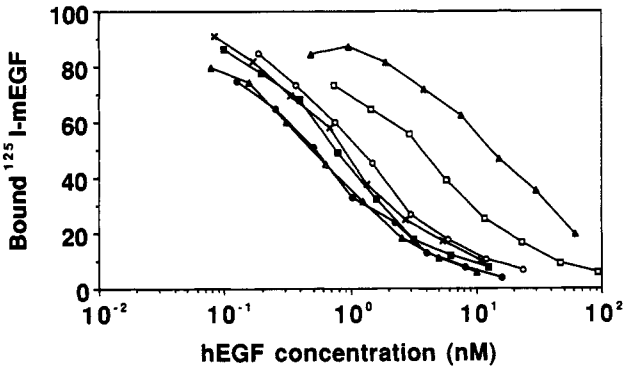

FIGURE 4: Competition receptor binding of [^{125}I]mEGF vs [Phe^{23}]-hEGF (○), [2aF^{23}]hEGF (●), [3aF^{23}]hEGF (△), [4aF^{23}]hEGF (▲), [4fF^{23}]hEGF (□), [His^{23}]hEGF (■), and [Tyr^{23}]hEGF (×). The data are representative of three separate experiments.

and 29 (Koide et al., 1990). Site-specific incorporation of 2-, 3-, and 4-azaphenylalanines in position 23 of hEGF was also confirmed by the analysis of the amino acid compositions of [2aF^{23}]hEGF, [3aF^{23}]hEGF (Table 1), and [4aF^{23}]hEGF, respectively.

Synthesis of [His^{23}]hEGF and [Tyr^{23}]hEGF. Mutations in the codon for Ile-23 (from ATT to CAC for His and to TAT for Tyr) were introduced into hEGF, and both [His^{23}]-hEGF and [Tyr^{23}]hEGF were obtained with nearly the same yields as that of the wild-type hEGF (about 0.3 mg/L from the *E. coli* culture).

Receptor-Binding Affinities of hEGF Variants. The purified preparations of [His^{23}]hEGF, [Tyr^{23}]hEGF, [2aF^{23}]-hEGF, [3aF^{23}]hEGF, [4aF^{23}]hEGF, and [4fF^{23}]hEGF were subjected to the radioreceptor competition assay with intact human KB cells (Koide et al., 1992a) (Figure 4). The assay was performed at 4 °C to prevent internalization of the complex of hEGF with the receptor. As shown in Figure 4, the concentrations of the hEGF variants at which the amount of cell-bound [^{125}I]mEGF was reduced to 50% of the maximum level (IC_{50}) were obtained. The binding affinities of the hEGF variants relative to the [Phe^{23}]hEGF were obtained from IC_{50}([Phe^{23}]hEGF)/IC_{50}(variant) (Figure 5).

The relative binding affinities of [His^{23}]hEGF (140%) and [Tyr^{23}]hEGF (140%) were slightly higher than that of [Phe^{23}]-

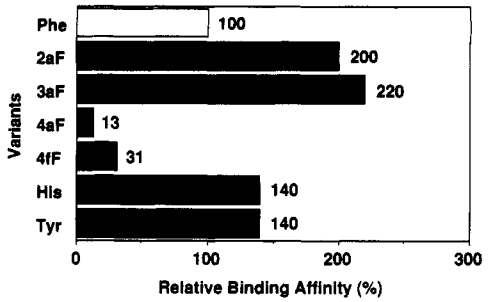

FIGURE 5: Receptor-binding affinities of hEGF variants (relative to [Phe^{23}]hEGF) obtained from IC_{50}([Phe^{23}]hEGF/IC_{50}(variant) in Figure 4.

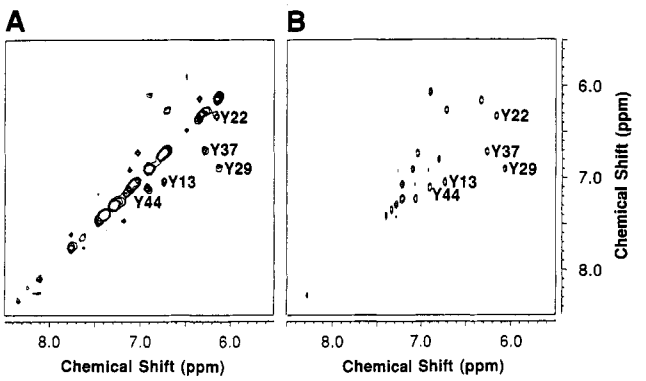

FIGURE 6: 2D HOHAHA spectra of (A) [2aF^{23}]hEGF and (B) [Phe^{23}]hEGF in the aromatic proton region.

hEGF (Figure 5). The [2aF^{23}]hEGF (200%) and the [3aF^{23}]-hEGF (220%) exhibited much higher affinities than the [Phe^{23}]hEGF. In contrast, the [4aF^{23}]hEGF (13%) and the [4fF^{23}]hEGF (31%) showed significantly lower affinities than the [Phe^{23}]hEGF. We performed this experiment three times independently and observed exactly the same tendency. Thus, the binding affinities of the hEGF variants in position 23 follow the order 2aF, 3aF > His, Tyr > Phe ≫ 4fF ≫ 4aF.

Conformation of [2aF^{23}]hEGF. In order to elucidate the conformational aspect of the enhancement of the receptor-binding affinity by the Phe → 2aF substitution, the tertiary structures of [2aF^{23}]hEGF and [Phe^{23}]hEGF in solution were compared by two-dimensional NMR spectroscopy. In the aromatic proton region of the 2D HOHAHA spectra, the cross peaks due to Tyr-13, Tyr-22, and Tyr-29 of [2aF^{23}]-hEGF (Figure 6A) were found at nearly the same chemical shifts as those of [Phe^{23}]hEGF (Figure 6B). The side chains of these Tyr residues are known to form a hydrophobic cluster, which maintains the tertiary structure of hEGF (Cooke et al., 1987, 1990). Thus, the Phe → 2aF substitution does not appreciably affect the tertiary structure.

DISCUSSION

Receptor-Binding Mechanisms of hEGF. The mechanisms of the receptor binding of hEGF have been studied extensively by site-directed mutagenesis. First, the importance of several amino acid residues, such as Arg-41 and Leu-47, in the carboxy-terminal domain was pointed out (Burgess et al., 1988; Ray et al., 1988; Engler et al., 1988, 1990; Moy et al., 1989; Dudgeon et al., 1990; Hommel et al., 1991). More recently, direct involvement in receptor recognition and binding has been indicated for several amino acid residues exposed on a wide area of one face of the major antiparallel β-sheet structure in the amino-terminal domain of hEGF: Ile-23 in a β-strand (Matsunami et al., 1990; Koide et al., 1992a), Ala-25 and

Leu-26 in a β-turn (Matsunami et al., 1990; Campion et al., 1990), and Ala-30 and Asn-32 on the other β-strand (Koide et al., 1992b; Campion et al., 1993). For position 23, the Ile residue of the wild-type hEGF has been substituted with a variety of natural amino acid residues (Matsunami et al., 1990; Koide et al., 1992a). As a result, it has been concluded that the receptor surface involved in binding with the major β-sheet of hEGF has a hydrophobic pocket for the side chain of Ile-23 (Koide et al., 1992a). This pocket for Ile-23 seems to be completely hydrophobic and in contact with the entire structure of the Ile-23 side chain (Koide et al., 1992a).

Receptor-Binding Affinities of hEGF Variants with Aromatic Residues in Position 23. If the Phe-23 side chain of [Phe^{23}]hEGF binds completely within the putative pocket for isoleucine, the substitution of Phe-23 with a less hydrophobic residue should reduce the receptor-binding affinity. However, the substitution of Phe-23 by 2-azaphenylalanine (2aF) and by 3-azaphenylalanine (3aF) was found in the present study to enhance the binding affinity up to 200% (Figure 5), although these azaphenylalanine residues are more hydrophilic than phenylalanine. This enhancement of the binding affinity could be due to changes in the direct interaction of the side chain in position 23 with the receptor, or in the tertiary structure around position 23 of the protein. In order to examine the possibility of a conformational change, the 2D HOHAHA spectra of [2aF^{23}]hEGF and [Phe^{23}]hEGF were compared (Figure 6). NMR analysis indicates that the Phe → 2aF substitution in position 23 has no appreciable effect upon the tertiary structure of [Phe^{23}]hEGF. This implies that the enhanced receptor-binding affinities of [2aF^{23}]hEGF and [3aF^{23}]hEGF, as compared with that of [Phe^{23}]hEGF, are due to favorable direct interactions of the pyridyl side chains of the unnatural amino acid residues, rather than to indirect conformational changes caused by the substitution. The receptor-binding affinity of [Phe^{23}]hEGF was also slightly enhanced by substitution of Phe-23 with His or Tyr, a more hydrophilic residue than Phe. In contrast, the replacement of Phe-23 by 4-azaphenylalanine (4aF) or 4-fluorophenylalanine (4fF) significantly reduced the receptor-binding affinity of [Phe^{23}]hEGF (Figure 5).

Interaction of the Aromatic Residue in Position 23 of hEGF Variants with the Receptor. On the basis of these data on the receptor-binding affinities of hEGF variants, including the allo-hEGFs with unnatural aromatic amino acid residues, we now speculate upon the receptor structure around the binding pocket for the isoleucine side chain of hEGF.

Under the conditions of the binding assay (at pH 7.4), the side-chain groups of azaphenylalanines (pK_a = *ca.* 5) and His [pK_a = 6.5–7.0 (Cantor & Schimmel, 1980)] are unprotonated. Therefore, the pyridyl groups of the 2aF and 3aF residues have proton-accepting nitrogen atoms. Correspondingly, the imidazole ring of the His residue has a proton-accepting nitrogen atom (and a proton-donating imino group), while the phenol ring of the Tyr residue has a hydroxyl group that serves as a proton acceptor (and/or as a proton donor). Therefore, the enhancement of the receptor-binding affinity upon substitution of Phe with 2aF, 3aF, Tyr, or His in position 23 of hEGF is probably due to the formation of a hydrogen bond between the hydrogen-acceptor group of the aromatic residue 23 of hEGF and some hydrogen-donor group near the putative hydrophobic pocket provided by the EGF receptor. In contrast to the Phe → Tyr substitution, which raises the receptor-binding affinity of [Phe^{23}]hEGF, the Phe → 4aF and Phe → 4fF substitutions significantly reduced the affinity. In the para position of their aromatic rings, the 4aF and 4fF residues have proton-accepting nitrogen and fluorine atoms, whereas the Tyr residue has a proton-donating/accepting hydroxyl group. These observations suggest that the EGF receptor has on its surface a proton acceptor in addition to the proton donor discussed above, both of which are located in close proximity to the putative hydrophobic pocket for binding the hydrophobic side chain of Ile-23 of the wild-type hEGF. Therefore, it is suggested that the large aromatic side chain introduced into position 23 of hEGF cannot be totally buried in the pocket for the isoleucine side chain, and therefore a part of the aromatic ring protrudes from this pocket toward the nearby hydrophilic portion of the surface of the receptor.

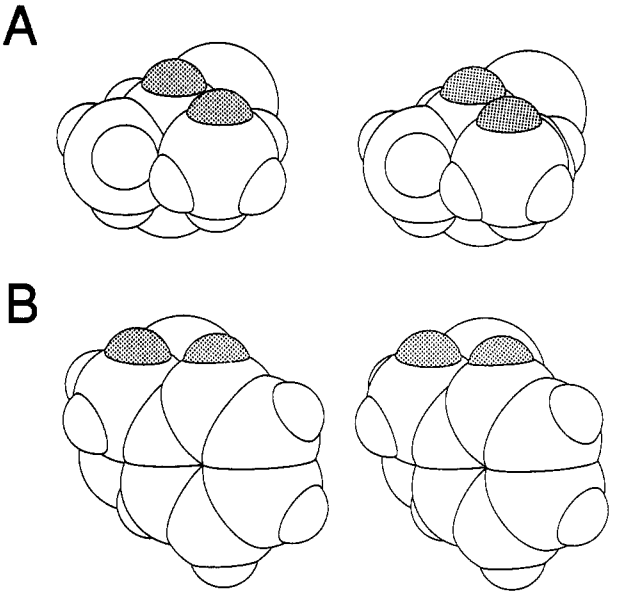

FIGURE 7: Models of (A) Ile-23 and (B) Phe-23 of hEGF, viewed from a direction perpendicular to the major β-sheet plane. The shaded hydrogens are supposed to be included in the hydrophobic "wall" and thereby brought into contact with the putative hydrophobic pocket on the receptor surface.

Interaction of the Isoleucine Side Chain of hEGF with the Receptor. To continue our discussion on the interaction between Ile-23 of hEGF and the putative hydrophobic pocket on the receptor, we examined in more detail the present and previous results of receptor-binding analyses of hEGF variants, as follows. The Ile → Leu and Ile → Val substitutions in position 23 of hEGF reduce the receptor-binding affinity of hEGF down to about 50% (Koide et al., 1992a), although the hydrophobicities of Leu and Val are not much different from that of Ile. This suggests that the hydrophobic recognition of the side chain in position 23 of hEGF by the pocket is somewhat strictly dependent on the well-defined shape of the side chain. The receptor-binding affinities of hEGF variants with natural aromatic amino acids, Phe, Trp, Tyr, and His, in position 23 are 22, 11, 31, and 31%, respectively (Koide et al., 1992a; the present study) and are not significantly lower than those of the variants with Leu and Val in this position (about 50%). This was unexpected, because these aromatic side chains are much larger than Ile, Leu, and Val.

However, we found by model building that a slight rotation of the phenyl ring about the χ_2 angle (C^α–C^β–C^γ–C^δ) within the favorable range causes the side chain of Phe-23 to partially mimic Ile-23 (Figure 7): a hydrophobic "wall" consisting of one half side of the β-methylene and phenyl-ring moieties of Phe mimics the major hydrophobic "wall" mainly consisting of the β-CH and CH_2CH_3 groups of Ile (the walls facing the top in Figure 7). Note that the side chains of Ile and Phe in Figure 7 are drawn as viewed from a direction perpendicular

to the major β-sheet plane of hEGF, which displays the side chains of Ile-23 and other important residues. Such a hydrophobic "asymmetric wall" can be formed by any of the aromatic amino acid residues, but not by Leu or Val. On the side opposite to the "wall" of the Ile side chain, recognition by the receptor does not appear to be strict. Therefore, the hEGF variants with large aromatic residues may bind to this site, allowing nearly half of the aromatic ring to protrude outside the pocket. Naturally, the receptor-binding affinities of these variants are lower than that of the wild-type hEGF. However, as discussed above, hydrogen bonding of the receptor with the protruding moiety of the aromatic residue in position 23 may partially restore the affinity.

In contrast, the hEGF variants with much smaller and/or more hydrophilic side chains (Ala, Thr, and Asp) in position 23 bind only weakly to the receptor (affinities of 6, 4, and 0.1%, respectively, of that of the wild-type hEGF) (Matsunami et al., 1990; Koide et al., 1992a). This is probably because these side chains cannot fit fully into the hydrophobic pocket that strictly recognizes the hydrophobic "wall" of the Ile side chain.

ACKNOWLEDGMENT

We are grateful to Dr. A. Nishimura (National Institute of Genetics, Mishima, Japan) for her generous gift of a phenylalanine auxotroph and to Dr. S. Mori (Morinaga Milk Industry Co., Ltd., Zama, Japan) for his helpful discussion on the chemical synthesis of phenylalanine analogs.

REFERENCES

Albertson, N. F., & Archer, S. (1945) *J. Am. Chem. Soc. 67*, 308–310.

Bain, J. D., Diala, E. S., Glabe, C. G., Dix, T. A., & Chamberlin, A. R. (1989) *J. Am. Chem. Soc. 111*, 8013–8014.

Bain, J. D., Switzer, C., Chamberlin, A. R., & Benner, S. A. (1992) *Nature 356*, 537–539.

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 393–402.

Bogosian, G., Violand, B. N., Dorward-King, E. J., Workman, W. E., Jung, P. E., & Kane, J. F. (1989) *J. Biol. Chem. 264*, 531–539.

Burgess, A. W., Lloyd, C. J., Smith, S., Stanley, E., Walker, F., Fabri, L., Simpson, R. J., & Nice, E. C. (1988) *Biochemistry 27*, 4977–4985.

Campion, S. R., Matsunami, R. K., Engler, D. A., & Niyogi, S. K. (1990) *Biochemistry 29*, 9988–9993.

Campion, S. R., Biamonti, C., Montelione, G. T., & Niyogi, S. K. (1993) *Protein Eng. 6*, 651–659.

Cantor, C. R., & Schimmel, P. R. (1980) *Biophysical Chemistry*, Part I, p 49, W. H. Freeman and Company, San Francisco, CA.

Chung, H.-H., Benson, D. R., & Schultz, P. G. (1993) *Science 259*, 806–809.

Cooke, R. M., Wilkinson, A. J., Baron, M., Pastore, A., Tappin, M. J., Campbell, I. D., Gregory, H., & Sheard, B. (1987) *Nature (London) 327*, 339–341.

Cooke, R. M., Tappin, M. J., Campbell, I. D., Kohda, D., Miyake, T., Fuwa, T., Miyazawa, T., & Inagaki, F. (1990) *Eur. J. Biochem. 193*, 807–815.

Dewitt, H. D., & Ingersoll, A. W. (1951) *J. Am. Chem. Soc. 73*, 3359–3360.

Dudgeon, T. J., Cooke, R. M., Baron, M., Campbell, I. D., Edwards, R. M., & Fallon, A. (1990) *FEBS Lett. 261*, 392–396.

Ellman, J., Mendel, D., Anthony-Cahill, S. J., Noren, C. J., & Schultz, P. G. (1991) *Methods Enzymol. 202*, 301–336.

Ellman, J. A., Mendel, D., & Schultz, P. G. (1992) *Science 255*, 197–200.

Engler, D. A., Matsunami, R. K., Campion, S. R., Stringer, C. D., Stevens, A., & Niyogi, S. K. (1988) *J. Biol. Chem. 263*, 12384–12390.

Engler, D. A., Montelione, G. T., & Niyogi, S. K. (1990) *FEBS Lett. 271*, 47–50.

Fowden, L., Lewis, D., & Tristram, H. (1967) *Adv. Enzymol. 29*, 89–163.

Gilles, A.-M., Marliere, P., Rose, T., Sarfati, R., Longin, R., Meier, A., Fermandjian, S., Monnot, M., Cohen, G. N., & Bârzu, O. (1988) *J. Biol. Chem. 263*, 8204–8209.

Heckler, T. G., Chang, L.-H., Zama, Y., Naka, T., Chorghade, M. S., & Hecht, S. M. (1984) *Biochemistry 23*, 1468–1473.

Hommel, U., Dudgeon, T. J., Fallon, A., Edward, R. M., & Campbell, I. D. (1991) *Biochemistry 30*, 8891–8898.

Judice, J. K., Gamble, T. R., Murphy, E. C., de Vos, A. M., & Schultz, P. G. (1993) *Science 261*, 1578–1581.

Kigawa, T., & Yokoyama, S. (1991) *J. Biochem. (Tokyo) 110*, 166–168.

Kohno, T., Kohda, D., Haruki, M., Yokoyama, S., & Miyazawa, T. (1990) *J. Biol. Chem. 265*, 6931–6935.

Koide, H., Yokoyama, S., Kawai, G., Ha, J.-M., Oka, T., Kawai, S., Miyake, T., Fuwa, T., & Miyazawa, T. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6237–6241.

Koide, H., Oishi, M., Oka, T., Miyake, T., Fuwa, T., Yokoyama, S., & Miyazawa, T. (1990) in *Amino Acids: Chemistry, Biology and Medicine* (Lubec, G., & Rosenthal, G. A., Eds.) pp 193–200, Escom, Leiden.

Koide, H., Muto, Y., Kasai, H., Kohri, K., Hoshi, K., Takahashi, S., Tsukumo, K., Sasaki, T., Oka, T., Miyake, T., Fuwa, T., Kohda, D., Inagaki, F., Miyazawa, T., & Yokoyama, S. (1992a) *Biochim. Biophys. Acta 1120*, 257–261.

Koide, H., Muto, Y., Kasai, H., Hoshi, K., Takusari, H., Kohri, K., Takahashi, S., Sasaki, T., Tsukumo, K., Miyake, T., Fuwa, T., Miyazawa, T., & Yokoyama, S. (1992b) *FEBS Lett. 302*, 39–42.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.

Matsunami, R. K., Campion, S. R., Niyogi, S. K., & Stevens, A. (1990) *FEBS Lett. 264*, 105–108.

Mendel, D., Ellman, J. A., & Schultz, P. G. (1991) *J. Am. Chem. Soc. 113*, 2758–2760.

Mendel, D., Ellman, J. A., Chang, Z., Veenstra, D. L., Kollman, P. A., & Schultz, P. G. (1992) *Science 256*, 1798–1802.

Moy, F. J., Scheraga, H. A., Liu, J.-F., Wu, R., & Montelione, G. T. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 9836–9840.

Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C., & Schultz, P. G. (1989) *Science 244*, 182–188.

Oka, T., Sakamoto, S., Miyoshi, K., Fuwa, T., Koda, K., Yamasaki, M., Tamura, G., & Miyake, T. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 7212–7216.

Ray, P., Moy, F. J., Montelione, G. T., Liu, J.-F., Narang, S. A., Scheraga, H. A., & Wu, R. (1988) *Biochemistry 27*, 7289–7295.

Richmond, M. H. (1962) *Bacteriol. Rev. 26*, 398–420.

Spirin, A. S., Baranov, V. I., Ryabova, L. A., Ovodov, S. Y., & Alakhov, Y. B. (1988) *Science 242*, 1162–1164.

Tasi, L. B., Lu, H. S., Kenney, W. C., Curless, C. C., Klein, M. L., Lai, P.-H., Fenton, D. M., Altrock, B. W., & Mann, M. B. (1988) *Biochem. Biophys. Res. Commun. 156*, 733–739.

Wheatley, D. N. (1978) *Int. Rev. Cytol. 55*, 109–169.

Wilson, M. J., & Hatfield, D. L. (1984) *Biochim. Biophys. Acta 781*, 205–215.

Woodward, R. B., Olofson, R. A., & Mayer, H. (1961) *J. Am. Chem. Soc. 83*, 1010–1012.