ELSEVIER

Journal of Chromatography A, 806 (1998) 47–60

JOURNAL OF CHROMATOGRAPHY A

Quantitative analysis of gene expression by ion-pair high-performance liquid chromatography

P.A. Doris[a,*], P.J. Oefner[b], B.S. Chilton[a], A. Hayward-Lester[a]

[a]*Deparment of Cell Biology and Biochemistry, Texas Tech University Health Sciences Center, Lubbock, TX 79430, USA*
[b]*Department of Biochemistry, Stanford University, Stanford, CA, USA*

Abstract

We have analyzed the utility of ion-pair reversed-phase HPLC for gene quantification by competitive reverse transcriptase polymerase chain reaction (RT-PCR). Competitive RT-PCR reactions employed various RNA competitors which shared high sequence similarity to the native transcripts for which they served as references. Competitive reactions resulted in the detection of two reaction products when reactions were analyzed by agarose gel electrophoresis, but three products when analyzed by HPLC. The third product was demonstrated to be a heteroduplex formed between mixed strands of native and competitor amplicons. Mathematical analysis of these competitive reactions indicated that identification and quantification of the heteroduplexes were essential to produce accurate gene quantification. PCR amplification efficiency was shown to be identical for native and competitor transcripts. However, RT efficiency differences were observed which may be sequence dependent. These differences were highly consistent across reactions for the same native and competitor inputs. Increasing the sequence similarity resulted in a competitor which had the same RT efficiency as the native transcript. Titration of various levels of competitor against native RNA resulted in the expected linear relationships which had slopes of unity. Quantitation could be performed with similar precision in single tube comparisons in which the initial abundance of the native transcript was calculated by knowledge of the final reaction product ratio and the initial competitor input level. The assay system is highly accurate, i.e. the measured level of gene expression reflected the actual copy number of the gene present in the sample. This was demonstrated by performing reactions in which known amounts of native transcript were quantified and the amount estimated by the assay was shown to be the same as the known amount added to the reaction. A similar approach has been devised for examining the relative levels of alternatively spliced isoforms. In this system, primers were selected to produce reaction products which served as their own internal competitors (by spanning the alternative splice site). Hormonal dependence of the ratio of abundance of two isoforms of the rabbit RUSH-1 gene was demonstrated.

Keywords: Gene expression; RNA

1. Introduction

Techniques to quantify gene expression by RT-PCR have been developed in part because of the limitations of other techniques such as solution hybridization, in situ hybridization and blotting approaches. These limitations include the absence of precise quantification (blotting and in situ hybridization) and the necessity of collecting relatively large amounts of RNA to perform the assay and generate sufficient signal for measurement (solution hybridization and blotting). A technique is needed which provides accurate quantification (the term ''absolute'' has been used in the past to distinguish

*Corresponding author.

0021-9673/98/$19.00
PII S0021-9673(97)00514-1

accurate from relative quantification, however, this use is inappropriate since ''accurate'' means that the estimate produced by a system reflects the actual, not relative, amount of material present in the sample) and which can be applied successfully to small tissue samples and biopsies.

2. Competitive RT-PCR

In order to overcome these limitations, an RT-PCR assay system using competitive amplification of deletion or insertion mutant RNA has been developed. This approach holds the same advantage as solution hybridization in allowing quantification of gene expression to be readily accomplished. It has the further advantage of offering far greater sensitivity so that it can be applied to analyze tissue at the microscopic level. Recent accomplishments with this technique to date requires an understanding of the principles of the recently developed technique of competitive RT-PCR using RNA competitors [1–3].

PCR is a powerful tool which allows the detection of very small amounts of DNA. Since the expression of a gene results in the production of mRNA, not DNA, it is necessary to adapt PCR to this situation by converting RNA extracts of the tissue to DNA. This is performed with reverse transcriptase which transcribes first-strand DNA from a single stranded RNA template. When the target of specific nucleic acid quantification (SNAQ) is DNA (for example, a DNA virus), the process is simplified and the exacting handling requirements of RNA (which is easily degraded) are removed.

3. Problems in quantitation arising from the exponential amplification of DNA in PCR

The next problem is to apply this technique in a way that allows quantitative measurements of gene expression. This is a difficult problem because PCR is a replicative technique in which DNA is detected because the original DNA in the sample is multiplied by numerous cycles of replication. With exponential increases occurring in repetitive cycles, small differences between samples in efficiency of amplification due to such things as temperature, buffer composition, volume, or presence of contaminants means that the comparison of the amount of DNA amplified from one sample with that amplified from another can become risky at best and meaningless if conditions between samples vary only trivially from identical. As can be seen from the equation below, the effect of variations in efficiency (E) can lead to dramatic variations in product yield.

$$N = N_0(1 + E)^n \quad (1)$$

Where N is the final amount of reaction product, N_0 is the initial amount of DNA in the reaction, n is the number of cycles and E is the efficiency of the reaction. Thus, if variations in reaction efficiency of only 5% occur between two samples being compared in two different reaction tubes, changes of product amount of more than 100% after 30 reaction cycles will result and will make meaningless comparison of gene expression levels between the samples.

Competitive PCR is designed to overcome this limitation. The idea is to amplify not only the DNA of interest in the PCR reaction, but also another DNA template (competitor) which can be amplified by exactly the same primers. If the amount of the competitor DNA in the starting sample is known or is always constant between samples it can be used as a quantitative reference. Of course, the primers were designed to amplify only DNA encoding the gene sequence of the gene of interest and amplification of competitor product requires engineering of a DNA sequence to serve in this role [1–3].

This engineering begins with the cDNA of the genes of interest. The work which has been accomplished to date with this system has advanced furthest with studies of the sodium pump (sodium, potassium-ATPase or NKA) genes of the rat. The sodium pump includes a catalytic alpha subunit which exists in three isoforms, each is the product of a separate gene, but each shares considerable sequence homology with one another. These isoforms provide a good illustration of our application of this technique to examine gene expression in very small tissue samples (isolated microdissected nephron segments). They also illustrate well the advantages of quantitative gene determinations over immunological approaches to detect and quantify specific proteins. An advantage of the SNAQ approach is that spe-

cificity (limitation of detection of products to one isoform at the exclusion of others) can be preserved by use of specific primers which have no complementarity with other members of the isoform family.

We have transferred cDNA's for three alpha NKA isoforms (alpha1, alpha2 and alpha3) into plasmid vectors which contain RNA polymerase promoters. The use of three isoforms, each the product of different genes illustrates well the advantage of template specificity that PCR amplification can have over other techniques by use of specific primers designed to complement heterologous regions of base sequence in each isofrom. The use of cDNA's in plasmid constructs allows us to synthesize alpha isoform RNA from each plasmid. Each cDNA was engineered by removing a short section or adding an additional insert of DNA in the middle of the sequence. Thus, the RNA synthesized from these engineered plasmids will be either shorter or longer than the native alpha isoform mRNA sequence but will be highly similar in sequence and at the same time retain the sequences required for binding the isoform specific primers. This RNA is referred to as the competitor RNA. The size difference provides a means of distinguishing the two reaction products.

RNA synthesized from these competitor DNA templates can be added in fixed, known amounts to the reverse transcriptase step in which RNA from tissue or cell extracts is converted into DNA for PCR. The resulting DNA will then contain both native sequence DNA and competitor DNA. The amplified PCR products derived from these templates can then be analyzed and the quantity of gene expression in any sample can be related to the known quantity of competitor RNA added to the starting tissue RNA preparation.

A critical element in the design of a reliable SNAQ system is the choice of method to separate and quantify the reaction products. Novel HPLC stationary phase material which is capable of separating double-stranded DNA rapidly, in a size-dependent manner and with very high efficiency, has recently been devised and is now commercially available (DNASep™, Sarasep, Santa Clara, CA, USA). These materials are reversed phase coatings which are applied over novel polymeric packing materials capable of tolerating high pressures and therefore permitting the high flow rates required to make HPLC a viable technique for routine separation of PCR reaction products [4,5].

4. HPLC separation of DNA by ion-pair reversed phase HPLC

Fig. 1 shows the HPLC separation of pBR322-HaeIII pBR322-MspI digest mixture. The use of on-line 254 nm UV detection provides very high sensitivity. Quantities of DNA as low as 5–10 ng can be detected, depending on size of the fragment. The resolution of the column is very high as indicated by the separation of closely-related size bands. Size-dependence of elution is another important characteristic of this separation method.

This technique lends itself ideally to the use of PCR for quantitative gene expression measurements. By employing internal standards as controls, the

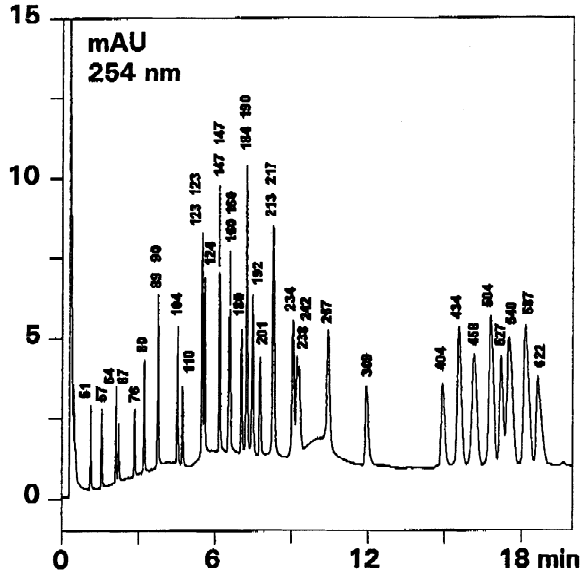

Fig. 1. Chromatogram of pBR322-HaeIII digest (0.75 μg) and pBR322-MspI digest (0.4 μg) on PS-DVB-C_{18} HPLC column (50×4.6 mm I.D.). Fragment sizes are labelled. Elution profile was modified to optimize separation of the large number of fragments present. Mobile phase composition: A=0.1 *M* triethylammonium acetate (TEAA), pH 7.0, B=0.1 *M* TEAA pH 7.0, 25% acetonitrile. Gradient: linear 37–55% B in 6 min; 55–65% B in 14 min; flow rate 1 ml/min; temp.=50°C; detection UV 254 nm.

level of gene expression in a tissue sample can be accurately related to the amount of internal control added. The PCR reaction products can be analyzed immediately, with no further processing and no requirement to incorporate radiolabels, make, load and run gels, worry about losses due to unequal gel loading, inefficient or variable gel blotting or autoradiographs outside the linear detection range. By automation of sample loading, the products of 100 PCR reactions can be analyzed easily within a day with no more labor needed than to load the PCR reaction tubes into a refrigerated autosampler tray. Other techniques which provide rapid separation of double-stranded DNA have been developed. The most advanced of these is capillary zone electrophoresis (CZE). However, HPLC offers some important advantages over CZE [6]. Gel scanning techniques provide sensitivity equal to that obtained by HPLC (with use of fluorescent primers [7]). However, the equipment is three or more times as expensive, does not provide the automation or flexibility available in HPLC and requires much greater sample handling.

Highly efficient DNA separations can be obtained on alkylated poly(styrene–divinylbenzene) particles using an ion-pairing reagent, typically tri- or tetraalkylammonium salts, in the mobile phase [5,7,8]. The ion-pairing reagent converts the stationary phase into a dynamic anion-exchanger. The number of methylene groups in the alkyl chains of the ion-pairing reagent also determines the extent to which the surface of the beads retains its hydrophobic properties, which in turn exerts a significant effect on the mode of separation. Short alkyl chains (number of methylene groups per alkyl chain$\leq$2) result in incomplete coverage and the stationary phase retains partially its hydrophobic or reversed phase properties. Longer alkyl chains create complete coverage and anion-exchange becomes the predominant mechanism of separation [7]. This also explains why the separation of single-stranded oligonucleotides is so dependent on the ion-pairing reagent chosen. Triethylammonium ions, for instance, cover the stationary phase only partially and therefore separation is not only governed by size, but also by base composition. Tetrabutylammonium ions, on the other hand, result in complete coverage of the stationary phase, consequently separation of single-stranded oligonucleotides is almost strictly size-dependent [7]. The reverse situation is observed in the separation of double-stranded DNA. Tri- and tetraalkylammonium salts with short alkyl chains are able to ion-pair specifically with AT-base pairs. This suppresses their preferential interaction with the dynamic anion-exchanger and enables a degree of size-dependent separation otherwise only observed in denaturing polyacrylamide slab gel electrophoresis [8].

Several important advantages are provided by ion-pair reversed phase HPLC technology. First, PCR reaction products can be analyzed without further treatment, i.e. crude reaction product can be injected directly into the column. Second, the sensitivity of UV detectors makes detection of reaction products simple and accurate. Third, the resolving power of these columns is sufficient to resolve small differences in size (in fact, the relationship between size and elution in HPLC appears to be more reliable than in agarose electrophoresis [8]). Fourth, quantitative PCR techniques have involved (prior to our introduction of HPLC technology) the generation of titration curves which result in large numbers of samples. The ability to automate sample loading and analysis using HPLC autoloaders and computerized data acquisition and analysis software provides a major advance over other methods in efficiency and productivity. This means that there is the potential to increase from 100 to about 200 analyses per day with minimal hands on involvement.

5. Application of HPLC to separation of competitive RT-PCR products

Fig. 2 illustrates the application of HPLC to separation of RT-PCR reaction products produced in a competitive reaction titration. Chromatograms of reaction products are inset in the corresponding lanes of an agarose gel which was also used to examine the same reaction products. The expected competition of native and competitor are clearly visible on the gel as competitor input increases progressively from reaction 1 to 8. However, densitometric analysis of such gels do not indicate that this reaction is meeting the mathematical criteria expected and discussed below.

The reason becomes apparent when the chromato-

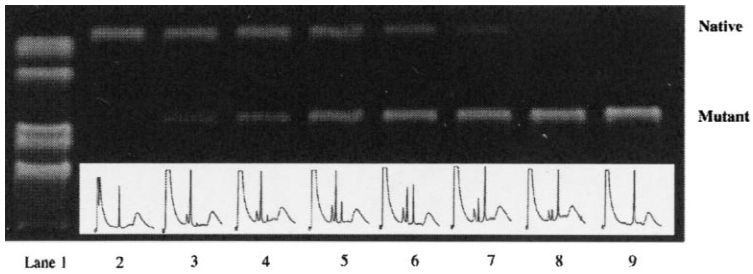

Fig. 2. Analysis by HPLC and agarose gel electrophoresis of products from a single RT-PCR titration series. Aliquots from RT-PCR (15 μl) were run on a 5% Metaphor gel (FMC Rockland, ME) at 50 V for 2.5 h in 1x Tris-phosphate-EDTA buffer. The gel was post-stained with ethidium bromide, photographed under UV illumination and analyzed by densitometry using NIH Image 1.55. For HPLC, aliquots of the same reaction products (6 μl) were injected sequentially onto the PS–DVB HPLC column at approximately 6.5 min intervals. The elution system was a gradient of acetonitrile in 0.1 *M* TEAA, pH 7, 1ml/min. UV detection was at 254 nm and the UV absorbance signal was analyzed by an on-line integrator. No processing of PCR products was required prior to injection. Each reaction in the titration contained 25 ng rat brain total RNA input with varying competitor input. On the gel, lane 1 is a pUC18 HaeIII ladder, lane 2 contains RT-PCR products of native RNA only, lane 3–8 contain competitive reaction products in which native RNA was subject to RT-PCR with 0.25, 0.5, 1, 2, 4 and 8 pg competitor, respectively, and lane 9 contains RT-PCR products of 4 pg competitor RNA only. All reverse transcription in these and other experiments described in this paper employed random hexameric priming. Only two bands are evident from the competition reactions on gel electrophoresis, whereas the HPLC chromatograms reveal the presence of a heteroduplex product in reactions in which both inputs are present. Subsequent regression revealed that the densitometric analysis of the gel provided an estimate of molecule number approximately one third that estimated by the HPLC analysis. This indicates that most of the heteroduplex product co-migrates with the competitor product during agarose electrophoresis, overestimating the amount of competitor product formed and causing erroneous quantification. It is an important property of this HPLC system to reliably resolve heteroduplex from homoduplex products because the formation of heteroduplexes invalidates the mathematical relationship between the ratio of homoduplex products. For accurate quantification, the peak area of heteroduplex products is divided in proportion to the UV absorbance contributed by each heteroduplex component (native strand – 39% and competitor strand – 61%) and each component is added back to its appropriate homoduplex peak prior to calculation of native/competitor product ratios.

gram is examined. Each chromatogram has an initial peak in the void volume of unretained reaction components (salts, water, free nucleotides, primers, proteins). Following this are the peaks corresponding to specific reaction products. Reactions which appear to produce two specific products on gel electrophoresis clearly produce three specific products on HPLC (for example, lanes 5 and 6). The third product (actually eluting first) is a heteroduplex made up of one strand (forward or reverse) of native and one strand of competitor DNA. We know this because if two reactions, each of which has produced only one (native or competitor) product, are mixed, heated and cooled to allow reannealing, then three HPLC products are observed on analysis of this mixture, though only two products are seen in this sample on gel electrophoresis. Analysis of mixtures which have not been heated to produce melting and reannealing result in only the two homoduplex products being observed in both HPLC and gel electrophoresis. We believe these observations are explained by the fact that the cross-sectional area of a heteroduplex (the attribute which principally determines gel migration) resembles closely that of a homoduplex made up of the two larger strands. In HPLC, retention is determined by the number of matched base pairs. In a heteroduplex, this will always be slightly less than the number occurring in the smaller of the two homoduplex products. The inability of gel electrophoresis to detect heteroduplexes has led to confusion about whether and when they are formed in competitive PCR reactions in which homologous internal standards are employed. HPLC analysis of these reactions has clarified this issue so that not only are heteroduplexes detected, but their impact on the analysis can be fully incorporated into data computation.

Benefits of this approach compared with conventional gel electrophoresis are that the HPLC column can be used for thousands of runs allowing the initial investment to be highly cost effective compared to the purchase of agarose or polyacrylamide needed to

give the same resolution in gel electrophoresis. Secondly, there is essentially no set up time, an important difference compared to time required to set up agarose or polyacrylamide gels. Thirdly, run times are very short, lasting 5 min per sample in the current version of our equipment. There is a 2 min re-equilibration of the column for the subsequent run. Fourthly, on line data acquisition from the UV detector means that an additional step is not required for quantification of bands. Using the established method of autoradiography, this step is time consuming and requires access to sophisticated and expensive image analysis workstations or phosphorimagers costing over $100,000. Finally, all reactions can be run without ^{32}P present, providing an added safety benefit.

Quantification of the amount of gene expression in an individual sample is accomplished as illustrated in the graphs below which show quantification of alpha 1 signal in RNA extracts from rat blood (Fig. 3a), quantification of the alpha 2 signal in rat brain (Fig. 3b), and quantification of the alpha 3 signal in rat brain (Fig. 3c). The sensitivity of the technique is well illustrated by the amplification of alpha 1 isoform mRNA from rat blood. All the RNA used in this titration came from a total of 100 μl of blood. Blood is comprised mostly of plasma and enucleate cells. Leukocytes represent the principal source of mRNA in this sample. Even at these low RNA input levels, analysis is performed far above the limits of sensitivity of the method.

Having made a rough determination of the amount of gene expression likely to be found in a given tissue type, accurate quantification is accomplished by RT-PCR reactions of uniform amounts of tissue RNA mixed with known quantities of competitor RNA. The resulting competition results in larger quantities of native message amplified when competitor RNA quantities are less and the reverse when the opposite is true. Since it is impossible to determine the exact amount of competitor which will be equal to the amount of native RNA message in the sample, this amount is estimated by plotting the logarithm of the ratio of the amount of PCR reaction products (native/competitor) against the log of the amount of competitor added to each reaction. When the log of the ratio is estimated at a value of 0 (log 1/1), the equation of the regression line can be

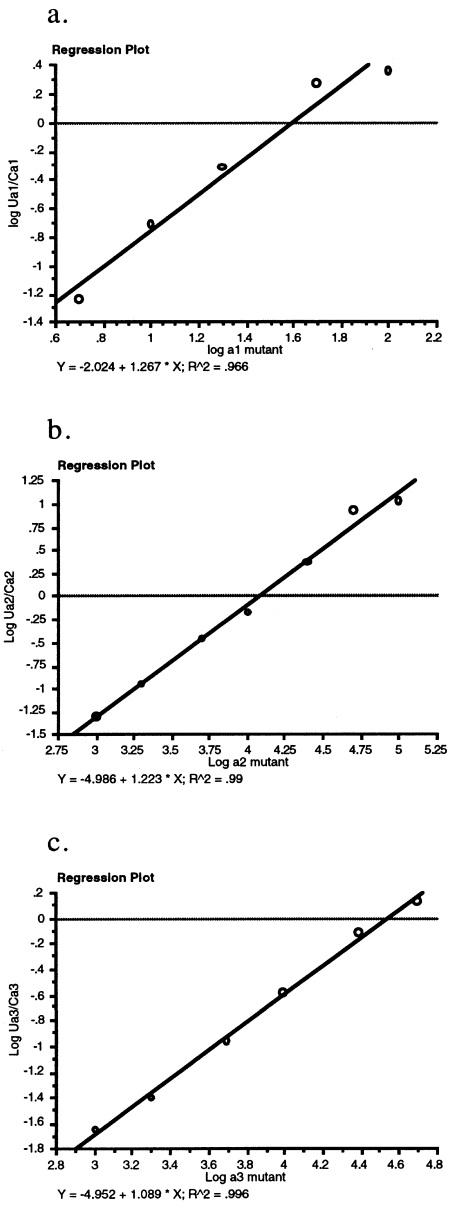

Fig. 3. (a) Estimate of alpha 1 gene expression in RNA extracted from rat blood. Plot of log of ratio of the amount of competitor to the log of the amount of native PCR reaction product determined as area under the curve in HPLC ($\log U_n/C_n$ in Eq. (5) below) against log of the initial amount of competitor RNA added to the RT-PCR reaction ($\log C_o$ in Eq. (5)). The amount of alpha 1 gene expression in this sample of rat blood total RNA was calculated as 263.7 fg/μg total RNA. (b) Estimate of alpha 2 gene expression in RNA extracted from rat brain. The data indicate that alpha 2 gene expression is 47.6 pg/μg total RNA. (c) Estimate of alpha 3 gene expression in RNA extracted from rat brain tissue. The data indicate that alpha 3 gene expression is 140.3 pg/μg total RNA.

used to determine the amount of gene expression as fg equivalents (or number of molecules) of competitor standard. Thus, the amount of specific mRNA in a sample can be determined. In fact, the real value of these competitive titrations is in demonstrating that the competitor and native amplicons are being amplified with equal efficiency. Once it has been demonstrated that the final product ratio of amplicons is identical to the initial product ratio in any given competitor/native system, repeated titration becomes unnecessary. A detailed discussion of the analysis of competitive reactions is presented below which is followed by further data demonstrating that accurate quantification can be obtained without titration in a system which has been shown to provide accurate SNAQ.

6. Theoretical questions concerning the reliability of the competitive RT-PCR approach

Although the deletion and insertion competitor strategy using competitor RNA in the RT reaction step is a major advance in permitting quantitative precision in determining gene expression levels, it is not without some theoretical problems. The most obvious of these are (1) that the processivity rate of Taq polymerase (about 50 bases/s) may be sufficiently low that during the 1 min elongation step in a typical PCR reaction, shorter target DNA's could be amplified with greater efficiency than longer ones and (2) there may be differences in the efficiency of reverse transcription which bias native or competitor signal generation. To understand the significance of these factors requires a mathematical approach to PCR amplification and the effects of variations in efficiency on accumulation of amplified products.

If the initial unknown amount of a gene U in a competitive RT-PCR reaction is U_o and that of its specific competitor RNA is C_o and these are subject to n reaction cycles in which the efficiency of amplification is E_u and E_c for the unknown and competitor respectively, then from Eq. (1) (above) we can describe the amount of reaction products at the end of n cycles by:

$$U_n = U_o \cdot (1 + E_u)^n \quad (2)$$

$$C_n = C_o \cdot (1 + E_c)^n \quad (3)$$

Making a ratio of Eq. (2) and Eq. (3) and taking the logarithm gives:

$$\log (U_n/C_n) = \log U_o - \log C_o + n \cdot \log [(1 + E_u)/(1 + E_c)] \quad (4)$$

A basic, but previously unproved, assumption of competitive RT-PCR is that E_u and E_c remain equal throughout the reaction. If true, Eq. (4) reduces to:

$$\log (U_n/C_n) = \log U_o - \log C_o \quad (5)$$

In calculating competitive RT-PCR reactions to obtain the unknown amount of a gene (U_o) present in a sample, a plot is made which relates $\log (U_n/C_n)$ to $\log C_o$, the known amount of starting competitor RNA (cf. Fig. 3a–c). This allows $\log U_o$ to be calculated. The estimated initial value for amount of native gene expressed in a sample (U_o) is the antilog of this value.

Eq. (5) indicates that such a plot will form a straight line having a slope of -1 (or of 1 if $\log(C_n/U_n)$ is plotted) because it is of the form $y = ax \pm b$ where the value of a must be 1 or -1 to follow the model. Most competitive PCR studies have paid attention to the property of linearity in this relationship, but less attention has been applied to the slope with numerous reports containing data in which the slope requirement fails, sometimes badly [9–11]. Our work amplifying alpha 1 NKA by RT-PCR from RNA extracts of blood has revealed correlation coefficients of 0.98 ± 0.01 (mean $\pm$ S.E.M., $n = 8$) when the data are fitted to a linear model. Thus, the property of linearity is met with very high confidence and consistency. The mean slope in these reactions was 1.04 ± 0.08 (mean $\pm$ S.E.M., $n = 8$). Thus, our reactions also appear to conform to theoretical predictions of slope.

Raeymaekers has modelled the effects of various factors which might influence the accuracy of competitive RT-PCR quantification [12]. He has pointed out that the basic assumption of $E_u = E_o$ may be preserved, even if the value of E_u and E_o decline as PCR reactions progress. This decline in efficiency could be due to thermal denaturation of DNA polymerase, reduction in primer concentration and possibly other factors. He has estimated the effect of

such a situation on accuracy of quantification and shown that both linearity and slope conditions can be preserved, while accurate (but not relative) quantification can be lost. Raeymaekers also proposes that other theoretical factors may invalidate accurate quantification. However, apart from our own work, there have been no data to test these theories either in Raeymaekers' analysis or in published reports of actual competitive RT-PCR quantifications.

One approach which has been employed to overcome some of the possible effects on amplification efficiency of size differences between native and competitor signals is the construction of a competitor by engineering in or removal of a unique restriction site. While this approach removes most, but not all, of the concern generated by size differences, it creates new concerns unique to this setting. These derive from the possibility that competitor DNA strands will hybridize to native DNA strands during the PCR reaction, thereby producing heteroduplex products which may be of maximum stability (single base difference between native and competitor strands). Since quantification depends on the resolution of reaction products by their ability to be cut by a restriction enzyme, the presence of heteroduplex products, undetectable on the basis of size, but not digested by the restriction enzyme is a major problem with this approach [13]. According to Raeymaekers' analysis, this type of competitor may also generate inaccurate calculations of levels of gene expression if amplification efficiency declines during the later cycles of PCR [12,14,15].

Having shown that our reactions meet requirements for relative quantitation when analyzed by HPLC, we have investigated whether we can make accurate quantification measurements by determining if our method gives the expected result when the reactions are performed using known starting amounts of both native and competitor RNA. Under usual circumstances, only the amount of competitor RNA added to the reaction is known and the amount of unknown message in the tissue RNA sample is calculated from the reaction products. However, by synthesizing native RNA by transcription from the plasmid cDNA insert, we have performed reactions which begin with known quantities of both competitor and native RNA sequences. If there is a decline in efficiency through the course of the RT-PCR amplification which is unequal for both native and competitor RNA signals and which results in an inaccurate estimate of amount of initial native RNA quantity, we will be able to detect it in these experiments.

We have performed these titrations over four orders of magnitude of RNA input (with replicate measurement at one level). We chose input levels to reflect the broad input range encountered from microscopic tissue samples up to more abundant RNA sources such as present in our rat brain RNA preparations. Yeast total RNA was added to the synthetic RNA solutions in order to mimic the background of non-specific RNA present in normal tissue RNA preparations. The regression lines of the resulting titrations are shown in Fig. 4.

These results indicate first that a series of 1/10 dilutions of native RNA yields the expected shift of the titration line by 1 log unit on the *x*-axis. However, Table 1 shows the relationship between the known input RNA amount and the estimated amount.

There is clearly an estimation error (3.79 fold) which, as expected from the parallel and uniform displacement of the titration lines in Fig. 4, is remarkably consistent regardless of input amount. Two questions are relevant to this error: first; what causes it? and second, can it be removed?

We have been able to generate an answer to the first of these questions. We were inclined to believe that the error was not due to a difference in PCR amplification efficiency because the mathematical model which Raeymaekers developed indicated that such amplification efficiency differences would produce titration lines which were curved, while our reactions generate remarkably straight lines over a wide range of input levels. However, we tested whether PCR efficiency differed between native and competitor inputs by amplifying from known amounts of starting DNA, thus avoiding the RT step. This was accomplished using DNA from the plasmid constructs which normally serve as templates for native and competitor RNA synthesis. Our tests demonstrated no difference in amplification efficiency comparing products from known amounts of native and competitor DNA input (Table 2). We conclude that RT efficiency differences can occur in RNA templates which share considerable sequence

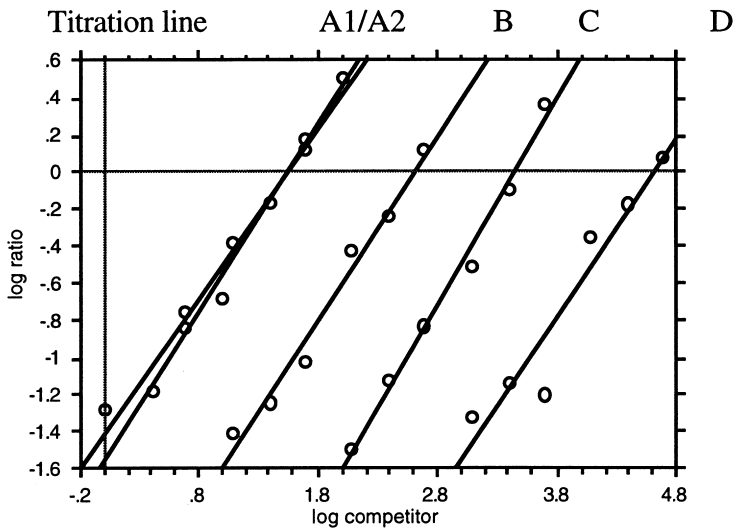

Fig. 4. Titration of known RNA inputs at four input levels (1.85×10^4 to 1850×10^4 molecules). The titration lines shifted leftwards as the input RNA was diluted over four orders of magnitude. Slopes and R^2 values were all indistinguishable from 1. Slopes of individual lines did not differ from one another (ANCOVA).

Table 1
Relationship between actual known input RNA amount and calculated amount over 4 log input amounts

	Mean	S.E.M.
Estimation discrepancy	3.79	0.2
Slope	1.01	0.013
R^2	0.99	0.005

identity, however, because the efficiency difference is constant over a very wide concentration range, once a correction factor is determined, accurate quantification can be achieved.

We have also obtained a partial answer to the second question posed above: can the correction factor be removed? The PS–DVB HPLC column permits size differences as small as 10 bases to be resolved. So, to investigate the possibility that the RT efficiency difference was due to the relatively large size difference between the original native and competitor templates (68% homology), a competitor was generated which differed by only 14 base pairs (internal deletion) from the native amplicon. When this mutant was used in competitive RT-PCR reactions with known amounts of transcribed native RNA a difference between the estimated amount of native RNA input compared with the known amount could not be detected (Table 3).

7. Detection of alpha 1 isoform gene expression in microdissected nephron segments

Individual nephron segments were microdissected from young adult rat kidneys which had been perfused with collagenase and sliced. The slices were further incubated for 15 min at 37°C in an oxygenated medium containing collagenase. Collagenase

Table 2
Estimation error when known DNA inputs are used instead of known RNA inputs

	Mean	S.E.M.
Estimation discrepancy	1.07	0.064
Slope	1.02	0.023
R^2	0.99	0.026

Table 3
RT-PCR quantification of known inputs of RNA using a highly homologus competitor RNA

	Mean	S.E.M.
Estimation discrepancy	0.945	0.064
Slope	0.978	0.032
R^2	0.989	0.005

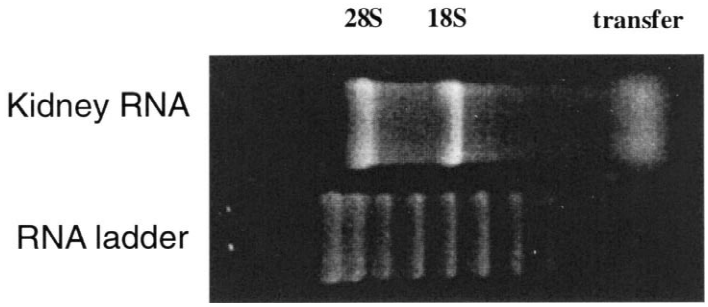

Fig. 5. Denaturing agarose gel electrophoresis of RNA extracted from collagenase-incubated slices of collagenase-perfused rat kidney.

was removed, tissue was rinsed and nephrons were dissected in the presence of vanadyl ribonucleoside (RNAse inhibitor). Individual nephron segments were identified, dissected free and their length was measured by a calibrated grid in the dissecting microscope lens. Total RNA was then immediately prepared from these measured segments using RNAzol extraction. Total RNA from proximal convoluted tubules was used to determine if RT-PCR reactions produced detectable product. Abundant products were obtained after 35–45 amplification cycles.

RNA quality has been evaluated to ensure that RT-PCR reactions using RNA from dissected nephron tissue did not employ degraded RNA. RNA was obtained by RNAzol extraction of total RNA from perfused, sliced and collagenase-dissociated renal tissue immediately after the addition of RNAse inhibitor. Fig. 5 shows a denaturing RNA gel which clearly indicates the presence of intact 18 and 28s RNA bands. Thus, we can safely conclude that our measurements of nephron gene expression are not affected by RNA deterioration or the possibility that RNAse activity is present in perfusion/incubation solutions.

8. Quantification of alpha 1 gene expression in nephron segments

To determine the feasibility of quantifying the alpha 1 gene expression detected in microdissected proximal tubule segments (corresponding to less than 1 μl of tissue), we performed quantitative titration analysis. The regression analysis of the titration obtained when increasing amounts of competitor alpha 1 RNA were added to total RNA from proximal tubules is shown in Fig. 6. Estimates of the

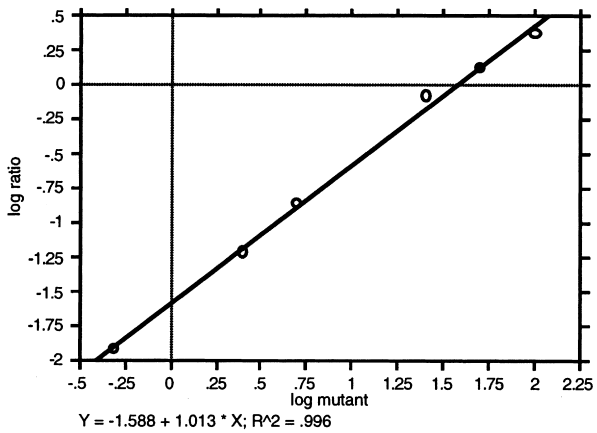

Fig. 6. Regression of competitor input (fg) against ratio of competitor to native proximal tubule alpha 1 RNA products.

Table 4
RT-PCR quantification of alpha 1 expression in nephron segments (molecules/0.0625 mm of nephron)

Nephron segment	Mean	S.E.M.
PCT	2761	483
CCT	617	144
MCT	314	104

quantity of alpha 1 gene expression in RNA extracted from several functionally distinct nephron segments is shown in Table 4.

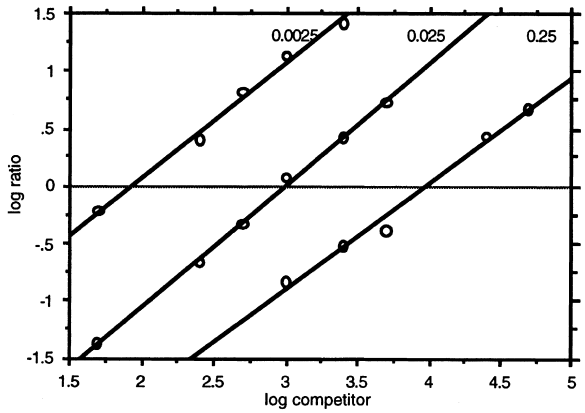

Fig. 7. Effect of varying amount of input RNA from brain tissue (bRNA) on quantitative accuracy. Brain RNA was analyzed by titration of competitor RNA against 0.25, 0.025 and 0.0025 μg bRNA. Effect of this dilution is to displace titration lines by one order of magnitude. Slopes of lines were not significantly different (Analysis of Covariance).

9. Precision and accuracy of measurement

Precision measures the ability to reproduce an estimate on repeated measurement(s). We have assessed precision in our system by repeatedly quantifying the amount of alpha 1 gene expression in a single sample of brain total RNA. The mean (±S.E.M.) obtained was 53.1±1.97 pg alpha 1 expression/μg total RNA which generates a coefficient of variation of 8.3%.

Accuracy is the ability to make a measurement which reflects the true level of the measured parameter. We have taken two approaches to determining accuracy. The first is to estimate accuracy at a known level of alpha 1 native RNA input. The results of this test of accuracy have been presented above.

In our second approach to establishing accuracy, we have tested varying levels of RNA input by dilution of brain total RNA over three orders of magnitude. Fig. 7 below shows the titration lines obtained for each dilution. Note that the point of intersection of the value of log ratio of 0 shifts by an order of magnitude with each order of magnitude dilution. The calculated amounts of gene expression determined (at each dilution) were 37.5 pg/μg brain RNA (0.25 μg brain RNA/reaction), 39.2 pg/μg brain RNA (0.025 μg RNA), 33.9 pg/μg brain RNA (0.0025 μg RNA), respectively. This indicates that the accuracy of the method when applied to total RNA extracted from tissue is highly independent of the level of input native RNA.

Precision describes the variation observed between replicate estimates of the same sample and can therefore vary depending on the sample, generally declining with the abundance of the signal in the sample. Fig. 8 (left bars) indicates that titration analysis of samples expected to contain a high and low abundance of specific target results in high levels of precision for both samples. However, such titrations involve multiple reactions to analyze a single sample and are therefore more expensive and time consuming than a single point measurement for each sample. We have investigated the possibility to simplify quantification by the elimination of titration. This approach is valid in a system which has been previously demonstrated by titration analysis to conform to the requirements of accurate quantification. If, in fact, the RT efficiency and PCR amplification efficiency are identical for competitor and native amplicons, then the final ratio of reaction products in a single reaction tube will always be identical to the initial ratio. Since one component of the initial ratio (amount of RNA competitor added) is known, the other component (the initial amount of unknown native RNA) can be estimated from the final product ratio. We have assessed the accuracy of single tube quantification in samples of brain and nephron RNA. The results are also shown in Fig. 8 above (right bars) and indicate that similar precision can be obtained using single tube quantification, even in low

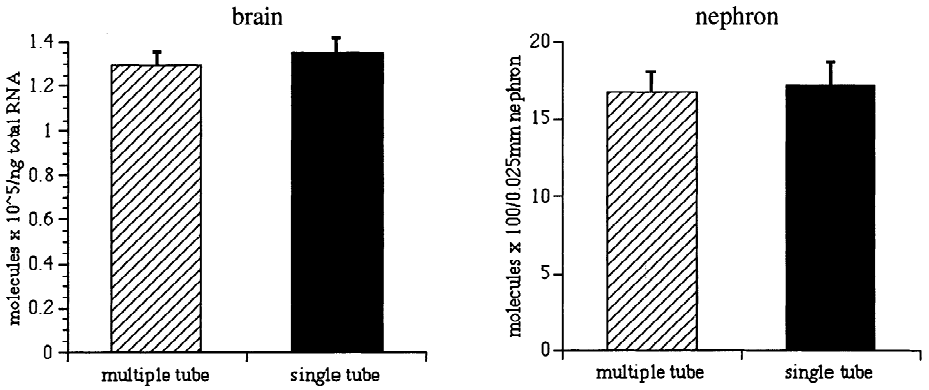

Fig. 8. Estimation of precision of measurement of A1NKA RNA comparing multiple tube titration with single tube quantification for a high abundance (brain) and low abundance (nephron) samples. For each tissue, multiple measurements were made from one RNA sample and the results are shown as mean±S.E.M.. This data confirms that, once the system is demonstrated to obey theoretical requirements for quantification, titration is not required and precise estimates of gene expression can be obtained from the analysis of products of a single tube reaction.

abundance samples such as microdissected nephron segments.

10. Application of HPLC to the detection and quantification of alternative splicing

We have developed the use of HPLC for the detection and quantification of alternatively spliced mRNA molecules. RT-PCR is performed using total or poly(A+) RNA and a pair of primers designed to encompass the sequence which is alternatively spliced. If more than one splice variant is expressed, HPLC will reveal both the specific homoduplex products and the heteroduplex products. The presence of heteroduplex products confirms that two specific PCR products amplified by the same primer pair share significant homology. RT-PCR and HPLC can then be used to quantitate changes in the relative levels of the two messages between different samples and under conditions of physiological and pharmacological manipulation. This can provide a measure of regulated splicing efficiency. We applied this technology to the study of alternatively spliced uteroglobin promoter-binding proteins in rabbit endometrial tissue.

In the rabbit uterus, uteroglobin (UG) is a progesterone-induced preimplantation secretory protein, which binds progesterone and has immunosuppressive and anti-inflammatory properties [16,17]. Its transcription is stimulated by progesterone [18–20]. UG is present in human endometrium [21], where like rabbit UG, the highest level of expression occurs in response to progesterone dominance and coincides with the events of embryo implantation. The recent localization of a high-affinity UG-binding protein on human trophoblast cells [22] suggests that UG may play a pivotal role in regulating cellular invasiveness.

The UG gene contains two strong and two weak progesterone receptor binding sites that are located between positions −2.7 kb and −2.3 kb [23]. Gel shift assays, Southwestern blots and UV-crosslinking in situ identify four proteins [24] that bind to the UG promoter (−194/+9). cDNA's for two of the UG promoter binding (RUSH-1β and RUSH-1α, M_r 95 000 and 113 000, respectively) were cloned by recognition site screening [25]. RUSH-1β is a truncated version of RUSH-1α that results from alternative splicing of a 57-bp exon. Northern analysis of uterine endometrial poly(A+) RNA revealed a single RUSH band of approximately 5.2-kb. Quantification of autoradiograms by computer-assisted image analysis indicated that progesterone increased the uterine content of message for the RUSH proteins in estrous rabbits. However, because the RUSH-1α and RUSH-1β proteins result from alternative splicing of a 57-bp exon, the small difference in their mRNA sizes could not be detected by Northern analysis.

There may be hormone-dependent differences in the relative levels of the alternatively spliced RUSH mRNA's. To assess this possibility, primers were designed to nucleotide sequences on either side of the 57-bp insert. This results in the amplification of two templates (282 and 225 bp, respectively) which differ only by the presence or absence of the alternatively spliced insert. Amplification of these templates simulates a competitive RT-PCR reaction. Initial ratios of these highly similar templates reflect the proportion of the alternatively spliced transcripts which are preserved during amplification. However, heteroduplex formation can occur between highly similar amplicons and must be both detected and quantified to provide an accurate estimate of initial ratios.

Analysis of competitive RT-PCR reaction products by agarose gel electrophoresis and ion-pair reversed phase HPLC is shown in Fig. 9. Metaphor agarose gel electrophoresis demonstrated the two principal reaction products visualized by ethidium bromide fluorescence, the ratios of which changed according to the hormonal condition of the animals. A third reaction product was also detectable by gel electrophoresis. HPLC analysis revealed that it eluted immediately prior to the smaller homoduplex product, which identified it as a heteroduplex from the annealing of the two principal reaction products. A spreadsheet was used to mathematically reallocate the absorbance due to the heteroduplex (A) and to calculate the ratio [(B) mRNA for RUSH-1α: (C) mRNA for RUSH-1β] of the two reaction products which is proportional to splicing efficiency.

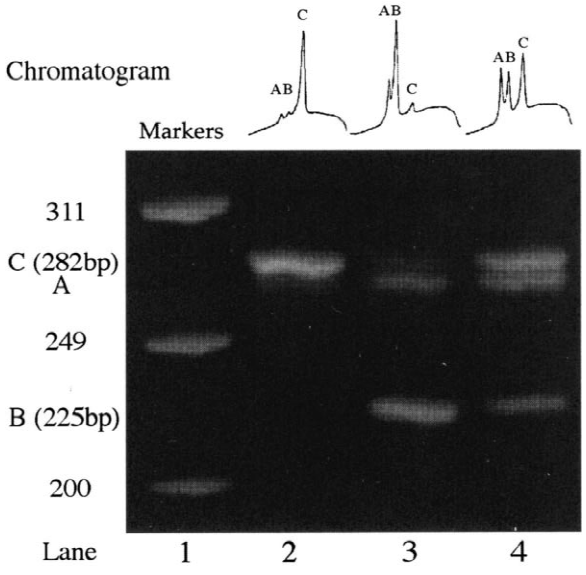

Fig. 9. Analysis of competitive RT-PCR products by HPLC and electrophoresis. Product A = heteroduplexes formed between RUSH-1α and RUSH-1β amplicons. Product B is from RUSH-1α amplification and product C is from RUSH-1β amplification. Lane 1 illustrates ø×174/Hinf I molecular size markers, and lanes 2–4 illustrate competitive RT-PCR products with mRNA from animals treated as follows: oestrous controls (2), oestrous + progesterone (3) and oestrous + progesterone + estradiol benzoate (4). (Reprinted with permission of Birkhauser Inc.).

The mRNA for RUSH-1β was preferentially expressed in oestrous rabbits. RUSH-1α was a minor component of the expressed transcripts in these animals. Ovariectomy resulted in the complete absence of detecable RUSH-1α. Treatment with progesterone resulted in a reversal of the ratio of RUSH-1α and RUSH-1β so that RUSH-1α becomes the predominant transcript present. Thus hormone-dependent changes in the abundance of alternatively spliced isoforms of RUSH-1 was clearly revealed by this competitive RT-PCR approach.

Acknowledgements

We are grateful to Dr. Aveline Hewetson for her contribution to this work. Dr. Sandra Sabatini graciously provided access to nephron tissue. Betty Lonis provided outstanding technical assistance. This work was supported by National Institute of Health grants DDK45538 (PAD), IPO1-HG00205 (PJO) and HD29457 (BSC).

References

[1] M. Piatak, K.-C. Luk, B. Williams, J.D. Lifson, Biotechniques 14 (1993) 70.
[2] A.M. Wang, M.V. Doyle, D.F. Mark, Biochemistry 86 (1989) 9717.
[3] M. Becker-André, K. Hahlbrock, Nucl. Acids. Res. 17 (1989) 9437.
[4] C.G. Huber, P.J. Oefner, E. Preuss, G.K. Bonn, Nucl. Acids Res. 21 (1993) 1061.
[5] P.J. Oefner, C.G. Huber, E. Puchhammer-Stockl, F. Umlauft, K. Grunewald, G.K. Bonn, C. Kunz, Biotechniques 16 (1994) 898.

[6] S. Nathakarnkitkool, P.J. Oefner, G. Bartsch, M.A. Chin, G.K. Bonn, Electrophoresis 13 (1992) 18.
[7] P.J. Oefner, C.G. Huber, F. Umlauft, G.-N. Berti, E. Stimpfl, G.K. Bonn, Anal. Biochem. 223 (1994) 39.
[8] C.G. Huber, P.J. Oefner, G.K. Bonn, Anal. Chem. 67(3) (1995) 578.
[9] G. Gilliland, S. Perrin, K. Blanchard, H.F. Bunn, Proc. Natl. Acad. Sci. USA 87 (1990) 2725.
[10] P.D. Siebert, J.W. Larrick, Nature 359 (1992) 557.
[11] V. Zachar, R.A. Thomas, A.S. Goustin, Nucl. Acids Res. 21 (1993) 2017.
[12] L. Raeymaekers, Anal. Biochem. 214 (1993) 582.
[13] A. Hayward-Lester, P.J. Oefner, S. Sabatini, P.A. Doris, Genome Res. 5 (1995) 494.
[14] L. Raeymaekers, Eur. Cytokine Network 5 (1994) 57.
[15] L. Raeymaekers, Genome Res. 5 (1995) 91.
[16] L. Miele, E. Cordella-Miele, A.B. Mukherjee, Endocr. Rev. 8 (1987) 474.
[17] L. Miele, E. Cordella-Miele, G. Mantile, A. Peri, A.B. Mukherjee, J. Endocrinol. Invest. 17 (1994) 679.
[18] T. Chandra, D.W. Bullock, S.L. Woo, DNA 1 (1981) 19.
[19] X.Z. Shen, M.J. Tsai, D.W. Bullock, S.L. Woo, Endocrinology 112 (1983) 871.
[20] R. Snead, L. Day, T. Chandra, M. Mace Jr., D.W. Bullock, S.L. Woo, J. Biol. Chem. 256 (1981) 11911.
[21] A. Peri, B.D. Cowan, D. Bhartiya, L. Miele, L.K. Nieman, I.O. Nwaeze, A.B. Mukherjee, DNA Cell Biol. 13 (1994) 495.
[22] G.C. Kundu, G. Mantile, L. Miele, E. Cordella-Miele, A.B. Mukherjee, Proc. Natl. Acad. Sci. USA 93 (1996) 2915.
[23] K. Jantzen, H.P. Fritton, T. Igo-Kemenes, E. Espel, S. Janich, A.C. Cato, K. Mugele, M. Beato, Nucleic Acids Res. 15 (1987) 4535.
[24] S. Kleis-SanFrancisco, A. Hewetson, B.S. Chilton, Mol. Endocrinol. 7 (1993) 214.
[25] A. Hayward-Lester, A. Hewetson, E.G. Beale, P.J. Oefner, P.A. Doris, B.S. Chilton, Mol. Endocrinol. 10 (1996) 1335.