

Cloning, Expression and Modeling Studies of Somatotropin cDNA of Local Buffalo Breed-Nili Ravi

¹SUMBUL KHALID, ¹SARDAR FAISAL MEHMOOD, ¹MIRZA IMRAN SHAHZAD,
¹QUDSIA BASHIR, ²SYEDA ZAREEN SABA, ²ZARTASHA MUSTANSAR,
³AYESHA FATIMA AND ⁴AZRA KHANUM*

¹Department of Biochemistry, Pir Mehr Ali Shah Arid Agriculture University,
 Murree Road, Rawalpindi, Pakistan.

²International Islamic University, Islamabad, Pakistan.

³Muhammed Ali Jinnah University, Islamabad, Pakistan.

⁴Faculty of Sciences, Pir Mehr Ali Shah Arid Agriculture University,
 Murree Road, Rawalpindi, Pakistan.

(Received 31st July 2007, revised 16th January 2008)

Summary: Pakistan is primarily an agricultural country and livestock is a very important means of livelihood resource for millions of people in Pakistan. It not only provides essentials for human life in the form of milk, meat and eggs but also wool, hair, skin, blood, bones and manure. Buffalo, one of the important domesticated animals in Pakistan, is not only popular for milk but also for meat. With an increasing human population and rising standards of living the demand for milk and meat is also increasing. With availability of recombinant technology, somatotropin (ST), a non-glycosylated protein hormone, can bring about ten to forty percent increase in milk yield when it is injected into the milking species. Due to its importance in milk and meat production ST cDNA has been cloned from a number of species. We have cloned, sequenced and expressed the ST cDNA from a local buffalo breed: *Nili ravi* and carried out comparative modeling study. For this purpose total cellular RNA was extracted from the pituitary gland of *Nili ravi*, cDNA was synthesized and ST cDNA was amplified with the help of sets of primers designed on the basis of sequence data available from NCBI Gene Bank. The PCR amplified products were cloned in T/A cloning vector, sequenced and then put in pLEX expression vector to carry out expression studies. The ST cDNA sequence was further subjected to various modelling softwares to obtain a validated 3D model of the protein.

Introduction

Livestock is an important sector of agriculture in national economy and accounts for 40% of the agriculture value added and about 9.4% of total gross domestic product (GDP). Its net foreign exchange earnings were Rs. 7000 million in 2004-05, which is almost 12 % of the overall export earnings of the country. Nearly 30-35 million of the rural population is engaged in live stock raising, deriving 30-40% of their income from it [1]. The total live stock population in country comprises 24.8 million buffalo, 23.3 million cows and oxen, 24.6 million sheep, 52.8 million goats and 0.8 million camels [1]. Buffalo is an important domesticated animal of Pakistan.

As mentioned earlier with increasing human population and rising living standards, the demand for meat and milk is also increasing. Since ages, man is trying to cope with these demands by different conventional methods like proper live stock management, cross breeding and animal food management. The development of recombinant

DNA technology has provided a new tool to address this problem. The techniques available in genetic engineering have provided the boost to dairy industry by the production of large quantities of recombinant bovine somatotropin (rbST) [2]. The purpose of rbST is to enable animal to produce more milk than they naturally would. It works by altering gene expression of glucose transporters in the animal's mammary gland, skeletal muscle and omental fat. The ST facilitates the repatriating of glucose to the mammary gland, which in turn produces more milk [3].

In addition to positive effects on the efficiency of food production, the benefits of ST include effects on environmental impact through reduction in animal waste products and expenditures for feed production, including fertilizer and other inputs associated with growing, harvesting, processing and sorting animal feed [4]. Low doses of ST during the transition period and early lactation period improves dry matter intake (DMI), body

*To whom all correspondence should be addressed.

weight (BW), or body condition score (BCS); provokes positive changes in concentrations of somatotropin, insulin, insulin-like growth factor-I (IGF-I), glucose, nonesterified fatty acids, and calcium; or improves milk yield (MY) response without obvious adverse effects on calving and health status of Holstein cows [5]. The administration of exogenous ST has been shown to increase lactation performance in mammals ranging from laboratory animals to humans [6]. In case of farm animals, an increase in milk yield response with ST has been demonstrated in pigs, sheep, goats and cows. All dairy breeds and animals of different parity and genetic potential have shown milk yield responses to exogenous administration of bST [7, 8]. Therefore, keeping in view the importance of this hormone in milk and meat production, somatotropin cDNA has been cloned from a local water buffalo breed: *Nili ravi*, sequenced and expressed in bacteria using pLEX expression system. The sequence obtained has also been used for comparative modeling. Because the structural characterization of proteins enables scientists to understand the relationship between amino acids and the finally folded protein, relationship among different protein structures and their functional characteristics. Several computational techniques have been developed to elucidate the structures of proteins [9-12]. Comparative modeling is an imperative technique for predicting structures of unknown proteins. The technique essentially uses the amino acid sequence of the target protein and aligns it with the template amino acid sequence. The aligned sequence produces picture that shows regions of similarity or dissimilarity between the two proteins and helps to elucidate the three dimensional structure (3D). Hence, we have predicted the structure of somatotropin cloned from *Nili ravi* using the predicted structure of bovine somatotropin (bST) and human growth hormone (hGH) as templates [13,14].

Results and Discussion

Cloning and Expression Studies

The bovine ST cDNA was amplified with three different primers *i.e.*, sequence specific, random hexamers and oligo dTs and a product of 700 bp was obtained with each (Fig. 1). Next amplification was done by nested set of primers having restriction enzyme sites NdeI and EcoRI which amplified a 592 bp fragment (Fig. 2). The

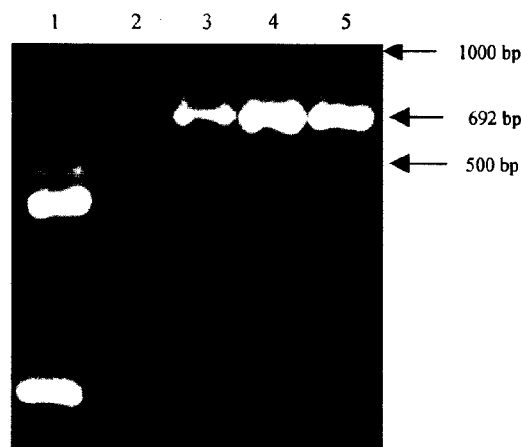


Fig. 1: Amplification of ST cDNA of *Nili ravi*. Lane 1: Positive control, Lane 2: GeneRuler™ 100bp DNA Ladder, Lane 3: cDNA with random hexamers, Lane 4: with oligo-dTs, and Lane 5: with gene specific primer.

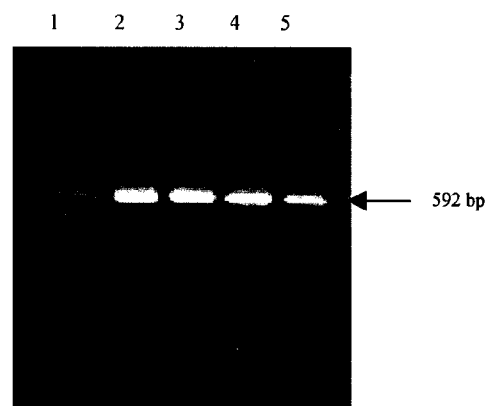


Fig. 2: PCR with nested primers. Lane# 1: GeneRuler™ 100bps DNA Ladder, Lane# 2 – 4: amplified fragment of ST with nested primers, Lane# 5: Positive Control.

result obtained is inconsistent with the results reported earlier [15].

The PCR product was then ligated in T/A cloning vector pTZ57R/T (Fermentas, K1214). The white colonies were selected from ampicillin resistant plates under blue and white screening method. Plasmid DNA was harvested and subjected to restriction digestion. The clones were finally confirmed by restriction digestion with NdeI and

EcoRI which produced a fragment of 592 bp (Fig. 3) and this fragment was further used for sub-cloning in pLEX expression vector which is a tryptophan-inducible expression system having the strong P_L promoter from bacteriophage lambda [16]. The vector was prepared for cloning by restricting with NdeI and EcoRI. Both vector and insert were purified from gel after restriction digestion by Perfectprep Gel Cleanup Kit (Eppendorf, Cat. No. 955152051) and ligation was done. The product was transformed in electrocompetant *E. coli* strain G 1724. The cells were grown on agar plate having ampicillin overnight at 37°C. Next day the several clones were picked and processed for confirmation of the presence of ST cDNA by digestion with NdeI and EcoRI REs. This double digestion step produced a fragment of 592 bp (Fig. 4). The positive clones were then used for expression of ST cDNA. After confirmation of the presence of ST cDNA in pLEX expression vector, the expression studies were carried out. The data obtained showed the expression of 22kD protein after the induction by tryptophan for 4 hours (Fig. 5). This study is in accordance to the expression studies carried out by Huai-jen *et al.* [17].

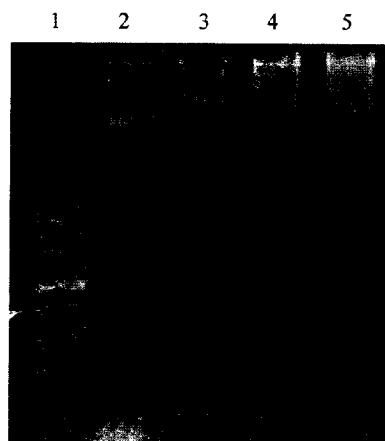


Fig. 3: T/A clone restricted with NdeI and EcoRI. Lane# 1: GeneRuler™ 100 bps Ladder, Lane# 2: unrestricted T/A vector, Lane#3: T/A vector restricted, Lane# 4 and 5: T/A vector having insert restricted with NdeI and EcoRI.

Modeling Studies

Blast Results

Basic Local Alignment Search Tool (BLAST) is an efficient tool that searches databases

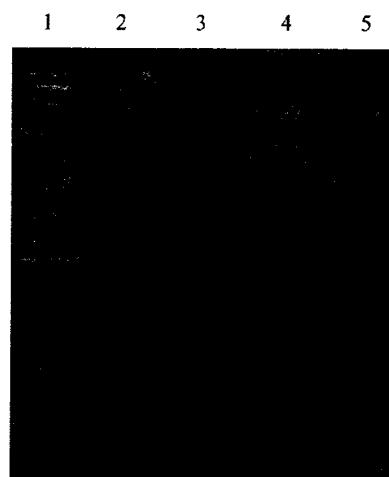


Fig. 4: pLEX clone restricted with NdeI and EcoRI. Lane# 1: 1 Kb DNA Ladder, Lane# 2: pLEX vector unrestricted, Lane# 3: pLEX vector restricted. Lane 4 and 5: pLEX vector with insert restricted with NdeI and EcoRI.

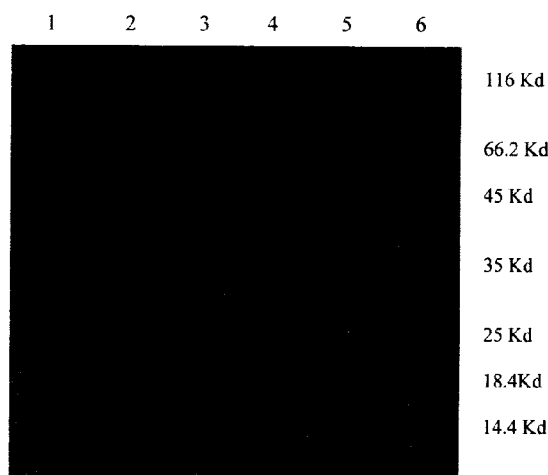


Fig. 5: SDS PAGE analysis of induction studies. Lane 1: Non transformed G1724 cells, uninduced, Lane 2: Non transformed G1724 induced, Lane 3: Transformed G1724 cells with pLEX vector having ST gene, uninduced, Lane 4: Transformed G1724 cells with pLEX vector having ST gene induced, Lane 5: Positive control of ST gene product, Lane 6: Standard protein markers.

gi|60651145|gb|AA31661.1| somatotropin [*Bubalus bubalis*] (Sadaf *et al.*, 2005) Length=191
 Score = 345 bits (886), Expect = 5e-94
 Identities = 189/191 (98%), Positives = 190/191 (99%), Gaps = 0/191 (0%)

```

Query 1  AFPAMSLSGLFANAVLRAQHLHQLAADTFKEFERTYIPEGQRYSIQNTQVAFCFSETIPA
Sbjct 1  AFPAMSLS L FANAVLRAQHLHQLAADTFKEFERTYIPEGQRYSIQNTQVAFCFSETIPA
AFPAMSLSSLFANAVLRAQHLHQLAADTFKEFERTYIPEGQRYSIQNTQVAFCFSETIPA

Query 61  PTGKNEAQQKSDLELLRISLLLIQSWLGPLQFLSRVFTNSLVFGTSDRVYEKLDLEEGI
Sbjct 61  PTGKNEAQQKSDLELLRISLLLIQSWLGPLQFLSRVFTNSLVFGTSDRVYEKLDLEEGI
PTGKNEAQQKSDLELLRISLLLIQSWLGPLQFLSRVFTNSLVFGTSDRVYEKLDLEEGI

Query 121 LALMRELEDGTPRAGQILKQTYDKFDTNMRSDDALLKNYGLLSCFRKDLHKTETYLVRMK
Sbjct 121 LALMRELEDGTPRAGQILK+TYDKFDTNMRSDDALLKNYGLLSCFRKDLHKTETYLVRMK
LALMRELEDGTPRAGQILKRTYDKFDTNMRSDDALLKNYGLLSCFRKDLHKTETYLVRMK

Query 181 CRRFGEASCAF 191
Sbjct 181 CRRFGEASCAF 191

```

gi|3660483|emb|CAA09679.1|growth hormone [*Bubalus bubalis*] (Tiwari and Garg, 1998) Length=191
 Score = 378 bits (971), Expect = 8e-104
 Identities = 191/191 (100%), Positives = 217/217 (100%), Gaps = 0/217 (0%)

```

Query 1  AFPAMSLSSLFANAVLRAQHLHQLAADTFKEFERTYIPEGQRYSIQNTQVAFCFSETIPA
Sbjct 1  AFPAMSLSSLFANAVLRAQHLHQLAADTFKEFERTYIPEGQRYSIQNTQVAFCFSETIPA
AFPAMSLSSLFANAVLRAQHLHQLAADTFKEFERTYIPEGQRYSIQNTQVAFCFSETIPA

Query 61  PTGKNEAQQKSDLELLRISLLLIQSWLGPLQFLSRVFTNSLVFGTSDRVYEKLDLEEGI
Sbjct 61  PTGKNEAQQKSDLELLRISLLLIQSWLGPLQFLSRVFTNSLVFGTSDRVYEKLDLEEGI
PTGKNEAQQKSDLELLRISLLLIQSWLGPLQFLSRVFTNSLVFGTSDRVYEKLDLEEGI

Query 121 LALMRELEDGTPRAGQILKQTYDKFDTNMRSDDALLKNYGLLSCFRKDLHKTETYLVRMK
Sbjct 121 LALMRELEDGTPRAGQILKQTYDKFDTNMRSDDALLKNYGLLSCFRKDLHKTETYLVRMK
LALMRELEDGTPRAGQILKQTYDKFDTNMRSDDALLKNYGLLSCFRKDLHKTETYLVRMK

Query 181 CRRFGEASCAF 191
Sbjct 181 CRRFGEASCAF 191

```

for similar DNA, mRNA or protein sequences [18]. It basically identifies the longest matching segment between our submitted sequence and the sequences in the database. The amino acid sequence deduced from the *Nili ravi* ST cDNA sequence was submitted to the National Commission for Biotechnology Information (NCBI, USA) server to look for possible similar or homologous proteins using BLAST. Our search returned several hits and showed close resemblance to the *Bubalus bubalis* growth hormone submitted by

Sadaf *et al.* [19] with E-value 5e-94, and Tiwari and Garg [20] with E-value 8e-104. Several *Bos taurus* sequences also resembled the *Nili ravi* somatotropin sequence. For example, the somatotropin of Pakistani breed Red Sindhi [21]. The results are given in Fig. 6.

Secondary Structure Prediction

The secondary structure refers to the geometrical arrangement of the polypeptide chain

gi|83416247|gb|ABC18166.1|growth hormone [*Bos taurus* Red Sindhi] (Mahmood et al., 2005) Length=191
 Score = 377 bits (967), Expect = 2e-103
 Identities = 190/191 (99%), Positives = 190/191 (99%), Gaps = 0/191 (0%)

```

Query 1  AFPAMSLSGLFANAVLRAQHLHQLAADTFKEFERTYIPEGQRYSIQNTQVAFCFSETIPA
Sbjct 1  AFPAMSLS L FANAVLRAQHLHQLAADTFKEFERTYIPEGQRYSIQNTQVAFCFSETIPA

Query 61  PTGKNEAQQKSDLELLRISILLIQSWLGPLQFLSRVFTNSLVFGTSDRVYEKLKDLEEGI
Sbjct 61  PTGKNEAQQKSDLELLRISILLIQSWLGPLQFLSRVFTNSLVFGTSDRVYEKLKDLEEGI

Query 121 LALMRELEDGTPRAGQILKQTYDKFDTNMRSDALLKNYGLLSCFRKDLHKTETYLVRMK
Sbjct 121 LALMRELEDGTPRAGQILKQTYDKFDTNMRSDALLKNYGLLSCFRKDLHKTETYLVRMK

Query 181  CRRFGEASCAF 191
Sbjct 181  CRRFGEASCAF 191

```

Fig. 6: BLAST results showing homologous sequences. The 'Query' represents the sequence submitted to the NCBI Databank, the 'Sbjct' is the sequence available in the Databank that match the query sequence, the central line indicates the presence or absence of a residue at the specific position. A 'blank' represents substitution of residue belonging to a different chemical group while '+' sign indicates substitution of amino acid residue within the same chemical group.

along one axis. The secondary structure prediction tools provide information about the local arrangement of the polypeptide backbone ignoring the conformation of the individual side chains. Neural Networking Prediction (NNPred) server was used to predict the secondary structure of the protein. NNPred algorithm is based on Neural Networking or training the server to locate the probable alpha helices or beta sheets by the probability of amino acids to occur in a particular secondary structure [22]. The results are depicted in Fig.7 which show that the ST protein of water buffalo has 191 amino acids folded is an all α -helical structure and four helical regions can be clearly identified. Hence the protein belongs to the four-helix bundle class of α -helices like the cytochrome, ferritin, and human growth hormone. The PsiPred server was used to confirm the results of NNPred and are presented in Fig. 8. The helical structures exist approximately between 7-35, 72-97, 105-127 and 152-183 residues.

3D Structure Prediction

Predicting 3D structure is the most important and tedious process of protein structure prediction as it accurately defines the structure-

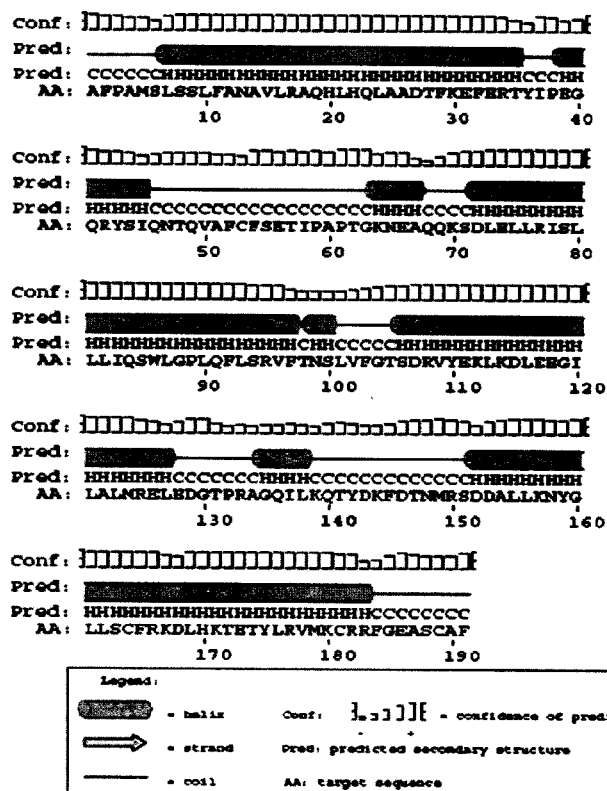
function relationship. Since the protein sequence of ST of *Nili ravi* has more than 25% identity with the available sequences, therefore comparative modeling was carried out. Comparative modeling involves the prediction of the structure of a sequence from the structures of one or more structural templates. The procedure of comparative modeling, as very well described by Goldsmith-Fischman and Honig [23], involves the identification from the reported structures to be used as possible template which has more than 25% sequence similarity to the query sequence, the assembly of the model, then the identification of the regions of the structure that are likely to have different conformations than the templates (e.g., loops) and finally the predicted structure to be refined keeping the inherent differences between the template and query structures. The deduced amino acid sequence was submitted to Protein Databank to identify possible templates. Our results returned several 3D growth hormone structures, from these results bST and hGH were identified as possible templates.

Clustal W ver 1.83 was used to align all the three sequences [11]. The results are given in Figs. 9 and 10 and showed a clear similarity between bST

```

1      10      20      30      40      50      60      70
AFPAMSLSSLFANAVLRAQHLHQLAADTFKEFERTYIPEGQRYSIQNTQVAFCFSETIPAPTGKNEAQKSDLELLRI
----H-----HHHHHHHHHHHHHHHHHHHHHHHHHHHH-----E-----EEEE-----HHHHHHHH
SLLLIQSWLGPLQFLSRVFTNSLVFGTSDRVYEKLDLEEGILALMRELEDGTPRAGQILKQTYDKFDTNMRSDALL
HHH---HHHEEHHH---EEEE---HHHHHHHHHHHHHHHHHHHH-----HHHHH-----HHHHH---
KNYGLLSCFRKDLHKTETYLVRVMKCRRFGEASCAF
-HHHHH-----HHHHHHH-----

```

Fig. 7: NNPred result of the predicted secondary structure of *Nili ravi*. H = helix, E = strand, - = no predictionFig. 8: PsiPred prediction of secondary structure of *Nili ravi* somatotropin.

```

1bst AFPAMSLSGLFANAVLRAQHLHQLAADTFKEFERTYIPEGQRYSIQNTQVAFCFSETIPAPTGKNEAQ
nili AFPAMSLSSLFANAVLRAQHLHQLAADTFKEFERTYIPEGQRYSIQNTQVAFCFSETIPAPTGKNEAQ
*****

1bst QKSDLELLRISLLLIQSWLGPLQFLSRVFTNSLVFGTSDRVYEKLDLEEGILALMRELEDGTPRAGQ
nili QKSDLELLRISLLLIQSWLGPLQFLSRVFTNSLVFGTSDRVYEKLDLEEGILALMRELEDGTPRAGQ
*****

1bst ILKQTYDKFDTNMRSDALLKNYGLLSCFRKDLHKTETYLVRVMKCRRFGEASCAF
nili ILKQTYDKFDTNMRSDALLKNYGLLSCFRKDLHKTETYLVRVMKCRRFGEASCAF
*****

```

Fig. 9: Clustal W alignment for ST of *Nili ravi* buffalo and bST

```

1HGU      --PTIPLSRLFQNAMLRAHRLHQLAFDTYEEFEEAYIPQKYS-FLQAPQASLCFSES IPTPSNREQAQ
nili      AFPAMSLSSLFANAVLRAQHLLHQLAADTFKEFERTYIPEGQRYSIQNTQVAFCFSETIPAPTGKNEAQ
          *   **  **  **  **  **    *****  **  **  **  **           *  *   ****  **  *       **

1HGU      QKSNLQLLRISLLLIQSWLEFPVGFLRSVFANSLSVYGASDSVDYDLLKDLEEIGIQTLMGRLEDGSPRTG
nili      QKSDELLLRISLLLIQSWLGPLQFLSRVFTNSLVFGTSD-RVYEKLKDLEEGILALMRELEDGTPRAG
          ***  *  *****  *****  *   **  **  ****  *  **  **  *****  **  ****  **  *

1HGU      QAFKQTYAKFDANSHND DALLKNYGLLYCFRKDMDKVETFLRIVQCRSVE-GSCG-
nili      QILKQTYDKFDTNMRSDDALLKNYGLLSCFRKDLHKTEYL RVMKCRRFGEASC AF
          *   ****  ****  *   *****  *****  *  *  **  **  **  **

```

Helix 1

Helix 3

Helix 2

NH₂

COOH

Cystine Bridge I
(Cys⁵³:Cys¹⁶⁴)

Cystine Bridge II
(Cys¹⁸¹:Cys¹⁸⁹)

and *Nili ravi* (NR) sequence with only difference of one amino acid Gly at position 9 in bST is replaced with Ser 9 in *Nili ravi* ST (NRST). The sequence of hGH was found to be 62% similar to both STs.

Comparative modeling was carried out with Modeller 9v1 which generated pdb files based on similarity with templates provided and was viewed in Rasmol [24] to confirm the structure. The energy profile generated by Modeller 9v1 provided insight into the most stable structure of NRST. Similarly another software Deep View was used for further confirmation of the predicted structure [12]. The results are presented in Figs. 11-13 with bST and Figs. 14 and 15 with hGH.

The model of NRST using bST showed a complete overlap of the template and target sequences. The results showed that the NRST (Fig 12 and 13) was remarkable 99.5% identical with bST (Fig.11). This result confirmed the result obtained from alignment studies. Both (Modeller and Deep View derived) developed models showed that the somatotropin has the topology of four helix bundle class of alpha helices. Coiled coils are building blocks within a domain. The four helices are arranged in a bundle with helical axes parallel to each other. The helices 1 and 2 are arranged parallel to each other upward whereas helices 3 and 4 are arranged parallel to each other but downward. Thus the



Fig. 12: Predicted structure of *Nili ravi* ST using Modeller 9v1. Helices are numbered.

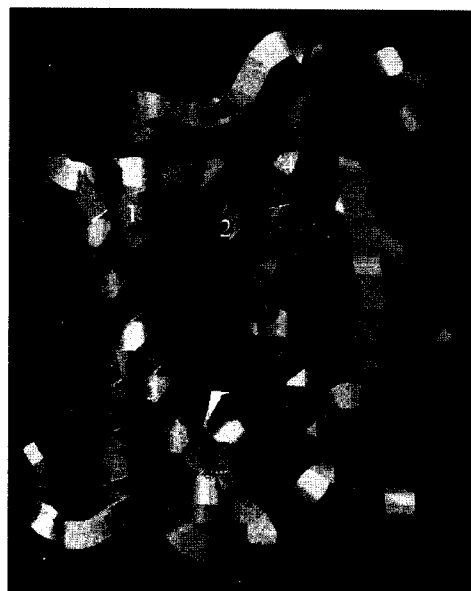


Fig. 13: Deep View predicted structure of *Nili ravi* ST. Superimposed structures of NRST and bST. Light gray represents bST and dark gray shows NRST

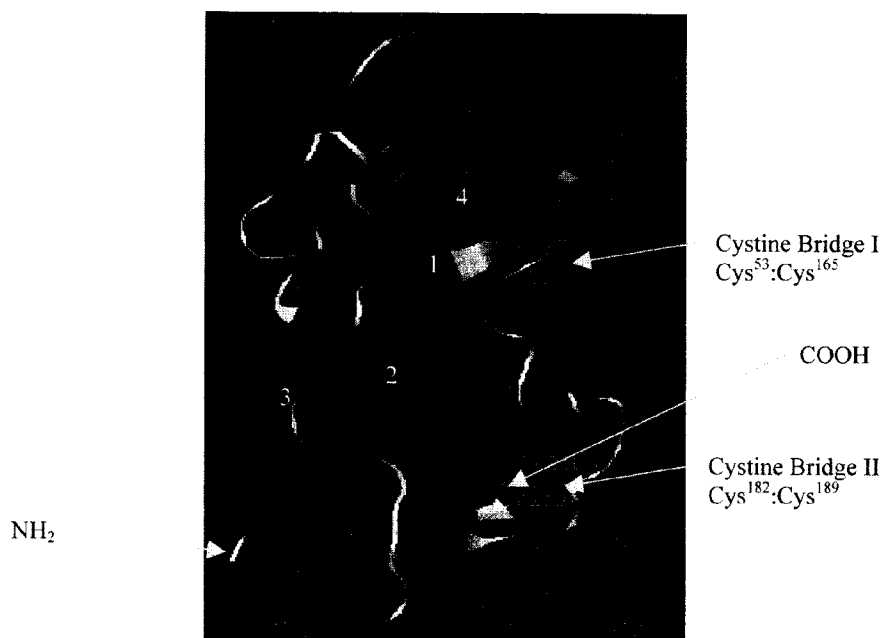


Fig. 14: Human growth hormone template, showing Cystine Bridge I and II.

topological arrangement of the four helices is up-up-down-down. The adjacent alpha helices are somewhat tilted at 20° angle between each of the adjacent pairs making the four alpha helices look like a left handed twisted handle. The helices are connected to each other with long loops between helices 1 and 2 and between helices 3 and 4 while a short loop exists between helices 2 and 3. The structure generally conforms to the structure of bST reported by Carlacci *et al.* [13].

A four helix bundle usually has a hydrophobic core. The predicted structure of NRST also showed that it had hydrophobic core (Fig.17) and polar side chains on the exposed surface that interact with the aqueous environment. Pairs of adjacent helices are additionally stabilized by salt bridges between charged amino acids (Fig. 18) and two Cystine bridges (Cys⁵³:Cys¹⁶⁴ and Cys¹⁸¹:Cys¹⁸⁹). It has been pointed out by Carlacci *et al.* [13] that the key role in electrostatic stabilization of protein is played by the interaction between helices and loops in the form of disulfide bonds. This interaction between helices and loops seemingly provide the main electrostatic energy which in turn provides stability to the protein structure. The electrostatic interaction between the helices and loops is 6 times stronger than the inter-helical electrostatic forces.

Predicted Structures of Somatotropin using Human Growth Hormone as Template

The model of NRST with human growth hormone (1HGU.pdb) modeled using Modeller 9v1 (Fig. 15) and Deep view (Fig. 16) showed that although the basic four helix bundle topology was conserved, there were differences in the amino acid composition and also in the helix composition which for both molecules is different. For example the composition of the first helix is different in number and amino acids

hGH: SALFQNAHLRAHRLHQLAFDTYEEFEE (27 AA)
NRST: SSLFANAVLRAQHLHQLAADTFKEFERTY (29 AA)

The model of NRST also showed helices 2 and 4 broken into two helices each while in hGH structure showed no such changes. Another difference was the presence of two β -sheets in NRST but not in hGH structure. The reason may be that the modeling program considered the amino acids in

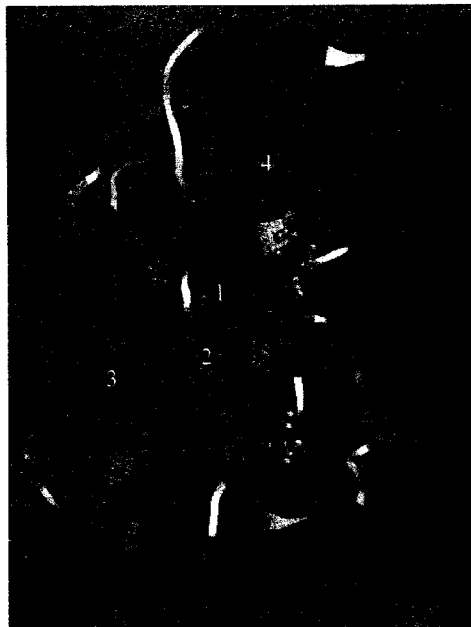


Fig. 15: Modeller predicted structure of *Nili ravi* ST.



Fig. 16: Deep View predicted structure of NRST (dark gray) superimposed over hGH template (light gray)

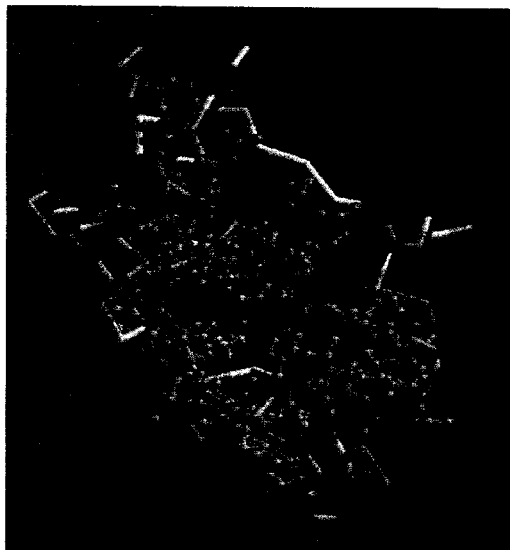


Fig. 17: Core of the *Nili ravi* ST. bST is template.

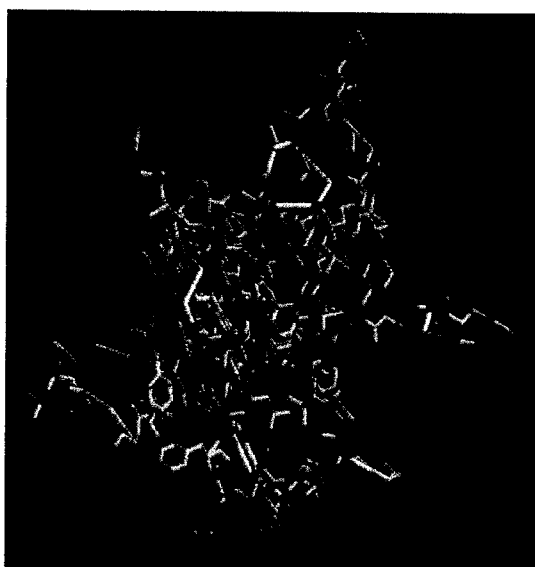


Fig. 19: Core of the *Nili ravi* ST. hGH is template.



Fig. 18: The alpha helices are additionally stabilized by hydrogen bonds and salt bridges.

these positions having high probability for β -sheets. The predicted structure of NRST when compared with hGH showed some similarity too. Like structure of hGH, the structure of NRST is also held together by Cystine bridges at same positions i.e. Cys⁵³:Cys¹⁶⁵ and Cys¹⁸²:Cys¹⁸⁹ (Fig.14 and Fig.15), hydrophobic

interaction between the helices and hydrogen bonds (Fig. 18). The hydrophobic core was also found to be conserved and is shown in Fig. 19.

Thereafter the predicted model was evaluated using two parameters: root mean square distance (rmsd) and energy studies. The rmsd can be calculated when Deep View is used as the visualization software [25]. When two molecules are superimposed for comparative modelling, the parameter of root mean square is calculated along with the number of atoms included that provides information of the similarity between the two molecules. A good model has distance of less than 2Å. The rmsd for bST generated model was found to be 0.31Å and that with hGH was 0.49Å. Similarly, we determined the stability of the molecule through energy profiles done on DOPE (discrete optimization profile energy). The DOPE profile obtained for the model of NRST with bST (Fig. 20) and hGH (Fig. 21) showed that the model we have generated was very stable and had low energy peaks than the template. Therefore, it can be assumed that the model generated by Modeller 9v1 using bST as template is the best predicted model.

We have also evaluated our generated model using Procheck_NT [9] which is a suite of programs that calculates Ramachandran plots and checks for

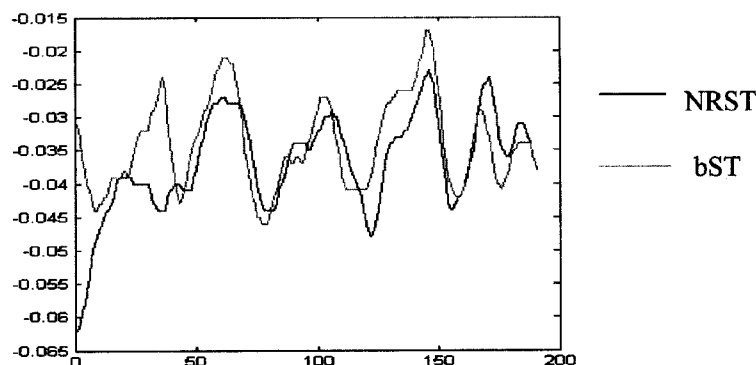


Fig. 20: Energy profile of the most stable predicted structure. bST is template.

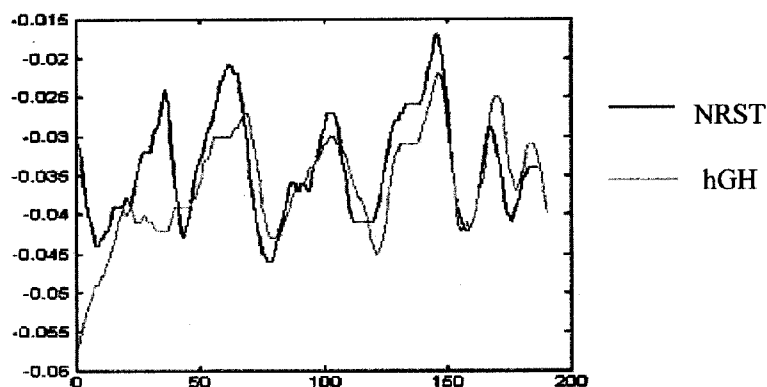


Fig. 21: Energy profile of the most stable predicted structure. hGH is template.

stereochemical parameters such as permitted bond lengths, permitted bond angles, side chain confirmations, etc. The Ramachandran plot showed that the predicted structure has approximately 76% of residues in its allowed regions (Fig. 22). The residues in unusual regions are highlighted. The summary file generated for each structure gives these valuable details (Fig. 23). It also provides a G factor which is a log-odds score showing how normal or unusual the results are. For example, it shows the normality or abnormality of residue location on the Ramachandran plot for the given residue type. When bST is used as template, 100% of bond lengths of the predicted structure were within limits whereas hGH based structure was only 64% within limits. Similarly 92% of bond angles are within limits in bST based structures whereas hGH based structures have only 44% angles within limits. It is therefore concluded that cloned NRST from local breed is very similar in structure to bST while different to hGH.

The cloning, expression and modeling studies of ST cDNA from a local breed provide a strong base for further exploration of this important product as it has high potential to be beneficially used in livestock specially to increase the milk and meat production of *Nili ravi*. This study will also be helpful in developing indigenous technology for production of ST.

Experimental

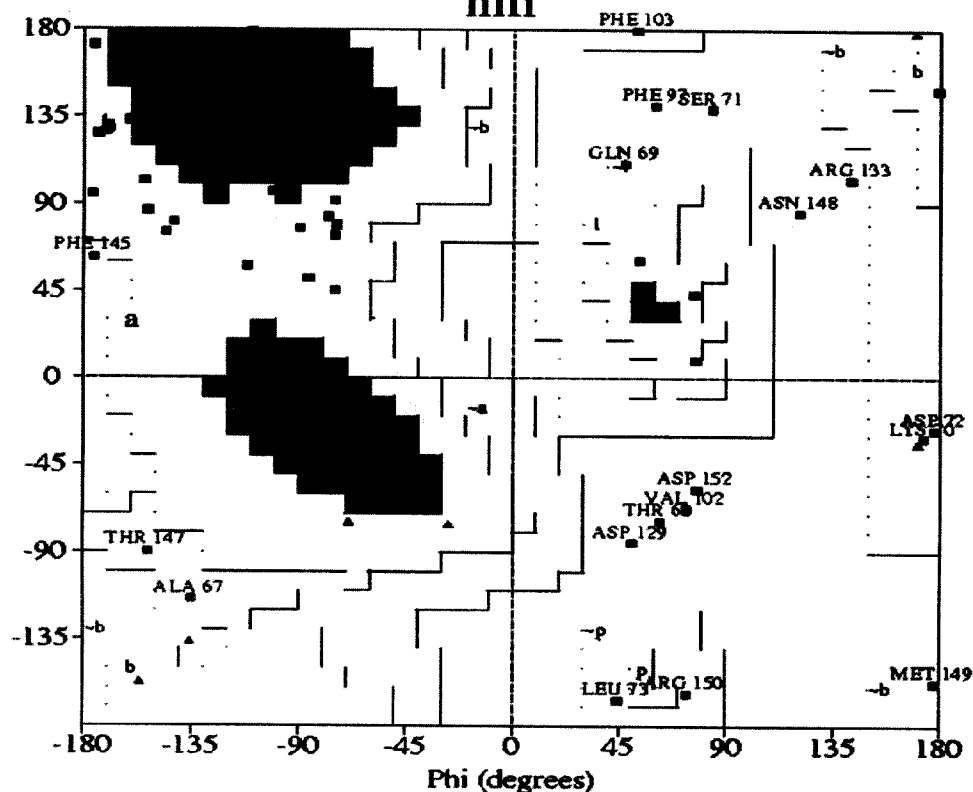
Cloning and Expression Studies

Pituitaries from freshly slaughtered animals were collected and carried under frozen conditions to the laboratory. RNA was then isolated from pituitary by using the method described by Chomczynski and Sacci [26] and used for cDNA synthesis by using RvertAid™ First Strand cDNA Synthesis Kit (Fermentas #K 1622). cDNA was

ROCHECK

Ramachandran Plot

nili



Plot statistics

Residues in most favoured regions [A,B,L]	135	78.0%
Residues in additional allowed regions [a,b,l,p]	20	11.6%
Residues in generously allowed regions [-a,-b,-l,-p]	10	5.8%
Residues in disallowed regions	8	4.6%
Number of non-glycine and non-proline residues	173	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	10	
Number of proline residues	6	
Total number of residues	191	

Fig. 22: Ramachandran plot for *Nili ravi ST*.

synthesized by using the sequence specific primer, random hexamer primer and oligo-dT primer. Thereafter PCR was carried out with specific primer set having restriction enzyme sites: NdeI and EcoRI, designed on the basis of sequences available in Gene Bank. PCR conditions were optimized for amplification of bST gene with use of various concentrations of primers, MgCl₂, template

concentrations and finally annealing temperature. In optimized PCR reaction 0.5 µl of 100 pM sequence specific primers, 1.6 µl of 25 mM MgCl₂, 0.4 µl of 10 mM dNTPs, 1 U of 500 U of Taq DNA polymerase and 5 µl of cDNA in total volume of 20 µl were used. PCR programme included, denaturation at 94°C for 5 minutes in first cycle, then for other cycles, denaturation at 94°C for 30 seconds,

```

+-----<<< P R O C H E C K   S U M M A R Y >>>-----+
|
| Nili ravi with bST                                     191 residues
|
| * Ramachandran plot: 78.0% core 11.6% allow 5.8% gener 4.6% disall
|
| * Gly & Pro Ramach: 2 labelled residues (out of 16)
| + Chi1-chi2 plots: 1 labelled residues (out of 129)
|
| Main-chain params: 6 better 0 inside 0 worse
| Side-chain params: 5 better 0 inside 0 worse
|
| * Residue properties: Max.deviation: 14.1      Bad contacts: 15
|                      Bond len/angle: 2.7  Morris et al class: 1 2 1
|
| G-factors      Dihedrals: -.05 Covalent: .42 Overall: .14
|
| M/c bond lengths: 100.0% within limits .0% highlighted
| M/c bond angles: 92.3% within limits 7.7% highlighted
| Planar groups: 100.0% within limits .0% highlighted
|
+-----+

+-----<<< P R O C H E C K   S U M M A R Y >>>-----+
|
| Nili ravi with hGH                                     186 residues |
|
| * Ramachandran plot: 61.1% core 32.9% allow 5.4% gener .6% disall
|
| * Gly & Pro Ramach: 3 labelled residues (out of 13)
| * Chi1-chi2 plots: 7 labelled residues (out of 116)
|
| + Main-chain params: 3 better 0 inside 3 worse
| + Side-chain params: 2 better 0 inside 3 worse
|
| * Residue properties: Max.deviation: 15.4      Bad contacts: 8
| *                      Bond len/angle: 75.6  Morris et al class: 3 4 3 |
| * G-factors      Dihedrals: -1.08 Covalent: -3.62 Overall: -1.97
|
| * M/c bond lengths: 63.7% within limits 36.3% highlighted 13 off graph
| * M/c bond angles: 43.9% within limits 56.1% highlighted 36 off graph
| + Planar groups: 90.9% within limits 9.1% highlighted 2 off graph
|
+-----+
+ May be worth investigating further. * Worth investigating further.

```

Fig. 23: Procheck summaries.

primer annealing at 54°C for 45 seconds followed by extension at 72°C for 90 seconds for 33 cycles. The final extension was done at 72°C for 25 minutes. The amplified product was electrophoresed onto a 1% agarose gel, the gel was stained with ethidium bromide and then visualized in UV transilluminator.

The amplified product was first ligated in a T/A cloning vector (Fermentas # K1214), transformed into DH5α strain of *E.coli*. The presence of ST cDNA was checked using NdeI and EcoRI restriction enzymes (REs). Thereafter, sequence was carried out using Beckman Coulter CEQ™ sequencer

with the reagent provided by suppliers of the system. The sequence obtained was submitted to Gene Bank (Accession # DQ307367). After the sequence analysis ST cDNA was directionally cloned in pLEX expression vector using NdeI and EcoRI. Further positive clones were checked with NdeI and EcoRI REs and then these clones were used for expression studies in the presence and absence of tryptophan as inducer for 4 hours at 37°C. The expression of protein product was then analyzed by SDS-PAGE using 12% polyacrylamide gel [27].

Model Development of Protein

The sequence of the ST cDNA was translated into the amino acid sequence. Secondary structure was predicted using NNPre and PsiPred servers. ClustalW ver. 1.83 created by Thompson *et al.* [11] was used to produce sequence alignments. The 3D structure of the protein was obtained using Deep View (Swiss PDB) viewer developed by Guex and Peitsch [12] and Modeller 9 ver.1. The model with minimum energy profile was selected as the best predicted model. The templates of bovine somatotropin (1bst.pdb) and human growth hormone (1HGU.pdb) were obtained from Protein Data Bank. The predicted structure was evaluated with Procheck NT as described by Laskowski *et al.* [9].

Acknowledgement

The authors are highly grateful to the Higher Education Commission, Pakistan, for generous financial assistance.

References

1. Ministry of Food, Agriculture and Livestock (MinFAL), Government of Pakistan (2005).
2. T. D. Etherton and D. E. Bauman, *Physiol. Rev.*, **78**, 745 (1998).
3. K. Paul, *The Ecologist*, **28**, 1 (1998).
4. D. E. Johnson, G. M. Ward and J. Torrent, *J. Environ. Quality*, **21**, 157 (1992).
5. M. S. Gulay, M. J. Hayen, M. Liboni, T. I. Belloso, C. J. Wilcox and H. H. Head, *Dairy Sci.*, **87**, 948 (2004).
6. A. J. Gunn, T. R. Gunn, D. L. Rabone, B. H. Breier, W. F. Bluem and P. D. Gluckman, *Pediatrics*, **98**, 279 (1996).
7. J. L. Burton, B. W. McBride, E. Block, D. R. Glimm and J. J. Kenelly, *Can. J. Anim. Sci.*, **74**, 167 (1994).
8. G. F. Hartnell, Bovine Somatotropin: Production, Management, and United States Experience. In: *Animal Sciences Research and Development: Moving towards a New Century*, M. Evan (Ed). Center for Food and Animal Research, Agriculture and Agri-Food, Ottawa, Canada. pp. 189 (1995).
9. R. A. Laskowski, M. W. MacArthur, D. S. Moss and J. M. Thornton, *J. Appl. Cryst.*, **26**, 283 (1993).
10. A. Sali and T. L. Blundell, *J. Mol. Biol.*, **234**, 779 (1993).
11. J. D. Thompson, D. G. Higgins and T. J. Gibson, *Nucl. Acids Res.*, **22**, 4673 (1994).
12. N. Guex and M. C. Peitsch, *Electrophoresis*, **18**, 14 (1997).
13. L. Caracci, K. C. Chou and G.M. Maggiora, *Biochemistry*, **30**, 4389 (1991).
14. L. Chantalat, N. D. Jones, F. Korber, J. Navaza and A. G. Pavlovsky, *Prot. Peptide Let.*, **2**, 333 (1995).
15. C. Guron, K. B. Rao, S. K. Jain, S. M. Totey and T. P. Talwar, *Indian J. Exp. Biol.*, **30**, 659 (1992).
16. E. R. LaVallie, A. E. DiBlasio, S. Kovacic, L. K. Grant, F. P. Schendel and M. J. McCoy, *BioTechnology*, **11**, 187 (1993).
17. T. Huai-jen, L. Ku-lin, K. Jen-chien and C. Shun-wen, *App. Environ. Microb.*, **61**, 4116 (1995).
18. S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller and D. J. Lipman, *Nucl. Acids Res.*, **25**, 3389 (1997).
19. S. Sadaf, M. A. Khan, D. B. Wilson and M. W. Akhtar, NCBI, USA, Accession no. AY940159 (2005).
20. G. Tiwari and L. C. Garg, NCBI, USA, Accession no. AJ011533, (1998).
21. S. F. Mahmood, M. J. Khan, S. Khalid, I. S. Mirza and A. Khanum, NCBI, USA, Accession no. DQ307370 (2005).
22. D. G. Kneller, F. E. Cohen and R. Langridge, *J. Mol. Biol.*, **214**, 171 (1990).
23. S. Goldsmith-Fischman and B. Honig, *Prot. Sci.*, **12**, 1813 (2003).
24. R. Sayle and E. James Milner-White, *Trends Biochem. Sci.*, **20**, 374 (1995).
25. O. Carugo and S. Pongor, *Prot. Sci.*, **10**, 1470 (2001).
26. P. Chomczynski and J. B. Sacchi, *Anal. Biochem.*, **162**, 156 (1987).
27. U. K. Lamellae, *Nature*, **227**, 818 (1970).