

Mistranslation of a TGA termination codon as tryptophan in recombinant platelet-derived growth factor expressed in *Escherichia coli*

Kan V. LU, Michael F. ROHDE, Arlen R. THOMASON, William C. KENNEY and Hsieng S. LU*

Amgen Inc., Amgen Center, Thousand Oaks, CA 91320, U.S.A.

The mature 109-amino-acid human platelet-derived growth factor B (PDGF-B) peptide is derived by intracellular processing from a 241-amino-acid precursor synthesized in mammalian cells, with removal of 81 N-terminal and 51 C-terminal amino acids. In order to produce directly the mature 109-amino acid PDGF-B peptide as a recombinant protein in *Escherichia coli*, a CGA codon at position 110 of a DNA sequence encoding the full-length precursor form of PDGF-B was converted into the translation termination codon TGA by *in vitro* mutagenesis. Expression of this DNA via a plasmid vector in *E. coli* resulted in production of two distinct PDGF-B proteins having apparent molecular masses of 15 and 19 kDa, with the latter species predominating. Structural characterization employing N- and C-

terminal amino acid sequencing and MS analyses indicated that the 15 kDa protein is the expected 109-amino-acid PDGF-B, and that the 19 kDa protein represents a C-terminal extended PDGF-B containing 160 amino acids. Characterization of a unique tryptic peptide derived from the 19 kDa protein revealed that this longer form of PDGF-B results from mistranslation of the introduced TGA termination codon at position 110 as tryptophan, with translation subsequently proceeding to the naturally occurring TAG termination codon at position 161. Owing to the high rate of translation readthrough of TGA codons in this and occasionally other proteins, it appears that the use of TGA as a translation termination codon for proteins to be expressed in *E. coli* should be avoided when possible.

INTRODUCTION

The impressive advances in recombinant DNA technologies have resulted in the development of highly efficient expression systems in a variety of host cells that are capable of producing important proteins in large quantities for human therapeutic use. Extremely high production yields can often be achieved by using bacterial expression systems. However, owing to its high rate of biosynthesis, the translation process in bacterial cells, especially in *Escherichia coli*, has been reported to be error-prone [1–3]. The estimated translational error could be as high as 10^{-3} – 10^{-4} [1]. These errors are more evident in the high-yield production of heterologous proteins, as reported for many recombinant products, including mis-sense translations (or mis-incorporations) of Nor-Leu for Met in interleukin-2 (IL-2) [4] or bovine somatotropin [5], Gln for His in granulocyte colony-stimulating factor [6], Arg for Lys in insulin growth factor-1 [7], transfer RNA 'hopping' in bovine placental lactogen [8], and Gln readthrough of a UAG stop codon in bovine somatotropin [8]. Most of these errors, as well as others such as aberrant initiation and frameshifting, have been observed in cells which are not of recombinant origin (for a review, see [9]).

Platelet-derived growth factor (PDGF) is a molecule involved in control of cell growth and tissue repair (for a review, see [10]). PDGF is a disulphide-linked dimer composed of A and B subunits. The three possible isoforms A–A, B–B and A–B, exhibit different cell and tissue distributions. Both the A chain and the B chain of PDGF are initially synthesized as long precursor polypeptides from which the mature, secreted, form is intracellularly produced by cleavage. PDGF-B is synthesized as a 241-amino-acid precursor, and the N-terminal 81 amino acids and the C-terminal 51 amino acids are subsequently removed, leaving a mature monomer of 109 amino acids that is secreted after dimerization [10,11]. To produce the mature 109-amino-

acid PDGF-B in *E. coli*, where N- and C-terminal processing is not expected to occur, we modified a DNA sequence coding for the full-length precursor of human PDGF-B. The region coding for the N-terminal 81 amino acids was removed, and the codon immediately following Thr-109 of mature PDGF-B was *in vitro* mutated to the termination codon TGA (corresponding to UGA in the mRNA). In the present paper we describe the detection of a C-terminal extended form of recombinant PDGF-B which is co-expressed at high levels along with the expected 109-amino-acid mature PDGF-B. Characterization of the extended form using in-gel peptide mapping, HPLC–MS and electroblotting in conjunction with both N- and C-terminal sequencing confirmed that it is derived from Trp translation termination bypass at the introduced TGA stop codon, and is a 160-amino-acid polypeptide terminated by the downstream natural TAG stop codon of the inserted PDGF-B gene.

EXPERIMENTAL

Materials

Pre-cast SDS-containing gels were products of Novex (San Diego, CA, U.S.A.). Trypsin (sequencing grade) and dithiothreitol were from Boehringer-Mannheim (Indianapolis, IN, U.S.A.). HPLC solvents were products of Burdick and Jackson (Muskegon, MI, U.S.A.).

Expression of PDGF-B in *E. coli*

Construction of a modified PDGF-B gene (see Figure 1) for *E. coli* expression was accomplished using site-directed mutagenesis of human *v-sis* essentially as described previously [12]. Codons that code for amino acids in *v-sis* which differ from human

Abbreviations used: PDGF, platelet-derived growth factor; IL-2, interleukin-2; TFA, trifluoroacetic acid; DTT, dithiothreitol; PVDF, poly(vinylidene difluoride); RF2, release factor 2.

* To whom correspondence should be addressed.

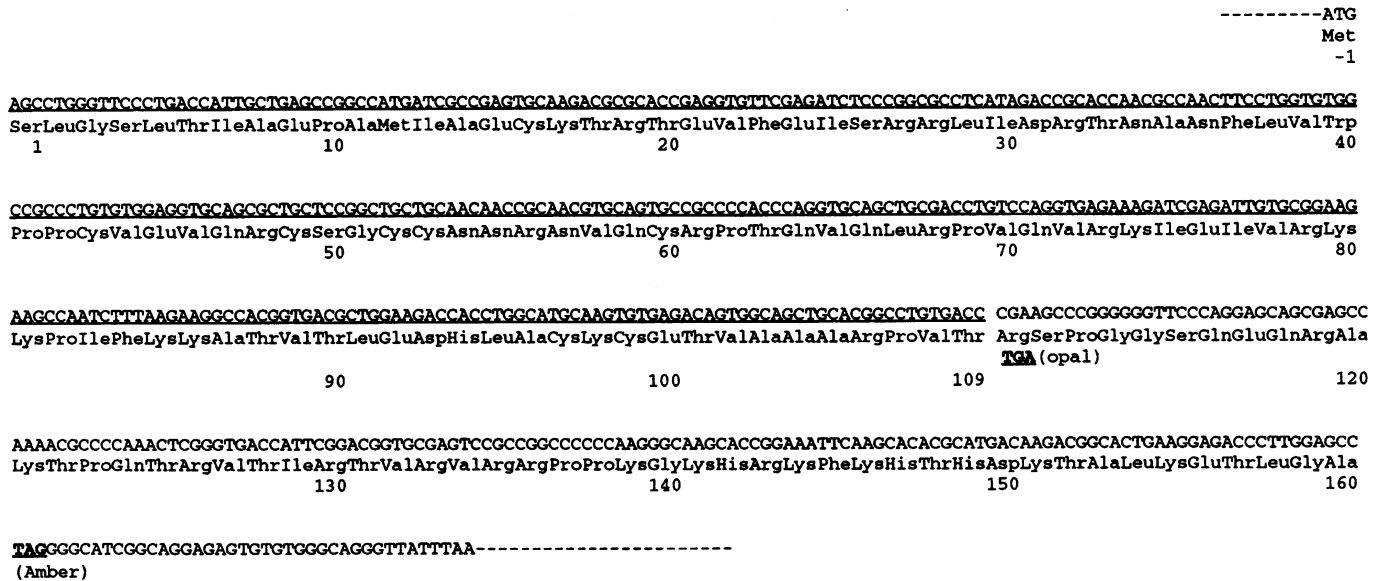


Figure 1 Coding sequence for human PDGF-B denoting the insertion of an ATG codon (for initiator Met) preceding the N-terminal Ser of mature PDGF-B and mutation of CGA (for Arg) into TGA (opal) for translational termination (bold face at position 110)

A second termination codon, TAG, located at position 161 is an authentic termination site for PDGF-B precursor protein synthesized by mammalian cells.

PDGF-B were converted into the human PDGF-B sequence. The DNA region upstream of Ser-1 of mature PDGF-B was replaced with a translation-initiating ATG codon, and the CGA encoding Arg-110 was *in vitro*-mutated to the termination codon TGA. The modified PDGF-B protein should begin with a translation-initiating methionine residue followed by 109 amino acids identical with amino acids 1–109 of human PDGF-B. The N-terminal methionine of PDGF-B is removed by intracellular processing in *E. coli*.

In separate studies of PDGF-B expression in *E. coli*, a second stop codon, TAA, in line with the TGA stop codon, was constructed by *in vitro* mutagenesis of the above-mentioned PDGF-B gene. We also constructed an analogous PDGF-B gene which encodes a 119-amino-acid [Ala³⁵ → His, ΔArg²⁸]PDGF analogue (see Figure 1) and contains TGA and TAA in-line stop codons for translation termination.

E. coli culture for production of recombinant PDGF-B proteins was carried out in defined media according to procedures described previously [13]. The cell paste was passed twice through a French pressure cell at 69 kPa (10000 lbf/in²). Broken-cell suspensions were centrifuged at 10000 *g* for 15 min at 4 °C, and the broken-cell supernatants were discarded. The pellet was resuspended in water and centrifuged at 10000 *g* for 15 min at 4 °C. The supernatants were discarded and the pellets were solubilized at room temperature in 50 mM Tris/HCl containing 2% dodecanoate, 5% ethanol and 50 mM dithiothreitol (DTT), pH 8.7. The solubilized protein was chromatographed on Sephadex G-75 as described below.

Purification of PDGF-B by gel filtration

The solubilized extract described above was then subjected to Sephadex G-75 (fine) gel filtration. Approx. 300 ml of G-75 medium was packed into a column (2.5 cm × 60 cm), which was then equilibrated with 50 mM Tris/HCl, pH 8.7, containing 2% sodium dodecanoate and 5% ethanol at a flow rate of 30 ml/h. After sample loading, proteins were eluted with the same buffer;

fractions (7.5 ml each) were monitored at 280 nm and collected. Fractions were analysed by SDS/PAGE to detect PDGF-B.

SDS/PAGE/electroblotting

Two PDGF-B samples recovered from gel filtration were each concentrated to 3 ml by using a Centriprep membrane filtration unit and centrifugation at 2060 *g*. Aliquots (50 μl) of samples were treated with equal volumes of SDS/PAGE sample buffer containing 1% β-mercaptoethanol and 1% SDS at 80 °C for 3 min.

All sample proteins were run on 16% pre-cast Novex Tris/glycine gels (1 mm thickness; 10 wells/gel) under the reducing conditions described by Laemmli [14]. Gels were stained with 0.05% Coomassie Brilliant Blue-G in methanol/acetic acid/water (40:1:159, by vol.) and destained with a solution of methanol/acetic acid/water (45:9:46, by vol.). The gels were then washed in water overnight. If the gels were to be electroblotted, freshly run samples were immediately electroblotted on to ProBlott poly(vinylidene difluoride) (PVDF) membranes (Applied Biosystems Inc.) by a semi-dry (MilliBlot-SDE) electroblotter using a previously described procedure [15]. After blotting, the membranes were stained using 0.1% Coomassie Blue R250 in methanol/acetic acid/water (40:1:159, by vol.) and destained with methanol/acetic acid/water (45:9:46, by vol.). Sample blots were then air-dried for subsequent analyses.

In-gel proteolytic digestion

In-gel digestion was performed by a combination of two reported methods [16,17] with some modifications [18]. Protein samples were first separated by SDS/PAGE as described above. The band to be digested was excised with an ethanol-cleaned razor blade. The excised band was then placed in an Eppendorf vial (1.5 ml) and washed with 150 μl of 0.1 M Tris/HCl/50% acetonitrile, pH 9.0, for 20 min using a shaker at ambient temperature. The supernatant was discarded and the 20 min

wash repeated. The excised band was then placed on to a clean glass plate and allowed to air-dry for 5–10 min. After rehydrating the slice with 5 ml of digestion buffer containing 100 mM Tris/HCl/0.01 % Tween-20, pH 9.0, 2 μ l (0.5 μ g) of sequencing-grade trypsin was added. After absorption of the enzyme, 3–5 μ l more of digestion buffer was added to return the gel slice to its original size. Two slices of the gel bands were then placed into a 0.5 ml Microfuge tube, immersed in 100–150 μ l of digestion buffer, and subsequently incubated at 30 °C for 48 h.

After digestion, the supernatant was pipetted into a clean 0.5 ml Eppendorf tube. Two extractions (both for 20 min in a Labquake shaker at ambient temperature) were performed using 100 μ l of 60 % acetonitrile/0.1 % trifluoroacetic acid (TFA). The supernatants of these two washes were combined with the digestion supernatant and then dried by vacuum centrifugation. The dried digest was redissolved in 100 μ l of 0.1 % TFA for subsequent HPLC peptide mapping.

Reverse-phase HPLC separation of PDGF-B and peptide mapping

HPLC separation of PDGF-B and generation of peptide maps were performed on a Hewlett-Packard 1090 M HPLC system equipped with a diode-array detector and ChemStation at room temperature.

To isolate PDGF-B by reverse-phase HPLC, concentrated samples (150 μ l) obtained from gel-filtration fractions were denatured by adding an equal volume of 8 M guanidinium chloride in 0.3 M Tris/HCl, pH 8.4, and 50 μ M of DTT, and then incubated at 37 °C for 20 min. Sample volumes of 250 μ l were injected into a Vydac C₄ column (4.6 mm \times 25 cm) equilibrated with 97 % mobile phase A/3 % mobile phase B at a flow rate of 0.7 ml/min. HPLC mobile phase A consisted of 0.1 % TFA in HPLC-grade water and mobile phase B was 0.1 % TFA in 90 % HPLC-grade acetonitrile. PDGF-B was eluted by the following linear gradient elution: 3–70 % B (0–70 min), 70–85 % B (70–80 min), and 85–95 % B (80–85 min). The eluant was monitored at 215 and 280 nm, and fractions were hand-collected for subsequent analyses.

In-gel peptide mapping was performed by injecting 100 μ l of each PDGF-B digest on to a narrow-bore Synchrom RP C₄ column (0.25 mm \times 5 cm) with a flow rate of 0.146 ml/min at ambient temperature. HPLC mobile phases A and B were identical with those described above. The gradient was 3 % B for (0–10 min), 3–18 % B (10–21 min), 18–70 % B (21–81 min), and 70–90 % B (81–100 min). Peptides were monitored at 215 and 280 nm, and fractions were collected for sequence analysis.

Electrospray MS

Molecular masses of PDGF-B preparations isolated by reverse-phase HPLC were determined by a Finnigan SSQ710C electrospray mass spectrometer operating in a single quadrupole mode. Samples collected from the HPLC (5 μ l) column were introduced by flow injection at 30 μ l/min. An orifice potential of 70 V was set for mass determination. The obtained multiple charged ion spectra were deconvoluted [19] to obtain the precise molecular masses.

Automated N- and C-terminal sequence analyses

Automated N-terminal sequence analysis of proteins and peptides was performed on a model 477 protein sequencer (Applied Biosystems) as reported [20]. Sequencing of samples prepared by electroblotting procedures was also performed as described by Fausset and Lu [15].

Automated C-terminal sequence analysis was carried out by

an Applied Biosystems C-terminal sequencer using the procedures and reagents described in [21].

RESULTS AND DISCUSSION

A PDGF-B DNA sequence was modified as described in the Experimental section to direct the *E. coli* expression of the 109-amino-acid protein shown in Figure 1, encompassing Ser-1–Thr-109. This DNA sequence, which still contained the region downstream of the TGA termination codon introduced at codon 110, was inserted into a bacterial expression vector and transformed into *E. coli*. When the proteins extracted from the transformed bacteria were analysed by SDS/PAGE under reducing conditions, two new bands at approx. 15 and 19 kDa were easily detected (results not shown). As shown in Figure 2(a), Sephadex G-75 gel filtration provided a partial purification of the solubilized sample. Two peaks eluted between fractions 23 and 34 were analysed by SDS/PAGE and found to contain the newly expressed 19 and 15 kDa proteins, with the former representing about 60 % of the new protein (Figure 2b). Migration of the 15 kDa protein was identical with that of a

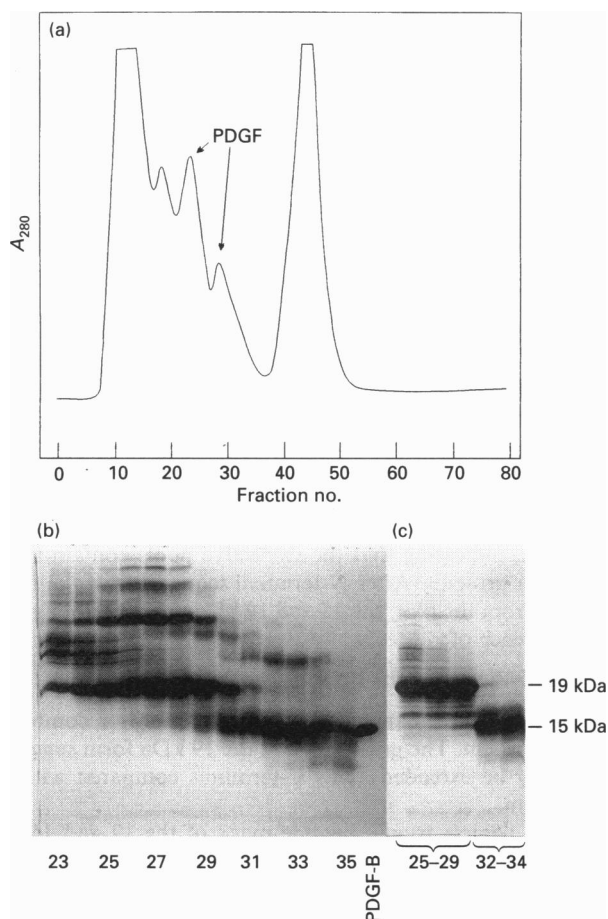


Figure 2 Partial purification and analysis of PDGF-B proteins synthesized in *E. coli*

(a) Sephadex G-75 gel-filtration chromatography of crude solubilized bacterial inclusion bodies containing reduced recombinant PDGF-B. Fractions from each of the two peaks marked 'PDGF' were separately pooled for further analysis. (b) Reducing SDS/PAGE of column fractions, containing 19 and 15 kDa PDGF bands, obtained by the gel filtration shown in (a). The PDGF-B standard is shown in the extreme right-hand lane. (c) Stained blots of two fraction pools from (a) re-subjected to reducing SDS/PAGE and electroblotted on to PVDF membrane.

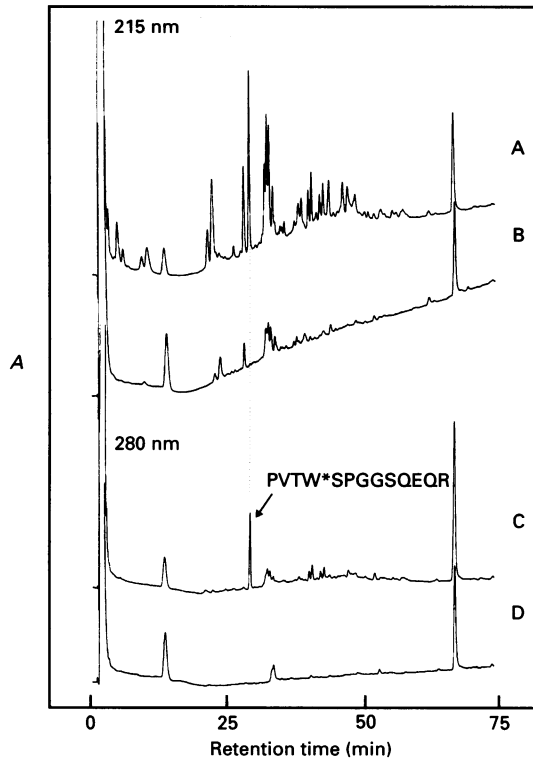


Figure 3 In-gel tryptic peptide maps derived from the 19 kDa and 15 kDa PDGF-B bands detected at 215 nm (chromatograms A and B respectively) and at 280 nm (chromatograms C and D respectively)

previously characterized mature PDGF-B standard, indicating that it may be the expected 109-amino-acid recombinant PDGF-B molecule.

To confirm that the new 15 kDa protein was the expected PDGF-B product, and to determine the nature of the 19 kDa band, fractions 24–26 and 32–34 were separately pooled, concentrated, and subjected to SDS/PAGE. The gel was then electroblotted on to a PVDF membrane. Figure 2(c) shows the stained PVDF blot, indicating that the respective pools contain the 19 and 15 kDa proteins. After N-terminal sequencing of the bands recovered from the blot, the 15 and 19 kDa proteins both gave a single sequence of:

Ser-Leu-Gly-Ser-Leu-Thr-Ile-Ala-Glu-Pro-Ala-Met-...

indicating that these bands both contain primarily recombinant PDGF-B protein. The greater size of the 19 kDa form suggested that it may be extended at its C-terminus compared with the 15 kDa form.

Figure 3 depicts tryptic peptide maps of the 19 and 15 kDa proteins, generated from each of the two pooled Sephadex fractions by SDS/PAGE and in-gel trypsin digestion of the excised bands (see the Experimental section). Figure 3, chromatograms A and C, are maps derived from the 19 kDa band and detected at 215 and 280 nm respectively. Figure 3, chromatograms B and D, are those derived from the 15 kDa band and similarly detected. A strong peak eluted at 65 min in all maps is derived from the elution of Coomassie Blue G and is thus not a peptide. When the 280 nm maps of the 19 and 15 kDa samples were compared, only the 19 kDa sample gave a strong peptide peak that was eluted at approx. 27 min. This peptide also gave a UV spectrum typical of that containing aromatic amino acids

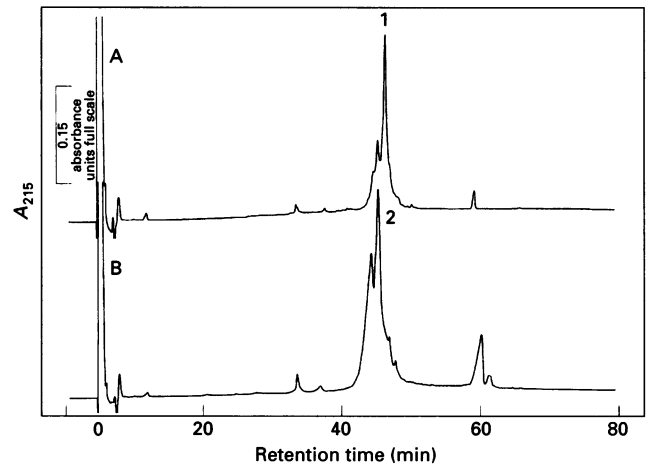


Figure 4 Reversed-phase HPLC of two reduced PDGF-B fractions obtained from gel filtration as shown in Figure 1

Chromatogram A corresponds to fractions containing the 19 kDa protein, and chromatogram B corresponds to the 15 kDa protein.

(results not shown), suggesting that Tyr or Trp may be present in the peptide. Complete amino-acid-sequence determination of this peptide yielded the sequence:

Pro-Val-Thr-Trp-Ser-Pro-Gly-Gly-Ser-Gln-Glu-Gln-Arg

By comparison with the predicted amino acid sequence of the PDGF-B DNA introduced into the *E. coli* cells (Figure 1), this peptide was found to be derived from trypsin cleavage at Arg-106 and Arg-119. The first three amino acids in this peptide overlap the expected C-terminus for mature PDGF-B. The Trp residue at the fourth position corresponds to the putative TGA stop codon introduced at codon 110, and is followed by the sequence predicted from the cDNA sequence after the TGA codon. The results thus confirmed that the 19 kDa protein is a C-terminal-extended form of PDGF-B resulting from Trp mistranslation at the TGA termination codon. Re-sequencing of the DNA in the expression vector confirmed the presence of TGA at this position. While the tryptic peptide mapping procedure was very useful in identifying an unusual Trp-containing peptide as described above, it also generated a number of small peptides from both the 15 and 19 kDa protein bands. These peptides were difficult to characterize owing to their low yields and recovery, rendering it impossible to precisely determine the exact C-terminus of each protein by isolation and identification of their C-terminal peptides. Therefore, MS and automated C-terminal sequencing of each protein was employed. As shown in Figure 4, the two pooled fractions obtained from gel filtration could be reduced with DTT and separated by reverse-phase HPLC, permitting the removal of contaminating proteins and excipients. The 19 kDa protein (peak 1 in Figure 4, chromatogram A) is eluted only slightly later than the 15 kDa protein (peak 2 in Figure 4, chromatogram B). Peaks 1 and 2 were subjected to electrospray MS analyses for precise mass measurement (Figures 5a and 5b respectively). Multiply-charged ion species were observed in both samples, as shown at the top of Figure 5; the deconvoluted molecular ions are shown at the bottom. Peak 1 gave a singly charged molecular ion (*M*) of mass 18126.0 ± 3.6 and peak 2 gave an ion of mass 12296.2 ± 1.3 . The determined molecular ion for peak 1 closely matches the calculated value ($MH^+ = 18125.7$) for an extended PDGF-B with 160 amino acids including a Trp

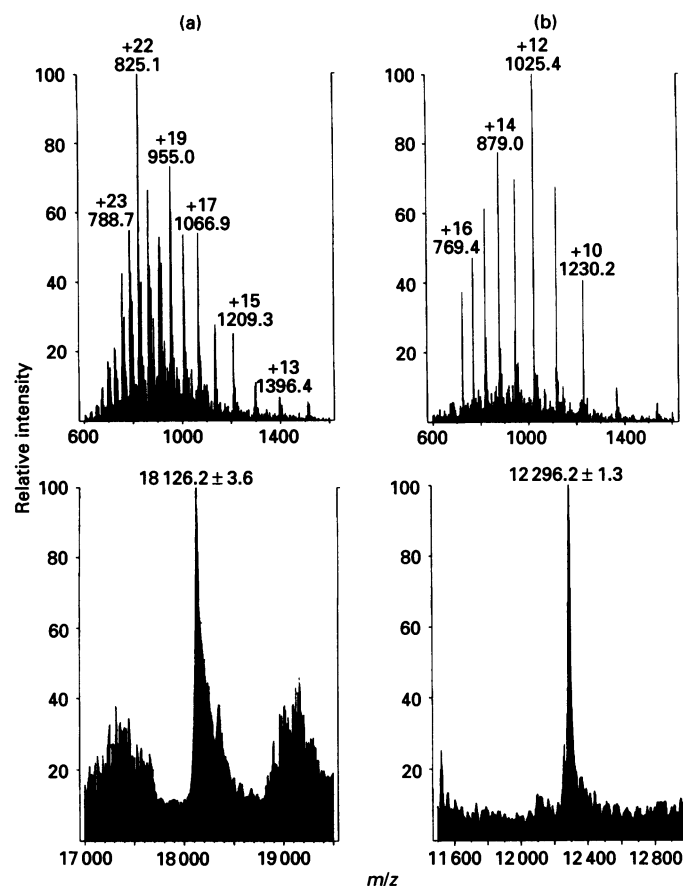


Figure 5 Electrospray MS analysis of HPLC-purified 19 kDa and 15 kDa PDGF-B samples (a and b respectively)

Spectra at the bottom represent those obtained from deconvolution of the spectra shown at the top.

residue at the TGA stop codon (see Figure 1). The molecular ion for peak 2 is also in good agreement with the calculated mass ($MH^+ = 12298.5$) for a PDGF-B molecule containing 109 amino acids. In both chromatograms A and B of Figure 4 there is a minor peak that is eluted slightly earlier than the main PDGF peak as analysed. These minor peaks are the contaminants containing the dimeric PDGF of the 15 and 19 kDa forms linked by an intermolecular disulphide bond (results not shown).

The 19 kDa PDGF-B blot shown in Figure 2(c) was further analysed by automatic C-terminal sequence analysis. Two amino acids, ...-Gly-Ala-COOH, were elucidated. The data are consistent with the conclusion that the 19 kDa PDGF-B form terminates at the second stop codon, TAG. No C-terminal sequence information for the 15 kDa blot was obtained, owing to destruction of the C-terminal Thr-109 and early termination at Pro-107 during C-terminal sequence analysis [21].

From the combination of data obtained, it is concluded that the 15 kDa form is the properly terminated PDGF-B protein containing 109 amino acids, and that the 19 kDa form represents a C-terminal-extended PDGF-B resulting from Trp readthrough at the TGA stop codon. The approaches described here employ partial purification by gel filtration, in-gel peptide mapping, HPLC-MS, and C-terminal sequencing for the determination of termination bypass and other translational errors in non-purified samples. Only minimal sample purification steps are required to

obtain precise data for structural characterization of crude samples.

Consistently accurate information transfer appears to be the rule at the levels of replication and transcription, since structural changes in recombinant proteins that derive from genetic instability have rarely been reported [22]. It is also often taken for granted that normal translation according to the 'universal genetic code' will occur to produce authentic and functional proteins during gene expression. Yet many recent reports have shown that a number of non-standard translation events may occur in various types of cells [9]. The types of translation errors that occur include aberrant initiation, mis-sense translation (i.e. mistranslation), frameshifting, tRNA hopping and translation of termination codons (often referred to as 'readthrough') [9]. These events may lead to generation of microheterogeneity in the translation products.

Termination of peptide-chain synthesis is signalled by any of the three termination codons, TAG (amber), TGA (opal) and TAA (ochre). The classically described means for overcoming translation termination at the level of tRNA is by nonsense suppression, where a mutation in a tRNA molecule (suppressor tRNA) permits a termination codon to be read as though it coded for an amino acid [23,24]. Suppressor tRNAs have been found in both prokaryotes and eukaryotes. However, termination codons can sometimes be read as sense codons by a normally

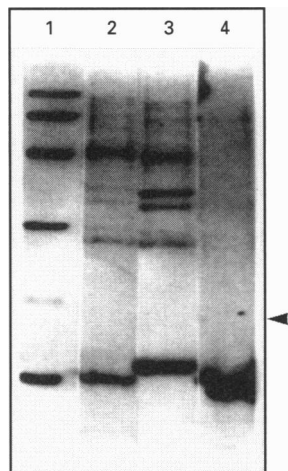


Figure 6 SDS/PAGE of crude extract of *E. coli* inclusion bodies derived from PDGF gene constructs with a second TAA stop codon in line with TGA stop codon

Lane 1, standard proteins (from the top, 90, 65, 45, 30, 21 and 14 kDa respectively); lane 2, a PDGF sample containing 109 amino acids; lane 3, a PDGF analogue containing 119 amino acids; lane 4, standard purified recombinant PDGF. The arrow indicates where the readthrough products migrate.

charged tRNA at a low level (about 1 to 10%), allowing the formation of an elongated peptide [25,26].

In bacteria, particularly *E. coli*, the codon-specific release factor 2 (RF2) is required for TGA-directed peptide chain termination, possibly via direct binding to the codon [27]. Protein and DNA sequence revealed that RF2 is encoded by two reading frames. Frameshifting from one reading frame to the next occurs at a higher rate when the TGA codon utilizes a minor tRNA species [27]. In previous studies, tRNA^{sec}_{UCA}, specified by the gene *selC* [28,29], was implicated as the natural tRNA involved with the TGA readthrough process in *E. coli*. tRNA^{sec}_{UCA} is directed by a TGA codon, which mediates the incorporation of the modified amino acid selenocysteine. Selenocysteine incorporation via in-frame TGA codon recognition and co-translational phosphorylation of tRNA^{ser} represents an unconventional protein-synthetic system to synthesize functionally important proteins or enzymes in bacteria and mammalian cells [29]. However, this is a rare translational event that is very specific to the synthesis of unusual selenocysteine-containing proteins and requires co-translational modification of a charged tRNA. In the case of PDGF-B, it appears that another tRNA, tRNA^{Trp}, is responsible for TGA readthrough resulting in synthesis of a PDGF-B extended form.

As determined from SDS-PAGE analyses, the synthesis of C-terminal extended PDGF-B reaches more than 60% of the total PDGF-B. Such a high level of termination codon readthrough has previously been detected only when a suppressor tRNA and a mutated RF2 were involved in the readthrough process [30]. However, from our available data we cannot determine whether tRNA^{Trp} involved in TGA translation in PDGF-B is a natural or suppressed tRNA. There may also be other factors responsible for the high readthrough rate of PDGF-B observed in the present study. For example, the codons of the PDGF-B gene used in the present study were not optimized for bacterial expression. The TGA stop codon is not one of the *E. coli*-preferred nonsense codons (i.e. TAA and TAG) [9]. In addition, it has been reported that readthrough of a TGA codon is affected by the adjacent 3' nucleotide, where the order of readthrough inhibition is T > C

> G > A [30]. In the PDGF-B gene sequence (Figure 1), the TGA stop codon is followed by A, which has the least effect on inhibiting readthrough. Another factor to consider is that recombinant proteins expressed in bacteria, including PDGF, are usually expressed at very high levels. It is possible that expression of RF2, which mediates TGA termination, may be limiting under these conditions.

Engineering PDGF gene with a second TAA or TAG codon in line with TGA codon can eliminate the readthrough problem. This is demonstrated in Figure 6, which shows the SDS/PAGE analysis of the translation product in the crude extract of *E. coli* inclusion bodies derived from a PDGF-B gene construct with a second TAA stop codon. The short PDGF-B form is detectable along with other *E. coli* proteins, while the higher-molecular-mass readthrough product (shown by the arrow) is apparently not present (lanes 2 and 3). A PDGF form that exhibited slightly higher molecular mass (Figure 6, lane 3) was obtained from the expression of a PDGF-B gene construct encoding a 119-amino-acid PDGF analogue (see the Experimental section). Studies with the *E. coli* expression of recombinant human neurotrophin-3 also clearly showed that replacing the TGA codon with the *E. coli*-preferred codon, TAA or TAG, can terminate the translation of neurotrophin efficiently [31]. Given these observations, together with the evidence that TGA termination is inefficient even for normal bacterial proteins [26], its use should be limited for large-scale production of recombinant proteins using bacterial expression systems.

The Trp readthrough of PDGF-B gene in *E. coli* as described has created a longer PDGF-B form of 160-amino-acid residues. The extended C-terminal portion found in this form is not required for the PDGF biological activity, since the mature form of PDGF was confirmed to be responsible for binding to the PDGF receptor and essential for PDGF activity [32,33]. Except the readthrough amino acid, Trp, at position 110, the long PDGF-B form is structurally identical with an endogenous PDGF-B form unprocessed at the C-terminus (see Figure 1 and [34]). This latter form extends its C-terminal sequence that is encoded by exon 6 and is posttranslationally processed to become mature PDGF-B. The extended C-terminal portion contains a stretch of basic amino acids (sequence 133–144; see Figure 1) which strongly binds to cell- and matrix-associated heparan sulphate proteoglycan [34]. This observation suggested that the translation of the exon 6-encoded sequence of PDGF results in compartmentalization of PDGF on extracellular binding sites, whereas mature forms lacking this sequence would be soluble and diffuse [34].

We are indebted to Dr. V. Katta in performing MS analysis, to G. Stearns for growing *E. coli* cells and collection of cell paste, and to J. Bennett for her help in typing this manuscript. Thanks are also due to Dr. Vicky Boyd and Dr. Meri Bozzini of Applied Biosystems Inc., Foster City, CA, U.S.A., for their help in setting up an automatic C-terminal sequencer in this laboratory. K.V.L., an undergraduate student at the University of California, Los Angeles, CA, U.S.A., worked as a summer student in 1994.

REFERENCES

- Bouadloun, F., Donner, D. and Kurland, C. G. (1983) *EMBO J.* **2**, 1351–1356
- Kurland, T. B. L. and Ehrenberg, M. (1987) *Annu. Rev. Biophys. Chem.* **16**, 291–317
- Johnston, T. C., Borgia, P. T. and Parker, J. (1984) *Mol. Gen. Genet.* **195**, 495–465
- Lu, H. S., Tsai, L. B., Kenney, W. C. and Lai, P. H. (1988) *Biochem. Biophys. Res. Commun.* **156**, 807–813
- Violand, B. N., Siegel, N. R., Bogosian, G., Workman, W. E. and Kane, J. E. (1989) in *Techniques in Protein Chemistry* (Hugli, T. E., ed.), vol. 1, pp. 315–326, Academic Press, San Diego
- Lu, H. S., Fausset, P. R., Sotos, L. S. et al. (1993) *Protein Express. Purif.* **4**, 465–472

- 7 Seetheram, R., Heeren, R. A., Wong, F. H. et al. (1988) *Biochem. Biophys. Res. Commun.* **155**, 518–523
- 8 Violand, B. N. (1994) *ABRF Workshops: 8th Protein Soc. Symp. 8th, San Diego*
- 9 Santos, M. A. S. and Tuite, M. F. (1993) *Trends Biotechnol.* **11**, 500–505
- 10 Heldin, C.-H. (1992) *EMBO J.* **11**, 4251–4259
- 11 Johnson, A., Heldin, C.-H., Wateson, A. et al. (1984) *EMBO J.* **3**, 921–928
- 12 Kenney, W. C., Haniu, M., Herman, A. C. et al. (1994) *J. Biol. Chem.* **269**, 12351–12359
- 13 Boone, T., Chazin, V., Kenney, W. C., Swanson, E. and Altrock, B., M. (1987) *Dev. Biol. Stand.* **69**, 157–168
- 14 Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685
- 15 Fausset, P. R. and Lu, H. S. (1991) *Electrophoresis* **12**, 22–27
- 16 Kawasaki, H., Emori, Y. and Suzuki, H. (1990) *Anal. Biochem.* **191**, 332–336
- 17 Rosenfield, J., Capdevielle, J., Guillemot, J. C. and Ferrara, P. (1992) *Anal. Biochem.* **203**, 173–179
- 18 Merewether, L. A., Clogston, C. L., Patterson, S. D. and Lu, H. S. (1995) in *Techniques in Protein Chemistry* (Crabb, J., ed.), Academic Press, San Diego, vol. 6, in the press
- 19 Mann, M., Meng, C. K. and Feng, J. B. (1989) *Anal. Chem.* **61**, 1702–1708
- 20 Lu, H. S., Boone, T. C., Souza, L. M. and Lai, P. H. (1989) *Arch. Biochem. Biophys.* **268**, 264–274
- 21 Boyd, V. L., Bozzini, M., Zon, G., Noble, R. L. and Mattaliano, R. L. (1992) *Anal. Biochem.* **206**, 344–352
- 22 Pharmaceutical Manufacturing Association (1991) *Biopharm.* **5**, 22–27
- 23 Hatfield, D. (1985) *Trends Biochem.*, May issue, 201–204
- 24 Eggertsson, G. and Soll, D. (1988) *Microbiol. Rev.* **52**, 354–374
- 25 Fox, T. D. (1987) *Annu. Rev. Genet.* **21**, 67–91
- 26 Parker, J. (1988) *Microbiol. Rev.* **53**, 273–298
- 27 Craigen, W. J., Lee, C. C. and Caskey, C. T. (1990) *Mol. Microbiol.* **4**, 861–865
- 28 Leinfelder, W., Zehelein, E., Mandrand-Berthelot, M.-A. and Bock, A. (1988) *Nature (London)* **331**, 723–725
- 29 Soll, D. (1988) *Nature (London)* **331**, 662–663
- 30 Kopelowitz, J., Hampe, C., Goldman, R., Reches, M. and Engelberg-Kulka, H. (1992) *J. Mol. Biol.* **225**, 261–269
- 31 Hui, J., Meng, S.-Y., Katta, V., Tsai, L., Rohde, M. F. and Haniu, M. (1995) in *Techniques in Protein Chemistry* (Crabb, J., ed.), vol. 6, Academic Press, San Diego, in the press
- 32 Clements, J. M., Bawden, L. J., Bloxidge, R. E. et al. (1991) *EMBO J.* **10**, 4113–4120
- 33 Oefner, C., D'Arcy, A., Winkler, F. K., Eggimann, B. and Hosang, M. (1992) *EMBO J.* **11**, 3921–3926
- 34 Raines, E. W. and Ross, R. (1992) *J. Cell Biol.* **116**, 533–543