1-1-1999

Towards the cloning of the human heat shock protein twenty-seven multigene family

Ann A Ohiaeri
University of Nevada, Las Vegas

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

TOWARDS THE CLONING OF THE HUMAN HEAT SHOCK PROTEIN TWENTY-SEVEN MULTIGENE FAMILY

by

Ann A. Ohiaeri

Bachelor of Science
University of Nevada, Las Vegas
1995

A thesis submitted in partial fulfillment
of the requirements for the degree of

Master of Science

in

Chemistry

Department of Chemistry
University of Nevada, Las Vegas
May 1999

UNLV
UNIVERSITY OF NEVADA LAS VEGAS

Thesis Approval

The Graduate College
University of Nevada, Las Vegas

November 20, 19 98

The Thesis prepared by

Ann A. Ohiaeri

Entitled

Towards The Cloning Of The Human Heat Shock Protein Twenty-seven Multigene Family

is approved in partial fulfillment of the requirements for the degree of

Master Of Science in Chemistry

Examination Committee Chair

Dean of the Graduate College

Examination Committee Member

Examination Committee Member

Graduate College Faculty Representative

ABSTRACT

Towards the Cloning of the Human Heat Shock Protein Twenty-Seven Multigene Family

by

Ann A. Ohiaeri

Dr. Stephen Carper, Examination Committee Chair
Associate Professor of Chemistry
University of Nevada, Las Vegas

We attempted to clone closely-related members of the human heat shock protein 27 (hsp 27) multigene family. Genomic DNA was obtained from human A549 lung cancer cells. Colony hybridization and the Polymerase chain reaction identified the hsp 27 genes. Positive clones were subcloned into pGEM-3Z. The DNA sequence information for three of the clones was determined, and analyzed using DNAsis and Blast Algorithms. DNA Alignments of the inserts contained in three positive clones showed a homology of 29-51% to the human hsp 27 gene. Our data suggests that these inserts are distant members of the human hsp 27 multigene family. Blast searches did not determine the identity of these inserts in GenBank and SwissProt databases (11/1/98). This result indicated that the inserts are novel sequences. Hence, the nucleotide information for the inserts was submitted to GenBank (Accession Numbers: AF120329, AF120330, AF120331, AF120332, and AF120333).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I am indebted to Dr. Stephen Carper, my academic advisor, for his inspiration, guidance, and encouragement throughout my graduate study at UNLV. When I began working in Dr. Carper's lab as an undergraduate student four years ago, I became fascinated with his breast cancer research. In fact, it was he who inspired me to go on to pursue a Masters degree in Chemistry. I have never met a professor who has mastered the art of teaching better than Dr. Carper. He has made the science of research both challenging and fun. I have learned so much from him - too much to mention here.

I wish to gratefully acknowledge Dr. Bryan Spangelo, Dr. Vernon Hodge, and Roberta Williams for being on my thesis committee, and for providing invaluable advice regarding my thesis. I have actually enjoyed taking graduate classes with Dr. Spangelo. He too, is a professor that I am fortunate to have known.

I wish to thank the faculty and staff of the Department of Chemistry, UNLV, for making my stay here both interesting and enjoyable. Special thanks to Ms. Juanita Lytel, for her support and assistance.

I would like to thank the Graduate College of UNLV for their scholarship and grant awards. Special thanks to Dr Harriet Barlow for her assistance during the last

stages of preparing my thesis.

I am grateful to my co-investigators, Michael Wright, Dr. Santosh Dubey, Daniel Mazur, and Will Rust, as well as to other members of our lab - Joe Stafford, Melissa Bracamontes, Trissa Miller, Eddie Kahl, and Charles Ohiaeri. It has been a great pleasure knowing Kathleen McClaren, Dr. Amina Sadik, Trena Tobin and Sindhu Padmanabhan.

I would like to thank my father, Dr. Benjamin Ohiaeri for his support. I am eternally grateful to my mother, Mrs. Julie Ohiaeri, for always believing in me, and supporting my dreams. Whatever I have achieved, and will ever achieve in life, I owe to my mother. She is an inspiration to me, a precious gem that I will always treasure.

Finally, I want to thank my family Kenneth Ohiaeri, Louis Ohiaeri, Charles Ohiaeri, Frank Green, Sara Ohiaeri, and Patrick Marshall for their love and constant encouragement. They have stood beside me during this challenging period in my life.

CHAPTER 1

INTRODUCTION

Cancer

Cancer is a disease characterized by the uncontrollable proliferation of cells. Cancer is the second leading cause of death in this country, today. According to studies conducted by the American Cancer Society in 1998, over 180,000 women will be diagnosed with breast cancer. Of these women, 44,000 will die from the disease. We desire to increase the number of women surviving breast cancer. However, chemotherapy has been hampered by the expression of certain proteins that enable cancer cells to survive the adjuvant therapies designed to destroy them. In order to develop better treatments for breast cancer, more studies examining the proteins involved in protecting breast cancer cells from death are required. Specifically, more information regarding the genes involved in the expression of these proteins is needed.

Human hsp 27 can be found at high levels in human breast tumor cells (1). Reports indicated that in these human breast tumors that express high levels of hsp 27, tumor cells come back sooner (2). Hence, identifying and studying all of the genes involved in the

expression of hsp 27 should ultimately lead to the development of more effective therapies for the treatment of breast cancer.

Heat Shock Proteins

Heat shock proteins are found in all living cells. Following the exposure of cells to stresses, there is an increase in the synthesis of these proteins. The production of heat shock proteins (hsps) protect cells from a wide variety of stresses such as a sudden increase in temperature, heavy metals, oxygen deprivation, and metabolic poisons (3-4). Other stresses such as the exposure to hormones, growth factors, and viruses also induce the heat shock response.

There are five major classes of heat shock proteins that are expressed in humans. These include hsp 110, hsp 90, hsp 70, hsp 60 and hsp 27. These heat shock proteins are classified according to their apparent molecular weight in kilodaltons. Hsp 110 is involved in the import of proteins into the mitochondrial compartment of a cell (5). Hsp 90 associates with steroid hormone receptors, and maintains them in an inactive state until they bind to their hormone (6). The ability of Hsp 70 to unfold proteins, via an ATP-dependent mechanism, allows for the import of these proteins into sub-celluar locations of the cell (7-8). Hsp 70 also associates with newly-synthesized proteins, assisting in their correct folding at desirable times. Hsp 60 possesses an ATPase activity. Hsp 60 and hsp 10 assist in the folding of proteins into the correct conformation through a barrel-shaped structure (9). Hence, the common feature amongst these classes of heat shock proteins is that they function as molecular chaperones in protein folding.

Human hsp 27 - Protein

Human hsp 27 is approximately 27,000 Da in size. Several functions for human hsp 27 have been reported. These include the development of thermotolerance following a heat shock (10-11), and the ability to function as an actin cap-binding protein (12). Human hsp 27 has also been reported to function as a molecular chaperone (13), and confers drug-resistance to cytotoxic drugs (14).

Human hsp 27 - Gene

Genes are located on chromosomes within the nucleus of a cell. A gene is a piece of double-stranded DNA that contains all the information necessary for the synthesis of a functional protein. A gene is made up of a promoter region (located at the 5'-end, and responsible for gene regulation), exons (that code for the synthesis of a protein), and introns (non-coding intervening regions). For a gene to synthesize a protein, it must first be converted into RNA (transcription). Transcription begins at a site known as the TATA box. RNA processing occurs, which involves the removal of introns, the joining of the exon regions (splicing), the addition of a 7-methylguanosine cap at the 5'-end, and the addition of a poly-A tail at the 3'-end. Following the poly-A tail, there is a 3'-untranslated region. The RNA is then further converted into a protein (translation). Translation usually begins at the first ATG site.

It has been suggested that human hsp 27 is made up of a multigene family. Studies conducted by Hickey et al. (15) indicated the presence of four genes encoding human hsp 27, including two linked genes and one pseudogene. Linked genes are located close to each other on a chromosome. Pseudogenes are genes that have evolved over time and lost their function.

Other findings by Mcguire et al. (16) reported the presence of at least three genes encoding for human hsp 27. Our research confirmed the presence of three hsp 27 genes in the human lung cancer cell line A549, and the presence of four hsp 27 genes in the human breast cancer cell line MCF-7 (17). The three human hsp 27 genes are located on chromosomes 3, 9 and X (16).

The genomic DNA and pseudogene that encode human hsp 27 have been sequenced (15). The human hsp 27 gene contains three exons, and two introns. The 5'-promoter region contains two TATA sequences (hence, two transcription start sites), and an SP1 binding site. The SP1 binding site is a sequence that a transcription factor binds to, resulting in an increase in transcriptional efficiency. The promoter region also contains a heat shock element (HSE). The HSE contains a pentameric DNA regulatory motif (GAAnnTTC), that is responsible for the induction of the gene following a heat shock. The pseudogene does not possess promoter elements or intervening sequences. The complementary DNA (cDNA), encoding hsp 27 has also been sequenced (17). Carper et al. (17) reports a mistake in the gene sequence of hsp 27, as reported by Hickey et al. (1986). The corrected amino acid sequence for human hsp 27 is 100% identical to the previous amino acid sequence in the first 193 amino acids.

Additional human hsp 27 promoter region information has been reported by Oesterreich et. al (18). Hickey et al. (1986) previously reported only 200 base pairs of 5'-information in the promoter region of the gene. In the promoter region there are also two CAAT boxes, a G/C region, an additional SP1 site, two AP2 sites, and an estrogen responsive element half-sites (ERE). The AP2 binding site is a sequence that a transcription factor binds to, resulting in an increase in transcriptional efficiency. The ERE is a palindromic, regulatory DNA motif (GGTCA), located near the first TATA box. It has been

proposed that the ERE, and the proximal TATA box are responsible for the regulation of human hsp 27 by the hormone, estrogen.

Other Genes Related To Human Hsp 27

In addition to the genes that make up the human hsp 27 multigene family, there are other genes that are related to human hsp 27. These include the alpha-crystallins, murine hsp 27, and rat hsp 27.

Alpha-crystallins are proteins which polymerize to form large aggregates. In mammals, alpha-crystallins are a major component of the eye lens. The aggregation undergone by alpha-crystallins, plays a crucial role in determining the properties of the eye lens. The alpha-crystallins can act as molecular chaperones (19). The hsp 27 multigene family is related to the alpha-crystallins (20).

The structure and organization of the genes encoding murine hsp 27 has been determined (21-22). Unlike human hsp 27 which has at least three genes encoding for hsp 27, only two genes have been reported to encode hsp 27 in the mouse genome - one gene, and a pseudogene. Murine hsp 27 contains one more HSE than is found in human hsp 27. However, murine hsp 27 is similar to human hsp 27, as it is also made up of three exons, and two introns. The amino acid sequence of murine hsp 27 is 81% homologous to the predicted amino acid sequence for human hsp 27 (17).

The structure and organization of the genes encoding rat hsp 27 has been determined (23). Unlike human hsp 27, only one gene has been reported to encode rat hsp 27. This gene has three exons, and two introns. The promoter region of rat hsp 27 possesses more heat

shock elements, and SP1 sites than in seen in human hsp 27. The amino acid sequence of rat hsp 27 is 81% homologous to the predicted amino acid sequence of human hsp 27 (17).

Objective Of The Study

As discussed previously, identifying and studying all the human hsp 27 genes should lead to better treatments for breast cancer. The first step in this process involves finding all the members of the multigene family. The second step involves studying the regulation of these genes. The final step will examine methods to decrease the level of expression of these genes, in order to lead to better treatments for breast cancer patients. At this point in time, only the human hsp 27 gene and pseudogene members of the family, have been identified and studied. The linked genes have been mentioned, but have not been reported on in any subsequent studies. The present study was aimed at cloning the closely-related, linked gene members of the human hsp 27 multigene family. Our results, however, indicated that we had cloned three distantly-related members of the family. Searches within the GenBank database indicated that these inserts were unique sequences. Hence, the DNA sequences for these three inserts have been deposited with GenBank.

CHAPTER 2

MATERIALS AND METHODS

General Overview

The individual steps in recombinant DNA technology that were used to clone the genes are outlined in the general overview diagram shown in Figure 1.1 (p. 8). The process of cloning the DNA fragments began with the construction of a DNA library. This step was carried out in our lab by Michael Wright. The genomic DNA Library used in this study contained all the DNA from the human lung cancer cell line A549. Figure 1.2 (p. 9) depicts the individual steps involved in the construction of the genomic library. The DNA library was generated by partially digesting the genomic DNA with a specific restriction enzyme, producing restriction fragments of different sizes. The DNA was subjected to partial rather than complete digestion so that the genomic library contained intact representatives of all human genes, including those whose sequences contained restriction sites. The DNA fragments were separated by centrifugation in a gradient. DNA fragments of 15-23 kilobase pairs (kb) were then packaged into a lambda phage vector, which had also been digested with the same restriction enzyme.

Subsequent steps involved in the cloning of the genes were carried out by joint efforts

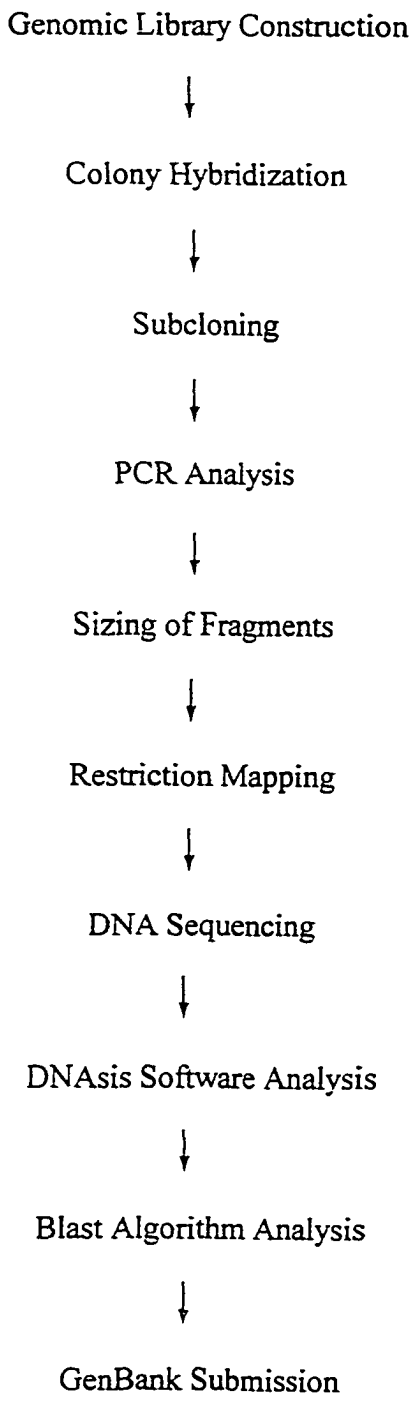


Figure 1.1 General Overview Flowchart. The steps involved in the cloning of the human hsp 27 multigene family are shown.

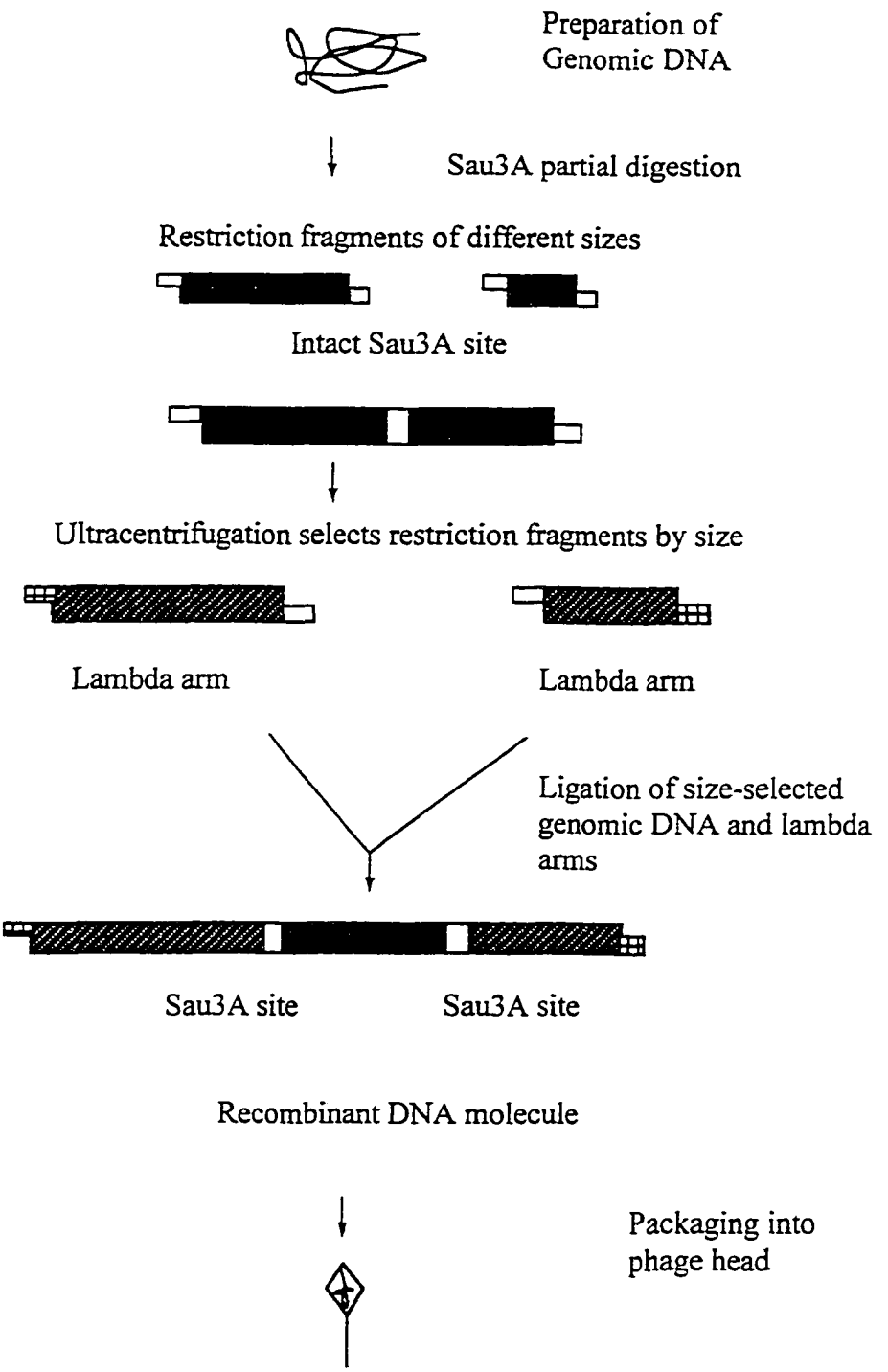


Figure 1.2 Construction Of A Genomic Library. The cloning of foreign DNA in lambda phage.

made by Dr. Santosh Dubey, Daniel Mazur, and myself. The human genomic library constructed in lambda phage was screened for hsp 27 genes. The lambda EMBL3 phage containing the genomic DNA from the A549 cell line was initially used to infect bacterial cells, which were plated on a petri dish. Colony hybridization was then used to identify the phage vectors that contained the desired DNA. Figure 1.3 (p. 11) depicts the individual steps involved in colony hybridization. Plaques (areas of clear lawn, signifying the location of the phages) were transferred to a filter. The filter was then treated to expose the DNA contained in the plaques. A probe, a DNA molecule specifically designed to bind to our desired DNA, was later allowed to anneal to the filter. Blackening on the filter identified plaques that had hybridized to our probe, and hence, contained our desired DNA.

The phage DNA containing the desired pieces of genomic DNA (the positive plaques), were isolated and purified. The phage DNA was removed, and the genomic DNA was enzymatically cut into smaller fragments. These small fragments were ligated into a plasmid vector for further analysis (subcloning). In this way, we narrowed our search for the hsp 27 genes from within larger pieces of DNA of 15-23 kb, to amongst smaller pieces of DNA (of about 5 kb in size). Figure 1.4 (p. 12) shows the individual steps involved in subcloning. The insert DNA fragment was cleaved with a specific restriction enzyme, to yield insert DNA fragments possessing this particular type of ends. The vector (the circular, double-stranded DNA that would hold the insert DNA), was also cleaved with the same restriction enzyme, producing vector DNA with identical ends. The complimentary ends of the insert DNA and the vector DNA, specifically associated under certain conditions, and covalently attached to each other via the action of the enzyme DNA ligase (ligation). This produced a recombinant plasmid.

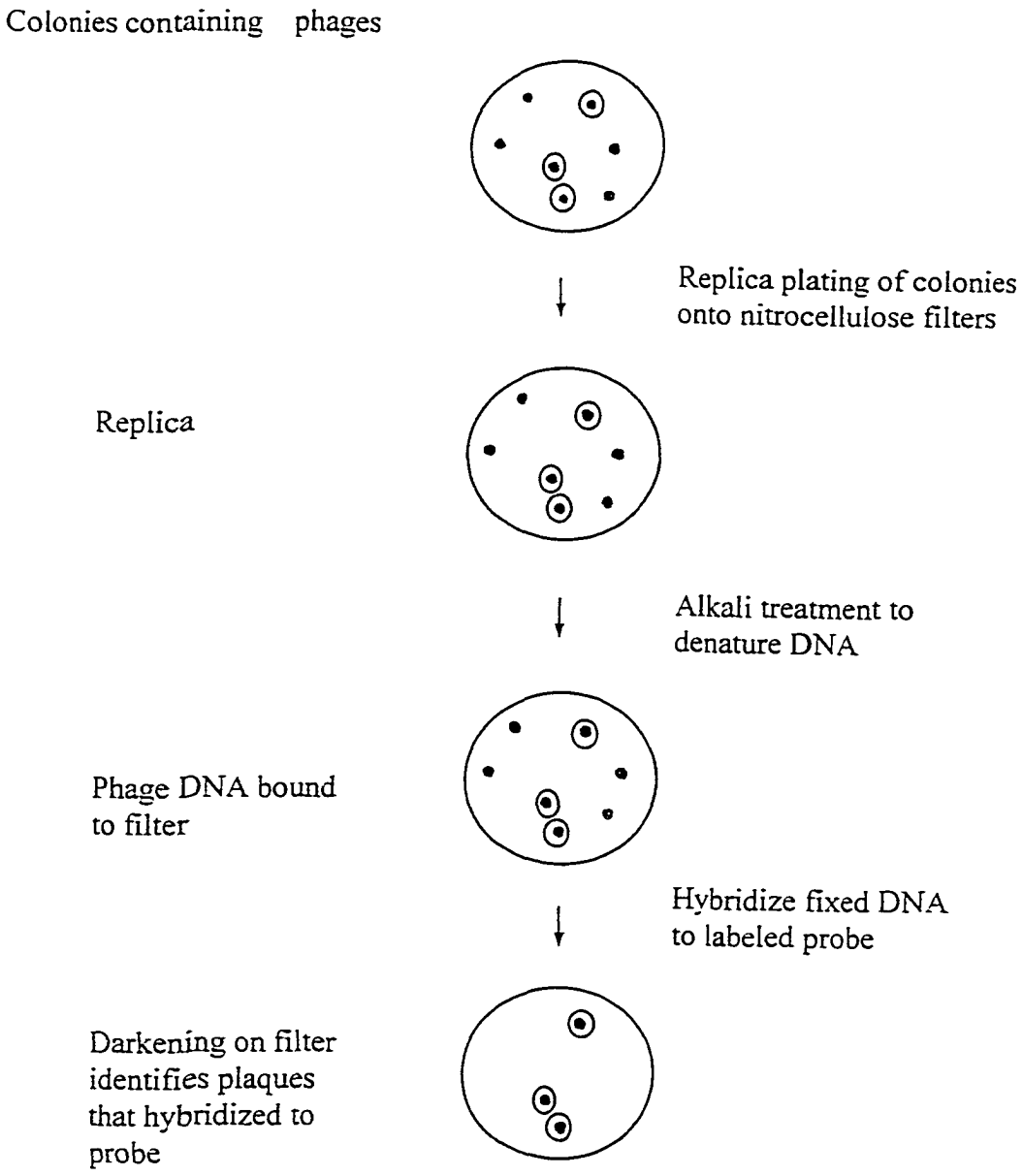


Figure 1.3 Colony Hybridization. This technique identifies the clones containing the desired DNA.

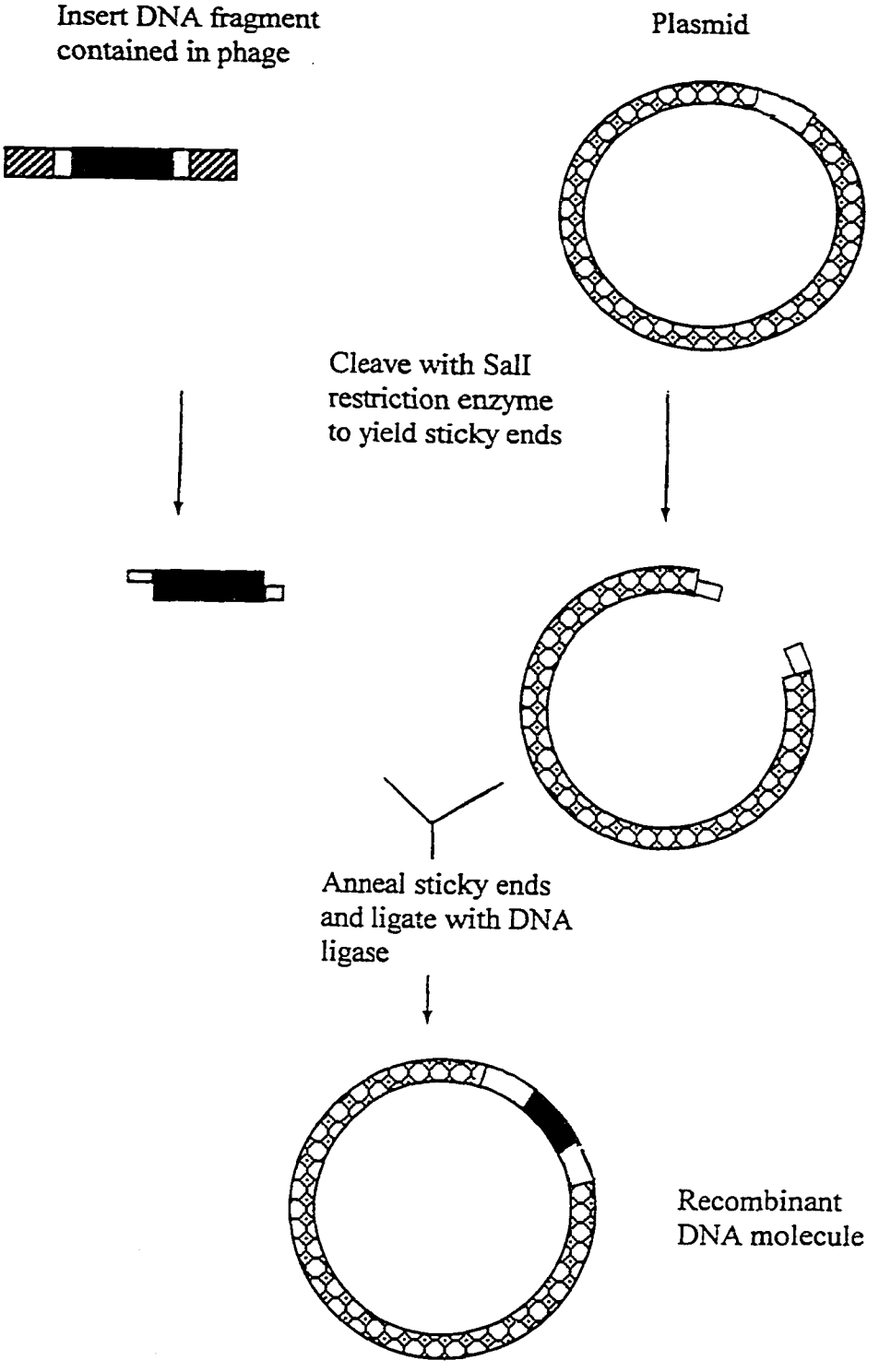


Figure 1.4 Subcloning. The construction of a recombinant DNA molecule through the insertion of an enzyme-digested-insert into a vector's corresponding restriction site.

The polymerase chain reaction (PCR), was used as a second technique to further determine if the DNA inserts in the recombinant plasmids were the desired hsp 27 genes. Figure 1.5 (p. 14) shows the individual steps involved in PCR. Heat was initially used to denature and separate the double-stranded DNA duplex into single strands. Then, a primer was added and allowed to anneal to the single-stranded DNA. The primer was a short DNA molecule specifically made to bind to a region on our desired DNA. When the enzyme DNA polymerase was later added, it catalyzed the extension of the primer, resulting in the synthesis of a new identical DNA molecule. In this way, thousands and thousands of copies of the DNA were made. It is important to note here that each cycle of PCR only amplified our desired hsp 27-related DNA. If the plasmid vector contained hsp 27 DNA, then a single 363 base pair (bp) PCR product was obtained, and the plasmid insert was further characterized.

Restriction endonucleases were used to determine the sizes of the DNA inserts contained within the positive clones. The positive clones were cleaved with restriction enzymes, and the fragments were separated on an agarose gel (gel electrophoresis). In gel electrophoresis, an electric field was used to separate our DNA molecules by allowing the molecules to migrate through a porous agarose gel. Smaller molecules migrated faster relative to the larger ones. A DNA size ladder was used to identify the molecular range of the fragments. Restriction maps showing the sites where particular enzymes cleaved the insert contained in the positive clones, were also generated with the information.

To determine the identity of the DNA sequence information, the positive clones were

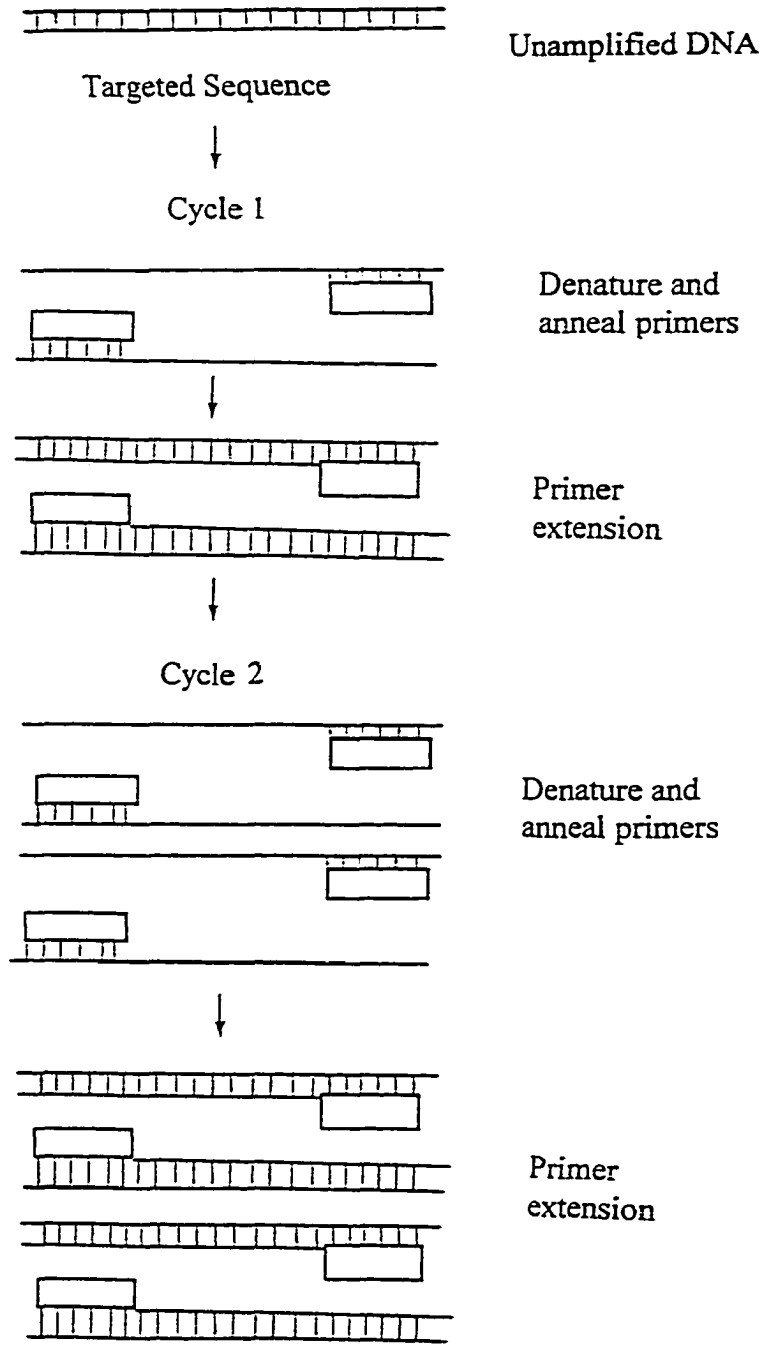


Figure 1.5 Polymerase Chain Reaction. In each cycle of the reaction, the double-stranded DNA is separated by heat denaturation, synthetic DNA primers anneal to a complementary segment on each strand, and the primers are extended by DNA Polymerase. The process is then repeated for numerous cycles.

transformed into bacterial cells, to make thousands of copies of the DNA. The plasmid DNA was extracted from the bacterial cells, purified and subjected to automated DNA sequencing.

The sequence information for the DNA inserts was initially analyzed using the DNAsis computer software (DNAsis Maximum Matching Alignment). This computer algorithm compared the similarity between two DNA sequences. The program was used to determine if the isolated DNA fragments were similar to hsp 27 in their nucleotide sequences.

In addition, we desired to compare the predicted amino acid sequences of the DNA inserts to the amino acid sequence of hsp 27. Therefore, we searched for sites in the DNA sequence that would indicate the coding strand for the amino acid sequence. This was accomplished using a modification of the DNAsis Restriction Enzyme Search Function. The determination of the open reading frame was performed using the DNAsis Open Reading Frame Function. The information was used to identify the corresponding predicted amino acid sequences using the DNAsis DNA-To-Protein Translation Algorithm. The amino acid sequences of the inserts were used for later amino acid comparisons between the inserts and hsp 27. This was done using the DNAsis Maximum Matching Alignment.

In order to further determine if the inserts were closely-related to human hsp 27, we wanted to identify the location where the PCR primers should have bound to the probe. Hence, the corresponding DNA sequence motifs for the forward and reverse PCR primers were searched for, within the DNA sequence of the inserts contained in the positive clones.

The DNA sequence information was also analyzed using the Blast Algorithm (Basic Local Alignment Search Tool). This computer algorithm searched for a query sequence in a database. The program was used to determine the identity of the isolated DNA inserts.

Finally, the nucleotide sequence information for the inserts contained in the positive clones was deposited with the GenBank database.

Bacterial Strains And Growth Conditions

Escherichia coli (*E. coli*) JM109 competent cells were utilized in the subcloning of the hsp 27 genes (Promega Corporation, Madison, WI). These bacterial cells carry certain genetic markers that were useful for our cloning purposes. These genetic markers included endA1 (an endonuclease mutation that improved the quality of plasmid DNA isolation), recA1 (a mutation in recombination that prevented the recombination of introduced DNA with host DNA, ensuring the stability of inserts), hsdR17 (which allowed cloning without the cleavage of DNA by endogenous restriction endonucleases), and lacZ-D-M15 on an F' (which when used with an appropriate vector, made it possible to distinguish between colonies that had a vector containing an insert from those that had a vector lacking an insert).

Bacterial cultures were grown with shaking in Luria Bertani Medium (LB) containing ampicillin (Amp; 50 mg/ml) at 37°C. LB was made using 5 g of yeast extract, 10 g of tryptone and 10 g of sodium chloride (all from Fisher Biotech., Santa Clara, CA). The final solution was made up to 1000 ml using deionized water (pH 7.0). LB + Amp plates were prepared by the addition of 7.5 g of Bacto agar (Difco, Detroit, Michigan) to 500 ml of LB. To this solution 0.5 ml of ampicillin (50 mg/ml; Sigma Chemical Company, St. Louis, MO), was added to select and maintain JM109 transformants. Cultures (100µl) were placed in 850 µl of 80% Glycerol stock solution. The 4 ml glycerol stock vials (VWR Scientific, San Francisco, CA) were stored at -70°C.

Parental Plasmids

Lambda EMBL3 phage was used as a vector for cloning large pieces of genomic DNA from the A549 lung cancer cell line (2 X 10^6 pfu/mg; Stratagene, La Jolla, CA). This vector was chosen as it is very useful for cloning Sau3A partial digests, since the BamHI sites in the vectors are flanked by EcoRI and SalI restriction sites. Lambda EMBL3 phage was supplied as a BamHI-digested lambda EMBL3 phage product. The BamHI-digestion removed the 13.7 kb stuffer fragment (which was unnecessary for the propagation of the vector), and allowed for the replacement of the stuffer fragment with an insert. Therefore, it was possible to insert a piece of Sau3A-digested genomic DNA into the phage via ligation to the BamHI-digested phage product. Also, it was possible to ultimately remove the cloned genomic DNA fragments from the recombinant phage when necessary, by digestion with SalI restriction enzyme. Recombinant phages were Spi-negative, and were therefore selected on the P2-lysogenic host, *E. coli* XL-1 blue cells.

pGEM-3Z, a 2,743 bp plasmid was used in subcloning and DNA sequencing experiments (Promega Corporation, Madison, WI). This vector carried the lacZ gene and multiple cloning site arrangement from pUC18. pGEM-3Z contained both the SP6 and T7 RNA polymerase promoters flanking the multiple cloning site. pGEM-3Z is capable of complementing the product of the lacZ-delta-M15 gene to produce functional beta-galactosidase. Hence, in *E. coli* JM109 cells (that contain the lacZ-delta-M15 on an F'), pGEM-3Z allowed for the visual screening of recombinant transformants when plated on ampicillin (50 mg/ml; Sigma Chemical Co., St. Louis, MO), X-Gal (5-bromo-4-chloro-3-indoyl-Beta-D-galactopyranoside in 2% DMF solution); (Promega Corporation, Madison, WI), and IPTG (isopropyl-Beta-D-thiogalactopyranoside; 100 mM; Promega Corporation,

Madison, WI). The parental pGEM-3Z vector contains a lacZ gene with a multiple cloning site for inserting foreign DNA. When the parental pGEM-3Z vector (with the intact lacZ gene) was grown in the presence of X-gal (5-bromo-4-chloro-3-indoyl-b-D-galactopyranoside), the lacZ gene encoded for beta-galactosidase, an enzyme which hydrolyzed the colorless X-gal, producing a blue product (5-bromo-4-chloro-3-hydroxyindole) and beta-D-galactose. IPTG (isopropyl-b-D-thiogalactopyranoside) was used with X-gal as it induced beta-galactosidase activity more than 10-fold above background levels in these *E. coli* cells. Ampicillin was also used as a selection marker in order to exclude bacteria that have not taken up any plasmid, which would otherwise also form white colonies in the presence of X-gal. Cells containing the recombinant vector will have an insert present in the lacZ gene, which interrupts the gene, preventing the production of beta-galactosidase and resulting in white colonies. Bacterial colonies containing the parental pGEM-3Z plasmid have no insert, which leaves the lacZ gene intact and produces beta-galactosidase, resulting in blue colonies. The plasmid map for the parental pGEM-3Z plasmid is shown in Appendix I (p. 109). The nucleotide sequence for the parental pGEM-3Z plasmid is shown in Appendix II (p. 111).

Enzymes

The enzymes SmaI, XbaI, PstI, ScaI, HincII, KpnI, EcoRI, BamHI, SalI and HindIII were obtained from New England Biolabs (Beverly, MA). Sau3A restriction enzyme was ordered from the Promega Corporation (Madison, WI). Vent DNA polymerase I large klenow fragment and calf intestinal alkaline phosphatase (CIP) were also obtained from New England Biolabs (Beverly, MA).

Genomic Library Construction

A genomic DNA library from the human A549 lung cancer cell line was previously constructed in our lab. The genomic DNA (100 μg) was partially digested with Sau3A restriction endonuclease (1 unit) at 37°C for 2 hours (24). The Sau3A-cut genomic DNA (0.127 μg/μl) containing the hsp 27-related sequences was recovered in 500 μl aliquots, by centrifugation over a 1-15% glycerol gradient at 24,300 rpm for 21.3 hours in a Sorvall AH-627 rotor (Dupont Company, Newton, CT). The 15-23 kb genomic DNA fragments (1 μl) were ligated into lambda EMBL3 phage (6.7 μl of 0.127 μg/μl) at 16°C for 6 hours. Specifically, BamHI-cut and dephosphorylated lambda EMBL3 phage was ligated to the Sau3A-cut genomic DNA using T4 DNA Ligase (4 units), 10X Ligase buffer (1.1 μl) and 10 mM ATP (1.1 μl). The lambda EMBL3 phages were amplified in *E. coli* XL-1 blue cells (Stratagene, La Jolla, CA), using a concentration of 1.2 X 10^{10} pfu/ml titre, and the recombinant phages were later used in colony hybridization.

Colony Hybridization

The human genomic library constructed in lambda EMBL3 phage was screened for hsp 27 genes, using full-length human hsp 27 cDNA obtained from our lab (17). First, the lambda EMBL3 phage containing the genomic library was plated on NZY agar plates, using 150 mm petri dishes (24). NZY agar plates (pH 7.0), were made using NZ amine (10 g), sodium chloride (5 g), yeast extract (5 g), and 950 ml of distilled water. The lambda phage lysate was subjected to a 1 X 10^{-4} dilution in SM Buffer before plating on the NZY agar plates. SM Buffer was made using 5.8 g of sodium chloride, 2 g of magnesium sulfate (Sigma Chemical Co., St. Louis, MO), 50 mg of 1 M Trizma-base (pH 7.5; Sigma Chemical

Co., St. Louis, MO), and 5 ml of gelatin (2% solution). The final volume was made up to 1,000 ml using deionized water. *E. coli* XL-1 blue cells were grown at 37°C for 3-4 hours until an optical density of 0.5 was obtained (600 nm). The diluted phage (200 µl) was mixed with the *E. coli* cells XL-1 blue cells (600 µl) and incubated at 38°C for 15 minutes. To this mixture, melted top agar was added (7.5 ml). Melted top agar was made using agarose (0.8 g), 1 M magnesium sulfate (1mg), and LB (100 ml).This solution was plated on NZY agar plates, which were later incubated at 37°C for 6-7 hours.

The plaques were then transferred to nitrocellulose filters that were obtained from Stratagene (La Jolla, CA) (24). Prior to colony hybridization, it was necessary to remove the human hsp 27 cDNA from the recombinant plasmid pTRI/62Y.16.7, via an EcoRI digestion, producing a 762 bp EcoRI-cDNA fragment to be used as the probe. The nucleotide sequence for human hsp 27 cDNA is assigned the GenBank accession number X54079 and is shown in Appendix II (p. 114).

A Nonradioactive DIG DNA Labeling and Detection System (Boehringer Mannheim Biochemicals, Indiananapolis, IN) was used to determine the sites of hybridization of the probe to the hsp 27-related sequences. The prehybridization step was carried out at 68°C for 1 hour, using 20 ml of hybridization solution (without labeled DNA). Hybridization solution was made using 0.25 g of blocking reagent, 50 mg of N-Lauroyl Sarcosine, 100 mg of SDS, 12.5 ml of 20X SSC, and 37.5 ml of distilled water. The hybridization step was carried out overnight at 68°C, using 26 ng of denatured probe, and 20 ml of hybridization solution. The hybridized nitrocellulose filters were washed twice, for 5 minutes at 65°C using 50 ml of wash buffer 1 (2X SSC and SDS 0.1% w/v).The filters were rewashed twice, for 15 minutes at 65°C, using 50 ml of wash buffer 2 (0.2X SSC and SDS 0.1% w/v).

Positive plaques were evident as dark spots on the nitrocellulose filter. These spots indicated the sites of hybridization of the human hsp 27 genes to the human hsp 27 cDNA probe. As a positive control for colony hybridization, we used *E. coli* cells with the plasmid pTRI/62Y.16.7, because they contained full-length hsp 27 cDNA and would hybridize to the probe.

The number of positive plaques was then determined. A pasteur pipette was then used to pick up the agar plugs containing the positive plaques, and these plaques were eluted in SM buffer (100 µl) overnight at 4°C. Three different rounds of colony hybridization were performed on three different sets of plates. After each of the screening steps, the positive plaques were eluted in SM Buffer, and the plates were stored at 4°C. These recombinant phages were later digested into smaller fragments (subcloning), and analyzed using PCR.

Subcloning

The positive plaques identified from colony hybridization were isolated following the protocol for the isolation of phage DNA (24). Subcloning was used to narrow the search for the hsp 27 genes from within larger fragments of DNA (15-23 kb), to within smaller fragments (of about 5 kb). In the subcloning step, the positive plaques were digested with SalI (1 unit) at 37°C for one hour to remove the phage arms, and to simultaneously cut the genomic DNA into smaller fragments. The SalI site was used as an insertion site instead of the BamHI or EcoRI sites, as these latter sites were lost once the larger genomic DNA fragments were previously inserted into BamHI-cut lambda EMBL3 phage. The SalI enzyme was then heat-inactivated at 65°C for 20 minutes. The pGEM-3Z vector DNA (Promega

Corporation, Madison, WI) was also digested with the SalI enzyme at 37°C for 2 hours. The SalI enzyme was once again heat-inactivated at 65°C for 20 minutes.

The digested pGEM-3Z vector DNA (16 µl) was dephosphorylated with 10X calf intestinal phosphatase enzyme (2 µl), and 10X alkaline phosphatase buffer (2 µl) at 37°C for 30 minutes (24). This was done to prevent the resealing of the digested parental pGEM-3Z vector ends before the insertion of a DNA fragment. The SalI-cut genomic DNA (10 µl) and the dephosphorylated-SalI-cut pGEM-3Z vector DNA (4 µl) were ligated together at 4°C overnight using T4 DNA ligase (4 unit/µl), 10X ligase buffer (5 µl), riboATP (5 µl) and distilled water (25.5 µl). The resulting recombinant plasmids contained smaller genomic DNA fragments from the lung cancer A549 cell line, than the 15-23 kb genomic fragments that were previously packaged into lambda EMBL3 phage. These smaller fragments of about 5 kb in length, were inserted into the SalI site, in the multiple cloning site of plasmid pGEM-3Z.

These resulting recombinant plasmids were transformed into *E. coli* JM109 cells, and the culture was plated on LB + Amp + X-gal + IPTG plates (24). This method of selecting recombinant plasmids was based on a selectable antibiotic marker, and a color-producing substrate. The white colonies were picked with a sterile toothpick, and grown overnight with shaking at 37°C in 3 ml of LB media containing 3 µl of ampicillin (50 mg/ml). The plasmid DNA of the positive clones were isolated from the *E. coli* JM109 host cells using a Plasmid DNA Extraction Miniprep Kit (Promega Corporation, Madison, WI). The manufacturer's instructions were followed, and the appropriate centrifugation steps were carried out in either a Micro Centrifuge Model 235C (Fisher Scientific, Santa Clara, CA), or in a Sorvall RC2B

Centrifuge (Dupont Company, Newton, CT). The plasmid DNA was resuspended in distilled water (50 µl), and further analyzed using PCR.

PCR Analysis

PCR was used to further identify positive clones containing the hsp 27 genes. In this reaction, synthetic oligonucleotide primers were used to amplify the first exon of hsp 27-related DNA. We initially deduced that the first exon region on the DNA would be a conserved region present in all members of the hsp 27 gene family. Figure 1.6 (p. 24) shows an illustration of the hybridization sites for the forward and reverse primers in the first exon region of the human hsp 27 gene sequence.

The primers were synthesized by an Applied Biosystems 380A DNA synthesizer (Integrated DNA Technologies Inc., Coralville, IA). The forward primer (19 bp) and the reverse primer (20 bp) contained the following nucleotide sequences:

Forward Primer- 5'-GAGCGCCGCGTCCCCTTCT-3'

Reverse Primer- 5'-ATCTCCACCACGCCATCCTT-3'

PCR was carried out in a Perkin Elmer Thermal Cycler (Perkin-Elmer Express, Foster City, CA). The 100 µl reaction mixture contained forward and reverse primers (50 pmoles each), dNTP's (2.5 mM each), 5X PCR reaction buffer (pH 10), and a single wax bead coated with magnesium chloride (1.5 mM), all obtained from Invitrogen (San Diego, CA). Also, the reaction mixture contained Vent DNA polymerase (1 unit; New England Biolabs, Beverly, MA). As a positive control for PCR, we used the plasmid pTRI/62Y.16.7 (5 µl), because it contained the full-length hsp 27 cDNA with the first exon region. As a negative control for

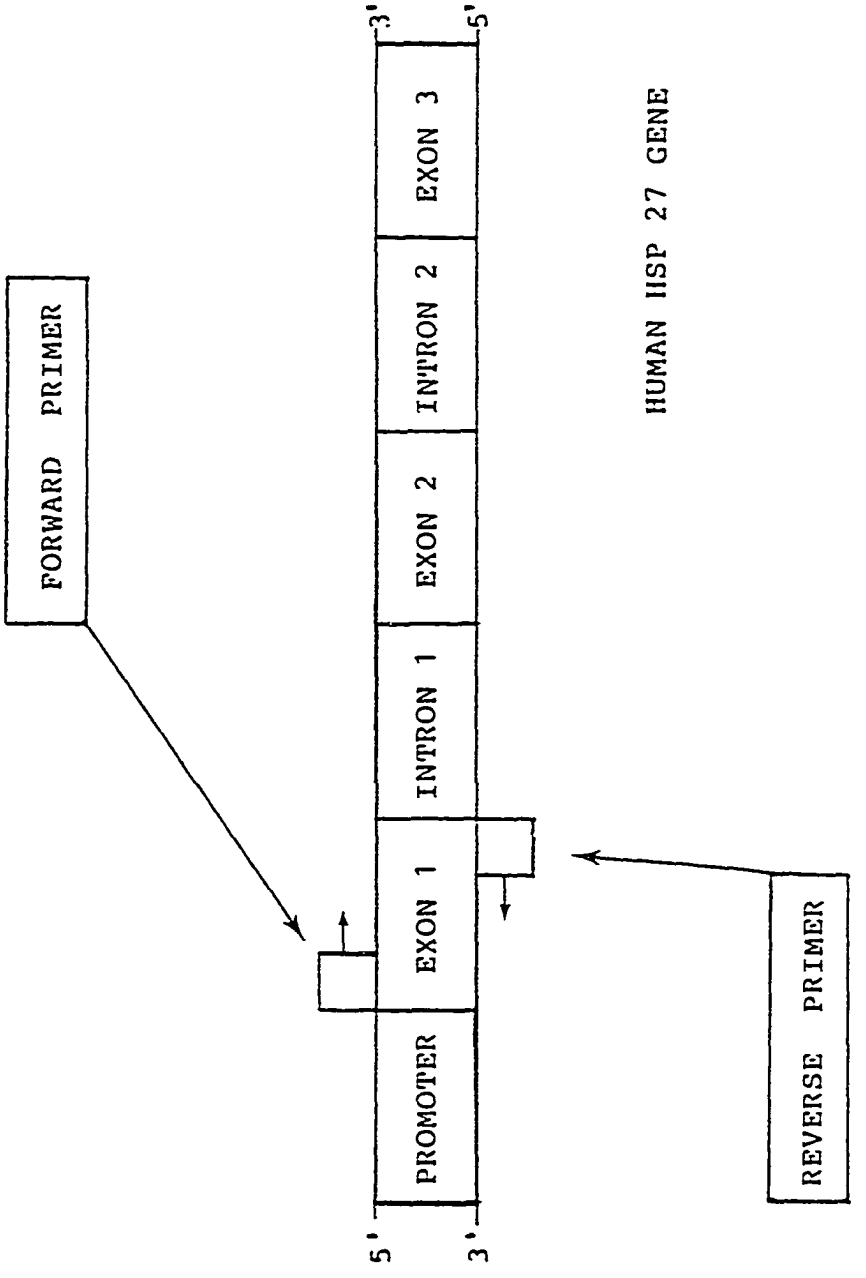


Figure 1.6 Hybridization Sites For PCR Primers. The forward and reverse PCR primers bind to the first exon region of the hsp 27 gene sequence.

PCR, we used wild-type lambda EMBL3 phage (5 μl), as it lacked the first exon region of hsp 27 cDNA.

The plasmid DNA sample (5 μl) from each clone that was analyzed underwent 30 cycles of denaturation, annealing and extension. In each cycle of the PCR, the DNA was denatured at 94°C for one minute, to separate the double-stranded DNA into single strands. The primers were then allowed to anneal to each of the single-stranded DNA at 60°C for one minute. Then, new double-stranded DNA was synthesized by DNA polymerase in the extension reaction that occurred at 72°C for one minute. To ensure that all of the DNA was in the double-stranded form, an extra 10 minute extension reaction was performed at 72°C.

The PCR products (5 μl) were separated by 2% agarose gel electrophoresis (3:1; low-melting agarose: normal agarose) in Tris acetate EDTA buffer (1X). The gel was run at 70 V for an hour, and visualized by ultraviolet transillumination of ethidium bromide-stained DNA (24). A 100 bp DNA size ladder (Gibco-BRL, Grand Island, NY) was used to identify the molecular range of the fragments. The presence of a 363 bp fragment (the size of the first exon of hsp 27 DNA), depicted a clone that was PCR-positive. The DNA was photographed using Polaroid 664K film (Polaroid Corporation, Cambridge, MA). The sizes of the inserts contained in these positive clones was then determined.

Sizing Of Fragments

The inserts contained in the positive clones were sized by cutting the DNA at specific sites using EcoRI, BamHI and HindIII restriction endonucleases. The DNA was digested at 37°C for 1.5 hours using a variety of endonucleases (1 unit each). The DNA

products were separated by 1% agarose gel electrophoresis in Tris acetate EDTA buffer (1X), at 78 V for 2 hours, and visualized by ultraviolet transillumination of ethidium bromide-stained DNA (24). A 1 kb DNA size ladder (Gibco-BRL, Grand Island, NY) was used to identify the molecular range of the fragments. The sizes of the fragments were compared to the 2,743 bp parental pGEM-3Z plasmid (Promega Corporation, Madison, WI). The sizes of the inserts in each positive clone were determined using the digestion patterns, and subtracting the total sizes of the fragments from the 2,743 bp parental pGEM-3Z plasmid. The DNA was photographed using Polaroid 664K film (Polaroid Corporation, Cambridge, MA). The restriction sites available on the inserts of the positive clones was then determined.

Restriction Mapping

The restriction maps (or plasmid maps), of the insert contained in the positive clones were determined by cutting the DNA at specific sites with the restriction enzymes SmaI, XbaI, PstI, ScaI, HincII, KpnI, EcoRI, BamHI, and HindIII. The DNA was digested at 37°C for 1.5 hours using 1 unit of these various endonucleases. The DNA products were separated by 1% agarose gel electrophoresis in Tris acetate EDTA buffer (1X), at 70 V for 2.5 hours, and visualized by ultraviolet transillumination of ethidium bromide-stained DNA (24). A 1 kb DNA size ladder (Gibco-BRL, Grand Island, NY) was used to identify the molecular range of the fragments. The sizes of the fragments were compared to the 2,743 bp parental pGEM-3Z plasmid (Promega Corporation, Madison, WI), and plasmid maps were generated on the basis of the restriction enzyme digestion patterns. The DNA was photographed using Polaroid 664K film (Polaroid Corporation, Cambridge, MA). The nucleotide sequences of the inserts contained in the positive clones were then determined.

DNA Sequencing

Automated DNA sequencing was chosen as the tool for determining the nucleotide sequences of the inserts contained in the positive clones. Hence, it was necessary to use plasmid DNA that was isolated using a Qiagen Maxiprep (Qiagen Incorporated, Chatsworth, CA), to provide a clean plasmid to be used in an automated DNA sequencer. Prior to using the Qiagen Maxiprep, the glycerol stock solutions containing the positive clones (10 µl) were grown overnight with shaking at 37°C in 3 ml of LB media containing 3 µl of ampicillin (50 mg/ml). This culture was used to inoculate a 200 ml LB media containing 200 µl of ampicillin (50 mg/ml). This 200 ml culture was further grown with shaking at 37°C overnight, and was then used to isolate plasmid DNA via the Qiagen Maxiprep. The plasmid DNA containing the positive clones was resuspended in distilled water (9.5 µl), prior to automated DNA sequencing.

Fluorescent dye dideoxy chain-terminating DNA sequencing was performed by a DNA Sequencing Laboratory (Biotechnology Division, University of Arizona, AZ), on an Applied Biosystems Automated Sequencer Model 377. The DNA sequence reaction was conducted using suitable synthetic oligonucleotide primers for the T7 and SP6 promoters, and 1 µg DNA samples for each reaction.

DNAsis Software Analysis

Nucleotide sequence comparisons were performed using the DNAsis computer software's Maximum Matching Alignment (Hitachi Software Engineering Company Limited, Yokohama, Japan). This function analyzed the similarity between two sequences according to the Needleman-Wunsch Algorithm (25). It aligned two sequences side by side, assigning

a score to the most similar alignment. If the score was less then a specific value, the function went back over the aligned sequence and inserted gaps, and rescored the new alignment. Gaps were typically inserted into the shorter of the two sequences, and a penalty was given for gaps. The highest score was determined and reported in relation to a homology percentage. In the illustrations of the alignments, gaps were designated as hyphens. Matching regions were indicated by vertical bars. If one of the sequences had been shifted relative to the other, vacant positions were indicated by dots.

In the initial analysis of the DNA sequence data, the nucleotide sequences of the inserts contained in the positive clones (and the complementary strand DNA sequences) were compared to the human hsp 27 gene using the Maximum Matching Alignment. The nucleotide sequence of the human hsp 27 gene is assigned the GenBank accession number XO3900, and is shown in Appendix II (p. 115).

We desired to compare the predicted amino acid sequences of the inserts to the amino acid sequence of the human hsp 27 gene. Therefore, we searched for sites in the DNA sequence that would indicate the coding strand that would translate the DNA sequence into a predicted amino acid sequence. This was accomplished using a modification of the Restriction Enzyme Search Function. This function was intended to be used to search for restriction enzyme sites within a nucleotide sequence, but was modified for our purposes. The excision sites for an intron 5'-splice site is AA!GUAAGU (26). The excision sites for an intron 3'-splice site is UNCAG!G, with a polypyrimidine string preceding the 3'-splice sites (26). The symbol, '!', denotes the site of actual bond cleavage. The 5'-splice site, and the 3'-splice site were searched for as a motif, within the DNA sequences of the inserts. The

sequence CCCCC, and TTTTT were used here to search for the string of polypyrimidines preceding the 3`-splice site.

If 5' and 3'-splice sites were found on a particular strand of an insert, this strand was used as the coding strand. Hence, the introns (the intervening regions between the 5' and 3'-splice sites), were removed from the DNA sequences of the inserts, and the remaining DNA information (exons) were joined together. The exon information from the coding strand, was translated into predicted amino acid sequences in three reading frames. If no introns were found within an insert, we could not determine the coding strand. Therefore, the insert was analyzed in all six reading frames, by obtaining a variety of open reading frames for that insert, that would result in the synthesis of a variety of predicted amino acid sequences.

The determination of the open reading frame was performed using the Open Reading Frame Function. Each strand of a DNA molecule can be read in one of three reading frames, for the synthesis of one of three possible proteins. Hence, a double-stranded DNA molecule can be translated into six possible proteins. The open reading frame is the one particular reading frame out of the six, that is the most likely reading frame for the synthesis of the appropriate protein.

The Open Reading Frame Function is based on the Fickett Algorithm (27). The Fickett Algorithm scans a query sequence for codons known to be associated with the initiation (ATG) and termination of transcription (TAA, TAG and TGA). In this way, the algorithm distinguishes true protein coding sequences from false open reading frames. This function indicated the predicted open reading frame by displaying the sizes of the proteins produced by each reading frame (in Daltons).

Once the open reading frames were determined, the translation of the nucleotide

sequences into predicted amino acid sequences was performed using the DNA-To-Protein Translation Function. The predicted amino acid sequences of the inserts contained in the positive clones were compared to the amino acid sequence of the human hsp 27 gene, using the Maximum Matching Alignment. The amino acid sequence information of human hsp 27 is assigned the GenBank accession number XO3900, and is shown in Appendix III (p. 131).

In order to further determine if our inserts were the closely-related, linked human hsp 27 genes, we searched for regions in the DNA sequence of the inserts, that would indicate sites to which the primers should have bound during PCR. If the inserts were closely-related to human hsp 27, we predicted that within the nucleotide sequences of the inserts, we would identify the motifs for the corresponding hybridization sites for the forward and reverse primer PCR primers. The motifs of the complementary sequences for the forward and reverse PCR primers (to the first exon of the human hsp 27 gene), were searched for using a modification of the Restriction Enzyme Search Function.

Blast Analysis

The Blast Algorithm was used to determine the identity of the inserts with known databases. Homology searches using the nucleotide sequences of the insert contained in the positive clones were conducted by using the Blast Algorithm at GenomeNet, Japan (Basic local alignment search tool at www.blast.genome.ad.jp.). This algorithm is derived from the Lipman-Pearson Algorithm (28), and facilitated the search for similarities between these newly determined sequences, and already available databases. The Blast Algorithm selected DNA sequences in two DNA databases, that were similar to the query sequence (Blastn). It also used the DNA sequence of a query, translated it into six possible amino acid sequences,

and searched for these six sequences in two protein database (Blastx). By translating each strand of double-stranded DNA into three reading frames for the synthesis of six possible proteins, comparisons were made at the protein level that allowed for mutations in DNA. Protein comparisons were used to identify weak similarities encoded in nucleic acids that were not apparent from just looking at similarities between two nucleotide sequences.

The sequences of the insert contained in the positive clones were searched for in the GenBank nucleotide database, GenBank nucleotide updated database, SwissProt protein database and SwissProt protein updated database. The BLOSUM62 scoring matrix was selected for the Blastx searches. No scoring matrix was required for the Blastn searches. The results were displayed as the names of the "sequences producing high scoring segment pairs", and shown as visual homology regions and identity percentages.

CHAPTER 3

RESULTS

Overview

Cloning was used to isolate specific DNA fragments of linked human hsp 27 genes. The process of cloning the DNA fragments began with the construction of a genomic library. The DNA fragments from the human lung cancer cell line A549 genomic library were ligated into lambda phage, a DNA virus that can accommodate large inserts. The human genomic library constructed in lambda phage was screened for hsp 27 genes. Colony hybridization was used to identify the phages that contained the hsp 27-related genes, via the use of a probe. The hsp 27-related genes contained in lambda phage (the positive plaques), were enzymatically cut into smaller fragments. These small fragments were ligated into a plasmid vector for further analysis (subcloning). In this way, we narrowed our search for the hsp 27 genes from within larger pieces of DNA to amongst smaller pieces of DNA.

PCR was used as a second technique, to further determine if the DNA inserts in the recombinant plasmids were positive clones containing the desired Hsp 27 genes. Restriction endonucleases were used to determine the sizes of the DNA fragments (inserts), contained in the positive clones. To determine the DNA sequence information of the inserts, the inserts

were subjected to automated DNA sequencing. To determine if the inserts contained in the positive clones were similar to hsp 27 in their nucleotide sequences and predicted amino acid sequences, the DNAsis software program, and the Blast Algorithm were used. Finally, the nucleotide sequence information for the inserts contained in three of the positive clones were deposited in the GenBank database.

In presenting the results of this study, I will initially show the specific results for a particular positive clone (DNA 4), from the stage of colony hybridization until the results for the Blast algorithm analysis. Then, I will discuss the results for four other positive clones.

Colony Hybridization

The human genomic library constructed in lambda EMBL3 phage was screened for the human hsp 27 genes, as outlined in the Materials and Methods section. A total of three rounds of colony hybridization screenings were performed. It was necessary to perform three screenings in order to isolate unique positive plaques that may contain the hsp 27 genes. Plaques were purified before further analysis (24). We obtained 100,000 plaques from each round of colony hybridization. Of these 300,000 total plaques, 50 were positive plaques.

One of the plates used in colony hybridization (Plate 9), contained 8 positive plaques, and is shown in Figure 2.1 (p. 34). This figure also shows the positive control that was used for colony hybridization. We used *E. coli* cells containing the plasmid pTRI/62Y.16.7 as a positive control, because this plasmid contained full-length hsp 27 cDNA and would hybridize to the probe. The positive plaques were then subjected to subcloning.

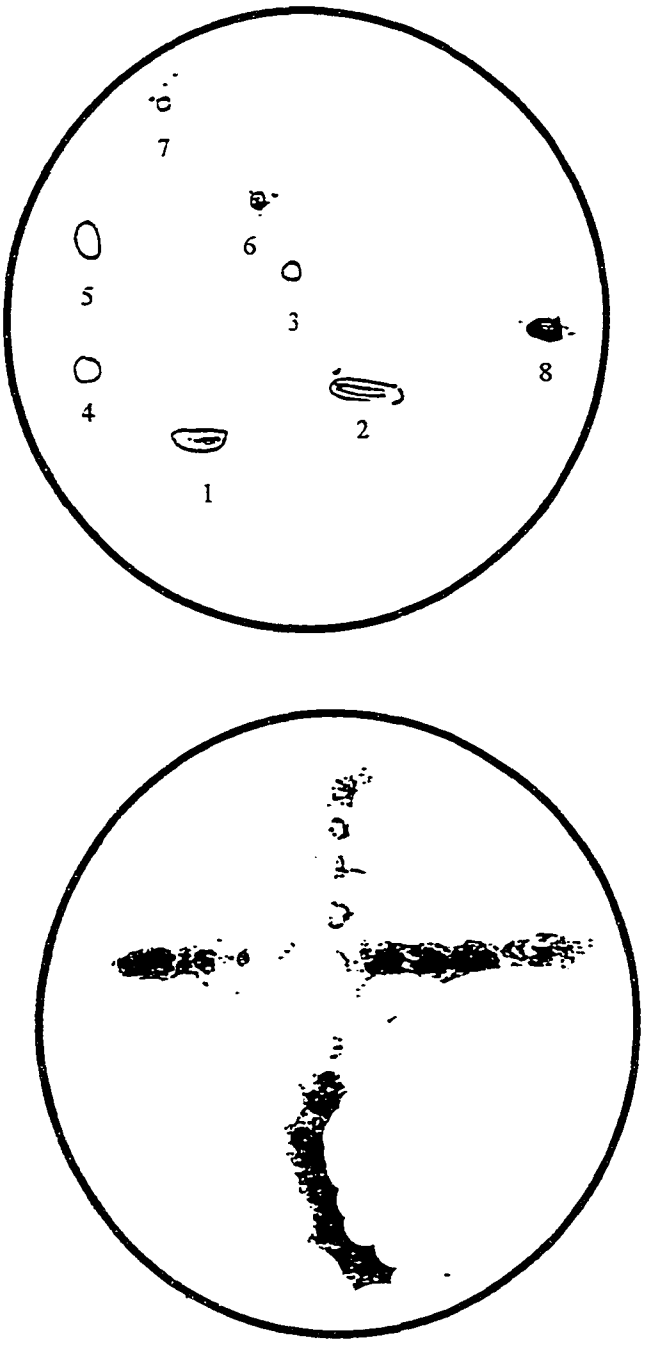


Figure 2.1 Colony Hybridization Results. The upper filter shows positive clones from plate 9. Dark spots on the nitrocellulose paper are the sites of hybridization to the human hsp 27 cDNA probe. Positive clones have been circled for further emphasis. The lower filter shows the positive control (the plasmid pTRI/62Y.16.7), containing human hsp 27 cDNA that served as the probe.

Subcloning

The positive plaques identified from colony hybridization were purified, and subcloned into pGEM-3Z, as stated in the Materials and Methods section. Using Colony hybridization, a positive plaque was chosen from Plate 9, and called 9.1. Following the purification of plaque 9.1, subcloning into pGEM-3Z, transformation into *E. coli* JM109 cells, and plating on LB + Ampicillin + X-gal + IPTG media (on Plate 9.1), 24 recombinants were selected as white colonies. These white colonies were selected as they indicated plasmids containing DNA inserts.

PCR Analysis

PCR was used as a second technique to further determine if the DNA inserts in the recombinant plasmids contained hsp 27 genes. We amplified the first exon region of human hsp 27 cDNA (363 bp), using forward and reverse primers to the first exon of human hsp 27 cDNA. If human hsp 27 genes were present in a particular recombinant plasmid, after gel electrophoresis analysis, a band indicative of the 363 bp first exon region of the hsp 27 gene would be present.

Of the 24 white recombinant colonies that were screened from plate 9.1 using PCR amplification, two colonies, 9.1/1 and 9.1/2 were found to be PCR-positive. Figure 2.2 (p. 36) shows the gel electrophoresis result for the positive clones obtained from the PCR analysis of plate 9.1. In this picture, a positive clone is identified as a 363 bp amplified band on the gel. Lane 2 is the positive control, which contains a 363 bp band, and is indicated by an arrow. Lane 2 also shows a smear of large DNA fragments that were also amplified. The

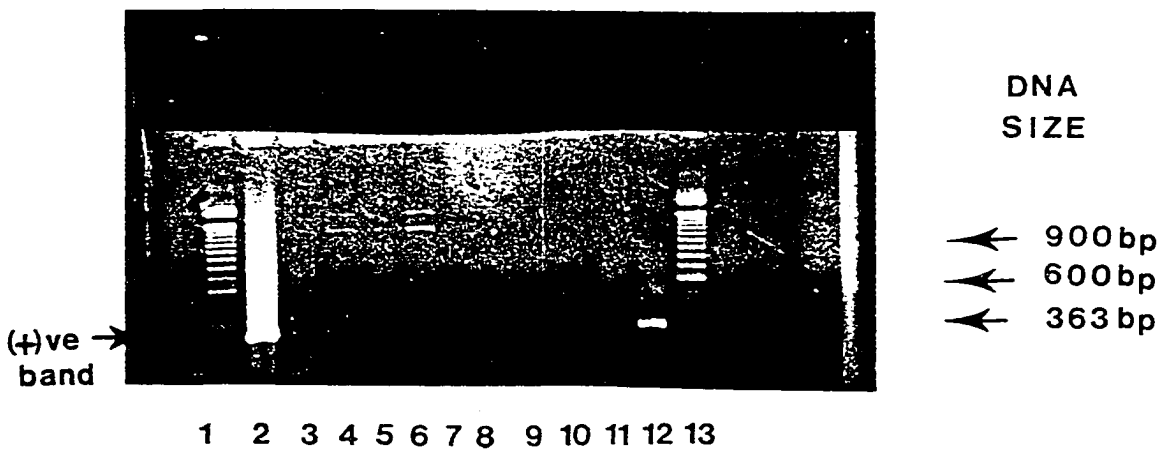


Figure 2.2 Gel Electrophoresis Results For PCR Analysis Of DNA 4. Lane 1 contains the 100 bp DNA mass ladder. Lane 2 is the positive control (the plasmid pTRI/62Y.16.7, containing human hsp 27 cDNA, and hence shows the 363 bp amplified band). Lanes 3-10 show the negative results for other clones. Lane 11 shows the amplified 363bp band from Plate 9.1/1 (a positive clone). Lane 12 shows the amplified 363bp band from Plate 9.1/2 (a positive clone). Lane 13 contains the 100 bp DNA mass ladder.

positive clone shown in Lane 12, 9.1/2, was picked and chosen for further study. This positive clone was named pDNA 4. pDNA 4 contained the insert, DNA 4.

We conducted two more rounds of colony hybridization, which yielded four other positive clones, that were chosen for further analysis. The four additional positive clones were designated as pDNA 1 (containing the insert, DNA 1), pDNA 2 (containing the insert, DNA 2) , pDNA 3 (containing the insert, DNA 3) and pDNA 5 (containing the insert, DNA 5). It was then necessary to obtain more information, regarding the inserts contained in the positive clones.

Sizing Of Fragments

The inserts contained in the positive clones were sized by cutting the DNA at specific sites using EcoRI, BamHI and HindIII restriction endonucleases and separating the fragments on an agarose gel. These restriction enzymes were chosen because they possessed cleavage sites within the multiple cloning site of pGEM-3Z. Hence, after cleavage with these enzymes, the total sizes of the fragments obtained could be used to determine the size of the insert. The sizes of the inserts in each positive clone were determined using the digestion patterns, and subtracting the total sizes of the fragments from the 2,743 bp parental pGEM-3Z plasmid. Figure 2.3 (p. 38) shows a representation of an agarose gel showing the digestion patterns for pDNA 4. In Lane 2, we observe pDNA 4 that has been cut with EcoRI. Lane 2 shows the single fragment that results from this digestion - a 3,900 bp fragment. In addition, Lane 3 shows the pDNA 4 digestion fragment from restriction enzyme BamHI. Lane 3 also contains the 3,900 bp fragment. The actual size of DNA 4 was calculated by subtracting the 2,743 bp pGEM-3Z vector DNA from the 3,900 bp size of the band in Lane 2. The restriction

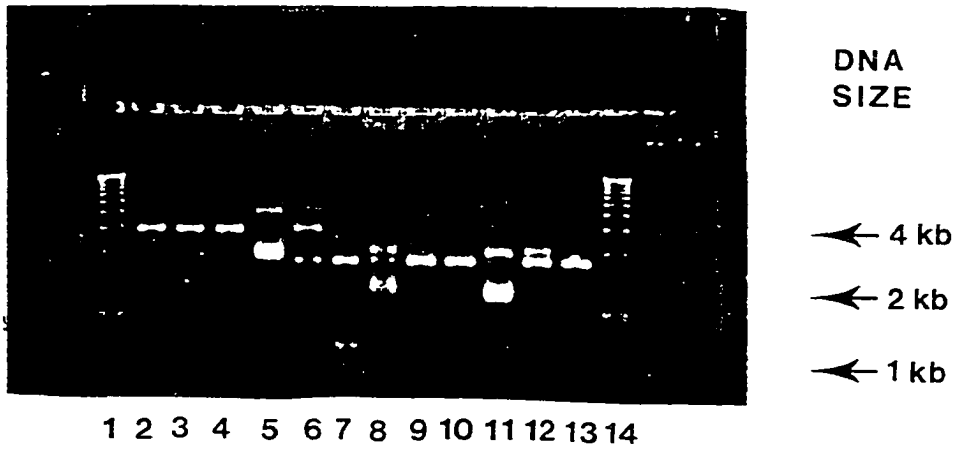


Figure 2.3 Gel Electrophoresis Results For Sizing Of DNA 4. Lane 1 = 1 kb DNA mass ladder. Lane 2 = EcoRI-cut pDNA 4. Lane 3 = BamHI-cut pDNA4. Lane 4 = HindIII-cut pDNA 4. Lane 5 = Undigested pDNA 4. Lane 6 = EcoRI-cut + HindIII-cut pDNA 4. Lane 7 = BamHI-cut + HindIII cut pDNA 4. Lanes 8-13 = pDNA 5 digests. Lane 14 = 1 kb DNA mass ladder.

analysis of pDNA 4 indicated that we had cloned a fragment, (DNA 4) that was approximately 1,200 bp long.

Restriction Mapping

Restriction digestion sites in pDNA 4 were examined by cutting the DNA at specific sites using the restriction enzymes EcoRI, PstI, ScaI, BamHI and HindIII. The restriction enzymes EcoRI and PstI were chosen because they cleaved within the human hsp 27 gene sequence. Since we were looking for hsp 27 genes, it was possible that one of these restriction enzymes would also cleave within the inserts of our positive clones. The restriction enzymes BamHI and HindIII were chosen because they cleaved within the multiple cloning site of pGEM-3Z. The restriction enzyme ScaI was chosen randomly.

If no restriction site was present in DNA 4, but was present in the multiple cloning site, a single fragment would be identified as a single band on the gel. This is because the enzyme would cut the plasmid in the multiple cloning site, thereby linearizing the plasmid, and producing a single fragment. If a single restriction site was present in DNA 4, and also present in the multiple cloning site, two fragments would be identified as two bands on the gel. This is because the enzyme would cut the plasmid in the multiple cloning site, as well as within the insert region. Hence, x bands identified on the gel would be indicative of (x-1) locations at which the particular restriction enzyme cleaved within the insert. This reasoning was true for all single digestions involving enzymes that cleaved within the multiple cloning site of the plasmid.

A plasmid map was generated for pDNA 4, on the basis of the restriction enzyme

digestion patterns. The plasmid map for pDNA 4 is shown in Figure 2.4 (p. 41). The plasmid map indicated that there were no sites in DNA 4 for EcoRI, PstI, ScaI, BamHI or HindIII.

DNA Sequence Analysis

Since the plasmid map for pDNA 4 showed that the size of DNA 4 was approximately 1,200 bp in length, and since automated DNA sequencing can determine the nucleotide sequence of about 1,000 bp in a single run, it was apparent that a large portion of the nucleotide sequence information for DNA 4 could be obtained from just two rounds of automated DNA sequencing (using the T7 and SP6 promoters). The sequencing strategy for regions of DNA 4 is shown in Figure 2.5 (p. 42).

An appropriate primer to the T7 RNA polymerase promoter was used to determine the nucleotide sequence of the top strand of DNA 4 (DNA 4/T7) shown in Appendix II (p. 118). DNA 4/T7 contains 942 bp of sequence information. To obtain more sequence information for DNA 4, another appropriate primer to the SP6 RNA polymerase promoter was used to determine the nucleotide sequence of the complementary strand of DNA 4. The nucleotide sequence encoding DNA 4 using the SP6 promoter (DNA 4/SP6) is shown in Appendix II (p. 119) and contains 944 bp of sequence information.

The DNA sequences for both DNA 4/T7 and DNA 4/SP6 contain a small number of unknown nucleotides designated as "N". Unknown nucleotides are sometimes present in sequence information obtained from automated DNA sequencing. Unknown nucleotides occur when the automated DNA sequencer is unable to distinguish the overlapping signals obtained as it encounters more than one nucleotide, during a single point of analysis.

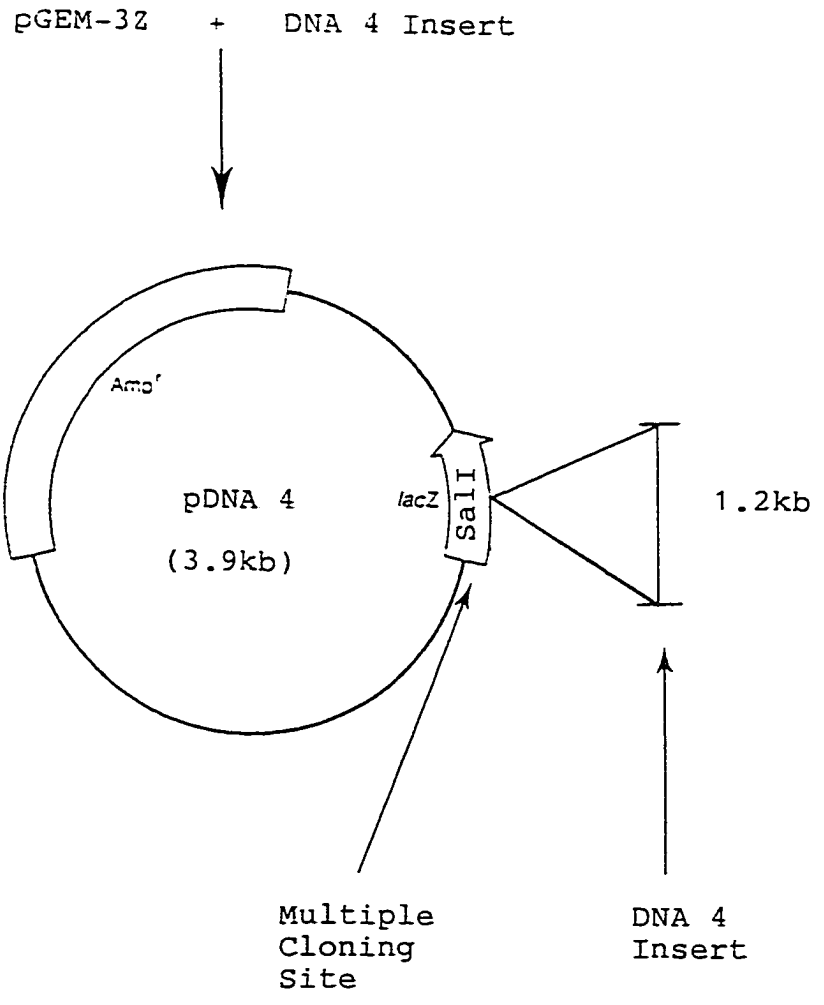


Figure 2.4 Plasmid Map For pDNA 4.

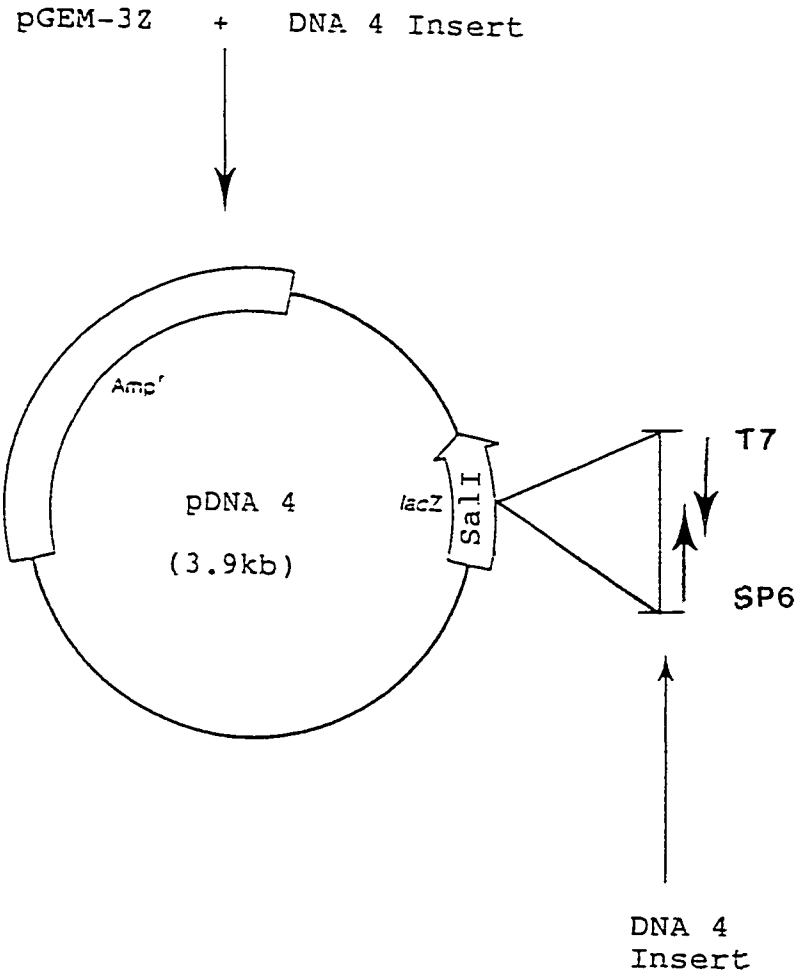


Figure 2.5 Sequencing Strategy For DNA 4.

DNAsis Analysis - DNA Level

Before DNAsis Software analysis was used, it was necessary to eliminate any pGEM-3Z vector nucleotide information that was present in the DNA 4/T7 nucleotide sequence. This was done by visually comparing the initial information of DNA 4/T7 with that of pGEM-3Z, and manually eliminating the short homologous region (1-29 bp). Once this was done, the DNA 4/T7 information that lacked any pGEM-3Z information was redesignated as DNA 4-1. The nucleotide information removed from DNA 4/T7 to obtain DNA 4-1 is underlined in Appendix II (p. 118). The DNA 4-1 nucleotide information was the partial sequence information for the insert contained in pDNA 4 using the T7 promoter. The nucleotide sequence for DNA 4-1 was used in subsequent analysis.

It was also necessary to eliminate any pGEM-3Z vector nucleotide information present in the DNA 4/SP6 nucleotide sequence. This was done by visually comparing the initial information of DNA 4/SP6 with that of pGEM-3Z, and manually eliminating the short homologous region (1-26 bp). Once this was done, the DNA 4/SP6 information that lacked any pGEM-3Z information was redesignated as DNA 4-2. The nucleotide information removed from DNA 4/SP6 to obtain DNA 4-2 is underlined in Appendix II (p. 119). The DNA 4-2 nucleotide information was the partial sequence information for the insert contained in pDNA 4 using the SP6 promoter. The nucleotide sequence for DNA 4-2 was used in subsequent analysis.

In order to obtain the complete sequence information for the approximately 1,200 bp DNA 4 insert, we required the 5'-sequence information for DNA 4-1 (normal strand), the 5'-information for DNA 4-2 (complementary strand), and the overlapping regions between these two sequences. To obtain the overlapping sequence information, the Maximum

Matching Alignment was used to compare the similarity between regions of the DNA 4-1 sequence information (normal strand), and regions of the DNA 4-2 sequence information (complementary strand). The results for this comparison are shown in Appendix IV (p. 133). A graphical illustration of the homologous regions between these two DNA sequences is shown in Figure 2.6 (p. 45).

Since we had more confidence in the 5'-DNA 4-1 information (as it was closer to the T7 sequencing primer), it was used to start the complete sequence of DNA 4. We had more confidence in the 5'-DNA 4-2 information (as it was closer to the SP6 sequencing primer), hence, it was used to finish the DNA 4 complete sequence. When a gap, or unknown nucleotide in either of the sequences was encountered, it was replaced with the appropriate nucleotide from the other strand. The new complete sequence information for DNA 4 was 1,147 bp in length, and is shown in Appendix II (p. 120).

The DNAsis computer software was used to compare the similarity between two DNA sequences. A homology of greater than 50% represented a good homology. In the positive control for the Maximum Matching Alignment, the DNA sequence for the human hsp 27 gene was aligned against itself and produced a 100% homology. Then, the DNA sequence of human hsp 27 cDNA (789 bp) was aligned against the DNA sequence of the entire human hsp 27 gene (2,496 bp). A 48% homology was evident, and is shown in Appendix IV (p. 135). This alignment of the human hsp 27 cDNA with the human hsp 27 gene, should have produced a homology percentage much greater than 48%, as these two sequences are definitely related. Visual analysis of the homology regions indicated that certain segments of the alignments showed better homologies than were indicated by the total homology percentage. Thus, when the cDNA of human hsp 27 was aligned against the most

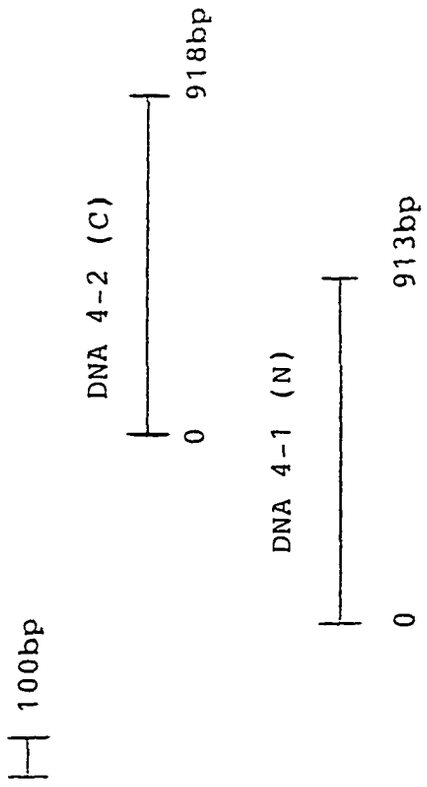


Figure 2.6 Homologous DNA Areas Between DNA 4-1 (Normal Strand), And DNA 4-2 (Complementary Strand).

homologous, similar-sized region of the human hsp 27 gene (198-418 bp region of the first exon), an increased homology of 100% was shown. This indicated that the pattern of alignment was just as important as the homology percentage, in determining relatedness.

The Maximum Matching Alignment was used to determine if the nucleotide sequence of DNA 4 (the normal strand) was similar to the nucleotide sequence of the human hsp 27 gene. In this comparison, a 38% homology was indicated by the results shown in Appendix IV (p. 139).

Alignments were performed using DNA 4 (normal strand), and similar-sized regions of the human hsp 27 gene. The alignment of the normal strand of DNA 4, and a similar-sized region of the human hsp 27 gene (1-1,147 bp), produced a 50% homology. The alignment of the normal strand of DNA 4, and another similar-sized region of the human hsp 27 gene (1,147-2,294 bp), produced a 51% homology. The alignment of the normal strand of DNA 4, and the most homologous, similar-sized region of the human hsp 27 gene (1,349-2,496 bp), is shown in Appendix V (p. 188). A homology percentage of 51% was indicated. A graphical illustration of this comparison is shown in Figure 2.7 (p. 47).

At this point, we were unaware of which strand of the double-stranded DNA 4 (the normal or complementary strand) would lead to the synthesis of the appropriate protein. Hence, the DNAsis Maximum Matching Alignment was used to determine if the complementary strand of DNA 4 was similar to the nucleotide sequence of the human hsp 27 gene. In this comparison, a 38% homology was indicated by the results shown in Appendix IV (p. 143).

Alignments were performed using DNA 4 (complementary strand), and similar-sized regions of the human hsp 27 gene. The alignment of the complementary strand of DNA 4,

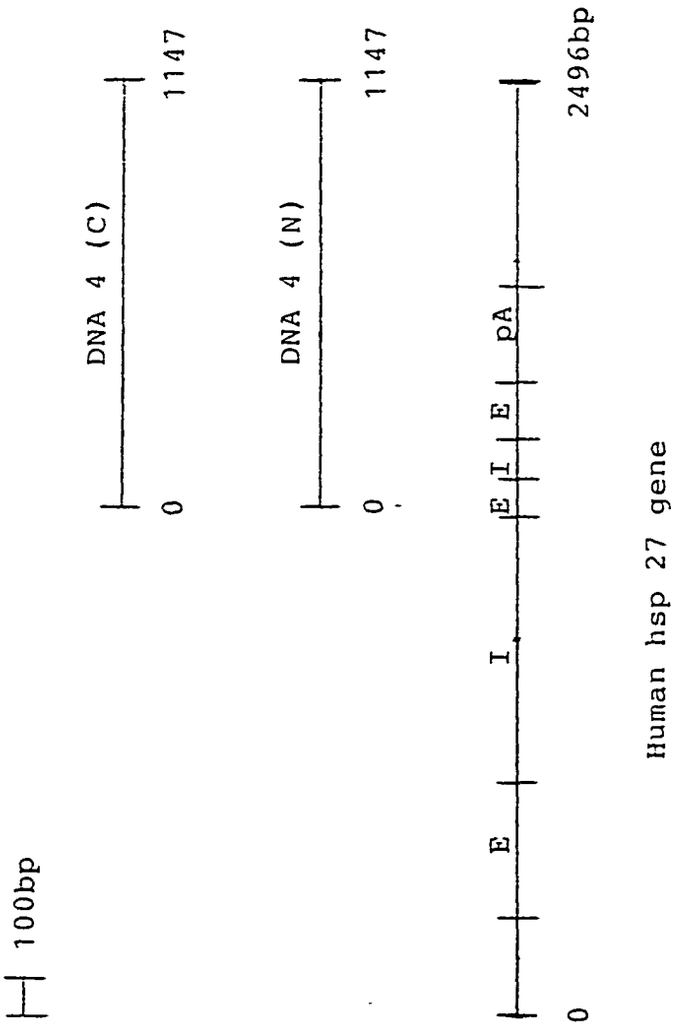


Figure 2.7 Homologous DNA Areas Between DNA 4 (Normal And Complementary Strands), And The Human hsp 27 Gene. The exon (E), intron (I), and polyA (pA) regions of the human hsp 27 gene are shown.

and a similar-sized region of the human hsp 27 gene (1-1,147 bp), produced a 49% homology. The alignment of the complementary strand of DNA 4, and another similar-sized region of the human hsp 27 gene (1,147-2,294 bp), produced a 51% homology. The alignment of the complementary strand of DNA 4, and the most homologous, similar-sized region of the human hsp 27 gene (1,349-2,496 bp), is shown in Appendix V (p. 190). A homology percentage of 51% was indicated. A graphical illustration of this comparison is shown in Figure 2.7 (p. 47).

Blast Analysis - DNA Level

In order to determine if DNA 4 was similar to any known DNA sequence, the DNA sequence information for DNA 4 was analyzed using the Blast Algorithm (Basic Local Alignment Search Tool). This computer algorithm searched for DNA 4 in the GenBank nucleotide database and the GenBank nucleotide updated database (Blastn).

An identity percentage of greater than 50% represented a good homology. In the control for the Blastn Algorithm, the DNA sequence for the human hsp 27 gene was searched for, and found to be 99% homologous to itself in the GenBank nucleotide database. The human hsp 27 gene was also found to be 75-90% homologous to the mouse hsp 27 nucleotide sequence, and 72-88% homologous to the rat hsp 27 nucleotide sequence.

We attempted to search for the nucleotide sequence of DNA 4 (normal strand), using the Blastn Algorithm. The results were displayed as the names of the "sequences producing high scoring segment pairs", and were shown as visual homology regions and identity percentages. The results for the Blastn searches using the nucleotide sequence of DNA 4 (normal strand) are shown in Table 1.1 (p. 49). Only 2 DNA sequences were identified as

Table 1.1 Blastn Results For DNA 4.

Homolog	GenBank Accession Number	Homology	Length
Homo sapiens cDNA clone NhHMPu	AI042521	75%	47 bp
Homo sapiens Chromosome 15q26.1 PAC clone	AC005319	75%	47 bp

as being related to this query sequence. The human hsp 27 gene sequence was not identified as being homologous to the nucleotide sequence for DNA 4 (normal strand).

DNAsis Analysis - Protein Level

Due to the fact there is a wobble in the genetic code, allowing for the translation of more than one codon into the same amino acid, it was possible that two DNA sequences that differed slightly at the nucleotide level, might show a greater homology at the protein level. Therefore, we looked for homologies at the protein level, that may not have been apparent at the DNA level. In order to compare the predicted amino acid sequence of DNA 4 to the amino acid sequence of the human hsp 27 gene, it was necessary to determine the coding strand for DNA 4, that would ultimately translate the DNA sequence into a predicted amino acid sequence. Therefore, we attempted to identify the coding strand by determining which strand of DNA 4 contained exons and introns, by looking at the 5' and 3'-splice sites, as outlined in the Materials and Methods section.

DNA 4 (normal strand) showed the presence of 2 potential 3'-splice sites at positions 84 and 841. The second potential splice site was preceded by a string of polypyrimidines (TTTTT) at position 487. No 5'-splice sites were identified. DNA 4 (complementary strand) showed the presence of one potential 3'-splice site at position 175. The potential splice site was not preceded by a string of polypyrimidines, nor was it preceded by any 5'-splice sites. Since there were 3'-splice sites on both strands of DNA 4, we were unable to determine which strand of DNA 4 was the coding strand. Therefore, we used a more random approach that involved the identification of predicted open reading frames that were present in all six reading frames.

The Open Reading Frame Function was used to translate the nucleotide sequence of DNA 4 into a variety of predicted amino acid sequences. This was done by translating the nucleotide sequence in both orientations, and in all three reading frames. The results for the open reading frames for both the normal and complementary strands of DNA 4 are shown in Table 1.2 (p. 52). In order to determine which DNA regions to translate into predicted amino acid sequences, we considered 20 amino acid residues (aa) to be unique. Since an amino acid is about 110 Daltons in size, a 20 amino acid sequence is 2,200 Da in size. Therefore, we picked at least one sequence from every reading frame (that did not contain an unknown nucleotide), and would result in a predicted amino acid sequence of greater that 2,200 Da in size. Then, the selected Open Reading Frames were translated into proteins using the DNA-To-Protein Translation Function. The Maximum Matching Alignment was used to compare these predicted amino acid sequences to the amino acid sequence for the human hsp 27 gene. Table 1.3 (p. 53) shows the results for these comparisons.

Blast Analysis - Protein Level

In order to determine if DNA 4 was similar to any known amino acid sequence, the amino acid sequence information for DNA 4 was analyzed using the Blast Algorithm (Basic Local Alignment Search Tool). This computer algorithm searched for DNA 4, as a protein in all six reading frames, in the SwissProt protein database and the SwissProt protein updated database (Blastx). The program also determined if the DNA insert in this positive clone would select hsp 27 as a homologous sequence in these databases.

As a control for the Blastx algorithm, the human hsp 27 gene sequence was searched for and found to be 100% homologous to the human hsp 27 amino acid sequence in the

Table 1.2 Open Reading Frames For DNA 4.

NORMAL

	1 Frame		
	Start	End	MW
	1	46	1704.85
	112	124	563.62
	160	265	3916.38
A	394	601	7751.61
	1054	1144	3333.62
	2 Frame		
	Start	End	MW
	2	8	271.31
	425	506	3094.49
B	524	788	8993.43
	713	788	2493.75
	896	935	1603.83
	1001	1145	5520.39
	3 Frame		
	Start	End	MW
C	3	426	15678.73
D	72	426	13438.36
	306	426	4442.85
	402	426	1041.15
	492	501	395.51
	540	558	706.78

COMPLEMENTARY

	1 Frame		
	Start	End	MW
E	1	40	1840.07
	616	646	1175.45
	841	850	347.42
	2 Frame		
	Start	End	MW
F	2	359	13402.20
	623	641	796.93
	878	911	1280.57
	1025	1145	4658.17
	3 Frame		
	Start	End	MW
G	3	150	5319.89
	6	150	5163.71
	18	150	4745.27
	48	150	3700.12
	726	828	4059.54

Table 1.3 Amino Acid Alignments Between DNA 4 And Human hsp 27.

Sequence	Strand	Homology To hsp 27	Length
A	Normal	17%	69 aa
B	Normal	25%	88 aa
C	Normal	30%	141 aa
D	Normal	28%	118 aa
E	Complementary	7%	13 aa
F	Complementary	27%	119 aa
G	Complementary	19%	49 aa

SwissProt protein database. It was also found to be 86% homologous to the mouse hsp 27 amino acid sequence, and 87% homologous to the rat hsp 27 amino acid sequence. The human hsp 27 gene sequence was also found to be 55% homologous to the human alpha-crystallin A-chain amino acid sequence, and 45% homologous to the human alpha-crystallin B-chain amino acid sequence.

The Blastx Algorithm was then used to search for the nucleotide sequence of DNA 4 (normal strand). Only 3 proteins were identified as being homologous to this query sequence. The results for the homologs are shown in Table 1.4 (p. 55). The human hsp 27 protein was not identified as being homologous to the nucleotide sequence of DNA 4 (normal strand).

PCR Primer Location Analysis

In order to further determine if DNA 4 was one of the closely-related hsp 27 genes, we attempted to identify sites where the PCR primers should have bound to the DNA 4 nucleotide sequence (normal and complementary strands). A modification of the DNAsis Restriction Enzyme Search Function was used to search for the motifs of the complementary strands of the forward and reverse PCR primers to the first exon of human hsp 27 gene. The results indicated that the PCR primers did not bind to any region on the DNA 4 nucleotide sequence.

Table 1.4 Blastx Results For DNA 4.

Homolog	GenBank Accession Number	Homology	Length
E.coli Phospho-ribosyl AMP	HIS2ECOLI	31%	7 aa
Rat Cytochrome P450	CPC7RAT	40%	23 aa
SHFL Phospho-	HIS2SHIFL	29%	41aa

Other Positive Clone Information

DNA 1 - Size And Restriction Map

The positive clone, pDNA 1, was sized by cutting the DNA at specific sites, using restriction endonucleases, and separating the fragments on an agarose gel. The restriction analysis of pDNA 1 indicated that we had cloned an insert, (DNA 1) that was approximately 4,000 bp long. Restriction digestion sites in the positive clone, pDNA 1, were examined by cutting the DNA at specific sites using all the restriction enzymes mentioned in the Materials and Methods section. A plasmid map was then generated for pDNA 1, and is shown in Figure 3.1 (p. 57).

DNA 1 - DNA Sequence

Since the plasmid map for the positive clone pDNA 1 showed that the size of DNA 1 was approximately 4,000 bp in length, it was apparent that only about 2,000 bp of the nucleotide sequence information for DNA 1 could be obtained from two rounds of automated DNA sequencing (using the T7 and SP6 promoters). The sequencing strategy for regions of DNA 1 is shown in Figure 3.2 (p. 58).

The nucleotide sequence encoding DNA 1 using the T7 promoter (DNA 1/T7) is shown in Appendix II (p. 122). DNA 1/T7 contains 1,122 bp of sequence information. The nucleotide sequence encoding DNA 1 using the SP6 promoter (DNA 1/SP6) is shown in Appendix II (p. 124). DNA 1/SP6 contains 958 bp of sequence information.

Manipulation of the positive clone pDNA 1 was necessary to obtain more nucleotide information for DNA 1. Since the plasmid map for pDNA 1 indicated the presence of a

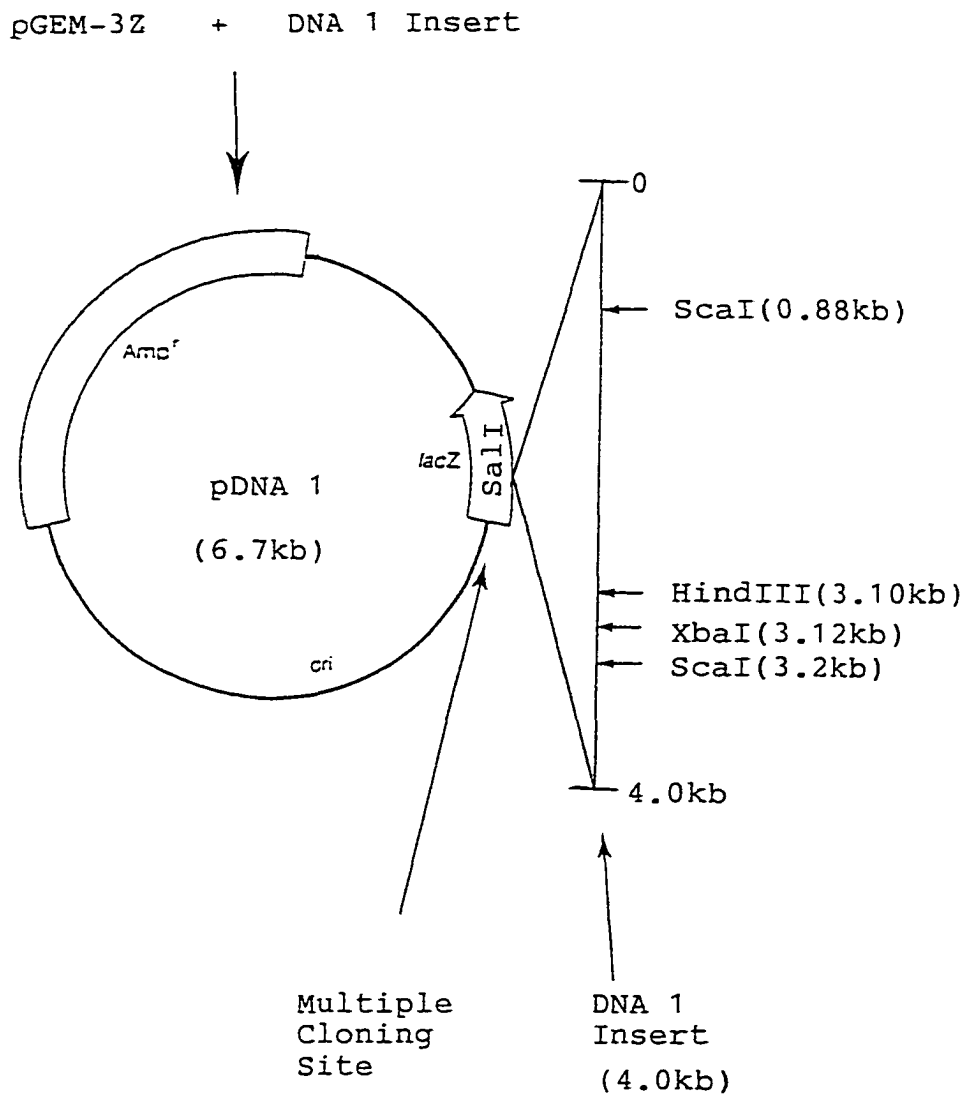


Figure 3.1 Plasmid Map For pDNA 1.

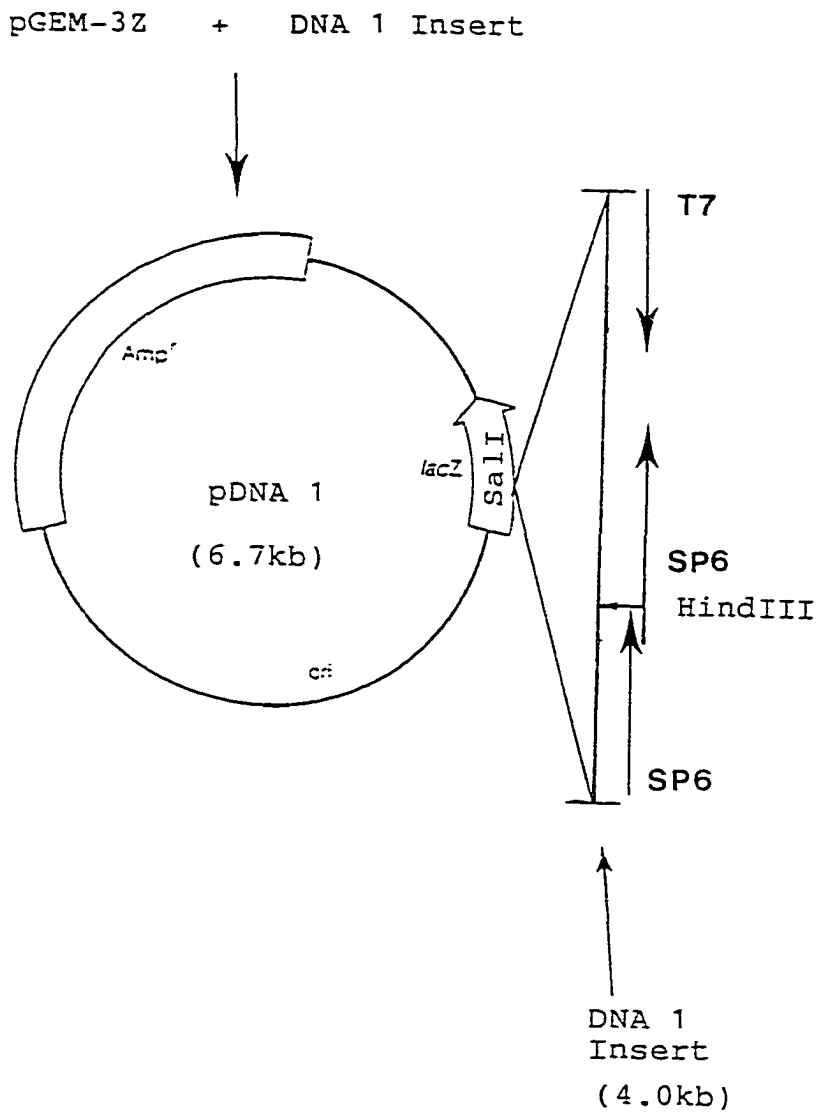


Figure 3.2 Sequencing Strategy For DNA 1.

HindIII site within DNA 1, it was apparent that a HindIII restriction digestion to remove a 1,000 bp fragment from DNA 1, followed by a ligation step to reseal the remaining 3,000 bp of information would enable more information to be determined from DNA sequencing via the SP6 promoter. The sequencing strategy for this region of DNA 1 is also shown in Figure 3.2 (p. 58).

Hence, following the required restriction digestion and ligation steps, an appropriate primer to the SP6 RNA polymerase promoter was used to determine the additional nucleotide information. This new partial DNA sequence for DNA 1 (DNA 1/HIND3/SP6) provided an additional 1,000 bp of information. The nucleotide sequence encoding DNA 1/HIND3/SP6 is shown in Appendix II (p. 125). DNA 1/HIND3/SP6 contains 1,032 bp of additional sequence information.

DNA 1 - DNAsis Analysis

(DNA Level)

Before DNAsis Software analysis was used, it was necessary to eliminate any pGEM-3Z vector nucleotide information that was present in the nucleotide sequence for DNA 1/T7, DNA 1/SP6, and DNA 1/HIND3/SP6. The nucleotide information for DNA 1/T7 that lacked any pGEM-3Z information (with 1-61 bp removed) was designated as DNA 1/Region 1. The nucleotide information removed from DNA 1/T7 to obtain DNA 1/Region 1 is underlined in Appendix II (p. 122).

The nucleotide information for DNA 1/SP6 that lacked any pGEM-3Z information (with 1-27 bp removed) was designated as DNA 1/Region 2. The nucleotide information removed from DNA 1/SP6 to obtain DNA 1/Region 2 is underlined in Appendix II (p. 124).

The nucleotide information for DNA 1/HIND3/SP6 that lacked any pGEM-3Z information (with 1-9 bp removed) was designated as DNA 1/Region 3. The nucleotide information removed from DNA 1/HIND3/SP6 to obtain DNA 1/Region 3 is underlined in Appendix II (p. 125).

The DNAsis Maximum Matching Alignment was used to determine if the nucleotide sequence of DNA 1 /Region 1, DNA 1/Region 2 and DNA 1/Region 3 (normal and complementary strands), was similar to the nucleotide sequence of the human hsp 27 gene. The alignments are shown in Appendix IV (p. 147, 151, 155, 159, 163, and 167). A summary of the alignments is shown in Table 2.1 (p. 61).

The DNAsis Maximum Matching Alignment was used to determine if the nucleotide sequence of DNA 1 /Region 1, DNA 1/Region 2 and DNA 1/Region 3 (normal and complementary strands), were similar to the nucleotide sequence of the most homologous, similar-sized region of the human hsp 27 gene. The alignments are shown in Appendix V (p. 192, 194, 196, 198, 200 and 202). A graphical illustration of these comparisons is shown in Figure 3.3 (p. 62).

In order to obtain partial sequence information for the approximately 4,000 bp DNA 1 insert, we required the 5'-sequence information for DNA 1/Region 3 (complementary strand), the 5'-information for DNA 1/Region 2 (complementary strand), and the overlapping regions between these two sequences. To obtain the overlapping sequence information, the Maximum Matching Alignment was used. The new partial sequence information for DNA DNA 1 (DNA 1/Region 3+2) was 991 bp in length, and is shown in Appendix II (p. 127).

Table 2.1 DNA Alignments Between Areas Of DNA 1 And The Human hsp 27 Gene.

Clone	Strand	Homology
DNA 1/Region 1	Normal	34%
DNA 1/Region 1	Complementary	35%
DNA 1/Region 2	Normal	31%
DNA 1/Region 2	Complementary	32%
DNA 1/Region 3	Normal	32%
DNA 1/Region 3	Complementary	32%

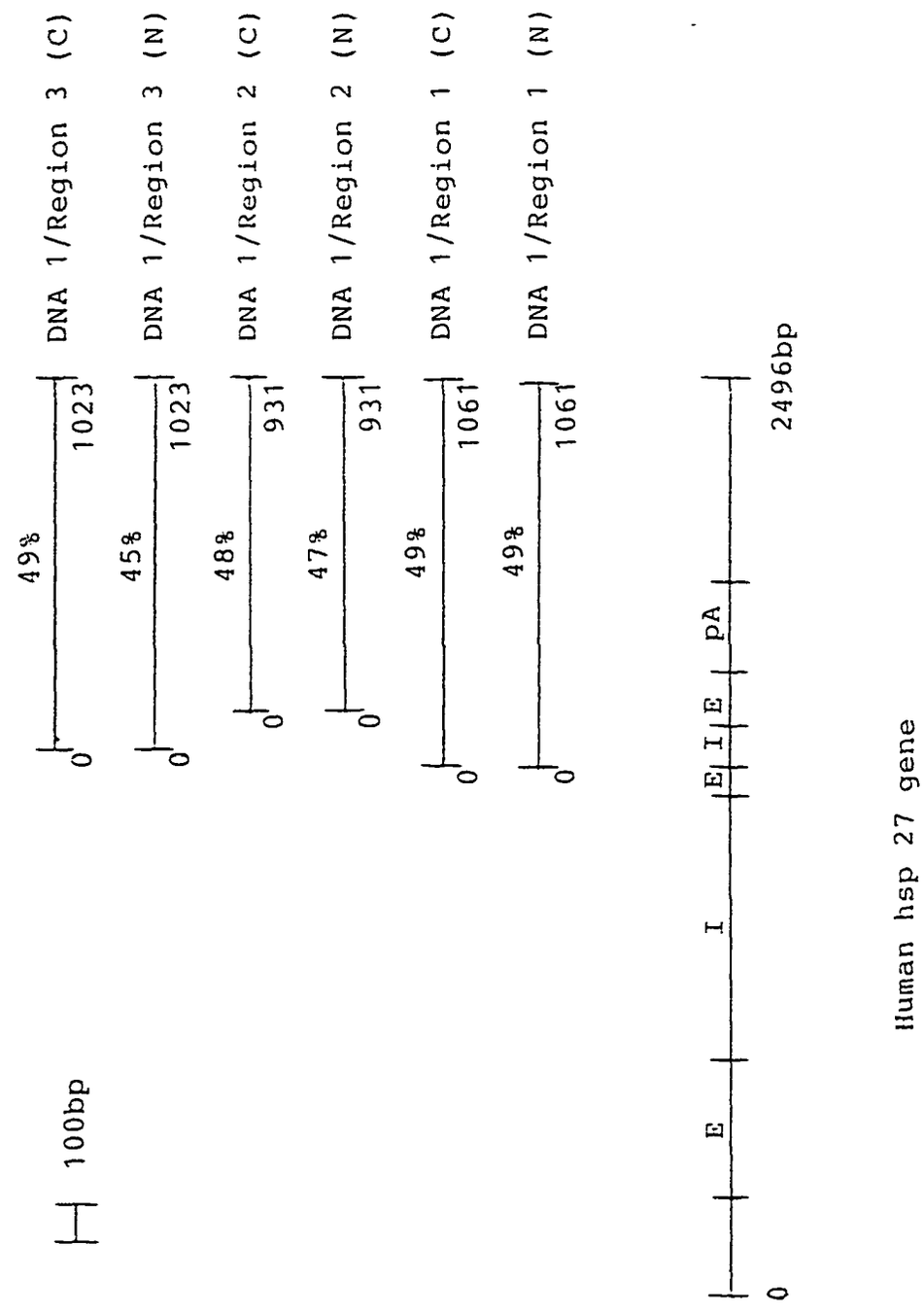


Figure 3.3 Homologous DNA Areas Between Regions Of DNA 1 (Normal And Complementary Strands), And The Human hsp 27 Gene. The exon (E), intron (I), and polyA (pA) regions of the human hsp 27 gene are shown.

DNA 1 - Blast Analysis

(DNA Level)

In order to determine if DNA 1 was similar to any known DNA sequence, the DNA sequence information for regions of DNA 1 was analyzed using the Blastn Algorithm. The results for the Blastn search using the nucleotide sequence of DNA 1/Region 1 (normal strand) are shown in Table 2.2 (p. 64). Some of the human sequences have been listed, out of the 46 total sequences identified for DNA 1/Region 1 (normal strand). The Blastn searches did not identify the human hsp 27 gene as a homologous sequence.

The results for the Blastn search using the nucleotide sequence of DNA 1/Region 2 (normal strand) are shown in Table 2.3 (p. 65). Some of the human sequences have been listed, out of the approximately 200 total sequences identified for DNA 1/Region 2 (normal strand). The Blastn searches did not identify the human hsp 27 gene as a homologous sequence.

The results for the Blastn search using the nucleotide sequence of DNA 1/Region 3 (normal strand) are shown in Table 2.4 (p. 66). Some of the human sequences have been listed, out of the 87 total sequences identified for DNA 1/Region 3 (normal strand). The Blastn searches did not identify the human hsp 27 gene as a homologous sequence.

DNA 1 - DNAsis Analysis

(Protein Level)

In order to compare the predicted amino acid sequences for DNA 1 to the amino acid sequence of the human hsp 27 gene, we attempted to identify the coding strands for regions of DNA 1, by determining which strands contained exons and introns. DNA 1/Region 1

Table 2.2 Blastn Results For DNA 1/Region 1.

Homolog	GenBank Accession Number	Homology	Length
Human DNA sequence from clone 757P12 on Chromosome Xp22.11-22	HS757P12	68%	59 bp
Homo sapiens NFKB1 gene	HSNFX1112	77%	43 bp
Homo sapiens Chromosome 17	AC005883	65%	59 bp
Homo sapien Chromosome 5	AC004629	68%	59 bp
Homo sapiens NFKB1 gene	HSNFX1112	77%	43 bp

Table 2.3 Blastn Results For DNA 1/Region 2.

Homolog	GenBank Accession Number	Homology	Length
Homo sapiens DNA sequence on Chromosme Xq25-q27	HS119E23	80%	39 bp
Homo sapiens Chromosome 16	HUAC004525	62%	59 bp
Human Lambda DNA for Immunoglobolin light chain	D87004	64%	59 bp
Homo sapiens LIPA gene	HSLIPA2	72%	64 bp

Table 2.4 Blastn Results For DNA 1/Region 3.

Homolog	GenBank Accession Number	Homology	Length
Homo sapiens Chromosome 17	AC004448	75%	51 bp
Homo sapiens Chromosome 21	AP000035	74%	36 bp
Homo sapiens T-cell Receptor	HUAE000658	60%	61 bp

(normal strand) showed the presence of one potential 5'-splice site at position 781. No 3'-splice sites were identified. DNA 1/Region 1 (complementary strand) showed the presence of 2 potential 3'-splice sites at positions 1154 and 174. The potential splice sites were preceded by a string of polypyrimidines. Hence, since there was a 5'-splice site on one strand, and a 3'-splice site on the other strand, we were unable to determine which strand of DNA 1/Region 1 was the coding strand. Therefore, we used a more random approach that involved the identification of a variety of Open Reading Frames, within all six reading frames.

The results for the open reading frames for both the normal and complementary strands of DNA 1/Region 1 are shown in Table 2.5 (p. 68). The DNAsis Maximum Matching Alignment was used to compare a number of these predicted amino acid sequences to the amino acid sequence for the human hsp 27 gene. Table 2.6 (p. 70) shows the results for these comparisons.

DNA 1/Region 2 (normal strand) showed the presence of 2 potential 3'-splice sites at positions 7 and 91. Neither of the potential splice sites were preceded by a string of polypyrimidines. No 5'-splice sites were identified. DNA 1/Region 1 (complementary strand) showed the presence of 8 potential 3'-splice sites at positions 1, 7, 9, 19, 158, 171, 588 and 598. The last five 3'-splice sites were preceded by a string of polypyrimidines. Hence, since there were 3'-splice sites on both strands, we were unable to determine which strand of DNA 1/Region 2 was the coding strand. This necessitated the search for a variety of open reading frames, within all six reading frames.

The results for the open reading frames for both the normal and complementary strands of DNA 1/Region 2 are shown in Table 2.7 (p. 71). The DNAsis Maximum Matching Alignment was used to compare a number of these predicted amino acid sequences

Table 2.5 Open Reading Frames For DNA 1/Region 1.

NORMAL

	1 Frame		
	Start	End	MW
	1	37	1450.67
A	472	556	3519.78
	481	556	3126.34
	538	556	807.75
	592	631	1468.65
	598	631	1238.33
	865	910	1792.28
	2 Frame		
	Start	End	MW
	2	32	1276.43
	614	620	335.41
	626	638	491.54
B	641	692	1962.66
	851	890	1674.15
	1046	1058	585.66
	3 Frame		
	Start	End	MW
C	3	42	1502.66
	243	267	904.01
	633	666	1371.61
	645	666	907.07

COMPLEMENTARY

	1 Frame		
	Start	End	MW
	1	16	527.27
	463	526	2670.93
	523	526	149.21
D	553	694	5510.89
	589	694	4183.40
	634	694	2416.40
	700	706	263.31
	712	784	2734.96
	829	832	149.21
	916	949	1351.57

	2 Frame		
	Start	End	MW
E	2	119	4522.78
	608	662	2085.12
	683	698	580.59
	749	845	3611.15
	812	845	1343.52
	3 Frame		
	Start	End	MW
	3	144	5507.02
	108	144	1466.38
	177	369	7548.13
F	210	369	6267.36
	549	561	524.66
	879	906	1075.24

Table 2.6 Amino Acid Alignments Between DNA 1/Region 1 And Human hsp 27.

Sequence	Strand	Homology To hsp 27	Length
A	Normal	11%	28 aa
B	Normal	11%	17 aa
C	Normal	7%	13 aa
D	Complementary	16%	47 aa
E	Complementary	17%	39 aa
F	Complementary	18%	53 aa

Table 2.7 Open Reading Frames For DNA 1/Region 2.

NORMAL

	1 Frame		
	Start	End	MW
A	1	112	4372.19
	25	112	3543.26
	70	112	1792.06
	151	178	1138.22
	253	346	3491.90
	301	346	1662.89
	340	346	246.32
	469	487	764.98
	568	580	537.70
	604	610	296.38
	670	685	709.80
	2 Frame		
	Start	End	MW
	2	26	793.85
B	182	398	8251.68
	227	398	6710.95
	680	728	1805.77
	3 Frame		
	Start	End	MW
C	3	99	3865.16
	102	105	149.21
	822	858	1410.19
	828	858	1160.12

COMPLEMENTARY

	1 Frame		
	Start	End	MW
	1	58	2186.44
D	109	238	5105.26
	331	355	964.97
	373	391	712.82
	421	472	1996.44
	667	679	392.41
	760	772	478.55

	2 Frame		
	Start	End	MW
	2	32	1219.12
E	251	344	3649.19
	410	422	537.70
	428	431	149.21
	515	551	1536.68
	671	761	3480.98
	719	761	1631.92
	3 Frame		
	Start	End	MW
	3	120	4435.66
	447	480	1271.46
	630	663	1353.49
	741	753	565.67
F	885	927	1761.18
	906	927	902.14

to the amino acid sequence for the human hsp 27 gene. Table 2.8 (p. 74) shows the results for these comparisons.

DNA 1/Region 3 (normal strand) showed the presence of one potential 3'-splice site at position 992. The potential 3'-splice site was preceded by a string of polypyrimidines. No 5'-splice sites were identified. DNA 1/Region 3 (complementary strand) showed the presence of 4 potential 3'-splice sites at positions 35, 98, 222 and 586. The last three 3'-splice sites were preceded by a string of polypyrimidines. Hence, since there were 3'-splice sites on both strands, we were unable to determine which strand of DNA 1/Region 3 was the coding strand. Therefore, we used a more random approach to determine the coding strand, that involved the identification of an open reading frame.

The results for the open reading frames for both the normal and the complementary strands of DNA 1/Region 3 are shown in Table 2.9 (p. 75). The DNAsis Maximum Matching Alignment was used to compare a number of these predicted amino acid sequences to the amino acid sequence for the human hsp 27 gene. Table 2.10 (p. 77) shows the results for these comparisons.

DNA 1 - Blast Analysis

(Protein Level)

In order to determine if DNA 1 was similar to any known amino acid sequence, the amino acid sequence information for regions of DNA 1 was analyzed using the Blastx Algorithm. i.e. the Blastx Algorithm was used to search for the nucleotide sequences of DNA 1/Region 1 (normal strand), DNA 1/Region 2 (normal strand), and DNA 1/Region 3 (normal

Table 2.8 Amino Acid Alignments Between DNA 1/Region 2 And Human hsp 27.

Sequence	Strand	Homology To hsp 27	Length
A	Normal	12%	37 aa
B	Normal	18%	72 aa
C	Normal	12%	32 aa
D	Complementary	15%	43 aa
E	Complementary	12%	31 aa
F	Complementary	6%	14 aa

Table 2.9 Open Reading Frames For DNA 1/Region 3.

NORMAL

	1 Frame		
	Start	End	MW
	1	46	1829.16
	40	46	305.39
	214	268	1982.85
	382	451	2651.23
A	622	686	2639.92
	2 Frame		
	Start	End	MW
	2	2	18.02
	65	71	277.38
	116	167	1970.39
	122	167	1740.07
B	236	317	3135.47
	260	317	2205.50
	461	494	1269.32
	482	494	390.48
	554	602	2147.29
	563	602	1715.80
	3 Frame		
	Start	End	MW
C	3	24	788.95
	285	288	149.21
	330	342	519.61
	546	564	580.70
	672	678	206.26
	855	1020	6463.49

COMPLEMENTARY

	1 Frame		
	Start	End	MW
	1	79	2770.05
D	460	496	1361.61
	493	496	149.21
	541	547	268.09
	604	616	523.61
	607	616	392.42
	796	829	1370.31
	820	829	425.53

	2 Frame		
	Start	End	MW
E	2	86	3085.56
	575	605	1129.00
	644	647	149.21
	653	716	2495.93
	857	914	2259.67
	3 Frame		
	Start	End	MW
	3	171	6297.79
	312	339	969.87
	591	642	2065.48
	615	642	1031.30
F	738	810	2838.45
	939	960	988.20
	945	960	709.84

Table 2.10 Amino Acid Alignments Between DNA 1/Region 3 And Human hsp 27.

Sequence	Strand	Homology To hsp 27	Length
A	Normal	15%	22 aa
B	Normal	10%	27 aa
C	Normal	9%	7 aa
D	Complementary	7%	12 aa
E	Complementary	18%	28 aa
F	Complementary	15%	24 aa

strand). The results for the Blastx searches using these nucleotide sequences did not identify any protein as being related to this query sequence, including human hsp 27.

DNA 1 - PCR Primer Location Analysis

In order to further determine if DNA 1 was a human hsp 27 gene, we attempted to identify sites where the PCR primers should have bound to the DNA 1/Region 1, DNA 1/Region 2 and DNA 1/Region 3 nucleotide sequences (normal and complementary strands). The results indicated that the PCR primers did not bind to any of these regions on the DNA 1 nucleotide sequence.

DNA 2 - Size And Restriction Map

The positive clone pDNA 2 was sized by cutting the DNA using specific restriction endonucleases and separating the fragments on an agarose gel. The restriction analysis of pDNA 2 indicated that we had cloned a fragment, (DNA 2) that was approximately 4,300 bp long. Restriction digestion sites in pDNA 2 were examined by cutting the DNA 2 using all the restriction enzymes mentioned in the Materials and Methods section. A plasmid map was then generated for pDNA 2, as shown in Figure 4.1 (p. 79).

DNA 2 - DNA Sequence

Since the plasmid map for the positive clone, pDNA 2, showed that the size of DNA 2 was approximately 4,300 bp in length, it was apparent that only about 2,000 bp of the nucleotide sequence information for DNA 2 could be obtained from two rounds of automated

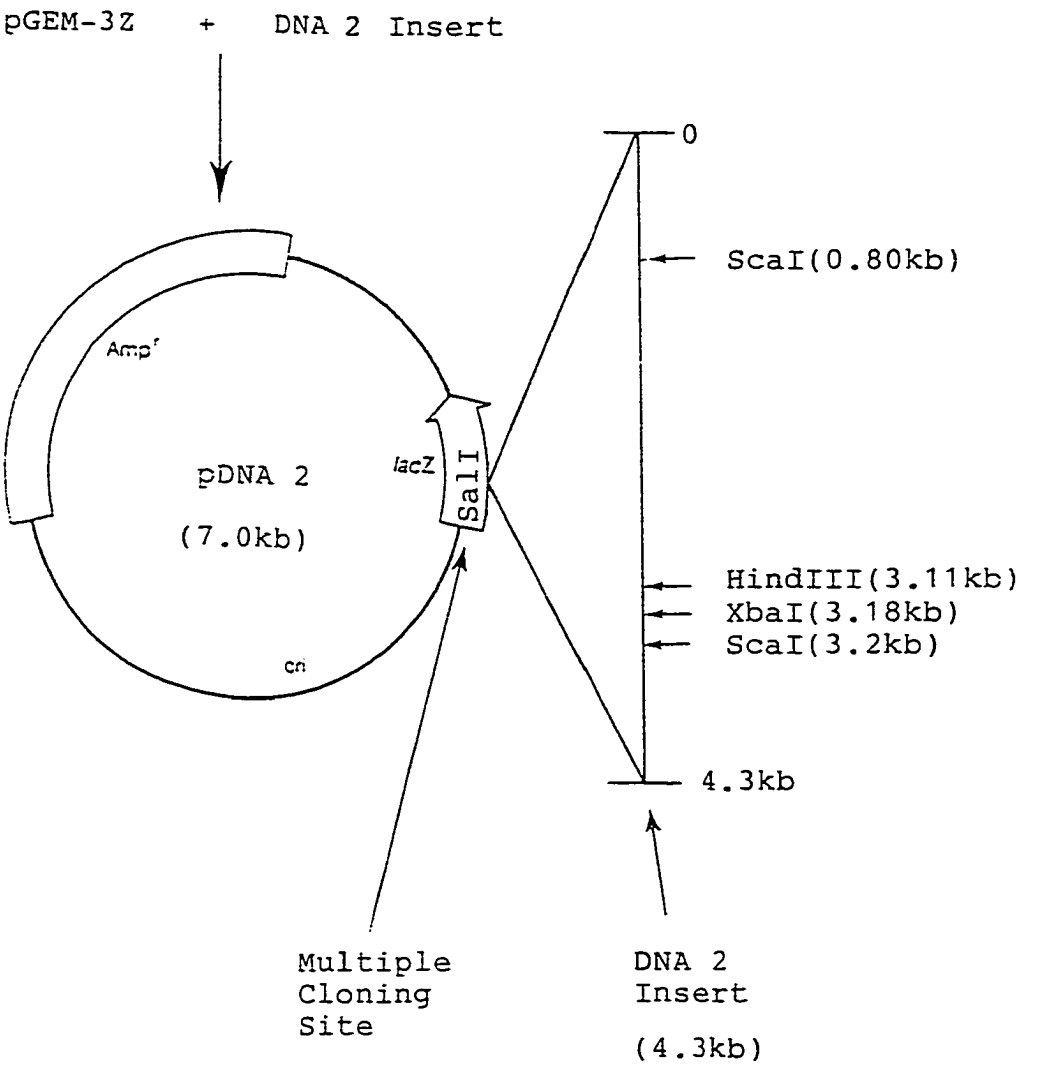


Figure 4.1 Plasmid Map For pDNA 2.

DNA sequencing (using the T7 and SP6 promoters). The sequencing strategy for regions of DNA 2 is shown in Figure 4.2 (p. 81).

The nucleotide sequence encoding DNA 2 (892 bp) using the T7 promoter (DNA 2/T7), is shown in Appendix II (p. 128). The nucleotide sequence encoding DNA 2 (977 bp) using the SP6 promoter (DNA 2/SP6), is shown in Appendix II (p. 129).

DNA 2 - DNAsis Analysis

(DNA Level)

Before DNAsis Software analysis was used, it was necessary to eliminate any pGEM-3Z vector nucleotide sequence that was present in the nucleotide sequence for DNA 2/T7 and DNA 2/SP6. The modified information of DNA 2/T7 that lacked any pGEM-3Z information (with 1-35 bp removed) was redesignated as DNA 2/Region 1. The nucleotide information removed from DNA 2/T7 to obtain DNA 2/Region 1 is underlined in Appendix II (p. 128).

The modified information of DNA 2/SP6 that lacked any pGEM-3Z information (with 1-27 bp removed) was redesignated as DNA 2/Region 2. The nucleotide information removed from DNA 2/SP6 to obtain DNA 2/Region 2 is underlined in Appendix II (p. 129).

The DNAsis Maximum Matching Alignment was used to determine if the nucleotide sequence of DNA 2 /Region 1 and DNA 2/Region 2 (normal and complementary strands), were similar to the nucleotide sequence of the human hsp 27 gene. The alignments are shown in Appendix IV (p. 171, 175, 179, and 183). A summary of the alignments is shown in Table 3.1 (p. 82).

The DNAsis Maximum Matching Alignment was used to determine if the nucleotide

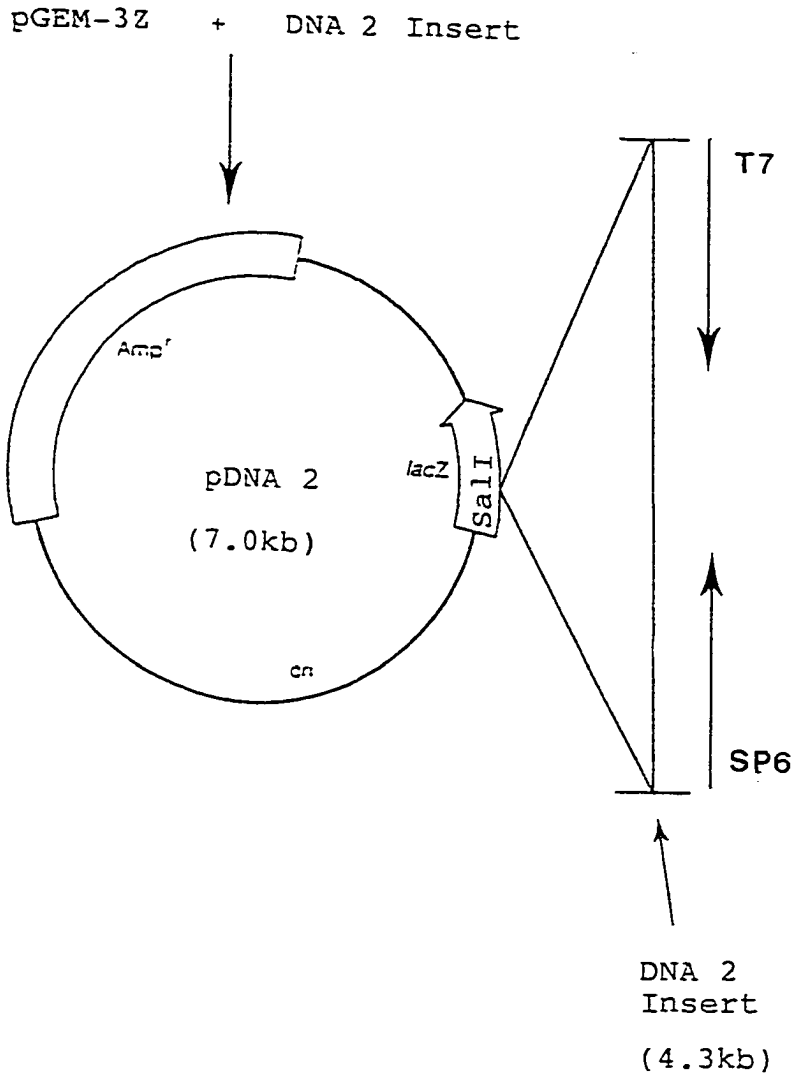


Figure 4.2 Sequencing Strategy For DNA 2.

Table 3.1 DNA Alignments Between Regions Of DNA 2 And The Human hsp 27 Gene.

Clone	Strand	Homology
DNA 2/Region 1	Normal	29%
DNA 2/Region 1	Complementary	29%
DNA 2/Region 2	Normal	31%
DNA 2/Region 2	Complementary	32%

sequence of DNA 2 /Region 1 and DNA 2/Region 2 (normal and complementary strands), were similar to the nucleotide sequence of the most homologous, similar-sized region of the human hsp 27 gene. The alignments are shown in Appendix V (p. 204, 206, 208, and 210). A graphical illustration of these comparisons is shown in Figure 4.3 (p. 84).

DNA 2 - Blast Analysis

(DNA Level)

In order to determine if DNA 2 was similar to any known DNA sequence, the DNA sequence information for regions of DNA 2 was analyzed using the Blastn Algorithm. The results for the Blastn search using the nucleotide sequence of DNA 2/Region 1 (normal strand) are shown in Table 3.2 (p. 85). Some of the human sequences have been listed, out of the 17 total sequences identified for DNA 2/Region 1 (normal strand). The Blastn searches did not identify the human hsp 27 gene as a homologous sequence.

The results for the Blastn search using the nucleotide sequence of DNA 2/Region 2 (normal strand) are shown in Table 3.3 (p. 86). Some of the human sequences have been listed, out of the approximately 125 total sequences identified for DNA 2/Region 2 (normal strand). The Blastn searches did not identify the human hsp 27 gene as a homologous sequence.

DNA 2 - DNAsis Analysis

(Protein Level)

In order to compare the predicted amino acid sequence for DNA 2 to the amino acid sequence of the human hsp 27 gene, it was necessary to determine the coding strands for

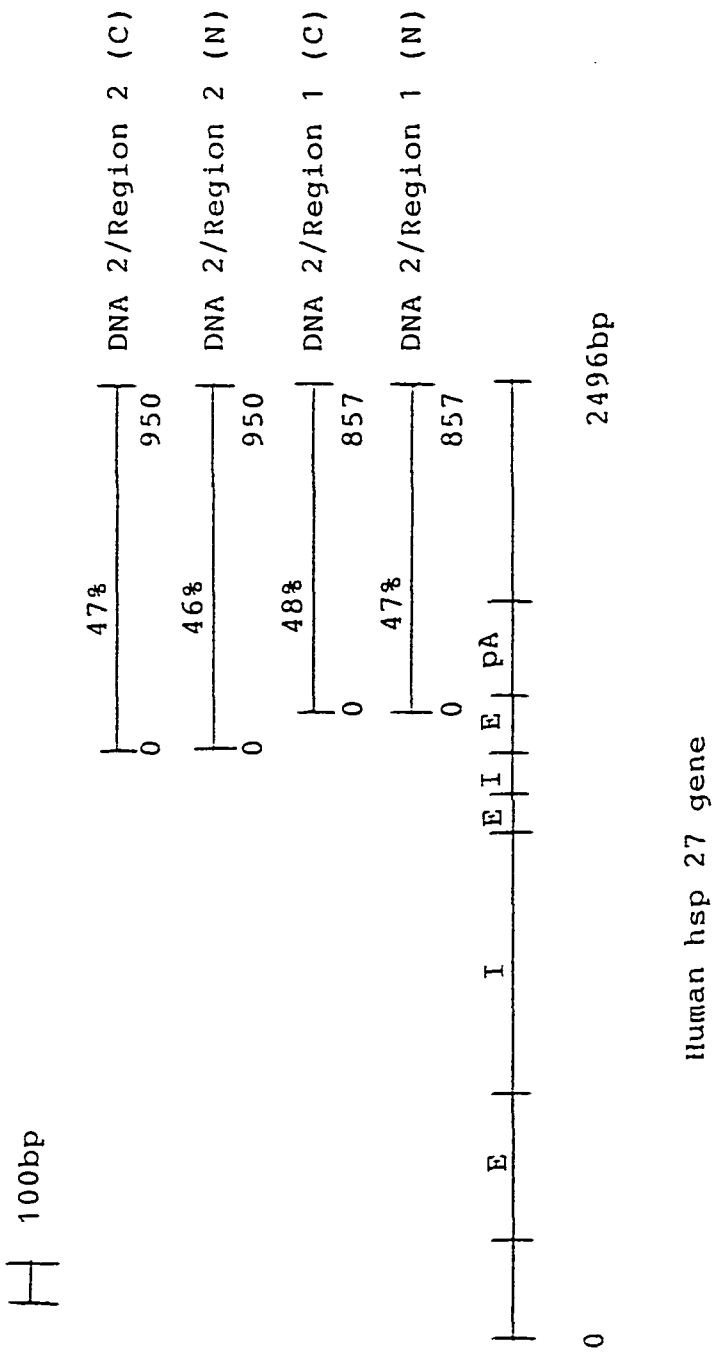


Figure 4.3 Homologous DNA Areas Between Regions Of DNA 2 (Normal And Complementary Strands), And The Human hsp 27 Gene. The exon (E), intron (I), and polyA (pA) regions of the human hsp 27 gene are shown.

Table 3.2 Blastn Results For DNA 2/Region 1.

Homolog	GenBank Accession Number	Homology	Length
Human DNA sequence region on ChromosomeX	HSU65A4	72%	50 bp
Homo sapiens NFKB1 gene	HSNFX1112	77%	43 bp
Human Genomic Sperm Library	AO170454	62%	39 bp
Homo sapiens Chromosome 20	AC004499	73%	52 bp

Table 3.3 Blastn Results For DNA 2/Region 2.

Homolog	GenBank Accession Number	Homology	Length
Homo sapiens Serine Threonine Kinase Receptor	HUMSKR2	82%	38 bp
Homo sapiens Chromosome 21 Down Syndrome	AP000015	76%	45 bp
Homo sapiens Chromosome 4	AC004056	70%	59 bp
Homo sapiens DNA sequence on Chromosme Xq25-q27	HS119E23	80%	39 bp

regions of DNA 2. Therefore, we attempted to identify the coding strands by determining which strands of regions of DNA 2 contained exons and introns.

DNA 2/Region 1 (normal strand) indicated the presence of 2 potential 3'-splice sites at positions 717 and 809. Both of the 3'-splice sites were preceded by a string of polypyrimidines. No 5'-splice sites were identified. Interestingly, DNA 2/Region 1 (complementary strand) showed the presence of one potential 5'-splice site at 16. There were also 4 potential 3'-splice sites at positions 10, 67, 68, and 253. Only the last three potential splice sites were preceded by a string of polypyrimidines. Hence, since there was a 5'-splice site and a 3'-splice site on this strand, we predicted that the normal strand of DNA 2/Region 1 was the coding strand, which would result in the synthesis of the predicted amino acid sequence. Therefore, an intron region of the normal strand of DNA 2/Region 1 (16-68 bp) was excised, and the remaining exons were joined together.

The exon information was used to determine the coding strand for the Open Reading Frame. Table 3.4 (p. 88) shows the open reading frames for the normal strand of DNA 2/Region 1. Then, the DNAsis Maximum Matching Alignment was used to compare three of these predicted amino acid sequences to the amino acid sequence for the human hsp 27 gene. Table 3.5 (p. 89) shows the results for these comparisons.

DNA 2/Region 2 (normal strand) showed the presence of 3 potential 3'- splice sites at positions 7 and 91 and 907. Only one of the potential splice sites was preceded by a string of polypyrimidines. No 5'-splice sites were identified. DNA 2/Region 2 (complementary strand) showed the presence of 2 potential 3'-splice sites at positions 607 and 617. Both of the 3'-splice sites were preceded by a string of polypyrimidines. Since there were 3'-splice sites on both strands, we were unable to determine which strand of DNA 2/Region 2 was the

Table 3.4 Open Reading Frames For DNA 2/Region 1.

NORMAL

	1 Frame		
	Start	End	MW
	1	43	1814.03
	427	460	1280.26
	493	583	3514.54
A	589	700	4364.23
	2 Frame		
	Start	End	MW
	2	11	418.47
	197	221	904.01
B	437	512	3006.77
	569	587	725.83
	3 Frame		
	Start	End	MW
C	3	15	403.46

Table 3.5 Amino Acid Alignments Between DNA 2/Region 1 And Human hsp 27.

Sequence	Strand	Homology To hsp 27	Length
A	Normal	15%	37 aa
B	Normal	11%	25 aa
C	Normal	26%	4 aa

coding strand. Therefore, we used a more random approach to that involved all six reading frames.

The results for the open reading frames for both the normal and complementary strands of DNA 2/Region 2 are shown in Table 3.6 (p. 91). The DNAsis Maximum Matching Alignment was used to compare a number of these predicted amino acid sequences to the amino acid sequence for the human hsp 27 gene. Table 3.7 (p. 93) shows the results for these comparisons.

DNA 2 - Blast Analysis

(Protein Level)

In order to determine if DNA 2 was similar to any known amino acid sequence, the amino acid sequence information for regions of DNA 2 was analyzed using the Blastx Algorithm. Initially, the Blastx Algorithm was used to search for the nucleotide sequences of DNA 2/Region 1 (normal strand). The results for the Blastx search using this nucleotide sequence identified only one homolog as being homologous to this query sequence. The homolog was the 40 aa Pyruvate Flavodoxin Oxidoreductase Protein (NIFJ ENTAG), and was found to be 32% homologous. The Blastx search did not identify the human hsp 27 protein as a homologous sequence. The Blastx Algorithm was used to search for the nucleotide sequences of DNA 2/Region 2 (normal strand). Only four proteins were identified as being homologous to the query sequence. Some of the human sequences have been listed in Table 3.8 (p.94). The Blastx searches did not identify the human hsp 27 protein as a homologous sequence.

Table 3.6 Open Reading Frames For DNA 2/Region 2.

NORMAL

	1 Frame		
	Start	End	MW
A	1	112	4372.19
	25	112	3543.26
	70	112	1792.06
	151	178	1138.22
	253	346	3491.90
	301	346	1662.89
	340	346	246.32
	469	487	764.98
	2 Frame		
	Start	End	MW
	2	26	793.85
B	182	398	8251.68
	227	398	6710.95
	569	572	149.21
	671	686	649.70
	824	836	496.33
	3 Frame		
	Start	End	MW
C	3	99	3865.16
	102	105	149.21
	603	609	296.38
	681	711	1151.06
	819	849	1309.68

COMPLEMENTARY

	1 Frame		
	Start	End	MW
	1	67	2772.64
D	391	436	1813.12
	466	499	1271.46
	649	682	1353.49
	760	772	565.67
	904	946	1761.18
	925	946	902.14

	2 Frame		
	Start	End	MW
	2	179	6855.27
	131	179	1806.68
	269	281	472.51
E	440	491	2016.19
	686	698	392.41
	779	791	478.55
	3 Frame		
	Start	End	MW
	3	147	5776.97
	63	147	3410.63
	159	363	7808.52
	351	363	537.66
	429	441	537.70
	447	450	149.21
	534	570	1536.68
F	690	780	3480.98
	738	780	1631.92

Table 3.7 Amino Acid Alignments Between DNA 2/Region 2 And Human hsp 27.

Sequence	Strand	Homology To hsp 27	Length
A	Normal	12%	37 aa
B	Normal	18%	72 aa
C	Normal	12%	32 aa
D	Complementary	10%	15 aa
E	Complementary	8%	17 aa
F	Complementary	12%	30 aa

Table 3.8 Blastx Results For DNA 2/Region 2.

Homolog	GenBank Accession Number	Homology	Length
Insulin	INS ANSAN	75%	16 aa
Insulin	INS LEPSP	60%	17 aa

DNA 2 - PCR Primer Location Analysis

In order to further determine if DNA 2 was a linked human hsp 27 gene, we attempted to identify sites where the PCR primers should have bound to the DNA 2/Region 1 and DNA 2/Region 2 nucleotide sequences (normal and complementary strands). The results indicated that the PCR primers did not bind to any of these regions on the DNA 2 nucleotide sequence.

Summary

Table 4.1 (p. 96) shows a summary of the homologies between DNA 4, and human hsp 27 that were obtained from the DNAsis Maximum Matching Alignment. The same table shows the summary of the homologies between regions of DNA 1 and human hsp 27, as well as the homologies between regions of DNA 2 and human hsp 27.

DNA 3

The positive clone pDNA 3 was sized by cutting the DNA using specific restriction endonucleases and separating the fragments on an agarose gel. The restriction analysis of pDNA 3 indicated that we had cloned a fragment, (DNA 3) that was approximately 1,200 bp long.

Restriction digestion sites in pDNA 3 were examined by cutting the DNA at specific sites using the restriction enzymes EcoRI, PstI, BamHI, HindIII, and ScaI. A plasmid map was then generated for pDNA 3, on the basis of the restriction enzyme digestion patterns. The plasmid map for pDNA 3 is shown in Figure 5.1 (p. 97). There were no restriction sites

Table 4.1 Summary Of DNA And Amino Acid Alignments Between Human hsp 27 And DNA 4, DNA 1 And DNA 2.

DNA Clone	Nucleotide Homology To hsp 27	Amino Acid Homology To hsp 27
DNA 4	38-51%	7-30%
DNA 1/Region 1	35-49%	7-18%
DNA 1/Region 2	31-48%	6-18%
DNA 1/Region 3	32-45%	3-18%
DNA 2/Region 1	29-48%	11-26%
DNA 2/Region 2	31-47%	8-18%

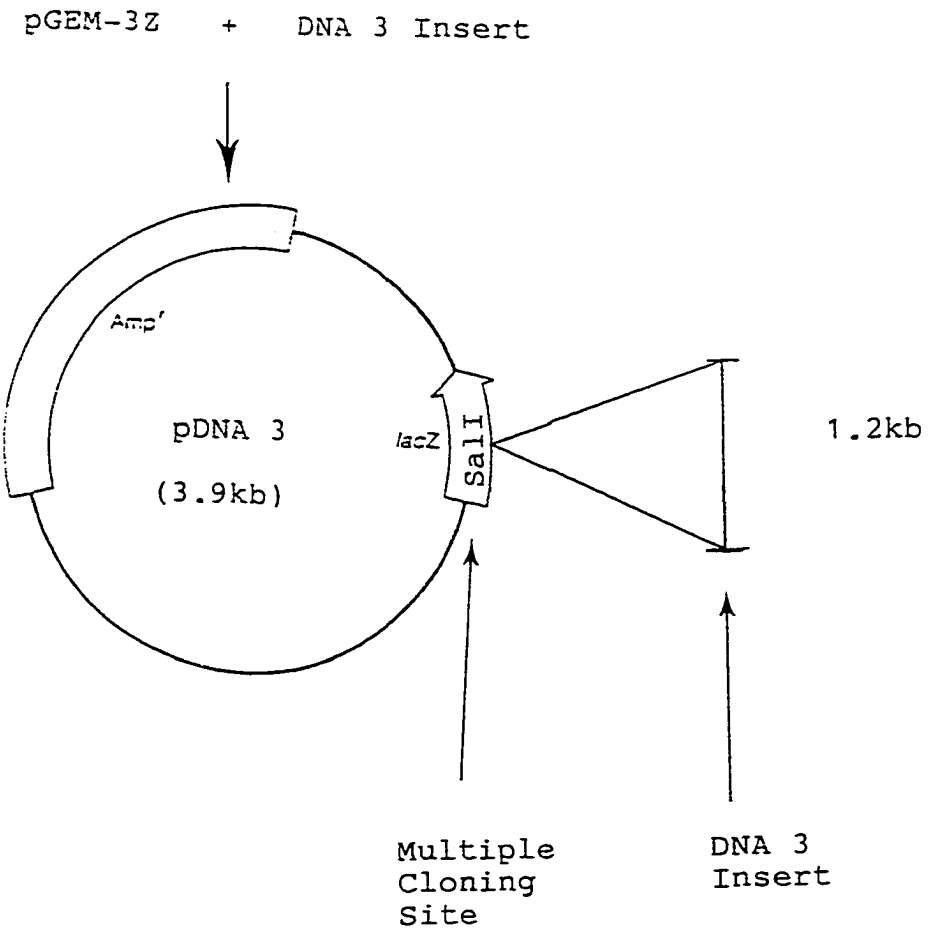


Figure 5.1 Plasmid Map For pDNA 3.

for any of these enzymes within the DNA 3 insert. The plasmid map for pDNA 3 showed that we had cloned a fragment that was approximately 1,200 bp in size.

Since the plasmid maps of pDNA 3 and pDNA 4 showed similarities in the digestion patterns for DNA 3 and DNA 4 (i.e. no restriction sites for five enzymes), and since DNA 3 was identical in size to DNA 4, this raised the possibility that DNA 3 and DNA 4 were identical pieces of DNA. Comparisons between the nucleotide information for DNA 3 and DNA 4 may have confirmed the similarity between these two inserts. However, at this point in time, no further analysis has been conducted using pDNA 3.

DNA 5

The positive clone pDNA 5 was sized by cutting the DNA using specific restriction endonucleases and separating the fragments on an agarose gel. The total size of the digestion product was approximately 2,700 bp. i.e. it was the size of the parental pGEM-3Z plasmid. The restriction analysis of pDNA 5 indicated that this clone contained no DNA insert. Hence, no further analysis was conducted using pDNA 5.

DNA 1 vs. DNA 2

The nucleotide sequence of DNA 2 was submitted to the GenBank database, because it was the largest-sized insert (of approximately 4,300 bp). Before the submission of DNA 1, it was necessary to determine if DNA 1 came from a larger-sized DNA 2, i.e. it was important to determine if DNA 1 and DNA 2 contained identical sequences. This possibility arose since the plasmid map for pDNA 2 seemed to contain identical restriction sites as in pDNA 1, and the inserts that these plasmids contained were both similar in size (DNA 1 -

4,000 bp and DNA 2- 4,300 bp). In order to determine if DNA 1 and DNA 2 contained similar sequences, the nucleotide sequences of regions of DNA 1 were compared to the nucleotide sequences of regions of DNA 2.

The DNAsis Maximum Matching Alignment was used to determine if the nucleotide sequence of DNA 1/Region 1 (normal strand), was similar to the nucleotide sequence of DNA 2/Region 1 (normal strand). In this comparison, a 70% overall identity was indicated (some regions are 100%), by the results shown in Appendix VI (p. 213). The DNAsis Maximum Matching Alignment was used to compare the nucleotide sequence of DNA 1/Region 2 (normal strand), with the nucleotide sequence of DNA 2/Region 2 (normal strand). In this comparison, an 86% identity was indicated by the results shown in Appendix VI (p. 215).

Hence, due to the high identities in the starting and ending regions of DNA 1 and DNA 2, it was possible that DNA 1 is related to DNA 2 (with DNA 2 containing about 300 bp of additional unsequenced information that was not contained in DNA 1). The extra 300 bp of information may lie in the unsequenced region between the SP6 and T7 information. Figure 6.1 (p. 100) shows the comparison between the insert DNA 1, and the insert DNA 2. Hence, it was possible that DNA 1 and DNA 2 were different pieces of DNA. Therefore, the nucleotide sequence of DNA 1 was submitted to the GenBank database.

DNA 4 vs. DNA 2/1

Before submission of DNA 4 to the GenBank database, it was necessary to determine if DNA 4 came from a larger-sized DNA 2 or DNA 1 (DNA 2/1). In order to determine if DNA 4 and DNA 2/1 contained identical nucleotide sequences, the nucleotide sequence of

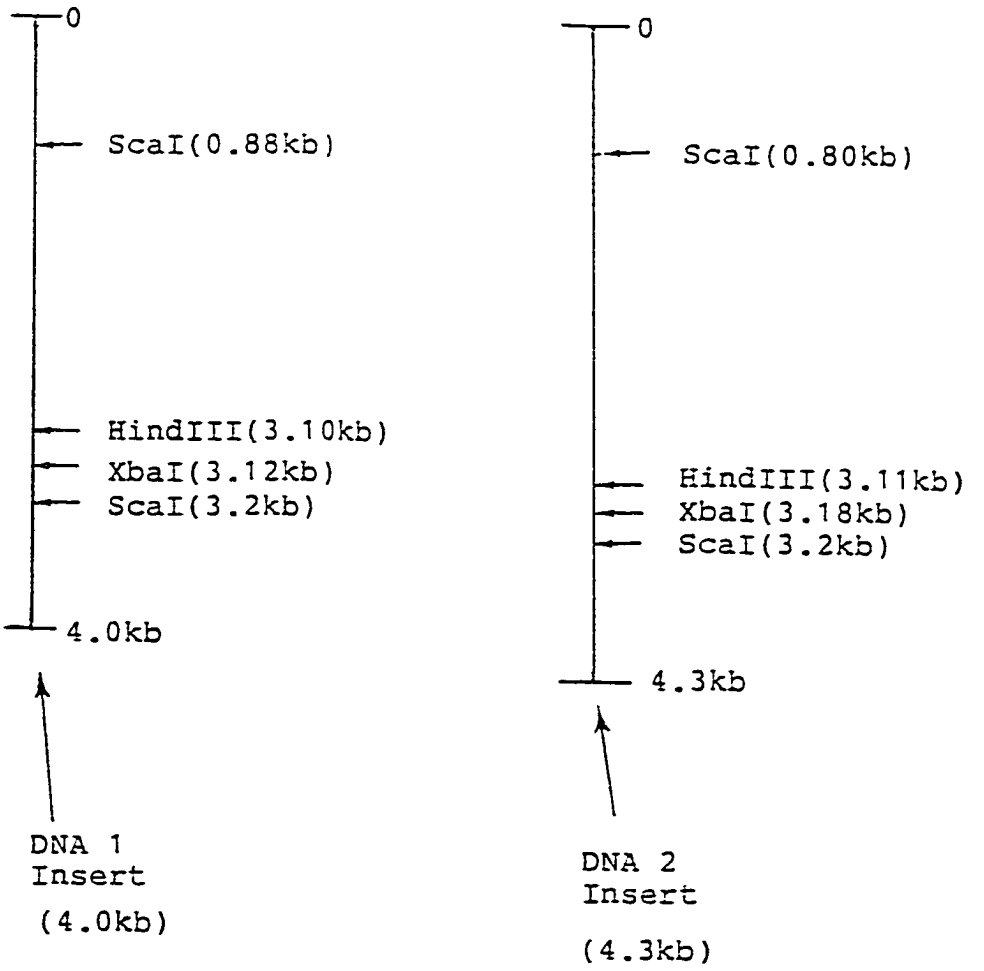


Figure 6.1 Comparison Between DNA 1 And DNA 2.

DNA 4 (normal strand) was compared to the nucleotide sequence for DNA 2/Region 1 (normal strand), using the DNAsis Maximum Matching Alignment. In this comparison, a 48% identity was indicated by the results shown in Appendix VI (p. 217).

The nucleotide sequence of DNA 4 (normal strand) was compared to the nucleotide sequence for DNA 2/Region 2 (normal strand), using the DNAsis Maximum Matching Alignment. In this comparison, a 47% identity was indicated by the results shown in Appendix VI (p. 219).

Due to these low identity results, it was possible that DNA 4 was a unique sequence. Hence, it was possible that DNA 4 and DNA 2/1 were different pieces of DNA. Therefore, the nucleotide sequence of DNA 4 was submitted to the GenBank database.

Nucleotide Sequence

Accession Numbers

The nucleotide sequence for DNA 1, the nucleotide information for DNA 2, and the complete nucleotide information for DNA 4 were deposited with GenBank under the accession numbers: AF120329, AF120330, AF120331, AF120332, and AF120333.

CHAPTER 4

DISCUSSION

In order to find the gene members of the human hsp 27 multigene family, five positive clones were identified following colony hybridization and PCR analysis (pDNA 1, pDNA 2, pDNA 3, pDNA 4, and pDNA 5). Four of these clones contained DNA inserts (DNA 1, DNA 2, DNA 3 and DNA 4). pDNA 5 did not contain an insert, and no further analysis was conducted on pDNA 5. The insert contained in pDNA 3 was also eliminated from further analysis, as discussed in the Materials and Methods section. Hence, the inserts DNA 4, DNA 1, and DNA 2 were subjected to subsequent analysis to determine if they were the linked gene members of the human hsp 27 multigene family.

Since human hsp 27 cDNA is 762 bp in length, and the promoter region of the human hsp 27 gene is minimally 200 bp in length (15), we predicted that the sizes of the DNA inserts contained in the inserts must be greater than 1,000 bp, in order to contain the complete information for the genes. The restriction digestion results indicated that the inserts DNA 1, and DNA 2, were better candidates for full-length members of the human hsp 27 multigene family, as they were both significantly greater than 1,000 bp in length (DNA 1 and DNA 2 were approximately 4,000 bp and 4,300 bp long). The size of the insert, DNA 4 (approximately 1,200 bp), was too small to be a full-length member of the family.

Further considerations determining if the inserts were the closely-related, gene members of the human hsp 27 multigene family, were based on the fact that any closely-related member should be homologous to hsp 27 in its amino acid sequence. Murine hsp 27 is closely-related to human hsp 27 (23). Blastx search results indicated that murine hsp 27 exhibited an 86% homology to human hsp 27 at the protein level. Rat hsp 27 is also closely-related to human hsp 27 (23). Blastx search results indicated that rat hsp 27 exhibited an 87% homology to human hsp 27 at the protein level. The distantly-related alpha-crystallins (21) were found to be 45-55% homologous to human hsp 27 at the amino acid level (Blastx search results). Therefore, we predicted that our inserts would be greater than 50% homologous to human hsp 27, at the protein level, if they were the closely-related, gene members of the human hsp 27 multigene family. Align algorithm results using the amino acid sequences of the inserts indicated that neither DNA 1, DNA 2, nor DNA 4 appeared to be closely-related gene members of the human hsp 27 multigene family. These inserts were less than 50% homologous to human hsp 27 at the amino acid level, from all predicted open reading frames. See Table 4.1 (p. 96).

It is important to note the manner in which we selected our amino acid sequences. The coding strand for the synthesis of an amino acid sequence was determined from the identification of 5' and 3'-splice sites. In the identification of a potential 3'-splice site from a preceding polypyrimidine string, only the strings CCCCC and TTTTT were used. An exhaustive search of potential 3'-splice sites would have included all possible polypyrimidine strings such as CTTTT, CCTTT, TCTCT and CTCTC. Hence, it may appear that the omission of a potential 3'-splice site may have affected our results. However, in our analysis, a coding strand was chosen when both 5' and 3'-splice sites were identified. If a strand

contained only a 5'-splice site, it was not chosen as the coding strand, and both strands of DNA were used to code for predicted amino acid sequences. Therefore, if we did indeed miss a potential 3'-splice site, it should not have affected our results.

PCR primer location analysis also indicated that one of the inserts (DNA 4), was not closely-related to the human hsp 27 multigene family. This was due to the fact that a sequence search for the hybridization sites of the PCR primers, could not identify the corresponding motif in the DNA 4 nucleotide sequence. Therefore, since DNA 4 underwent PCR analysis specifically to amplify this region, this indicated that DNA 4 was a false positive. Hence, DNA 4 could not be one of the members of the gene family that we desired. When the sequenced regions of DNA 1, and DNA 2 were analyzed for the corresponding PCR primer motif, no such motif was found. However, we cannot conclude that DNA 1 and DNA 2 are also false positives, because the motif may be present in the unsequenced regions of these inserts. DNA 5 was also a false positive, as it contained no insert, and should not have been amplified by PCR. It is possible that PCR false positives may have been avoided, if we had used pGEM-3Z as a negative control, rather than the wild-type lambda EMBL3 phage.

We went on to determine if the inserts were distantly-related to the human hsp 27 multigene family at the DNA Level. An alignment of a closely-related member of the human hsp 27 multigene family, should be homologous to hsp 27 in its nucleotide sequence. Murine hsp 27 is closely-related to human hsp 27, and was found to exhibit an 75-90% homology to hsp 27 at the DNA level (Blastn search results, data not shown). Rat hsp 27 is also closely-related to human hsp 27, and was found to exhibit an 72-88% homology to hsp 27 at the DNA level (Blastn search results, data not shown). Hence, we determined that closely-related

members of the human hsp 27 multigene family, would be greater than 50% homologous to human hsp 27, at the DNA level. Therefore, we predicted that distantly-related members of the human hsp 27 multigene family would be less than 50% homologous to the human hsp 27 gene.

Align algorithm results using the nucleotide sequences of the inserts contained in the positive clones indicated that neither DNA 1, DNA 2, nor DNA 4 appeared to be closely-related members of the human hsp 27 multigene family. These inserts were not greater than 50% homologous to human hsp 27 at the DNA level. However, these inserts were very close to 50% homologous to hsp 27 at the DNA level. See Table 4.1 (p. 96). Hence, these results indicated that our inserts may be distantly-related to the human hsp 27 multigene family.

It is interesting to note that the nucleotide sequences of both the normal and complementary strands of an insert, aligned with the same region of the human hsp 27 gene. These similarities should not have occurred, as the normal and complementary strands of a DNA sequence contain different nucleotide information. However, the graphical illustrations shown in the Results section indicated that some of the normal and complementary strand information (for a particular insert), lined up with the 3'-untranslated region of the human hsp 27 gene, as well as other regions near the 3'-end of the gene. Hence, it seems that both the normal and complementary strands contain some similarity in the 3'-regions, which caused the entire sequences of the strands, to align with that region of the human hsp 27 gene (with the greatest homology percentage). The 3'-untranslated region of the human hsp 27 gene contains Alu sequences. Alu sequences are short repetitive sequences (SINES), that range from 100-500 bp in length. They are present in about 9% of the human genome , and have no specific function. The presence of these repetitive sequences can cause an artificial

increase in the homology percentage obtained for alignments (29). Hence, the presence of repetitive Alu sequences in the 3'-end of the human hsp 27 gene may have caused the normal and complementary strands of the inserts to line up with the same regions of the human hsp 27 gene. Therefore, these alignments showed an artificially higher homology percentage than alignments with other regions of the human hsp 27 gene.

We attempted to determine the identity of the proposed distantly-related human hsp 27 genes (i.e. the inserts). Blast Algorithm searches of the GenBank nucleotide database, and the SwissProt protein database indicated that the inserts were novel sequences. This is because the DNA sequences, and the amino acid sequences of the inserts showed homologies to very short regions of other known genes (regions of about 36-64 bp). These short regions of other genes, showed high homologies to our inserts, however, the regions of similarity might not actually be true homologies, but instead sites of short repeating secondary structures (such as alpha-helices and beta-pleated sheets), that most proteins typically possess (29).

In conclusion, in our attempts to isolate the closely-related members of the human hsp 27 multigene family (i.e. the linked-genes)(15), three positive clones containing genomic DNA from the human A549 cell line were identified. Subsequent Align algorithm analysis of the nucleotide sequences and the predicted amino acid sequences of the positive clones suggested that we had in fact, isolated distantly-related members of the human hsp 27 multigene family. A closer examination of the individual steps employed in our research may explain our results.

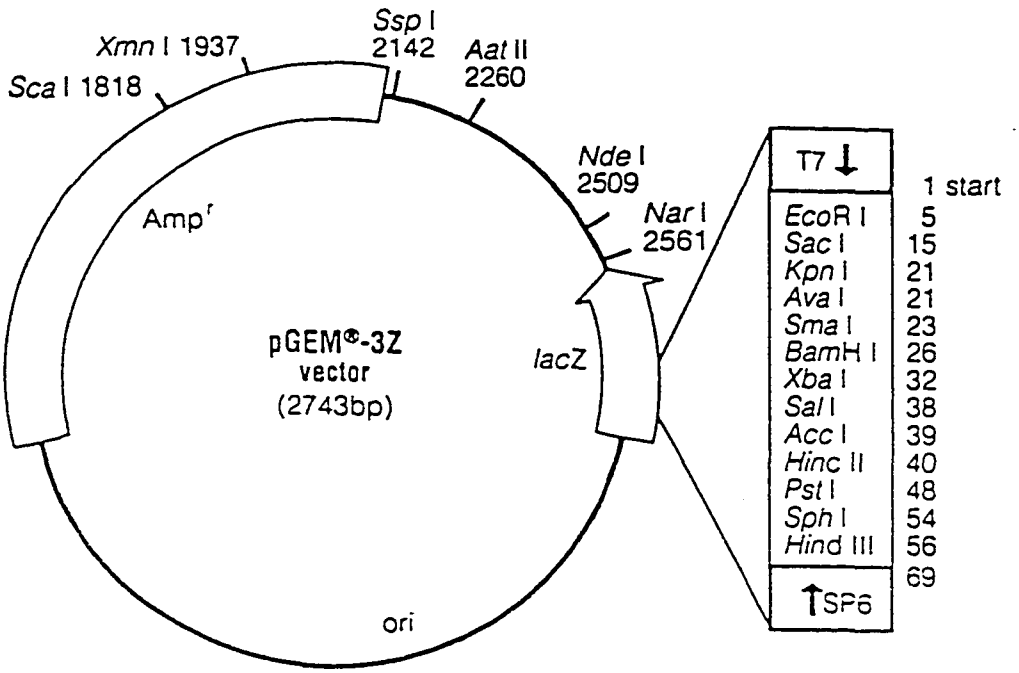
Firstly, it is possible that the distantly-related genes are present in a higher copy number in the genomic DNA (from the human A549 cell line), than the linked human hsp

27 genes. In this case, a greater proportion of the recombinant plaques selected from colony hybridization would contain the distantly-related genes. We obtained 300,000 total plaques from colony hybridization. Since we only analyzed 5 out of 50 positive plaques, it may appear that we may have statistically reduced our chances of isolating the linked genes. Hence, it may seem that further screenings of more plaques, may have enabled us to obtain the linked human hsp 27 genes. However, the human genome contains 6 X 10^9 bp of information. Since our plaques contained fragments of approximately 20 kb in length, it would require the examination of 300,000 plaques in order to find a single copy gene. Therefore, our examination of 5 of 50 positive plaques, obtained from 300,000 total plaques, should have enabled us to find at least one human hsp 27 gene.

Secondly, it is possible that the conditions chosen for the PCR analysis, may also have reduced our chances of finding the linked genes. The PCR conditions (i.e. the ideal reaction mixtures), were designed to amplify human hsp 27 cDNA. These conditions may have allowed our primers to hybridize to non-specific sites in the distantly-related genes, which had less than 100% homology to human hsp 27. It is possible that if the project is repeated, and the conditions for PCR amplification are designed to work best for the human hsp 27 gene sequence, we may isolate the desired linked genes. However, the focus of our laboratory has shifted from cloning these genes to other areas in apoptosis. Therefore, at this time, we have suspended further attempts to clone, and study the regulation of the linked human hsp 27 genes. Our efforts have concluded with the submission of the novel nucleotide sequences for the distantly-related genes to the GenBank nucleotide database. We will resume our studies involving the regulation of the linked gene members of the human hsp 27 multigene family, after the genes have been cloned by the Human Genome Project.

APPENDIX I

Plasmid Map For pGEM-3Z



APPENDIX II

DNA Sequence For pGEM-3Z

(The T7 and SP6 promoters are underlined).

```
        10          20          30          40          50          60
GGGCGAATTC GAGCTCGGTA CCCGGGGATC CTCTAGAGTC GACCTGCAGG CATGCAAGCT

        70          80          90         100         110         120
TGAGTATTCT ATAGTGTCAC CTAAATAGCT TGGCGTAATC ATGGTCATAG CTGTTTCCTG
           SP6 PROMOTER
       130         140         150         160         170         180
TGTGAAATTG TTATCCGCTC ACAATTCCAC ACAACATACG AGCCGGAAGC ATAAAGTGTA

       190         200         210         220         230         240
AAGCCTGGGG TGCCTAATGA GTGAGCTAAC TCACATTAAT TGCGTTGCGC TCACTGCCCG

       250         260         270         280         290         300
CTTTCCAGTC GGGAAACCTG TCGTGCCAGC TGCATTAATG AATCGGCCAA CGCGCGGGGA

       310         320         330         340         350         360
GAGGCGGTTT GCGTATTGGG CGCTCTTCCG CTTCCTCGCT CACTGACTCG CTGCGCTCGG

       370         380         390         400         410         420
TCGTTCGGCT GCGGCGAGCG GTATCAGCTC ACTCAAAGGC GGTAATACGG TTATCCACAG

       430         440         450         460         470         480
AATCAGGGGA TAACGCAGGA AAGAACATGT GAGCAAAAGG CCAGCAAAAG GCCAGGAACC

       490         500         510         520         530         540
GTAAAAAGGC CGCGTTGCTG GCGTTTTTCG ATAGGCTCCG CCCCCCTGAC GAGCATCACA

       550         560         570         580         590         600
AAAATCGACG CTCAAGTCAG AGGTGGCGAA ACCCGACAGG ACTATAAAGA TACCAGGCGT

       610         620         630         640         650         660
TTCCCCCTGG AAGCTCCCTC GTGCGCTCTC CTGTTCCGAC CCTGCCGCTT ACCGGATACC

       670         680         690         700         710         720
TGTCCGCCTT TCTCCCTTCG GGAAGCGTGG CGCTTTCTCA TAGCTCACGC TGTAGGTATC

       730         740         750         760         770         780
TCAGTTCGGT GTAGGTCGTT CGCTCCAAGC TGGGCTGTGT GCACGAACCC CCCGTTCAGC

       790         800         810         820         830         840
CCGACCGCTG CGCCTTATCC GGTAACTATC GTCTTGAGTC CAACCCGGTA AGACACGACT

       850         860         870         880         890         900
TATCGCCACT GGCAGCAGCC ACTGGTAACA GGATTAGCAG AGCGAGGTAT GTAGGCGGTG

       910         920         930         940         950         960
CTACAGAGTT CTTGAAGTGG TGGCCTAACT ACGGCTACAC TAGAAGGACA GTATTTGGTA
```

```
       970         980         990        1000        1010        1020
TCTGCGCTCT GCTGAAGCCA GTTACCTTCG GAAAAAGAGT TGGTAGCTCT TGATCCGGCA

      1030        1040        1050        1060        1070        1080
AACAAACCAC CGCTGGTAGC GGTGGTTTTT TTGTTTGCAA GCAGCAGATT ACGCGCAGAA

      1090        1100        1110        1120        1130        1140
AAAAAGGATC TCAAGAAGAT CCTTTGATCT TTTCTACGGG GTCTGACGCT CAGTGGAACG

      1150        1160        1170        1180        1190        1200
AAAACTCACG TTAAGGGATT TTGGTCATGA GATTATCAAA AAGGATCTTC ACCTAGATCC

      1210        1220        1230        1240        1250        1260
TTTTAAATTA AAAATGAAGT TTTAAATCAA TCTAAAGTAT ATATGAGTAA ACTTGGTCTG

      1270        1280        1290        1300        1310        1320
ACAGTTACCA ATGCTTAATC AGTGAGGCAC CTATCTCAGC GATCTGTCTA TTTCGTTCAT

      1330        1340        1350        1360        1370        1380
CCATAGTTGC CTGACTCCCC GTCGTGTAGA TAACTACGAT ACGGGAGGGC TTACCATCTG

      1390        1400        1410        1420        1430        1440
GCCCCAGTGC TGCAATGATA CCGCGAGACC CACGCTCACC GGCTCCAGAT TTATCAGCAA

      1450        1460        1470        1480        1490        1500
TAAACCAGCC AGCCGGAAGG GCCGAGCGCA GAAGTGGTCC TGCAACTTTA TCCGCCTCCA

      1510        1520        1530        1540        1550        1560
TCCAGTCTAT TAATTGTTGC CGGGAAGCTA GAGTAAGTAG TTCGCCAGTT AATAGTTTGC

      1570        1580        1590        1600        1610        1620
GCAACGTTGT TGGCATTGCT ACAGGCATCG TGGTGTCACG CTCGTCGTTT GGTATGGCTT

      1630        1640        1650        1660        1670        1680
CATTCAGCTC CGGTTCCCAA CGATCAAGGC GAGTTACATG ATCCCCCATG TTGTGCAAAA

      1690        1700        1710        1720        1730        1740
AAGCGGTTAG CTCCTTCGGT CCTCCGATCG TTGTCAGAAG TAAGTTGGCC GCAGTGTTAT

      1750        1760        1770        1780        1790        1800
CACTCATGGT TATGGCAGCA CTGCATAATT CTCTTACTGT CATGCCATCC GTAAGATGCT

      1810        1820        1830        1840        1850        1860
TTTCTGTGAC TGGTGAGTAC TCAACCAAGT CATTCTGAGA ATAGTGTATG CGGCGACCGA

      1870        1880        1890        1900        1910        1920
GTTGCTCTTG CCCGGCGTCA ATACGGGATA ATACCGCGCC ACATAGCAGA ACTTTAAAAG

      1930        1940        1950        1960        1970        1980
TGCTCATCAT TGGAAAACGT TCTTCGGGGC GAAAACTCTC AAGGATCTTA CCGCTGTTGA

      1990        2000        2010        2020        2030        2040
GATCCAGTTC GATGTAACCC ACTCGTGCAC CCAACTGATC TTCAGCATCT TTTACTTTCA
```

```
      2050        2060        2070        2080        2090        2100
CCAGCGTTTC TGGGTGAGCA AAAACAGGAA GGCAAAATGC CGCAAAAAAG GGAATAAGGG

      2110        2120        2130        2140        2150        2160
CGACACGGAA ATGTTGAATA CTCATACTCT TCCTTTTTCA ATATTATTGA AGCATTTATC

      2170        2180        2190        2200        2210        2220
AGGGTTATTG TCTCATGAGC GGATACATAT TTGAATGTAT TTAGAAAAAT AAACAAATAG

      2230        2240        2250        2260        2270        2280
GGGTTCCGCG CACATTTCCC CGAAAAGTGC CACCTGACGT CTAAGAAACC ATTATTATCA

      2290        2300        2310        2320        2330        2340
TGACATTAAC CTATAAAAAT AGGCGTATCA CGAGGCCCTT TCGTCTCGCG CGTTTCGGTG

      2350        2360        2370        2380        2390        2400
ATGACGGTGA AAACCTCTGA CACATGCAGC TCCCGGAGAC GGTCACAGCT TGTCTGTAAG

      2410        2420        2430        2440        2450        2460
CGGATGCCGG GAGCAGACAA GCCCGTCAGG GCGCGTCAGC GGGTGTTGGC GGGTGTCGGG

      2470        2480        2490        2500        2510        2520
GCTGGCTTAA CTATGCGGCA TCAGAGCAGA TTGTACTGAG AGTGCACCAT ATGCGGTGTG

      2530        2540        2550        2560        2570        2580
AAATACCGCA CAGATGCGTA AGGAGAAAAT ACCGCATCAG GCGCCATTCG CCATTCAGGC

      2590        2600        2610        2620        2630        2640
TGCGCAACTG TTGGGAAGGG CGATCGGTGC GGGCCTCTTC GCTATTACGC CAGCTGGCGA

      2650        2660        2670        2680        2690        2700
AAGGGGGATG TGCTGCAAGG CGATTAAGTT GGGTAACGCC AGGGTTTTCC CAGTCACGAC
                                T7  PROMOTER
      2710        2720        2730        2740        2750        2760
GTTGTAAAAC GACGGCCAGT GAATTGTAAT ACGACTCACT ATA..... ..........
```

DNA Sequence For Human hsp 27 cDNA

(GenBank Accession Number X54079)

```
        10          20          30          40          50          60
GCCCAGCGCC CCGCACTTTT CTGAGCAGAC GTCCAGAGCA GAGTCAGCCA GCATGACCGA

        70          80          90         100         110         120
GCGCCGCGTC CCCTTCTCGC TCCTGCGGGG CCCCAGCTGG GACCCCTTCC GCGACTGGTA

       130         140         150         160         170         180
CCCGCATAGC CGCCTCTTCG ACCAGGCCTT CGGGCTGCCC CGGCTGCCGG AGGAGTGGTC

       190         200         210         220         230         240
GCAGTGGTTA GGCGGCAGCA GCTGGCCAGG CTACGTGCGC CCCCTGCCCC CCGCCGCCAT

       250         260         270         280         290         300
CGAGAGCCCC GCAGTGGCCG CGCCCGCCTA CAGCCGCGCG CTCAGCCGGC AACTCAGCAG

       310         320         330         340         350         360
CGGGGTCTCG GAGATCCGGC ACACTGCGGA CCGCTGGCGC GTGTCCCTGG ATGTCAACCA

       370         380         390         400         410         420
CTTCGCCCCG GACGAGCTGA CGGTCAAGAC CAAGGATGGC GTGGTGGAGA TCACCGGCAA

       430         440         450         460         470         480
GCACGAGGAG CGGCAGGACG AGCATGGCTA CATCTCCCGG TGCTTCACGC GGAAATACAC

       490         500         510         520         530         540
GCTGCCCCCC GGTGTGGACC CCACCCAAGT TTCCTCCTCC CTGTCCCCTG AGGGCACACT

       550         560         570         580         590         600
GACCGTGGAG GCCCCCATGC CCAAGCTAGC CACGCAGTCC AACGAGATCA CCATCCCAGT

       610         620         630         640         650         660
CACCTTCGAG TCGCGGGCCC AGCTTGGGGG CCCAGAAGCT GCAAAATCCG ATGAGACTGC

       670         680         690         700         710         720
CGCCAAGTAA AGCCTTAGCC CGGATGCCCA CCCCTGCTGC CGCCACTGGC TGTGCCTCCC

       730         740         750         760         770         780
CCGCCACCTG TGTGTTCTTT TGATACATTT ATCTTCTGTT TTTCTCAAAT AAAGTTCAAA

       790         800         810         820         830         840
GCAACCACC. .......... .......... .......... .......... ..........
```

DNA Sequence For The

Human hsp 27 Gene

(GenBank Accession Number X03900)

```
        10          20          30          40          50          60
GAATTCATTT GCTTTTCCTT AACGAGAGAA GGTTCCAGAT GAGGGCTGAA CCCTCTTCGC

        70          80          90         100         110         120
CCCGCCCACG GCCCCTGAAC GCTGGGGGAG GAGTGCATGG GGAGGGGCGG CCCTCAAACG

       130         140         150         160         170         180
GGTCATTGCC ATTAATAGAG ACCTCAAACA CCGCCTGCTA AAAATACCCG ACTGGAGGAG

       190         200         210         220         230         240
CATAAAAGCG CAGCCGAGCC CAGCGCCCCG CACTTTTCTG AGCAGACGTC CAGAGCAGAG

       250         260         270         280         290         300
TCAGCCAGCA TGACCGAGCG CCGCGTCCCC TTCTCGCTCC TGCGGGGCCC CAGCTGGGAC
           EXON 1
       310         320         330         340         350         360
CCCTTCCGCG ACTGGTACCC GCATAGCCGC CTCTTCGACC AGGCCTTCGG GCTGCCCCGG

       370         380         390         400         410         420
CTGCCGGAGG AGTGGTCGCA GTGGTTAGGC GGCAGCAGCT GGCCAGGCTA CGTGCGCCCC

       430         440         450         460         470         480
CTGCCCCCCG CCGCCATCGA GAGCCCCGCA GTGGCCGCGC CCGCCTACAG CCGCGCGCTC

       490         500         510         520         530         540
AGCCGGCAAC TCAGCAGCGG GGTCTCGGAG ATCCGGCACA CTGCGGACCG CTGGCGCGTG

       550         560         570         580         590         600
TCCCTGGATG TCAACCACTT CGCCCCGGAC GAGCTGACGG TCAAGACCAA GGATGGCGTG

       610         620         630         640         650         660
GTGGAGATCA CCGGTGAGCC CCCCTGCTCC TGCAGGGGAG AGGAGGAGGC TAGCAGGGCG

       670         680         690         700         710         720
GGCAGGGCCG GGGGCGTGCG GTTGAAACGG GGGTCCCGGG GGCCTGGGGA GTTAAACGTT

       730         740         750         760         770         780
GGCCCAGCAC CGGGAAAAAC AGGACTCCTG ATTCCCTTGC TCAGGAATTG GGAGTGCGGG

       790         800         810         820         830         840
TCGCTTCTAA GGGCGCTTTC TGCTCTGTAA TCCCAGCGCT TTGGGAGGCC GAGACGGGAG

       850         860         870         880         890         900
GATCGCTTGA GGCCAGGAGT TCAAGACTAG CCTGGGCAAC ATAGCGAGAC GCGCCCCCCC

       910         920         930         940         950         960
GCCCCGACCC CGCGCCATTA CAAAAAAAAA GCAAACAAAA ATTTTTTTAA AGATCATCGA
```

970 980 990 1000 1010 1020
TGAAGAGAGA AAATGCGCTT TTCTACAGAG TCCCCTTCCC ACCCACAGCC CCATCCCCAG

1030 1040 1050 1060 1070 1080
ATAAGCGGGG AGTTCCCTGG CGCGGTGCCA GTTTCTAGCC GCTGAGTGGG CGTGTGCGCG

1090 1100 1110 1120 1130 1140
GCTCCAAGTG CGCCTGCGTA CTGCTCACTC CCCAGCTCCG CGCCCTGCTC CGTTCCTCCC

1150 1160 1170 1180 1190 1200
AAAACTCTGA ATCGAAGAAC TTTCCGGAAG TTTCTGAGAG CCCAGACCGG CGGGCACGCC

1210 1220 1230 1240 1250 1260
CCCATCCCCA ACCCCCTCTG TTAATCCCTA CCAGCCTGCA GTCCTGGCTG CTTCCAAGCA

1270 1280 1290 1300 1310 1320
GGAGGTGGGG CCTCTGGCTA GCGGGGCCGA AAAAGTCCCC TCCCCCGCAT GTCTGATTTC

1330 1340 1350 1360 1370 1380
CCTCTTCCCC CCAAAGGCAA GCACGAGGAG CGGCAGGACG AGCATGGCTA CATCTCCCGG

EXON 2

1390 1400 1410 1420 1430 1440
TGCTTCACGC GGAAATACAC GTGAGTCCTG GCGCCAGGTC GGGGTGGGTG GGTGGCGTGG

1450 1460 1470 1480 1490 1500
GGGTGGGGTC AGGGAAGAGG GCACAGGGAC CCACCCGGTG TGTAATGTAA CGCTTGCCTT

1510 1520 1530 1540 1550 1560
TCCTCTCTGC ACGTCCAGGC TGCCCCCCGG TGTGGACCCC ACCCAAGTTT CCTCCTCCCT

EXON 3

1570 1580 1590 1600 1610 1620
GTCCCCTGAG GGCACACTGA CCGTGGAGGC CCCCATGCCC AAGCTAGCCA CGCAGTCCAA

1630 1640 1650 1660 1670 1680
CGAGATCACC ATCCCAGTCA CCTTCGAGTC GCGGGCCCAG CTTGGGGGCA GAAGCTGCAA

1690 1700 1710 1720 1730 1740
AATCCGATGA GACTGCCGCC AAGTAAAGCC TTAGCCCGGA TGCCCACCCC TGCTGCCGCC

1750 1760 1770 1780 1790 1800
ACTGGCTGTG CCTCCCCCGC CACCTGTGTG TTCTTTTGAT ACATTTATCT TCTGTTTTTC

1810 1820 1830 1840 1850 1860
TCAAATAAAG TTCAAAGCAA CCACCTGTCA CTGGCCCAGG CCCTGGTGTT TGTGGAAGGA

1870 1880 1890 1900 1910 1920
AGCCTCAGGC ACCTGCCATT TGCTGGCTTT CAGGAGTCAT CTTTGCTCAG GCCCGTGCTG

1930 1940 1950 1960 1970 1980
GGCCATGTGG GTACACTGGT GTAGGTTGCT GGACACAGGC TGACTCACAT CCATAAAGAC

1990 2000 2010 2020 2030 2040
AGAGGTCTTA GGGCCGGGCG CAGTGGCTCA TACCTACAAT CCCAGCACTT TGGGGGGTTG

```
      2050        2060        2070        2080        2090        2100
AAGCAGGAGG AGTGCTTGAA GCCAAGAGTT CTAGACCAGC CTGGACAACA TAGTAAGACT

      2110        2120        2130        2140        2150        2160
GTCTCTAAAA AATAAAAATT AGGCAGGGTG GTACTGCACG CCTGTAGTCC CAGCTACTCA

      2170        2180        2190        2200        2210        2220
GGAGGCTGAG GCAGGAGGAT CGCTTGAGCC CAGAGTTGTG AAGGTACAGT GAGCTAACAT

      2230        2240        2250        2260        2270        2280
CGTGCCATTG CACTCCAGCC TGGGCAACAG AACAAGATCC TGTCTCAAAA CAACCAAAAG

      2290        2300        2310        2320        2330        2340
CCCAGAGAGA AAGAGTGAGA CCCCATCTTT AAAAGAAAAA AAAAAAAGGT CATGATTGCA

      2350        2360        2370        2380        2390        2400
AGGTCACGAT TGCAATTAAA ACTGTAAGGT GGGGAAGGAG GAGGAAATAA GAGAAGCACC

      2410        2420        2430        2440        2450        2460
TGAGGCTTGA GTTCTCAGGA GCACCTAGGT TGGGTCCCAG GTGAAGGGGC ACAGAGGTAA

      2470        2480        2490        2500        2510        2520
TTGCACCTCA GAGCTGATGG GAGGATTACT ATGTCA.... .......... ..........
```

DNA Sequence For DNA 4/T7

(The pGEM-3Z information that was removed is underlined).

```
        10          20          30          40          50          60
CATTCGAGTC GGTACCGGGG ACCTCTAGAT CGACCTTGAA TACGGAGCTT CAAGCCAGGC

        70          80          90         100         110         120
TCCAGCTGGC GTTTTAGGTG TGGTGCAAGC AANGCTGGGT ANGTATTTCC AGGAAGCAGT

       130         140         150         160         170         180
AGANGAGCTC GCCGTGCAAT ATGAATTATC ATTAGCAGCA AGGAAATTCG TCAACGATTT

       190         200         210         220         230         240
GTTGGTAAAT GAATTTCCTG TCAGATATCC AGTTGCTTGG GAGAACGCTG CCTTTGCGCC

       250         260         270         280         290         300
TCCTGCGGAC GGCTCAATCT GGCTTAAGTA CGACTACACA GAAGTTGACA CAGTAACATA

       310         320         330         340         350         360
CGGGTTAAGC AGGAAGTGCA AATACTACGT TGGCATGGTG CAGATTTCAG TGTTCTTCAG

       370         380         390         400         410         420
TCCGGGGACT GGGATTGATA AGCCGAGACA AATAGCTAAT CAATTGGCAG AATCTATCGT

       430         440         450         460         470         480
TGATGGTACA ATGCTTGACA GCGGGACCAT TTATGANTCT GGAGTTGTTA ACCCGGTTAT

       490         500         510         520         530         540
CAAATCCAAG TCTGGGTGGT TTATCCCGGT TCGTTTTTAT GTTCGTCTAN ACTAACAAAA

       550         560         570         580         590         600
GGAAAAATAC ATGGCCCATC TCAGCAATGG CACGCAGGTC TTCGTNNAAA GCTCTCGCGG

       610         620         630         640         650         660
AACTGCAATC GACNTAACTG CAATTTCTAA CGCAGTTACC CCTGTTTTAA CTGTATCTGA

       670         680         690         700         710         720
CCNTCTGGTT TGGTCNTTGG TGAATACCTG CTTTTCANCT CTTCTGCTTC AACGCTTCTA

       730         740         750         760         770         780
ACTGANAANC AANTTCNAAT GACTGCAATC CCTGGTACNT CAGTAACTGT TGAAAGGTTT

       790         800         810         820         830         840
GATACCTCCC ACCACCNCAA AATTCCCNGG TNGGCTTAAC TGGGTGAAAT TTTCAAAATC

       850         860         870         880         890         900
ACCCCCGGGT TCCAAATTTC CNTGCGNTTC AGGAANTTTC NACNGAAAGG GGGGGAAANA

       910         920         930         940         950         960
ANCANTTCCC TTAACTTCCC ATTGCTTGTT CNAANNAAAC GG........ ..........
```

DNA Sequence For DNA 4/SP6

(The pGEM-3Z information that was removed is underlined).

```
        10         20         30         40         50         60
CAAGCTTGCA TGCCTGCAGG GTCGACCACG AATGGATACC GCGATGGATA CCGTTTCGAT

        70         80         90        100        110        120
TTCGTTAACA CCAATGGTAG GGGTGTCGTT GAAAGCGATA GTACCGGACT GCAAGCGCAT

       130        140        150        160        170        180
TTCGCTAGCT CGAGGTACGA ACAGACGAAT CGCAACAACC TGGCCAGACT CATCGTAGTT

       190        200        210        220        230        240
ACGCAGAACC GGATATACAG GGTTGGTGTA CTCGTGAGCG AAGGTGAAGG TGTTAGTTAC

       250        260        270        280        290        300
CGCAGATTTA TAGGTCGGAA TCTGCTGTTC ACGGTCATCG GACAAGCACT GGAAGTTAAC

       310        320        330        340        350        360
GAATTGCTGT TCGCCACCGT CAGTAGATAC GTCCTGAACG CACGGAACTT CGAACCAGGA

       370        380        390        400        410        420
GGTGATCTTG ACAACTTCAC CAGTTAAGCC AGCCGGGAAC TTTGTGGTGC TGGAGGTATC

       430        440        450        460        470        480
AATACCTTCA ACAGTTACTG ACGTACCAGT GATTGCAGTC ACTCNAACTT GCTTGTCAGC

       490        500        510        520        530        540
TAGAANCGTT GAAGCAGAAN AAGTGAAAAG CAGGTAATCA CCAACGACCA AACCAGACGC

       550        560        570        580        590        600
GTCAGATACA GTTAAAACAG GGGTAACTGC GTTAGAAATT GCAGTTACGT CNATTGCAGC

       610        620        630        640        650        660
TCCNCNNAN CCTTCTACNA AAAACTGCGT GCCATTGCTG AAATGGGCCA GTTTTTTTCC

       670        680        690        700        710        720
CTTTGTTAAT CTAGACAACA TAAAAACGAA CCNGGAATAA ACCCCCNAAC TGGATTTGAN

       730        740        750        760        770        780
AACCGGGTTA CAACNCCAAA CNCATAAATG GTCCCCTGTT CAACATGTTC CCNCACNAAA

       790        800        810        820        830        840
AATCCGCCCT TGAATAACTA TTTTTTCGNN TACAANCCAT CCCCGGACTG AAAAACTGAA

       850        860        870        880        890        900
ATCGCCCCNG CCACTTTTTT TTCCCTCCCC CTNACCCTTT TTCCGNGTCA CTCCGTTTTN

       910        920        930        940        950        960
CCTACTNAAC CNAATAACCC CCNGGAAGCC AANGGGNTTC NCCT...... ..........
```

Complete DNA Sequence For DNA 4

```
        10         20         30         40         50         60
TCGACCTTGA ATACGGAGCT TCAAGCCAGG CTCCAGCTGG CGTTTTAGGT GTGGTGCAAG

        70         80         90        100        110        120
CAAGGCTGGG TATGTATTTC CAGGAAGCAG TAGANGAGCT CGCCGTGCAA TATGAATTAT

       130        140        150        160        170        180
CATTAGCAGC AAGGAAATTC GTCAACGATT TGTTGGTAAA TGAATTTCCT GTCAGATATC

       190        200        210        220        230        240
CAGTTGCTTG GGAGAACGCT GCCTTTGCGC CTCCTGCGGA CGGCTCAATC TGGCTTAAGT

       250        260        270        280        290        300
ACGACTACAC AGAAGTTGAC ACAGTAACAT ACGGGTTAAG CAGGAAGTGC AAATACTACG

       310        320        330        340        350        360
TTGGCATGGT GCAGATTTCA GTGTTCTTCA GTCCGGGGAC TGGGATTGAT AAGCCGAGAC

       370        380        390        400        410        420
AAATAGCTAA TCAATTGGCA GAATCTATCG TTGATGGTAC AATGCTTGAA CAGCGGGACC

       430        440        450        460        470        480
ATTTATGAGT CTGGAGTTGT TAACCCGGTT ATCAAATCCA AGTCTGGGTG GTTTATTCCC

       490        500        510        520        530        540
GGTTCGTTTT TATGTTCGTC TAGACTAACA AAAGGAAAAA TACATGGCCC ATCTCAGCAA

       550        560        570        580        590        600
TGGCACGCAG GTCTTCGTAG AAAGCTCTCG CGGAACTGCA ATCGACGTAA CTGCAATTTC

       610        620        630        640        650        660
TAACGCAGTT ACCCCTGTTT TAACTGTATC TGACGCGTCT GGTTTGGTCG TTGGTGAATA

       670        680        690        700        710        720
CCTGCTTTTC ACCTCTTCTG CTTCAACGCT TCTAACTGAC AANCAAGTTC GAATGACTGC

       730        740        750        760        770        780
AATCCCTGGT ACGTCAGTAA CTGTTGAAGG TTTGATACCT CCAGCACCAC AAAGTTCCCG

       790        800        810        820        830        840
GCTGGCTTAA CTGGTGAAGT TGTCAAGATC ACCTCCTGGT TCGAAGTTCC GTGCGTTCAG

       850        860        870        880        890        900
GACGTATCTA CTGACGGTGG CGAACAGCAA TTCGTTAACT TCCAGTGCTT GTCCGATGAC

       910        920        930        940        950        960
CGTGAACAGC AGATTCCGAC CTATAAATCT GCGGTAACTA ACACCTTCAC CTTCGCTCAC
```

```
       970         980         990        1000        1010        1020
GAGTACACCA ACCCTGTATA TCCGGTTCTG CGTAACTACG ATGAGTCTGG CCAGGTTGTT

      1030        1040        1050        1060        1070        1080
GCGATTCGTC TGTTCGTACC TCGAGCTAGC GAAATGCGCT TGCAGTCCGG TACTATCGCT

      1090        1100        1110        1120        1130        1140
TTCAACGACA CCCCTACCAT TGGTGTTAAC GAAATCGAAA CGGTATCCAT CGCGGTATCC

      1150        1160        1170        1180        1190        1200
ATTCGTG... .......... .......... .......... .......... ..........
```

DNA Sequence For DNA 1/T7

(The pGEM-3Z information that was removed is underlined).

```
        10          20          30          40          50          60
GNNNNNGNNN NNNNNNNNNN NNNCATTNGA GCTCGGGTAC CGGGGGACCT CTAGAGTCGA

        70          80          90         100         110         120
CTATCTGGTA CTCTATACCT CTACACACCT TTTAAAGTGA ATTGACCTAT TTGCTTTTTC

       130         140         150         160         170         180
TAGAATTTTT AAAATTTTTG TTCCACCGTT AAACCCCAGT TGATTCTCGA CTCAACTGAA

       190         200         210         220         230         240
AACTATTCTT CTTTCGGAAA CCTTCATTGA CTATTCAAAG TTCTTCTTTA TTCTTTAGTG

       250         260         270         280         290         300
CCATCCTCTG TATATTACCC TTATTACCCT TCAAACTACC TATTGTAACT ACATATTGTT

       310         320         330         340         350         360
TTTATGTCCA TCTCCCCTAA TAAACAGTGA GCTCCTTCAA GAAAGAGTAT CCTTCCACCT

       370         380         390         400         410         420
CCCTTTTTCC GCATCACTAG AACAGGGTAG GTAGGCACTG ATTAAGTACA TTATTCAGTT

       430         440         450         460         470         480
CATCACTCAG TATTGGACAT ATATTTCTTG AGAATTTACT AAAGGCCTTG CTTGCGTGGT

       490         500         510         520         530         540
TCTANGCATC GTGAATTGAG TGGTGAACAA AACATTTAAA CCCCTTGTCC TCATGGACTT

       550         560         570         580         590         600
TATGTTCCTG GGGACAGAAT TAGATAACAT ACATATAAAA ACACAGTGTT ATTTTTACAT

       610         620         630         640         650         660
GGTCNTAAGA AGAAAATAAA GATAGATAAG GAAGTGGAAG TATTTAGGTG GAATGGTCAT

       670         680         690         700         710         720
GGAAGGTCTC TTTAATGTGG TGACAAATGA TTAATGACTG AATGTATGTT AGTGAGAAAT

       730         740         750         760         770         780
CAATTTTGAN AACACCTGTT GATAGANTGT CCTAATTTGA AGAAACAGCA AGCAGAAGGN

       790         800         810         820         830         840
AGAAAAAGCA GTGGGGTTTG AAGTTATTTT GAAAANGGTC ACTGTATCTG GGANACANAA

       850         860         870         880         890         900
AGTNANTTAA AGGTCGAANA AAAAAACAGA AGGTCAGATT TCCNNACGGT ACAAGGGAAT
```

```
       910         920         930         940         950         960
TTGAAATCCA CATGTTCCCG AATAAATGGA GAAANGAATT NANCATCCCC TGAATTTGTT

       970         980         990        1000        1010        1020
GAATAATANC TGAAAGTTAC CCCCTTTCAA CTGGTTCCNA ATTAAAAATT CTCATTAACC

      1030        1040        1050        1060        1070        1080
TTGAACCCCT CAANCNTTTT TTTGAAAAGT GGATAAAATT CAAGAACTTC NCCCNGGCCA

      1090        1100        1110        1120        1130        1140
AAAATTTGGG AAACCCCCCT ATNTTTATGC CCCAAAATCC NC..... ..........
```

DNA Sequence For DNA 1/SP6

(The pGEM-3Z information that was removed is underlined).

(The HindIII site is double-underlined).

```
        10         20         30         40         50         60
CTCNAGCTTG CATGCCTGCA GGTCGACCTG CAGGTCAACG GGATCTTTGG CATGATATTT

        70         80         90        100        110        120
ATACACTGTC TCATTTTCTT CGTTTGTATA TTGGGGATGT TTATAATTAT TTTGCAGGAT

       130        140        150        160        170        180
TATTTTGAAT GTGAGAAATA ATATCTTTAA TTTGATTTTC TTTTGTTAGC TAATTAAATG

       190        200        210        220        230        240
GTAACTAACT TTTATTCATA TAATTAGTAT GGTCTCATCA CACAAGGCCC TTTGCAACAT

       250        260        270        280        290        300
TGCTCCTACT TCTATGTCTA AGTTCTGCTT CACTTTCCTA TGCTCCATCC ATTTCAATCA

       310        320        330        340        350        360
CCCCAGACGT CTCTGTATAA TTTGGCCATG TCCTATCACA ACTCTAATCT TGTTTGTATT

       370        380        390        400        410        420
CCTGGAATGC CCTGAAAAAC TCCTACTCAG ACTTAAAAAG CCAGCTCAAA TATTGTGTTC

       430        440        450        460        470        480
TTCGTGATTC TTTTCCTTCT TCATTGANGT AGANTCTAAC ACTTCTTCAC ACACCACTAA

       490        500        510        520        530        540
CAAACTAGGT TTGAAATGTT ACTTTTGTAC ATTTAAGAAC CTTTTTTACA TTTTTCATTT

       550        560        570        580        590        600
TAAAAACATT TTAATAGCAA AAATATTATA ACAGAGAAGA ATACATACTA ACGAATGTTA

       610        620        630        640        650        660
TTTAAATAAG ANTTATTGGT ATAAACATAG ATGTTTTAAT AACTGGTAAT CTGGAAAAAG

       670        680        690        700        710        720
AAATAGAACA CAGCCTGGTA TCACTGAAAC TCCTCTATGT TCTTCCATGA ATAACCACCT

       730        740        750        760        770        780
ATCCCCGTGC CTCCTTGGAA TTAATAACNA ATTTTGAATT TTGGTGAANA TTTTTTTTAA

       790        800        810        820        830        840
AAANTTGCCT NAATTTTAAT CCNGAAACCT TTGGAAAAAA TACGTTTTTT CTATTTTTAA

       850        860        870        880        890        900
ACNGGCNCAT GNAAATGGGA AGGTNTGCTG GAAAAAAATT TTTTTGANAT AATTTCCTTT

       910        920        930        940        950        960
AACTGGGTCT NGNAAAAAAT NTTTTTAANN AAANNTTNCC CTGAACACCC CNNTNTNN..
```

DNA Sequence For DNA 1/HIND3/SP6

(The pGEM-3Z information that was removed is underlined).

```
        10          20          30          40          50          60
CNCAAGCTTA TAAAAGCAAA GTTTCTATTA AGTGACCAGT GTATTTATAT GCGATAAGAA

        70          80          90         100         110         120
GATACTACTT TTAATGAAAT AAATACATAA ACATTTTGAA CTGTCTACAG AAAAACTCAG

       130         140         150         160         170         180
AGAGATGGTA ATGAAAACAC TCTGTTTATA TAAAAGAACT TCTTCACTAG TCCCATAGAA

       190         200         210         220         230         240
TAGAGTGAAT AGAGTCTTGT CTTAATACAG CATTGATTTA TTATGTNATA TATCCATATT

       250         260         270         280         290         300
CCCGATGATC AAGATCTTGC TGGTGTNTAT GGAGGCTGAG GGCCGCTTGG CTCATGTGAA

       310         320         330         340         350         360
GAAAGATAAA TTTTTTATTG TGTCTTGAGT GTTTTAATAT GAAACAAAAC TAAAAAAGAA

       370         380         390         400         410         420
AAAGCCTCTG TCTTAAGCAT CACTTACATT ATGTNGAAAG AAGCAATTTT AAAATCATTC

       430         440         450         460         470         480
GTTCATCATA TAAGAGTATC ATCAACNCNG AATCCCATNT AGAGTGATAA TGAAGATCTG

       490         500         510         520         530         540
TTTTCACNCC ATGGGAGCTA TTTGAGTATT CTTTCCTTAA CTGCTTCTGG CAGTTCACAT

       550         560         570         580         590         600
TGTCAGCGAA GTTAATGCTG GGATGTGGAA CATGATTCCT CTCTTTTTTT TCCNCCAATT

       610         620         630         640         650         660
TATTTTTTTT TAGATCCACA GATAAAATTT ATGTTTTTCT CNNGTTTGAC TGAAGGTTTG

       670         680         690         700         710         720
GAAATNTNTT ACNTTGTTGA ATGGGTTAAA GTTAGCTAAT TNACTTTGTT TTACCTGACA

       730         740         750         760         770         780
TNNTTATCCT TTTTGAAGGT NGGGACACTT TNAATTGAAT TNGGTTNNAA NAANNTTGAC

       790         800         810         820         830         840
CTTNAANAAA ANANTNNTAN GGGCCTTTGC CCNNAATNGG GAAANTTNAC TGCCCCGGGG

       850         860         870         880         890         900
AAAAAANANT NCAATTACCT TAAATGCCCC CNTTNTTCCN CCCNAANTTN CCCTTGAATN
```

```
       910         920         930         940         950         960
TTTNAAGNTT TTTNAAAAAT NNTTNCCNCN TTNCCCNNAN NANTTTAANT CCTAAAANCC

       970         980         990        1000        1010        1020
CCCGNCCCNN NTTGGCGGGG NGAATTTCCC NNTTTTTNNN NTANGGGTNA TTCCGGGGGN

      1030        1040        1050        1060        1070        1080
TNACCCGGGN TN........ .......... .......... .......... ..........
```

Partial DNA Sequence For

DNA 1/Region 3+2

```
        10          20          30          40          50          60
CTGCAGGTCA ACGGGATCTT TGGCATGATA TTTATACACT GTCTCATTTT CTTCGTTTGT

        70          80          90         100         110         120
ATATTGGGGA TGTTTATAAT TATTTTGCAG GATTATTTTG AATGTGAGAA ATAATATCTT

       130         140         150         160         170         180
TAATTTGATT TTCTTTTGTT AGCTAATTAA ATGGTAACTA ACTTTTATTC ATATAATTAG

       190         200         210         220         230         240
TATGGTCTCA TCACACAAGG CCCTTTGCAA CATTGCTCCT ACTTCTATGT CTAAGTTCTG

       250         260         270         280         290         300
CTTCACTTTC CTATGCTCCA TCCATTTCAA TCACCCCAGA CGTCTCTGTA TAATTTGGCC

       310         320         330         340         350         360
ATGTCCTATC ACAACTCTAA TCTTGTTTGT ATTCCTGGAA TGCCCTGAAA AACTCCTACT

       370         380         390         400         410         420
CAGACTTAAA AAGCCAGCTC AAATATTGTG TTCTTCGTGA TTCTTTTCCT TCTTCATTGA

       430         440         450         460         470         480
GGTAGANTCT AACACTTCTT CACACACCAC TAACAAACTA GGTTTGAAAT GTTACTTTTG

       490         500         510         520         530         540
TACATTTAAG AACCTTTTTT ACATTTTTCA TTTTAAAAAC ATTTTAATAG CAAAAATATT

       550         560         570         580         590         600
ATAACAGAGA AGAATACATA CTAACGAATG TTATTTAAAT AAGAATTATT GGTATAAACA

       610         620         630         640         650         660
TAGATGTTTT AATAACTGGT AATCTGGAAA AAGAAATAGA ACACAGCCTG GTATCACTGA

       670         680         690         700         710         720
AACTCCTCTA TGTTCTTCCA TGAATAACCA CCTATCCCCG TGCCTCCTTG GAATTAATAA

       730         740         750         760         770         780
CAAATTTTGA ATTTTCGGTG AATATTTTTA TATTACATAA AAATCAATGC CTGAATTTTA

       790         800         810         820         830         840
ATCCAAGACT CTATTCACTC TATTCTATGG AACTAGTGAA GATACGTTTT TTCTATTTTT

       850         860         870         880         890         900
AAACAGAGCG TTTTCATGTA ACATGGCTCT GAAGGTTTGC TGGAGACAGT TCAAAATTTT

       910         920         930         940         950         960
TTTGTATATA ATTTCCTTTA ACTGGGTCTA GTAAAAAACT TCTTATCGCA TATAAATACA

       970         980         990        1000        1010        1020
TTGCTCACTT AATAGAACAC CCCGCTTTTA T......... .......... ..........
```

DNA Sequence For DNA 2/T7

(The pGEM-3Z information that was removed is underlined).

```
        10         20         30         40         50         60
CGATTCGAGC TCGGTACCGG GGATCCTCTA GAGTCGACTA TCTGGTAAGT ACTCTATACC

        70         80         90        100        110        120
TCTACACACC TTTTAAAGTG AATTGACCTA TTTGCTTTTT CTAGAATTTT TAAAATTTTT

       130        140        150        160        170        180
GTTCCACCGT TAAACCCCAG TTGATTCTCG ACTCAACTGA AAACTATTCT TCTTTCGGAA

       190        200        210        220        230        240
ACCTTCATTG ACTATTCAAA GTTCTTCTTT ATTCTTTAGT GCCATCCTCT GTATATTACC

       250        260        270        280        290        300
CTTATTACCC TTCAAACTAC CTATTGTAAC TACATATTGT TTTTATGTCC ATCTCCCCTA

       310        320        330        340        350        360
ATAAACAGTG AGCTCCTTCA AGAAAGANTA TCCTTCCACC TCCCTTTTTC CGCATCACTA

       370        380        390        400        410        420
AAACAGGGTA GGTAGGCACT GATTAAGTAC ATTATTCAGT TCATCACTCA GTATTGGACA

       430        440        450        460        470        480
TATATTTCTT GAGAATTTAC TAAAGGCCTT GCTTGCGTGG TTCTAGGCAT CGTGAAATTG

       490        500        510        520        530        540
AGTGGTGAAC AAAACATTTA AACCCCTTGT CCTCATGGGA CTTTATGTTC CTGGGGANAG

       550        560        570        580        590        600
AATTAGATAA CATACATATA AAAACACAGT GTATTTTTAC ATGGTCCTNN GAAGAAAATT

       610        620        630        640        650        660
AAGATANATA AGGAAGTGGA AGTATTTAGG TGGAANTGTC CTGGAAGGTC NCTTNAATGT

       670        680        690        700        710        720
NGGNGACCAN TGANTAATGA ACTTGAANTG TTTTGTNTTG AAGAAACNAC TTTTTGAAAA

       730        740        750        760        770        780
CCCCGGTGGA ANAAAATTTC CNCNTTTGNA GGAAACCNCC CCCCAAAGGN NAAAAAACCC

       790        800        810        820        830        840
NTGGGTTTGA AANTTCTTTT NAAAAANGNC CNCGTNNCTG GGGANNNNNA ANCTCATTTC

       850        860        870        880        890        900
NAGGTCNAAN AAAAAANANA ANNCTNATNT CNCNNACNNT GCCNNGGAAN CN........
```

DNA Sequence For DNA 2/SP6

(The pGEM-3Z information that was removed is underlined).

```
        10         20         30         40         50         60
NCCAAGCTTG CATGCCTGCA GGTCGACCTG CAGGTCAACG GGATCTTTGG CATGATATTT

        70         80         90        100        110        120
ATACACTGTC TCATTTTCTT CGTTTGTATA TTGGGGATGT TTATAATTAT TTTGCAGGAT

       130        140        150        160        170        180
TATTTTGAAT GTGAGAAATA ATATCTTTAA TTTGATTTTC TTTTGTTAGC TAATTAAATG

       190        200        210        220        230        240
GTAACTAACT TTTATTCATA TAATTAGTAT GGTCTCATCA CACAAGGCCC TTTGCAACAT

       250        260        270        280        290        300
TGCTCCTACT TCTATGTCTA AGTTCTGCTT CACTTTCCTA TGCTCCATCC ATTTCAATCA

       310        320        330        340        350        360
CCCCAGACGT CTCTGTATAA TTTGGCCATG TCCTATCACA ACTCTAATCT TGTTTGTATT

       370        380        390        400        410        420
CCTGGAATGC CCTGAAAAAC TCCTACTCAG ACTTAAAAAG CCAGCTCAAA TATTGTGTTC

       430        440        450        460        470        480
TTCGTGATTC TTTTCCTTCT TCATTGAGGT AGAGTCTAAC ACTTCTTCAC ACACCACTAA

       490        500        510        520        530        540
CAAACTAGGT TTGAAATGTT ACTTTTGTAC ATTTAAGAAN CTTTTTTACA TTTTTCATTT

       550        560        570        580        590        600
TAAAAACATT TTANTAGCNA AAATATTNTA ACCGAANAGG AGTACATNCT AACGAATGTA

       610        620        630        640        650        660
ATTTAAATAA GATTATTGTA TAAACATAGA TGTTTTAATA ACTGGTAATC TGGAAAAAAG

       670        680        690        700        710        720
AAATANAACA CAGCCTGGTA TCACTGAAGC NCCTCTAATG TCCTTCCATG AATAACCACC

       730        740        750        760        770        780
TAATCCCGNT GCCTCCTTGA ANTTAANAAA TAATTTTGAN TTTGGTGAAN ANTTTTTTTT

       790        800        810        820        830        840
ANAAATTCGC CNANNTTTAA TCCGAATCTT TGGAAACATA CGTTTTTTCT ATTTTNAAAC

       850        860        870        880        890        900
TGCCCATGNA ATGGAGGTTN GCTGAANNAA ATTTTTGAAA ANTTTCTTTN ACTGGTCTGN

       910        920        930        940        950        960
AAAAANTTTT AACATTNTTC NCCGACCNNC CANNTNCTCC TNAGCCAAAA AAAAANGNAA

       970        980        990       1000       1010       1020
ATNNTTTTNT NNCNNTT... .......... .......... .......... ..........
```

APPENDIX III

Amino Acid Sequence For The

Human hsp 27 Gene

(GenBank Accession Number X03900)

```
        10         20         30         40         50         60
MTERRVPFSL LRGPSWDPFR DWYPHSRLFD QAFGLPRLPE EWSQWLGGSS WPGYVRPLPP

        70         80         90        100        110        120
AAIESPAVAA PAYSRALSRQ LSSGVSEIRH TADRWRVSLD VNHFAPDELT VKTKDGVVEI

       130        140        150        160        170        180
TGKHEERQDE HGYISRCFTR KYTLPPGVDP TQVSSSLSPE GTLTVEAPMP KLATQSNEIT

       190        200        210        220        230        240
IPVTFESRAQ LGGRSCKIR. .......... .......... .......... ..........
```

APPENDIX IV

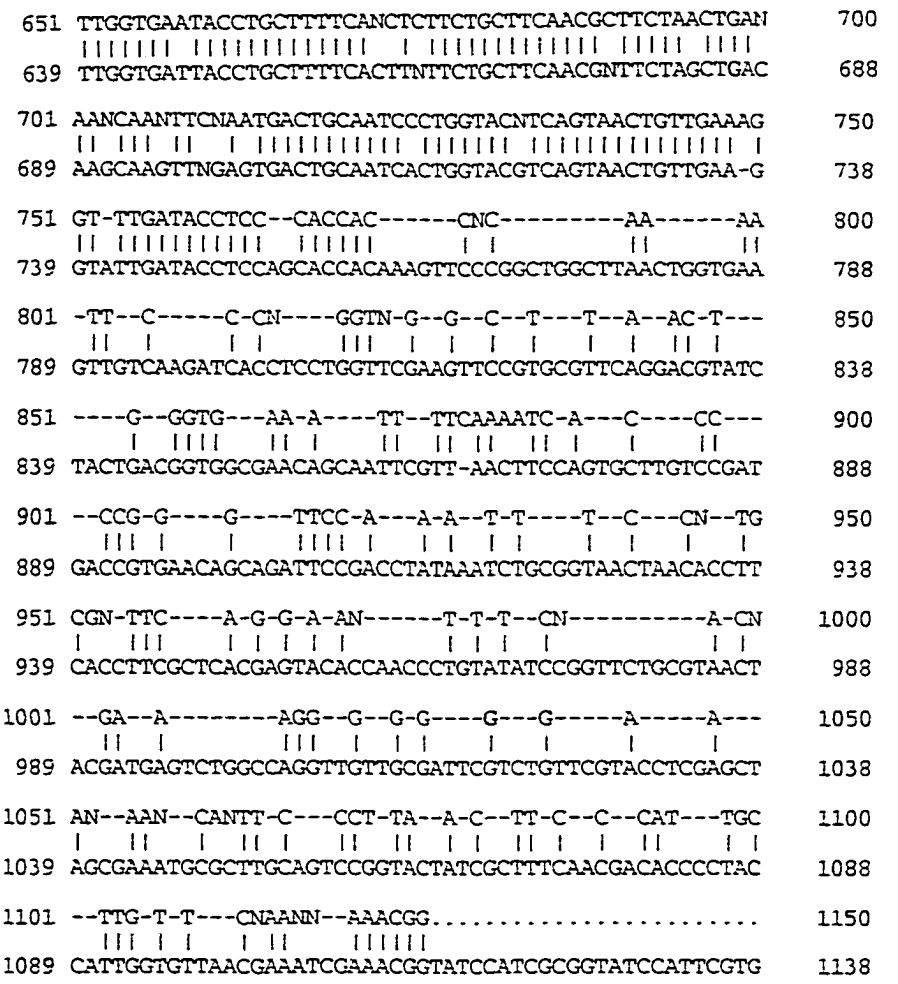
Alignment Of DNA 4-1 (Normal Strand),

And DNA 4-2 (Complementary Strand)

(The top row shows the DNA 4-1 nucleotide sequence).

(The bottom row shows the DNA 4-2 nucleotide sequence).

```
  1 TCGACCTTGAATACGGAGCTTCAAGCCAGGCTCCAGCTGGCGTTTTAGGT      50
                | || |    || ||    |     |||| ||
-11 ............A-GGNG----AANCCCN--T-----TGGC-TT------      38

 51 GTGGTGCAAGCAANGCTGGGTANGTATTTCCAGGAAGCAGTAGANGAGCT     100
          |   |  | ||||   |||    ||      |     || |
 39 ------C---CN-GG--GGGT-T--ATTN---GG------T---TNAG-T      88

101 CGCCGTGCAATATGAATTATCATTAGCAGCAAGGAAATTCGTCAACGATT     150
     |  |   || |  |     |   |  | | |    |  | |  ||
 89 AG--GN--AA-A--A-----C----G--G-A-G----T--GACN-CG---     138

151 TGTTGGTAAATGAATTTCCTGTCAGATATCCAGTTGCTTGGGAGAACGCT     200
     |     |||  ||      |  |       ||         ||   |
139 -G-----AAA--AA------G---G-------GTN-------AG---G--     188

201 GCCTTTGCGCCTCCTGCGGACGGCTCAATCTGGCTTAAGTACGACTACAC     250
     |     | |          | ||        |    ||  |  |  | |
189 G-----G-G----------A-GG--------G----AA--A--A--A-A-     238

251 AGAAGTTGACACAGTAACATACGGGTTAAGCAGGAAGTGCAAATACTACG     300
     | | || |     |   |   |||    || | |  |     |  | |
239 A-A-GT-G-----G---CN---GGG----GC-G-A--T-----T--T-C-     288

301 TTGGCATGGTGCAGATTTCAGTGTTCTTCAGTCCGGGGACTGGGATTGAT     350
       | | |     |||   |  |   ||||||||| |||  ||| |
289 -----A-G-T-----TTT---T---C---AGTCCGGGGA-TGGN-TTG-T     338

351 AAGCCGAGACAAATAGCTAATCAATTGGCAGAATCTATCGTTGATGGTAC     400
     |  | || | |||||| || ||||  ||| || | | | || |  || ||
339 ANNC-GAAA-AAATAGTTATTCAAG-GGCGGATTTT-TNGT-GNGGGAAC     388

401 AATGCTTGA-CAGCGGGACCATTTATGANTCTGGAGTTGTTAACCCGGTT     450
     | || |||| ||| |||||||||||||  | ||| |||||  |||||||||
389 A-TG-TTGAACAG-GGGACCATTTATGNGTTTGGNGTTGT-AACCCGGTT     438

451 ATCAAATCCAAGTCTGGGTGGTTTAT-CCCGGTTCGTTTTTATGTTCGTC     500
     |||||||||  ||  ||| |||||||  || ||||||||||||||| |||
439 NTCAAATCCA-GTTNGGG-GGTTTATTCCNGGTTCGTTTTTATGTT-GTC     488

501 TANACTAACAAAAGGAAAAATACATGGCCCATCTCAGCAATGGCACGCAG     550
     ||  | |||||||  |||||||  ||  ||||||||  ||||||||||||||||
489 TAGATTAACAAAGGGAAAAAAAC-TGGCCCATTTCAGCAATGGCACGCAG     538

551 GTCTTCGTNNAAAGCTCTCGCGGAACTGCAATCGACNTAACTGCAATTTC     600
     |  ||  ||   ||  |  |  |||  |||  |||||||  |||  ||||||||||||
539 TTTTTNGTAGAAGGNTNTCGNGGAGCTGCAATNGACGTAACTGCAATTTC     588

601 TAACGCAGTTACCCCTGTTTTAACTGTATCTGAC-CNTCTGGTTTGGTCN     650
     |||||||||||||||||||||||||||||||||  | ||||||||||||
589 TAACGCAGTTACCCCTGTTTTAACTGTATCTGACGCGTCTGGTTTGGTCG     638
```



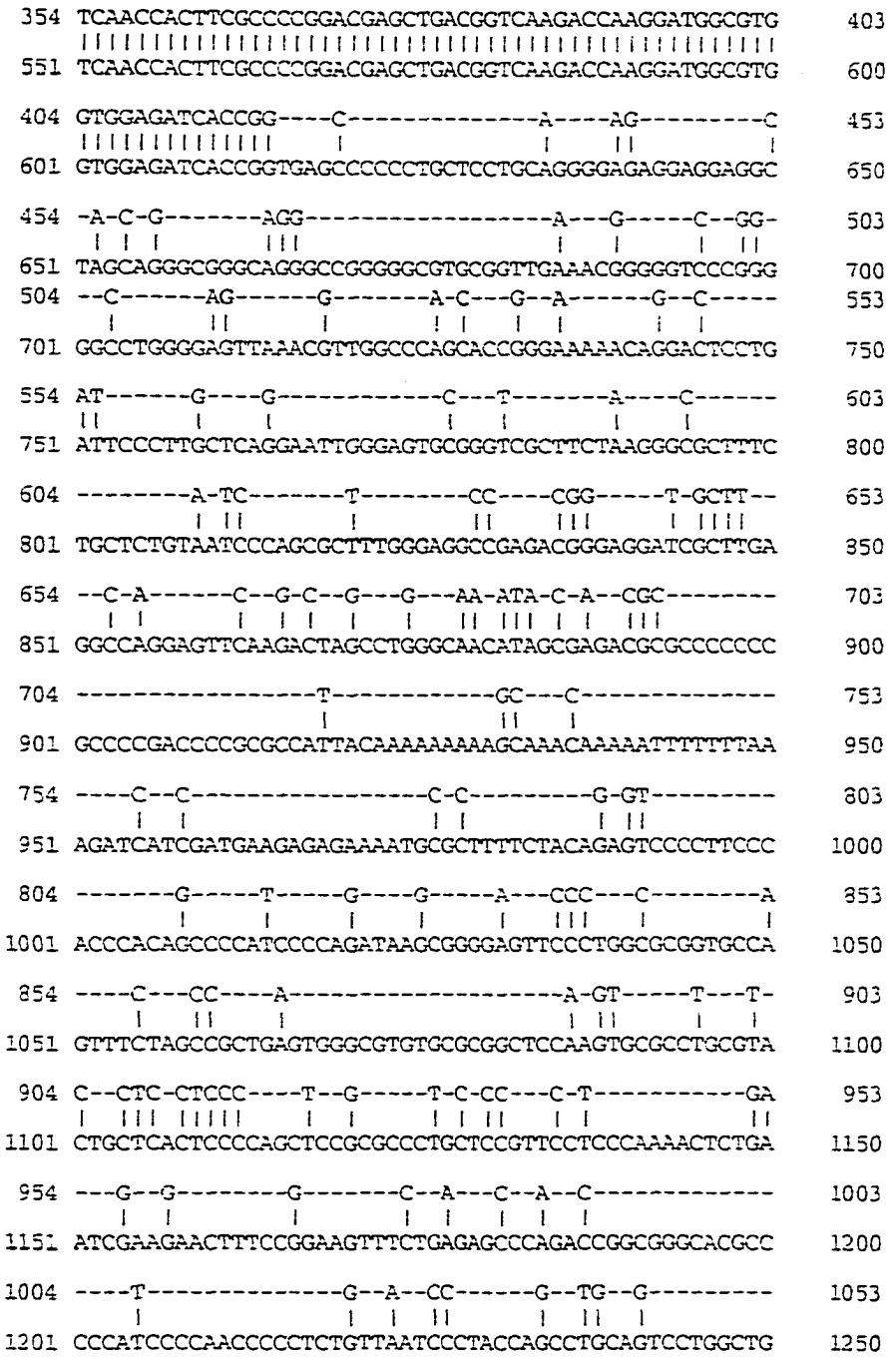
Alignment Of Human hsp 27 cDNA,

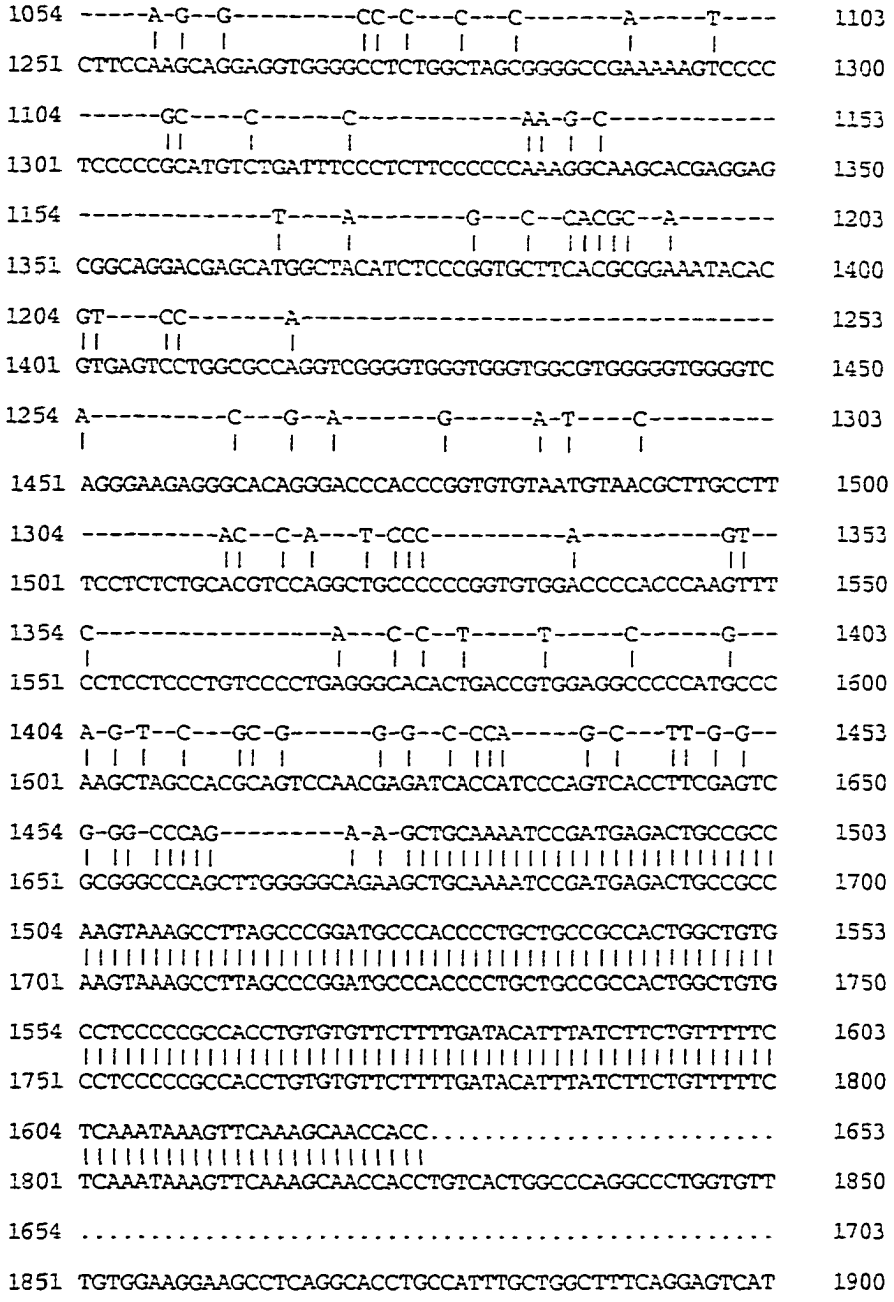
And The Human hsp 27 Gene

(The top row shows the human hsp 27 cDNA nucleotide sequence).

(The bottom row shows the human hsp 27 gene sequence).

```
-196 ..................................................  -147

   1 GAATTCATTTGCTTTTCCTTAACGAGAGAAGGTTCCAGATGAGGGCTGAA   50

-146 ..................................................   -97

  51 CCCTCTTCGCCCCGCCCACGGCCCCTGAACGCTGGGGGAGGAGTGCATGG  100

 -96 ..................................................   -47

 101 GGAGGGGCGGCCCTCAAACGGGTCATTGCCATTAATAGAGACCTCAAACA  150

 -46 ..............................................GCC     3
                                                  |||
 151 CCGCCTGCTAAAAATACCCGACTGGAGGAGCATAAAAGCGCAGCCGAGCC  200

   4 CAGCGCCCCGCACTTTTCTGAGCAGACGTCCAGAGCAGAGTCAGCCAGCA   53
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 201 CAGCGCCCCGCACTTTTCTGAGCAGACGTCCAGAGCAGAGTCAGCCAGCA  250

  54 TGACCGAGCGCCGCGTCCCCTTCTCGCTCCTGCGGGGCCCCAGCTGGGAC  103
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 251 TGACCGAGCGCCGCGTCCCCTTCTCGCTCCTGCGGGGCCCCAGCTGGGAC  300

 104 CCCTTCCGCGACTGGTACCCGCATAGCCGCCTCTTCGACCAGGCCTTCGG  153
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 301 CCCTTCCGCGACTGGTACCCGCATAGCCGCCTCTTCGACCAGGCCTTCGG  350

 154 GCTGCCCCGGCTGCCGGAGGAGTGGTCGCAGTGGTTAGGCGGCAGCAGCT  203
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 351 GCTGCCCCGGCTGCCGGAGGAGTGGTCGCAGTGGTTAGGCGGCAGCAGCT  400

 204 GGCCAGGCTACGTGCGCCCCCTGCCCCCCGCCGCCATCGAGAGCCCCGCA  253
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 401 GGCCAGGCTACGTGCGCCCCCTGCCCCCCGCCGCCATCGAGAGCCCCGCA  450

 254 GTGGCCGCGCCCGCCTACAGCCGCGCGCTCAGCCGGCAACTCAGCAGCGG  303
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 451 GTGGCCGCGCCCGCCTACAGCCGCGCGCTCAGCCGGCAACTCAGCAGCGG  500

 304 GGTCTCGGAGATCCGGCACACTGCGGACCGCTGGCGCGTGTCCCTGGATG  353
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 501 GGTCTCGGAGATCCGGCACACTGCGGACCGCTGGCGCGTGTCCCTGGATG  550
```





```
1704 ..................................................  1753
1901 CTTTGCTCAGGCCCGTGCTGGGCCATGTGGGTACACTGGTGTAGGTTGCT  1950
1754 ..................................................  1803
1951 GGACACAGGCTGACTCACATCCATAAAGACAGAGGTCTTAGGGCCGGGCG  2000
1804 ..................................................  1853
2001 CAGTGGCTCATACCTACAATCCCAGCACTTTGGGGGGTTGAAGCAGGAGG  2050
1854 ..................................................  1903
2051 AGTGCTTGAAGCCAAGAGTTCTAGACCAGCCTGGACAACATAGTAAGACT  2100
1904 ..................................................  1953
2101 GTCTCTAAAAAATAAAAATTAGGCAGGGTGGTACTGCACGCCTGTAGTCC  2150
1954 ..................................................  2003
2151 CAGCTACTCAGGAGGCTGAGGCAGGAGGATCGCTTGAGCCCAGAGTTGTG  2200
2004 ..................................................  2053
2201 AAGGTACAGTGAGCTAACATCGTGCCATTGCACTCCAGCCTGGGCAACAG  2250
2054 ..................................................  2103
2251 AACAAGATCCTGTCTCAAAACAACCAAAAGCCCAGAGAGAAAGAGTGAGA  2300
2104 ..................................................  2153
2301 CCCCATCTTTAAAAGAAAAAAAAAAAAGGTCATGATTGCAAGGTCACGAT  2350
2154 ..................................................  2203
2351 TGCAATTAAAACTGTAAGGTGGGGAAGGAGGAGGAAATAAGAGAAGCACC  2400
2204 ..................................................  2253
2401 TGAGGCTTGAGTTCTCAGGAGCACCTAGGTTGGGTCCCAGGTGAAGGGGC  2450
2254 ..................................................  2303
2451 ACAGAGGTAATTGCACCTCAGAGCTGATGGGAGGATTACTATGTCA....  2500
```

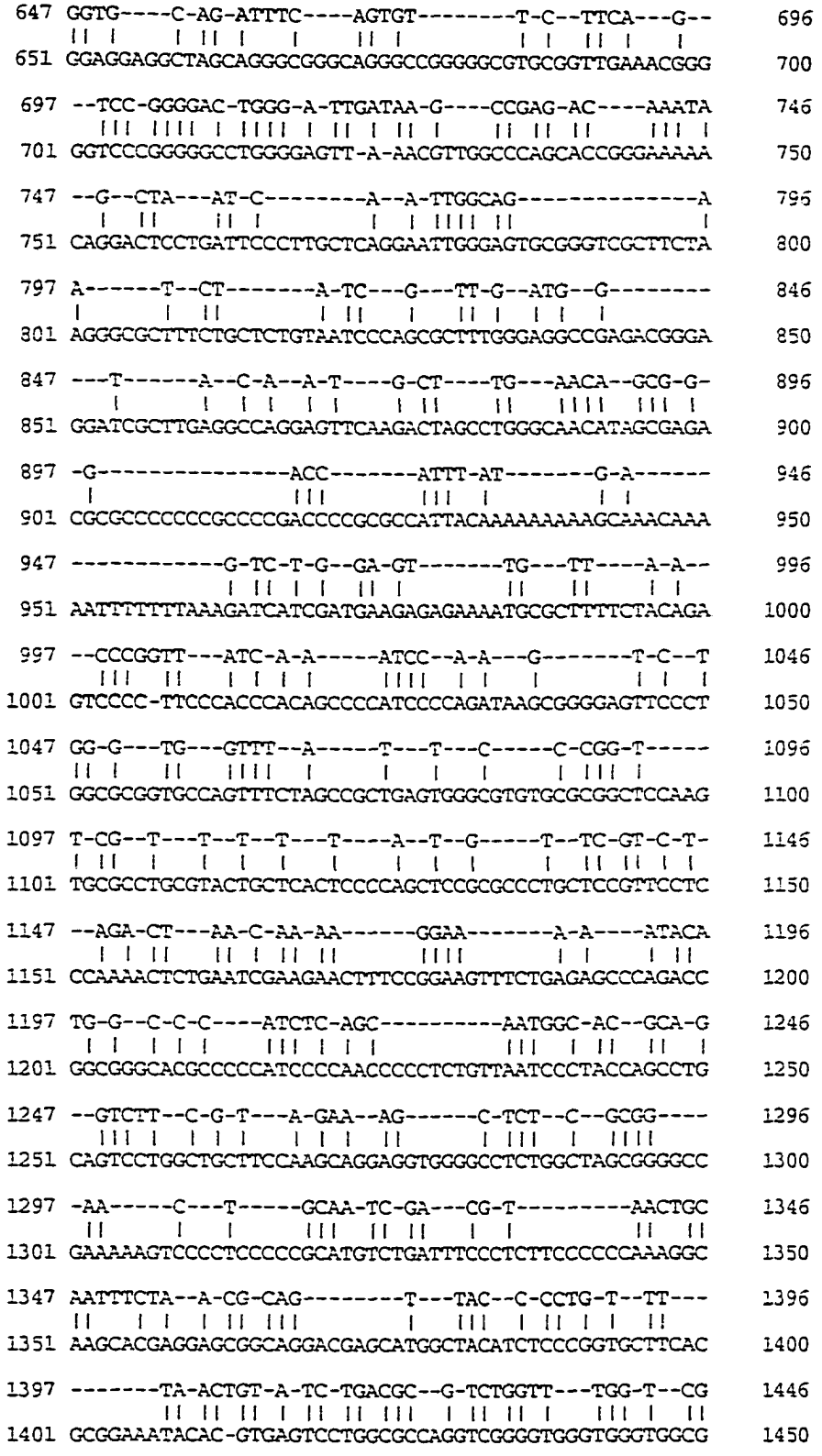
Alignment Of DNA 4 (Normal Strand),

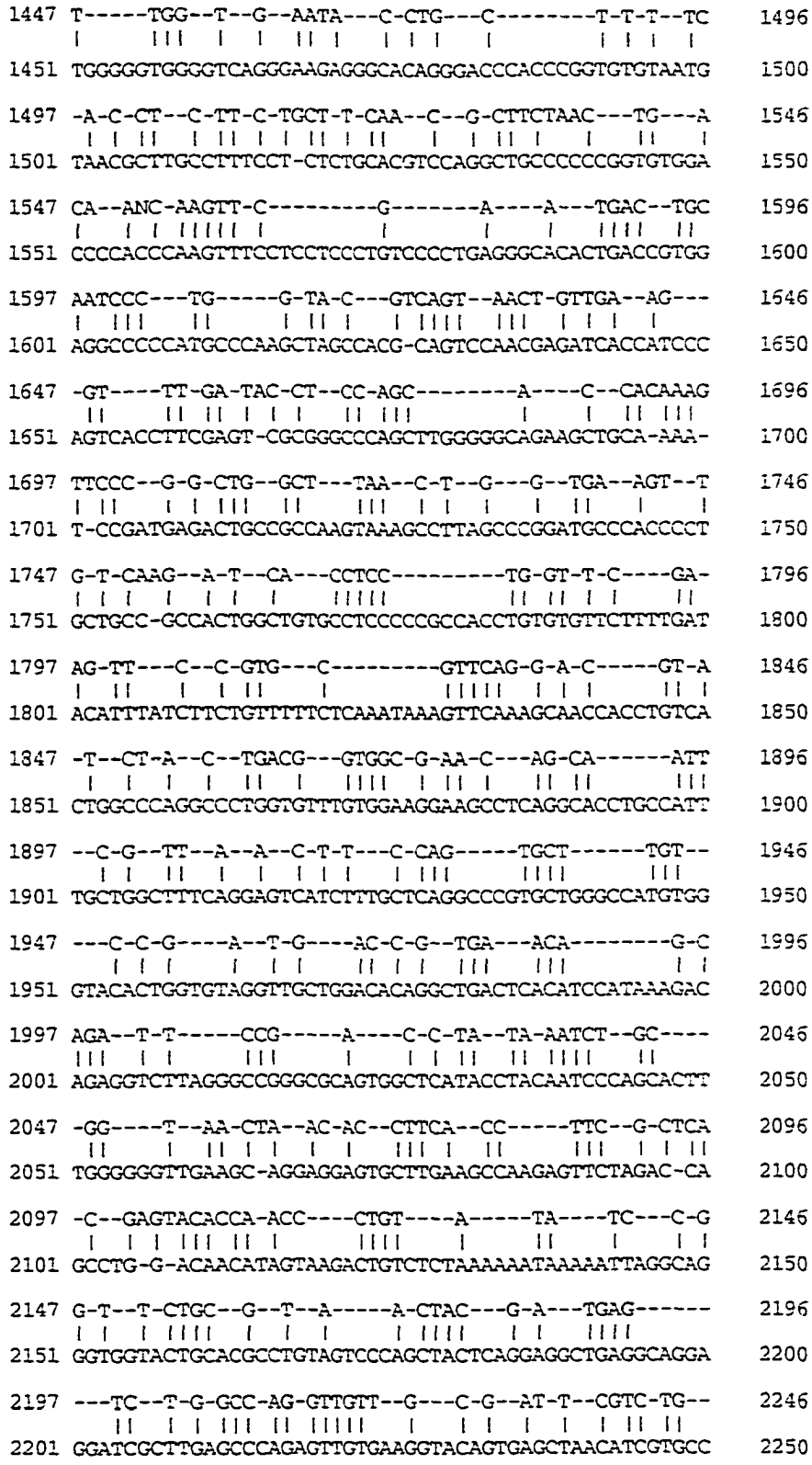
And The Human hsp 27 Gene

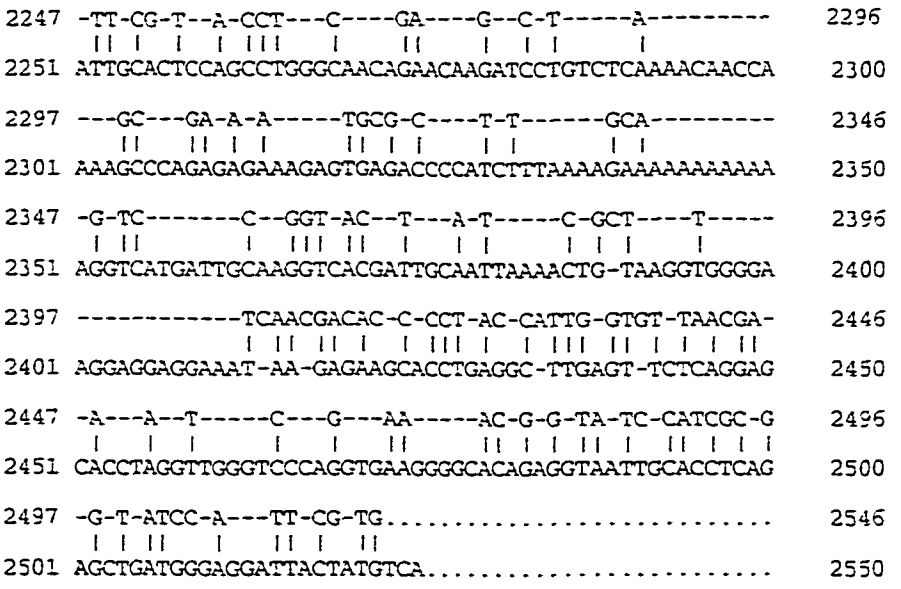
(The top row shows the DNA 4 nucleotide sequence).

(The bottom row shows the human hsp 27 gene sequence).

```
 -3 ....TC----GA----CCTT---GA-ATA-----C--G--GAG--CT---      46
 1 GAATTCATTTGCTTTTCCTTAACGAGAGAAGGTTCCAGATGAGGGCTGAA      50

 47 ---TCAA-GCC------A-GGCTCC--A--GCTGGCGTTTTAGG--TG--      96
 51 CCCTCTTCGCCCCGCCCACGGCCCCTGAACGCTGGGGG---AGGAGTGCA     100

 97 TGGTGCAAG--C-------AA--GG--C-TGGG--T-A-T-G-----T-A     146
101 TGGGG-AGGGGCGGCCCTCAAACGGGTCATTGCCATTAATAGAGACCTCA     150

147 ---------T--T-----T-CC--A--GGA--AGCAGTA---G---AN--     196
151 AACACCGCCTGCTAAAAATACCCGACTGGAGGAGCA-TAAAAGCGCAGCC     200

197 GAGCTC-GC-CGT-GCAATAT---GA--ATTA--TC-ATTAGCAG---CA     246
201 GAGCCCAGCGCCCCGCACTTTTCTGAGCAG-ACGTCCAG-AGCAGAGTCA     250

247 ---AGGA--A---A---------T----T-C--G-TCAA-CG-------A-     296
251 GCCAGCATGACCGAGCGCCGCGTCCCCTTCTCGCTCCTGCGGGGCCCCAG     300

297 -T--------TT--GT---TGGTA-----A-A------T----GA--AT-     346
301 CTGGGACCCCTTCCGCGACTGGTACCCGCATAGCCGCCTCTTCGACCAGG     350

347 --TTC---CTG-------T-C---AG-A-TA-TC-CAGTTGCTT-GG-GA     396
351 CCTTCGGGCTGCCCCGGCTGCCGGAGGAGTGGTCGCAGTGG-TTAGGCG-     400

397 GAA-C-GCTG-CCT---T---TGCGCCTCCTGC-----G--G--A-CG-G     446
401 GCAGCAGCTGGCCAGGCTACGTGCGCCCCCTGCCCCCCGCCGCCATCGAG     450

447 --CTCA--A-T--CTG-GCT----TA-AG--------T-A-C-G--A-CT     496
451 AGCCCCGCAGTGGCCGCGCCCGCCTACAGCCGCGCGCTCAGCCGGCAACT     500

497 -A-CA-C----------AGA------A----G--------T------TGA     546
501 CAGCAGCGGGGTCTCGGAGATCCGGCACACTGCGGACCGCTGGCGCGTGT     550

547 CAC---A-GT-AAC-A-TACG----GG------T-----T-AAG-C-A-G     596
551 CCCTGGATGTCAACCACTTCGCCCCGGACGAGCTGACGGTCAAGACCAAG     600

597 GAAG---TGCA--A-AT-AC---T-A-C------G-T--TG--GC--AT-     646
601 GATGGCGTGGTGGAGATCACCGGTGAGCCCCCCTGCTCCTGCAGGGGAGA     650
```





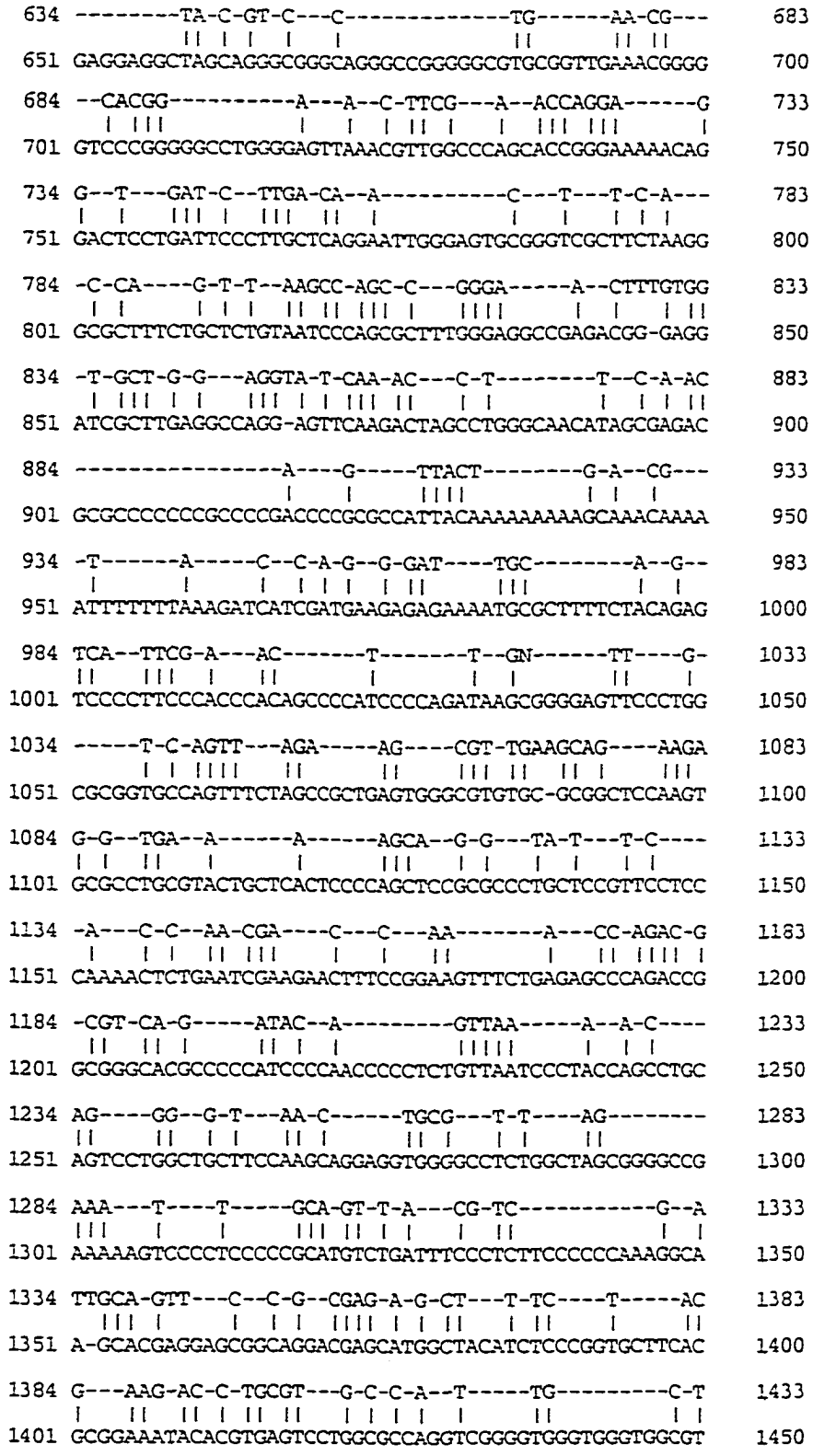


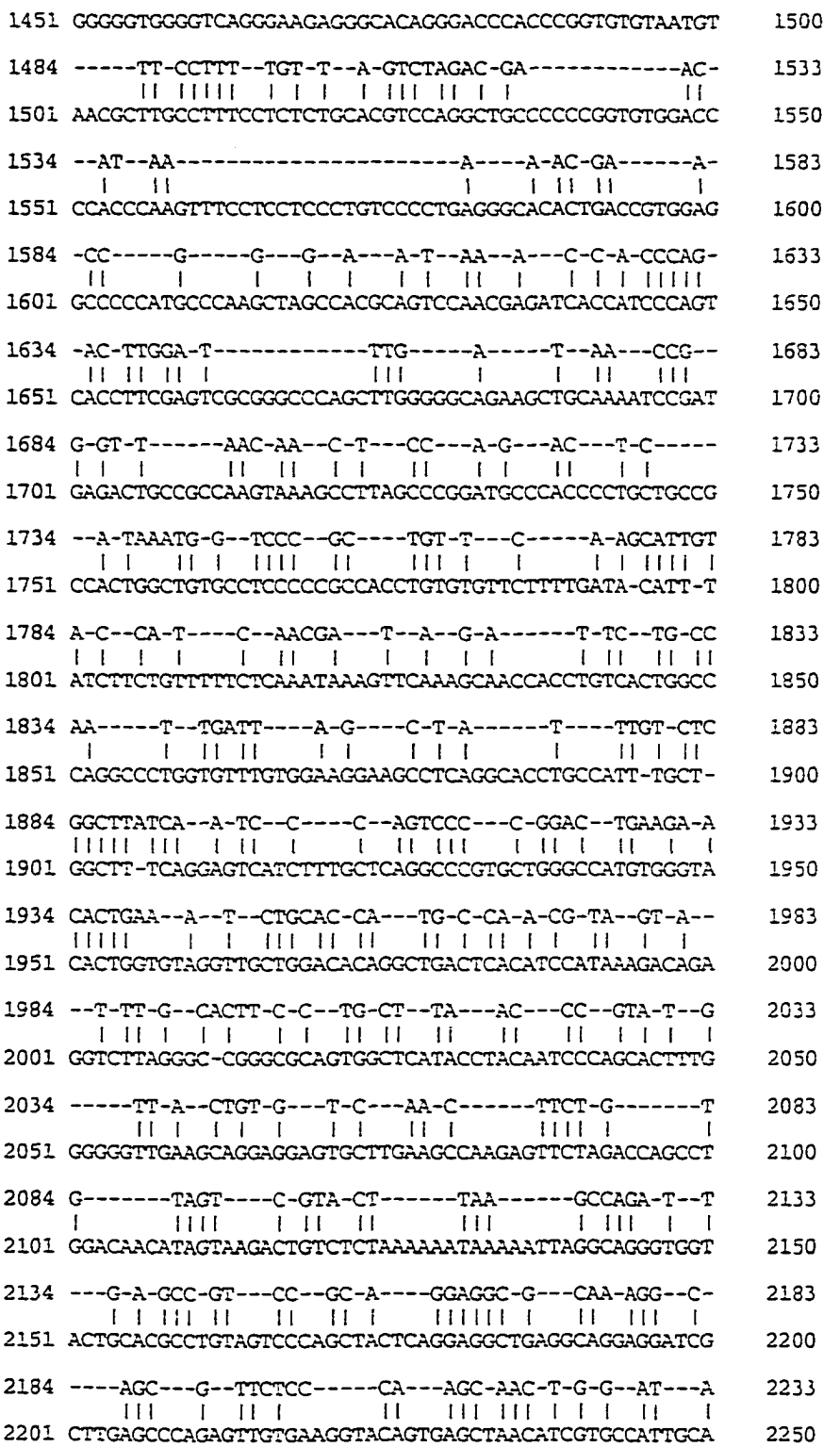
Alignment Of DNA 4 (Complementary Strand), And The Human hsp 27 Gene

(The top row shows the DNA 4 nucleotide sequence).

(The bottom row shows the human hsp 27 gene sequence).

```
-16 .................C--A-CGA-AT--GGATACC-GC-GATGGAT-A      33
                      |  | ||| |   || | || |  || || | |
  1 GAATTCATTTGCTTTTCCTTAACGAGAGAAGGTT-CCAGATGAGGGCTGA      50

 34 -CCGT-TTCG--------ATTTCGT----T-AACACC-----A--A-TGG      83
     || | ||||         |   ||     | ||| |       |  | ||
 51 ACCCTCTTCGCCCCGCCCA---CGGCCCCTGAACGCTGGGGGAGGAGTGC     100

 84 -T----AGGGGTG----TC----G--T--T-G--A--AA--GCGA--T-A     133
     |    ||||| |    ||    |  |  | |  |  ||  | ||  | |
101 ATGGGGAGGGGCGGCCCTCAAACGGGTCATTGCCATTAATAGAGACCTCA     150

134 ------G--T---A------CCGGACTG------CA-A---GCGCATTTC     183
          |  |   |      || |||||      || |   |||||   |
151 AACACCGCCTGCTAAAAATACCCGACTGGAGGAGCATAAAAGCGCAGC-C     200

184 G--CT-AGC---------T---C-GAG--GTACGAACAGA-C-GAATC-G     233
    |  |  |||         |   | |||  | |||  |||| | || || |
201 GAGCCCAGCGCCCCGCACTTTTCTGAGCAG-ACGTCCAGAGCAGAGTCAG     250

234 CAA-CA--ACCT-G-GCCA-GAC---T-CATCG-T-------------AG     283
    | | ||  |||  | |||  | |   | | ||| |             ||
251 CCAGCATGACCGAGCGCCGCGTCCCCTTC-TCGCTCCTGCGGGGCCCCAG     300

284 TT---AC-------GC-A--GA-ACCGG-ATA------T-----AC-AGG     333
     |   ||       || |  |  ||| | |||      |     || |||
301 CTGGGACCCCTTCCGCGACTGGTACCCGCATAGCCGCCTCTTCGACCAGG     350

334 G-TT-GG--TG-------TAC--------T---CG---TG---AG-CGA-     383
      || ||  ||       | |        |   ||   ||   || ||
351 CCTTCGGGCTGCCCCGGCTGCCGGAGGAGTGGTCGCAGTGGTTAGGCGGC     400

384 AG--G-TGA--AGG-T--GTTAG-----T--------------T--A---     433
    ||  | ||   ||| |  ||  |     |              |  |
401 AGCAGCTGGCCAGGCTACGTGCGCCCCCTGCCCCCCGCCGCCATCGAGAG     450

434 CC--GCAGATT------------TATAG--G-----TC-G--G--AA-TC     483
    ||  |||| |             || ||  |     || |  |  || ||
451 CCCCGCAG-TGGCCGCGCCCGCCTACAGCCGCGCGCTCAGCCGGCAACTC     500

484 TGCTG-----T-TC--A---C-GG-----T-C--A----T--CG-G----     533
     || |     | ||  |   | ||     | |  |    |  || |
501 AGCAGCGGGGTCTCGGAGATCCGGCACACTGCGGACCGCTGGCGCGTGTC     550

534 -----A---CAAGCACT--G----G-A--AG-T-----T-AA--CGAAT-     583
         |   ||| ||||  |    | |  || |     | ||  | ||
551 CCTGGATGTCAACCACTTCGCCCCGGACGAGCTGACGGTCAAGACCAAGG     600

584 -TG-C-TG-T-----TCGCC----A-CC-----G-TC----AGT--AGA-     633
     || | || |     || ||    | ||     | ||    ||   |||
601 ATGGCGTGGTGGAGATCACCGGTGAGCCCCCCTGCTCCTGCAGGGGAGAG     650
```





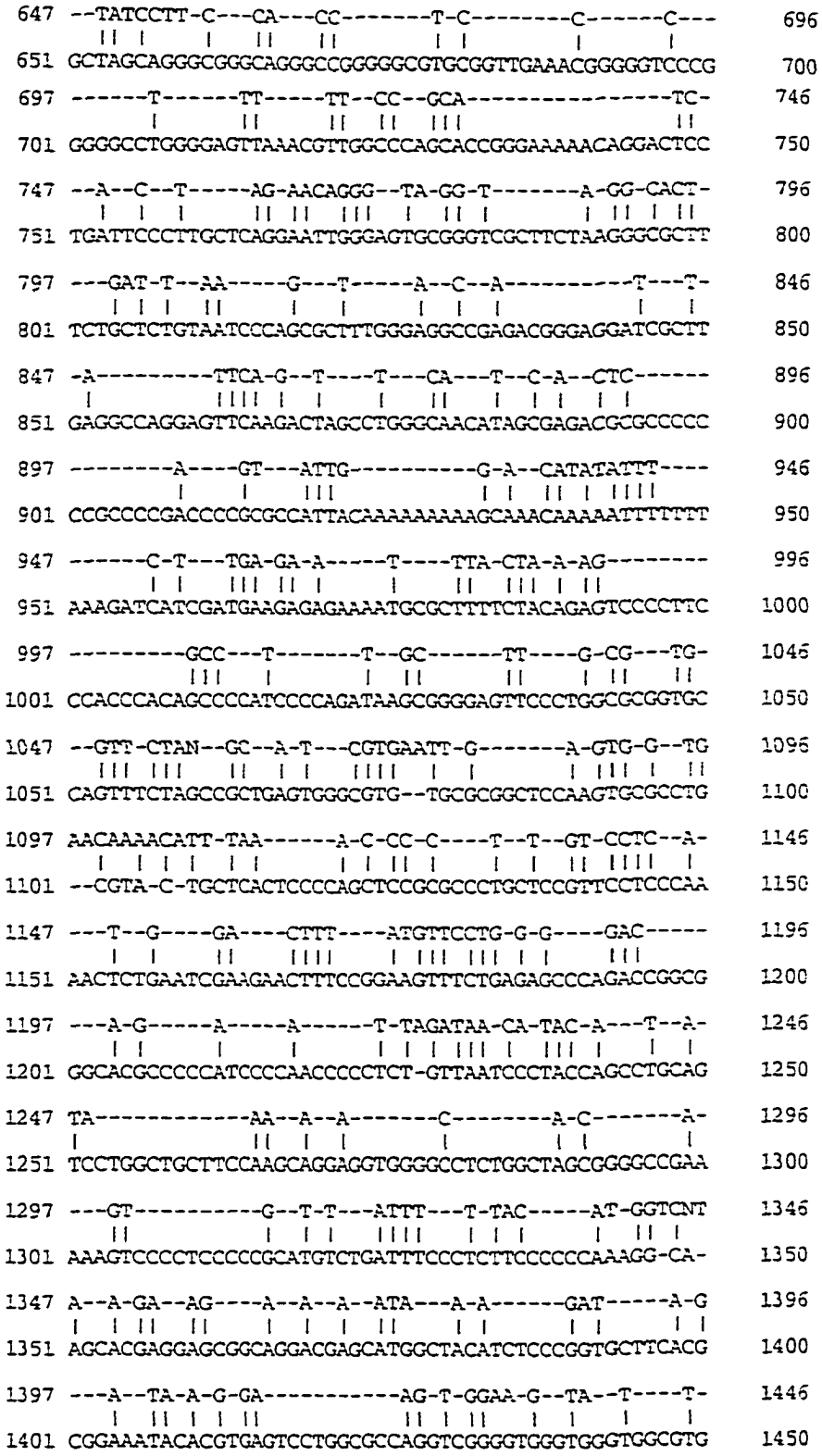
```
2234 -TC-----TG---A-CAGGA-AAT-TCAT-T-T-AC--CAAC-AAAT-C-    2283
      ||      ||    | ||| | ||  || | | | |   |||| |||  |
2251 CTCCAGCCTGGGCAACAGAACAAGATCCTGTCTCAAAACAACCAAAAGCC    2300

2284 --GTTGACGAATT--T----CCT--TGCTGCTAAT-GATAAT--------    2333
       |  || |||    |    ||   | ||  |||  || ||
2301 CAGA-GA-GAAAGAGTGAGACCCCAT-CTT-TAAAAGAAAAAAAAAAAAG    2350

2334 -TCAT-ATTGCACGG-CGA-GCT--CN-TCTA---CTGCTTCC-TGG--A    2383
      |||| |||||| || | | | |  | | ||   ||| |   |||  |
2351 GTCATGATTGCAAGGTC-ACGATTGCAAT-TAAAACTG-TAAGGTGGGGA    2400

2384 A--AT-AC-A--TA-------C-CC--AGCCTTGC-TTG-CACCA-CACC    2433
     |  |  |  |  ||       | ||  || |||| || || | ||||
2401 AGGAGGAGGAAATAAGAGAAGCACCTGAGGCTTGAGTTCTCAGGAGCACC    2450

2434 TA----------A----AACG--C-CAGCTGG-A---GC-C-T--G-GC    2483
     ||          |   || |  | |||  || |   || | |  | ||
2451 TAGGTTGGGTCCCAGGTGAAGGGGCACAGA-GGTAATTGCACCTCAGAGC    2500

2484 T--TGA-AGCTCCGTATT-CAAGGTCGA.....................    2533
     |  ||  ||    | ||| | | ||| |
2501 TGATGGGAG----G-ATTACTATGTC-A.....................    2550
```

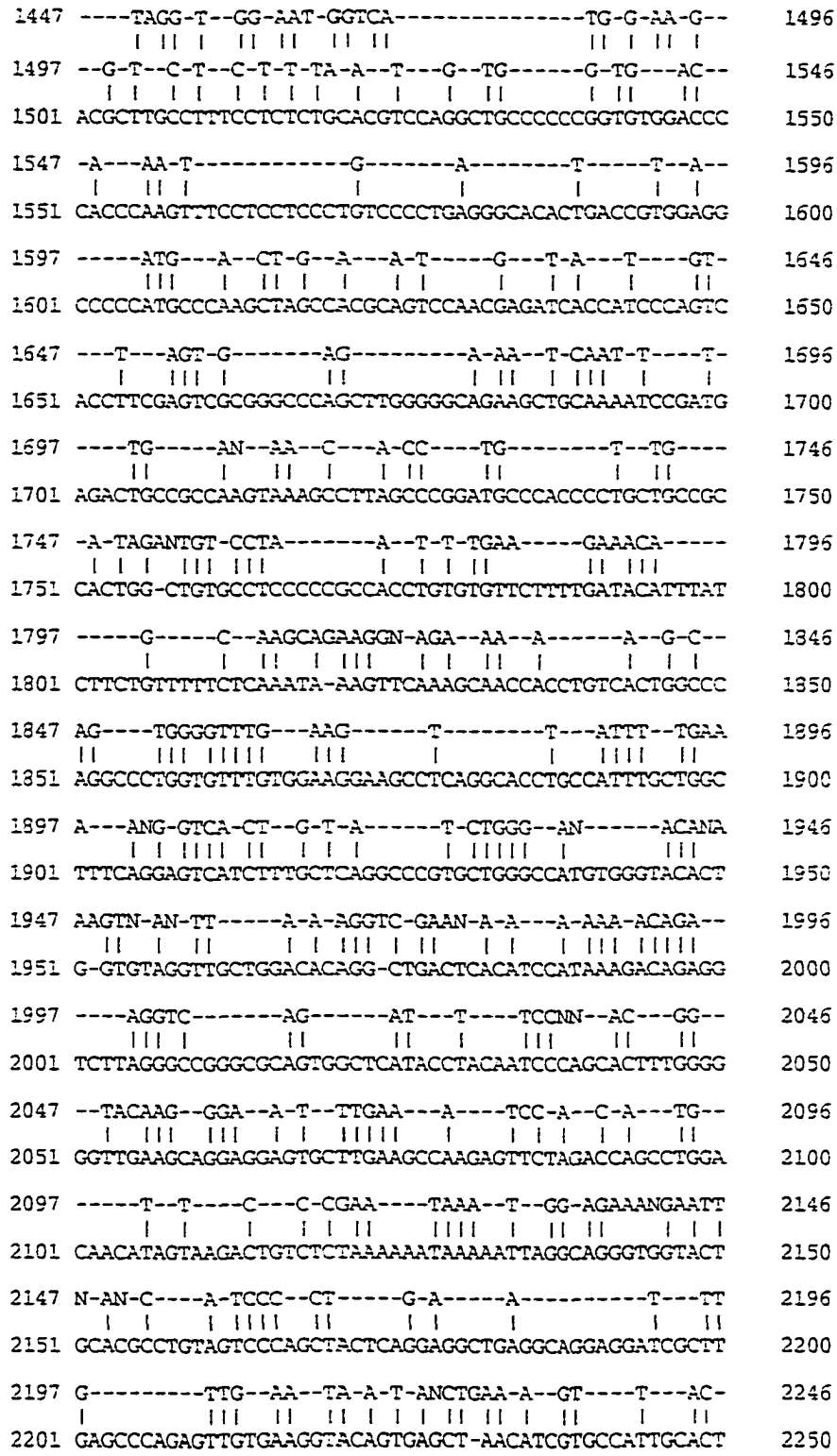
Alignment Of DNA 1/Region 1 (Normal Strand), And The Human hsp 27 Gene

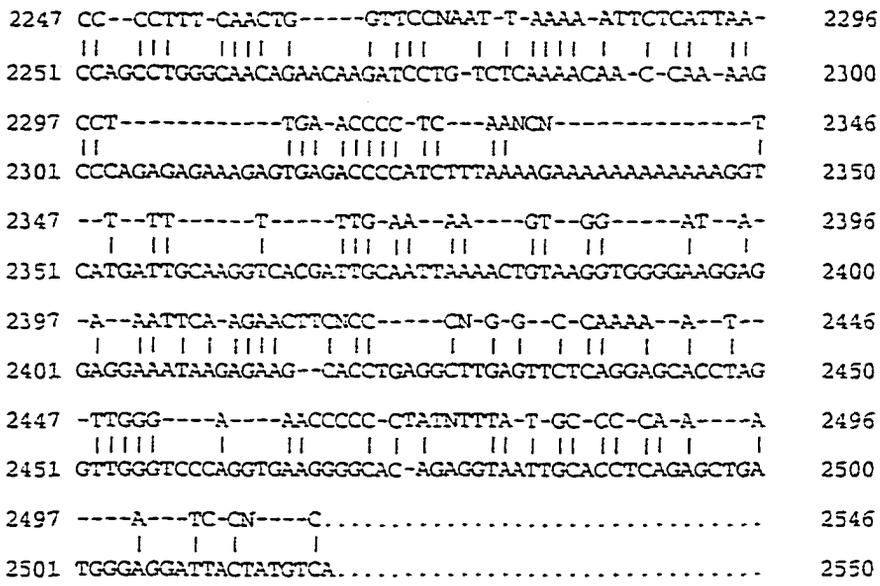
(The top row shows the DNA 1/Region 1 nucleotide sequence).

(The bottom row shows the human hsp 27 gene sequence).

```
 -3 ....T-ATCTGGT-------A-C---------T-C----T-A----T-A-      46
        |  ||  ||  |        |  |          |  |    |  |     |  |
  1 GAATTCATTTGCTTTTCCTTAACGAGAGAAGGTTCCAGATGAGGGCTGAA      50

 47 CC-TCTACAC-------AC--CTTTTAAA-G-TG----A--A-T---TG-      96
    ||  |||  |  |        ||   |    |  ||  |  ||     |   |  |    ||
 51 CCCTCTTCGCCCCGCCCACGGCCCCTGAACGCTGGGGGAGGAGTGCATGG     100

 97 --A----C--C--T-A------T--TTGCTTTTTCTAGA-ATTTTTAAA-     146
      |    |  |  | |       |  ||||  ||  |||| |  |  |||
101 GGAGGGGCGGCCCTCAAACGGGTCATTGCCATTAATAGAGACCTC-AAAC     150

147 ATTTT-TG-T-----T-CC--ACCGT-------TAAA--C-C--CC-AGT     196
    |     ||  |     |  ||  ||  |         ||||  |  |  ||  ||
151 ACCGCCTGCTAAAAATACCCGACTGGAGGAGCATAAAAGCGCAGCCGAGC     200

197 T--G--------A-TT--CT---C-GAC-TC-A-A-CTGAA--A---A-C     246
       |        | ||  ||    | ||| || | | | ||    |   | |
201 CCAGCGCCCCGCACTTTTCTGAGCAGACGTCCAGAGCAGAGTCAGCCAGC     250

247 -T-AT-----------TC---TTCT---T--T-CGGA-----A------A     296
     | |             ||   ||||   |  | |||       |      |
251 ATGACCGAGCGCCGCGTCCCCTTCTCGCTCCTGCGGGGCCCCAGCTGGGA     300

297 CC--TTC----ATTG--AC-----TA------T-T-CAA--AGTTCTTCT     346
    ||  |||    | ||  ||      ||       | | | |  ||  ||||
301 CCCCTTCCGCGACTGGTACCCGCATAGCCGCCTCTTCGACCAGGCCTTCG     350

347 ---T--------T-----A----T--TC----T--TTAGT-G-C--CA--     396
       |        |     |    |  ||    |  ||||  | |  ||
351 GGCTGCCCCGGCTGCCGGAGGAGTGGTCGCAGTGGTTAGGCGGCAGCAGC     400

397 T--CC----T-C-TG-------T-------------AT--ATTA-CCCTT     446
    |  ||    | | ||       |             ||  |  | |||
401 TGGCCAGGCTACGTGCGCCCCCTGCCCCCCGCCGCCATCGAG-AGCCCCG     450

447 -A-T-------------TAC--CCT-----TCA------AACT-A-C--C     496
     | |             |||  ||      |||      |||| | |  |
451 CAGTGGCCGCGCCCGCCTACAGCCGCGCGCTCAGCCGGCAACTCAGCAGC     500

497 TA--T-T-GTA-A-CT---ACA-T----ATTGTT------T-T---T--A     546
        | | | | | |     ||| |    |  | |      | |   |  |
501 GGGGTCTCGGAGATCCGGCACACTGCGGACCGCTGGCGCGTGTCCCTGGA     550

547 TGTC--C-A-T-CTCCCCT-A--A--T-A-----AA--C-A-G--TG---     596
    ||||  | | | | ||||  |  |  | |     ||  | | |  ||
551 TGTCAACCACTTCGCCCCGGACGAGCTGACGGTCAAGACCAAGGATGGCG     600

597 ------AG--C-----T---CC----T--TCA---AG---A-A--AG-AG     646
          ||  |     |   ||    |  ||    ||   | |  || ||
601 TGGTGGAGATCACCGGTGAGCCCCCCTGCTCCTGCAGGGGAGAGGAGGAG     650
```







Alignment Of DNA 1/Region 1

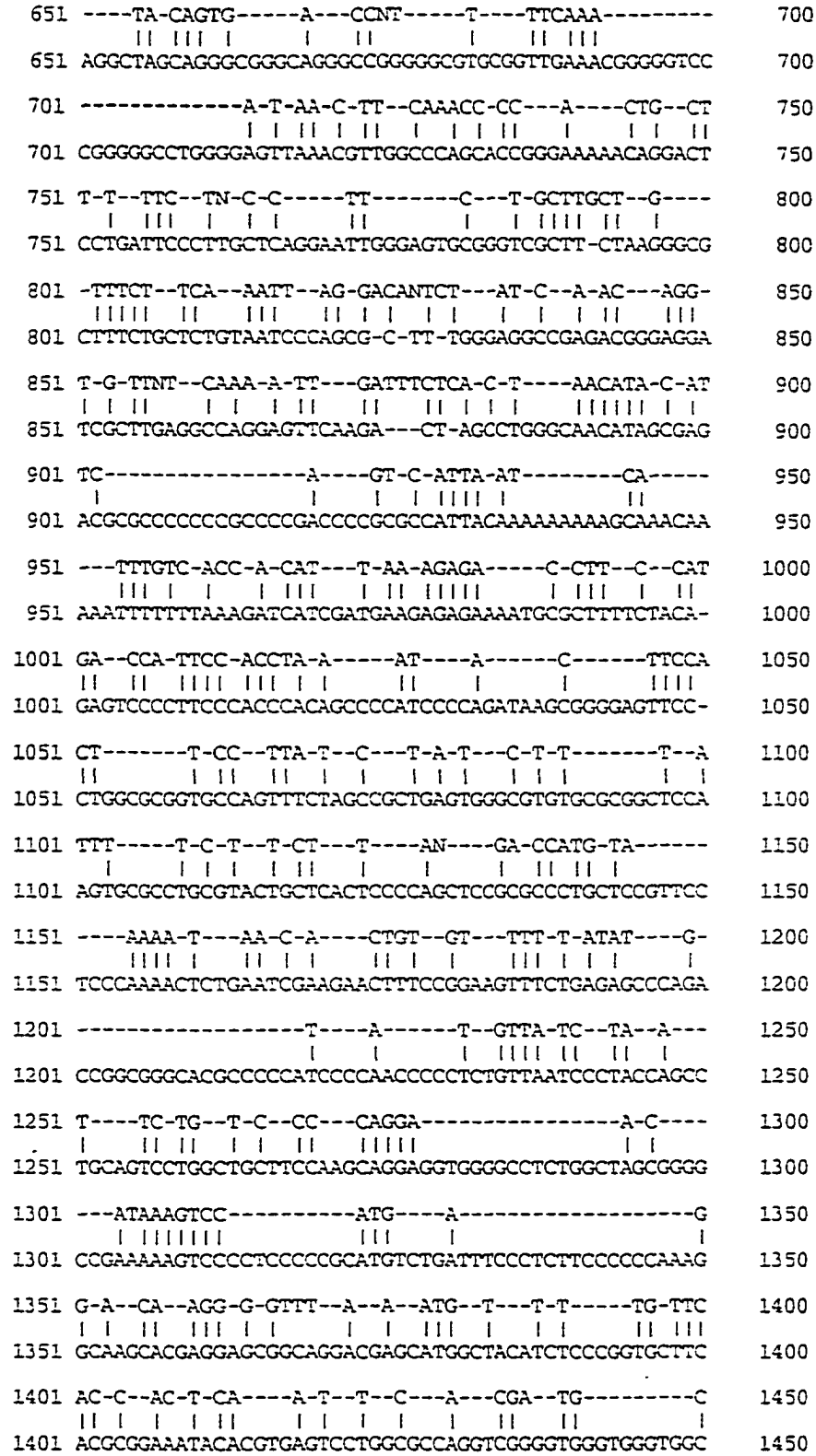
(Complementary Strand), And

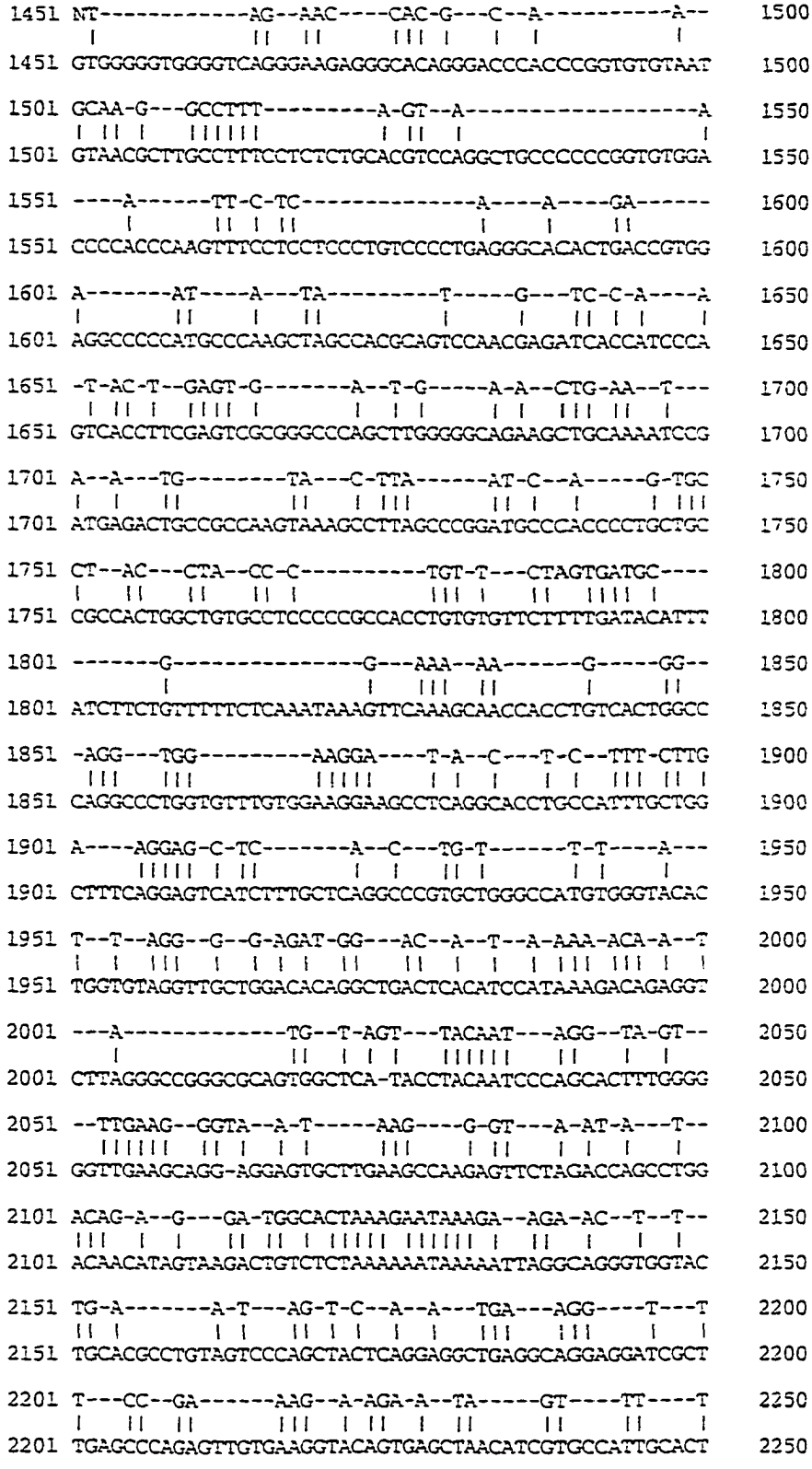
The Human hsp 27 Gene

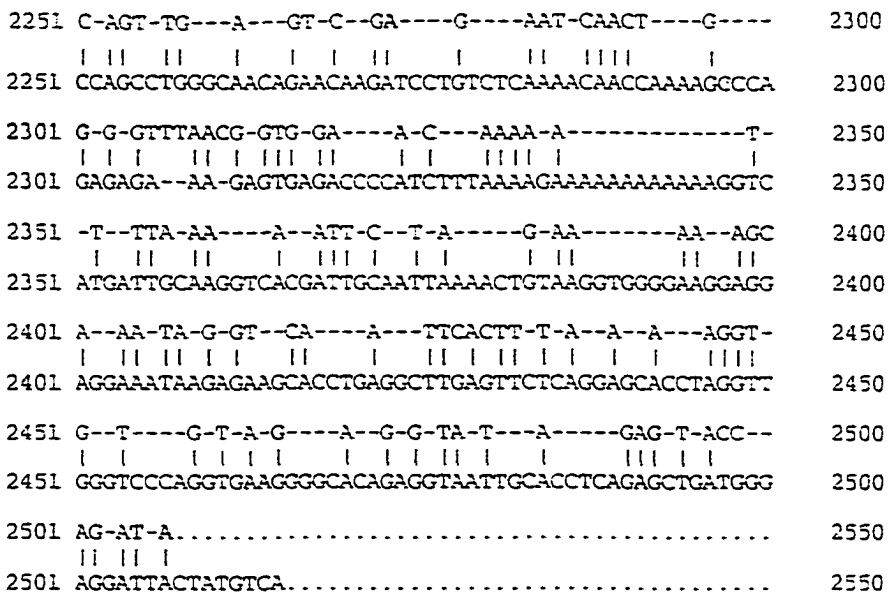
(The top row shows the DNA 1/Region 1 nucleotide sequence).

(The bottom row shows the human hsp 27 gene sequence).

```
 51 -----T-TTC-CC------A---------AATT-T----------T---T     100
       | ||| ||       |          ||   |          |   |
 51 AACCCTCTTCGCCCCGCCCACGGCCCTGAACGCTGGGGGAGGAGTGCAT     100

101 GGCCNGGGN-G------AA--GT-TC-TTGA-ATTT-TA------TC---     150
    ||   |||  |       ||  |  || |||  |||  ||       ||
101 GGGGAGGGGCGGCCCTCAAACGGGTCATTGCCATTAATAGAGACCTCAAA     150

151 CACTTT-T-CAAAAAA-AN--GNTTG-AGGGGT-TCAA-G-GTTAAT-GA     200
    |||    | | ||||| |   |  || ||| |  | || | |  |   ||
151 CACCGCCTGCTAAAAATACCCGACTGGAGGAGCATAAAAGCGC-AGCCGA     200

201 G---A---------A-TTTT-T-A--AT---TN--G-G-A-A--C--C-A     250
    |   |         | |||| | |  |    |   | | | |  |  | |
201 GCCCAGCGCCCCGCACTTTTCTGAGCAGACGTCCAGAGCAGAGTCAGCCA     250

251 GT-TGAAAG-G-G--G-GTAAC-TT-TC---------------AGNT--     300
    |  |||  | | |  | ||  | || ||               || |
251 GCATGACCGAGCGCCGCGTCCCCTTCTCGCTCCTGCGGGGCCCCAGCTGG     300

301 -A----TT-----A-T--T-CAA-CAAA------T-T-C-A---GG----     350
     |   ||     | |  | |  || |       | | | |   ||
301 GACCCCTTCCGCGACTGGTACCCGCATAGCCGCCTCTTCGACCAGGCCTT     350

351 -GGA-TGN------TN----A--A-T--TCN---T--TT---CT-C--CA     400
     ||  ||       |     |  | |  ||    |  ||   |  |  ||
351 CGGGCTGCCCCGGCTGCCGGAGGAGTGGTCGCAGTGGTTAGGCGGCAGCA     400

401 T-T--------TA--T--------T-C-----G--G-----GA-A-C---     450
      |        ||  |        | |     |  |     || | |
401 GCTGGCCAGGCTACGTGCGCCCCCTGCCCCCCGCCGCCATCGAGAGCCCC     450

451 --A-TGT--G-GAT----T-----------TCA------AATTC--C--     500
      | ||   | |      |            |||      || ||  |
451 GCAGTGGCCGCGCCCGCCTACAGCCGCGCGCTCAGCCGGCAACTCAGCAG     500

501 CTTG-TAC-CG----TNNGGAA-A-T-CTGACCT-T--CT-GTTTT--T-     550
    |  | | | ||    |  || | | | | ||||  |  |  || |   |
501 CGGGGT-CTCGGAGATCCGGCACACTGCGGACCGCTGGCGCGTGTCCCTG     550

551 --T-TN------TTCG------AC---CTT----T-AANTN-A-------     600
      | |       ||||      ||   ||     | ||    |
551 GATGTCAACCACTTCGCCCCGGACGAGCTGACGGTCAAGACCAAGGATGG     600

601 C-T--T-----TN-----TG--------TN-TCC--CAG---A-------     650
    | |  |     |      ||        |  |||  |||   |
601 CGTGGTGGAGATCACCGGTGAGCCCCCCTGCTCCTGCAGGGGAGAGGAGG     650
```





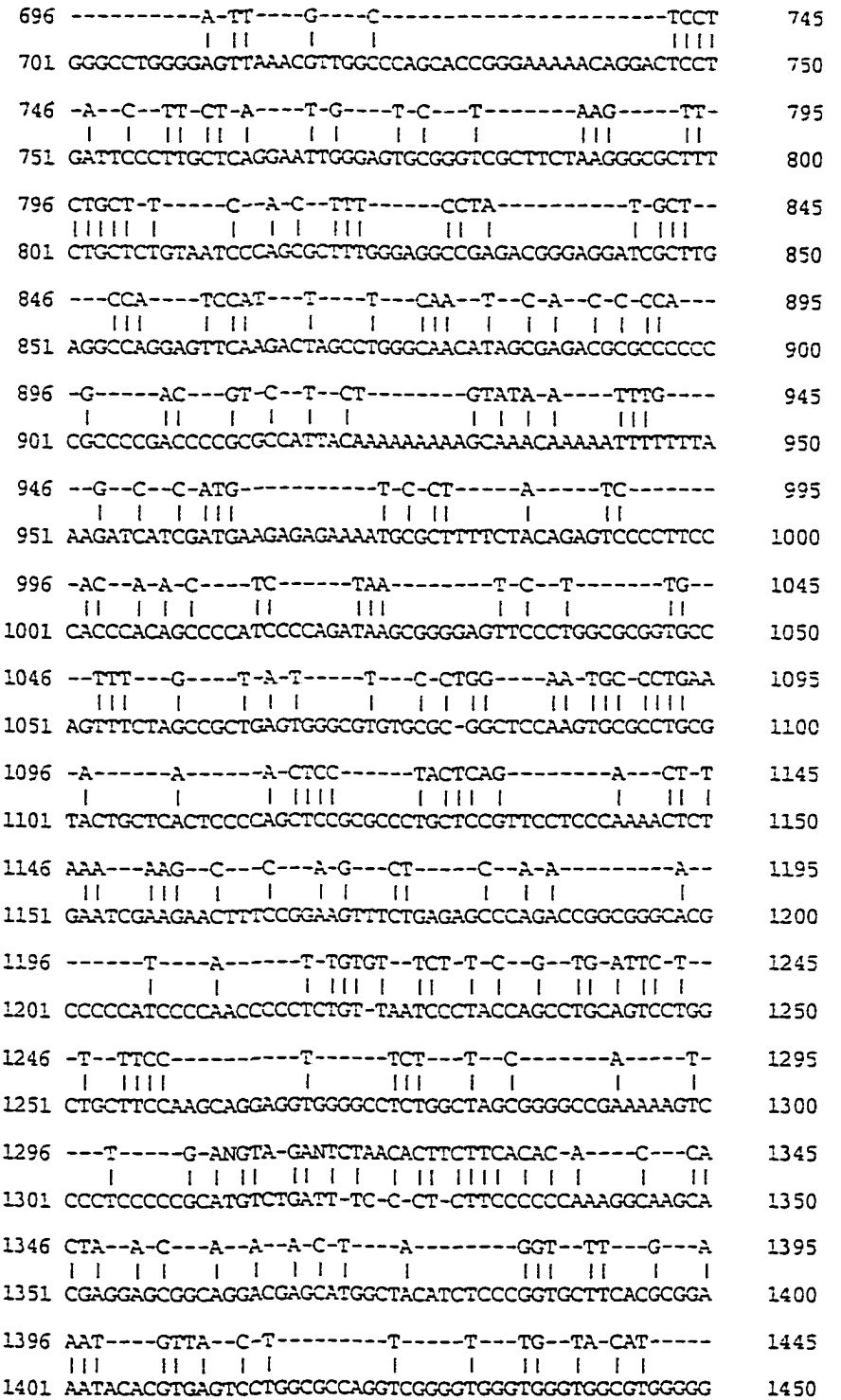


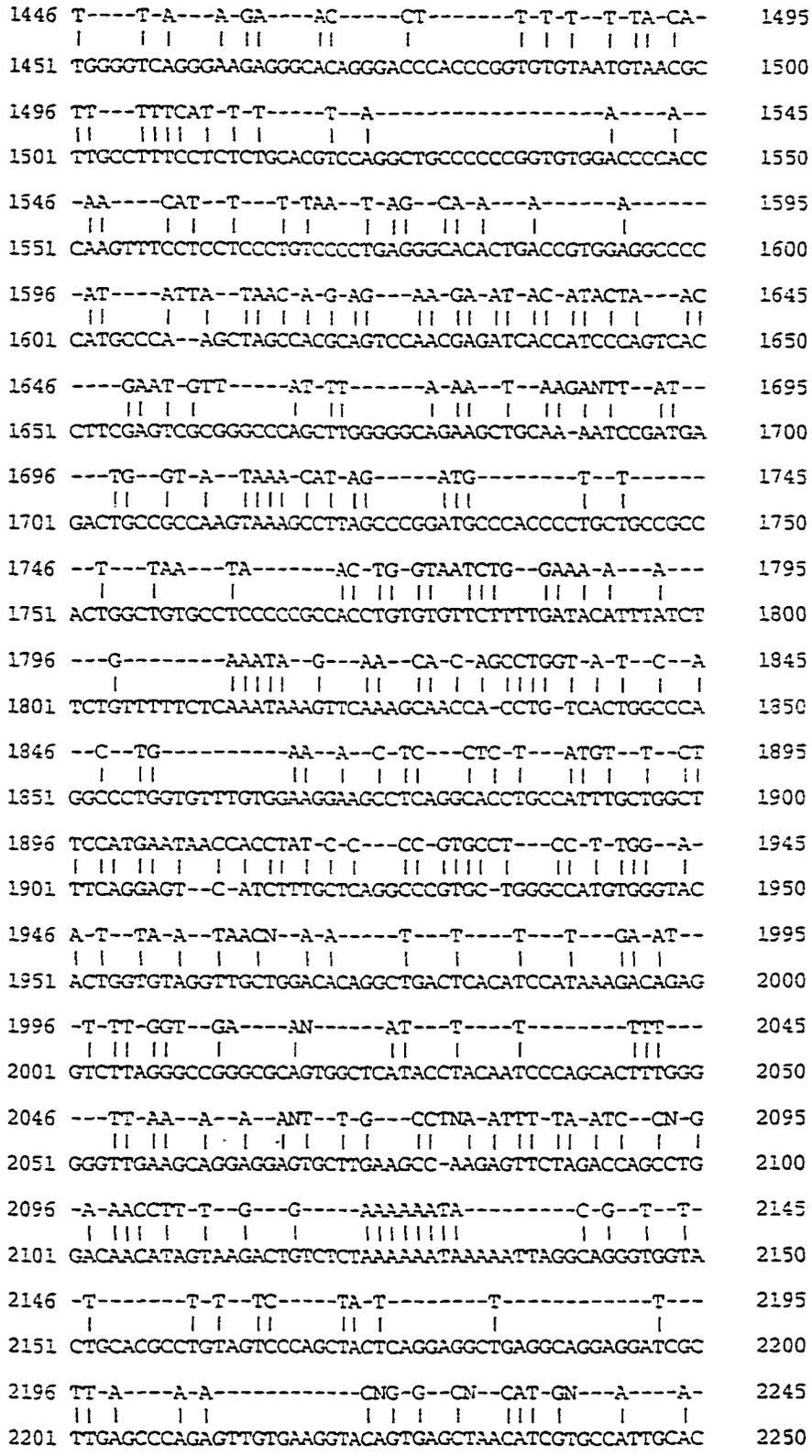
Alignment Of DNA 1/Region 2 (Normal Strand), And The Human hsp 27 Gene

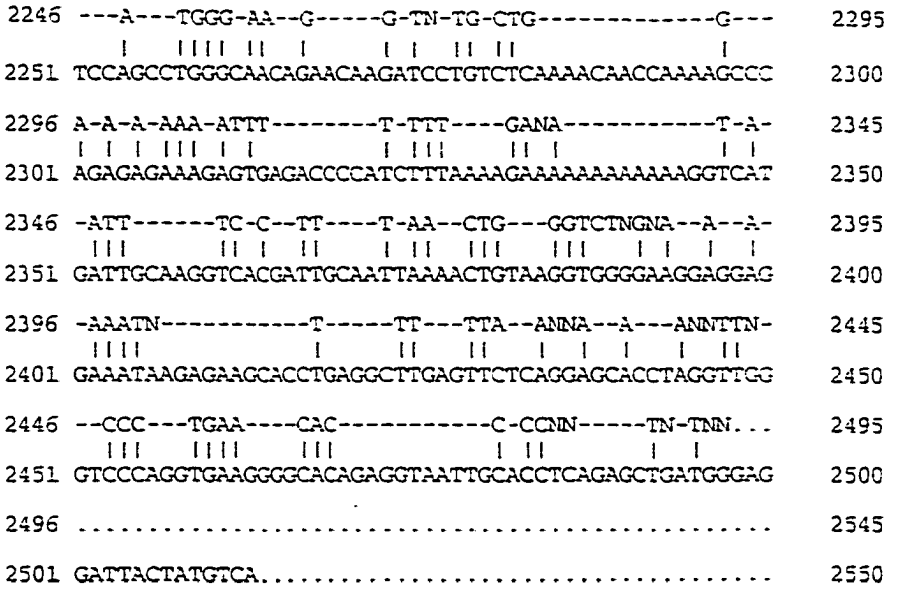
(The top row shows the DNA 1/Region 2 nucleotide sequence).

(The bottom row shows the human hsp 27 gene sequence).

```
 -4 .....C-T--GC--------A--G-G------T-C-A-ACG-GGA-T---      45
 1 GAATTCATTTGCTTTTCCTTAACGAGAGAAGGTTCCAGATGAGGGCTGAA      50

 46 C--T-TT-G----GC--A-------TGA----T-----ATT--T--AT--      95
 51 CCCTCTTCGCCCCGCCCACGGCCCCTGAACGCTGGGGGAGGAGTGCATGG     100

 96 --A----CA-C--T-----G--TC-T--C-ATTT-T-----CTTC-----     145
101 GGAGGGGCGGCCCTCAAACGGGTCATTGCCATTAATAGAGACCTCAAACA     150

146 --GTTTG-TATA--T-------TGG-GGA---T----GTTTAT---A---     195
151 CCGCCTGCTAAAAATACCCGACTGGAGGAGCATAAAAGCGCAGCCGAGCC     200

196 -ATT-------A-TTTT--G--CAG--G---A--------T---------     245
201 CAGCGCCCCGCACTTTTCTGAGCAGACGTCCAGAGCAGAGTCAGCCAGCA     250

246 T-A----------T----TT-T-GA--------------A--TG----     295
251 TGACCGAGCGCCGCGTCCCCTTCTCGCTCCTGCGGGGCCCCAGCTGGGAC     300

296 ---T---G--A--G--A-----A-A------T-----A--A----TATCT     345
301 CCCTTCCGCGACTGGTACCCGCATAGCCGCCTCTTCGACCAGGCCT-TCG     350

346 ---T--------T-----A--ATT--T-G-A-T--TT-------------     395
351 GGCTGCCCCGGCTGCCGGAGGAGTGGTCGCAGTGGTTAGGCGGCAGCAGC     400

396 T--C-----T---T--------T--------------T-G----------     445
401 TGGCCAGGCTACGTGCGCCCCCTGCCCCCCGCCGCCATCGAGAGCCCCGC     450

446 --T----------TA--GC-------T-A------AT-TAA--A---     495
451 AGTGGCCGCGCCCGCCTACAGCCGCGCGCTCAGCCGGCAACTCAGCAGCG     500

496 ---T---GG---T-----A-ACTA---AC---T------T-T---T--AT     545
501 GGGTCTCGGAGATCCGGCACACTGCGGACCGCTGGCGCGTGTCCCTGGAT     550

546 -TCA-----T---------A-----T-A-----A----------T----T     595
551 GTCAACCACTTCGCCCCGGACGAGCTGACGGTCAAGACCAAGGATGGCGT     600

596 -----AG-T-A----TG-G------T-CTCAT-CA---------------     645
601 GGTGGAGATCACCGGTGAGCCCCCCTGCTCCTGCAGGGGAGAGGAGGAGG     650

646 C-A-CA--------AGG-CC-----C-T----TTGCAAC-----------     695
651 CTAGCAGGGCGGGCAGGGCCGGGGGCGTGCGGTTGAAACGGGGGTCCCGG     700
```







Alignment Of DNA 1/Region 2

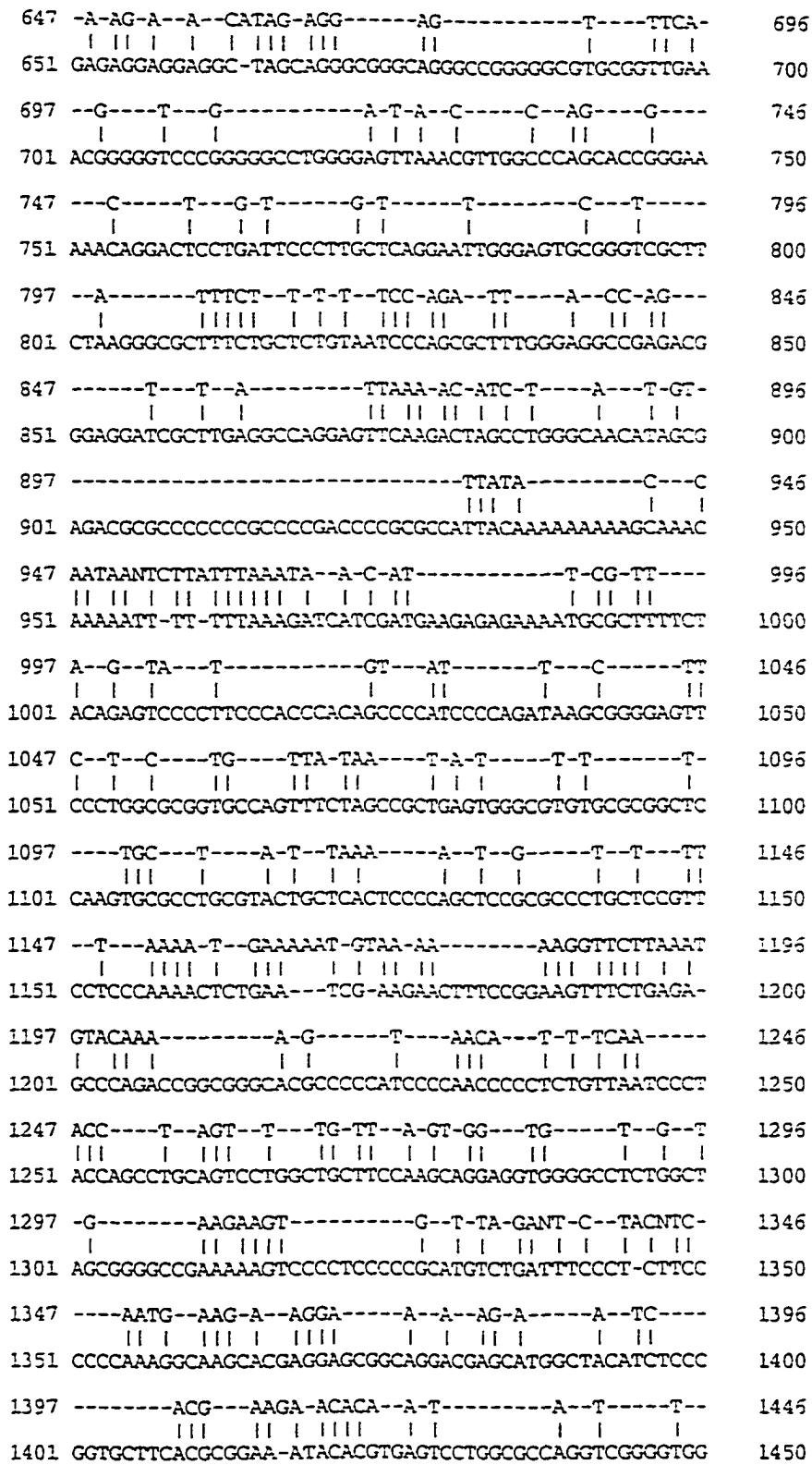
(Complementary Strand), And

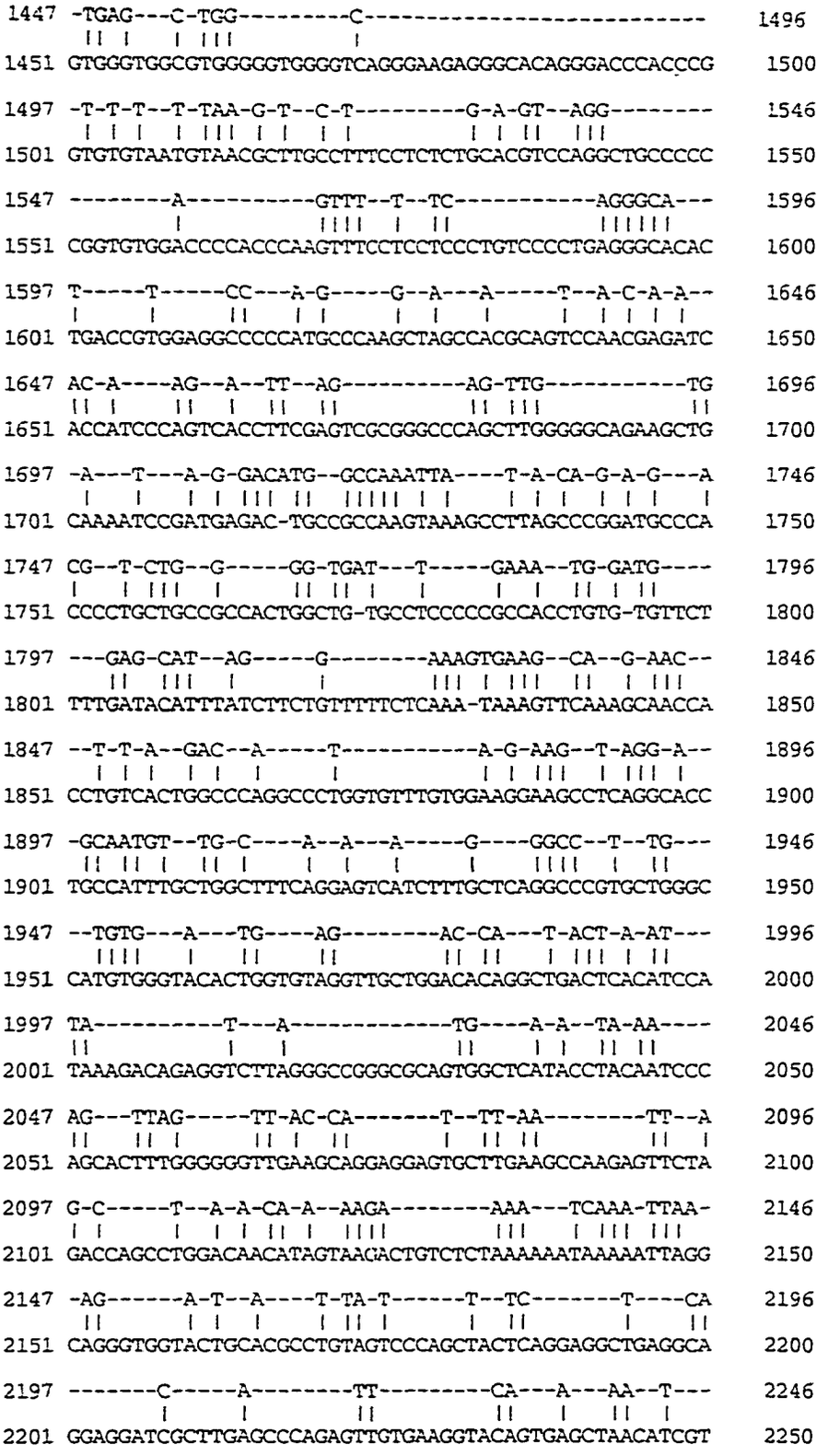
The Human hsp 27 Gene

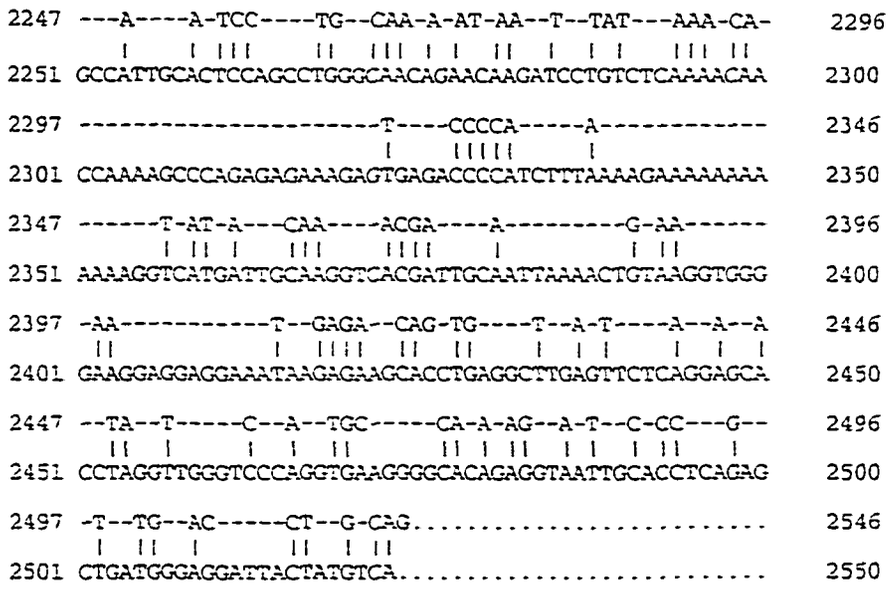
(The top row shows the DNA 1/Region 2 nucleotide sequence).

(The bottom row shows the human hsp 27 gene sequence).

```
 -3 ....NNAN------------ANNG-G-G--GTGTTC-AG--G-GN----A      46
          |               |  | | |  | |||| ||  | |     |
  1 GAATTCATTTGCTTTTCCTTAACGAGAGAAG-GTTCCAGATGAGGGCTGA      50

 47 ANN-T-TTNN----------------T------T-----A--A----A--      96
    |  | ||                  |      |     |  |    |
 51 ACCCTCTTCGCCCCGCCCACGGCCCCTGAACGCTGGGGGAGGAGTGCATG     100

 97 ---A------------ANA----T--TT----TT--TNCNAGACC-CAGT     146
       |            | |    |  ||    ||  |   |||||  ||
101 GGGAGGGGCGGCCCTCAAACGGGTCATTGCCATTAATA-GAGACCTCA--     150

147 TAA-A--G---G--AAAT-TATNT--C---A--A--A-AAAATTTTTTTTC     196
     || |  |   |  |||  ||      |   |  |  | ||||        |
151 -AACACCGCCTGCTAAAAATACCCGACTGGAGGAGCATAAAAG------C     200

197 -CAGC--AN---A-C-CTTCC-CA-TTTNCATGNGCCN--GTTTAAA--A     246
     ||||  |    | | |  || || ||| | || ||    ||  | |  |
201 GCAGCCGAGCCCAGCGCC-CCGCACTTTTC-TGAGCAGACGTCCAGAGCA     250

247 -A-T-AGAAA--A--A---A-CGTA---T----TT-T---T--T-C----     296
     | | ||  |  |  |   | ||     |    || |   |  | |
251 GAGTCAGCCAGCATGACCGAGCGCCGCGTCCCCTTCTCGCTCCTGCGGGG     300

297 CAA-AGGT--------TTCNG-GA-T--TAA----A-A------T-TN--     346
    |   || |        ||| | || |  ||     | |      | |
301 CCCCAGCTGGGACCCCTTCCGCGACTGGTACCCGCATAGCCGCCTCTTCG     350

347 A---GGC----------------------A--ANT--T-----T--TTA     396
    |   |||                    |  | |  |     |  |||
351 ACCAGGCCTTCGGGCTGCCCCGGCTGCCGGAGGAGTGGTCGCAGTGGTTA     400

397 ------A--A-------A----A-------------------------A-     446
          |  |       |    |                         |
401 GGCGGCAGCAGCTGGCCAGGCTACGTGCGCCCCCTGCCCCCCGCCGCCAT     450

447 --A-A---------TN-----------T-----------TCA-CC---     496
     | |          |           |           |||  ||
451 CGAGAGCCCCGCAGTGGCCGCGCCCGCCTACAGCCGCGCGCTCAGCCGGC     500

497 AA---A--A------T-TC--A-A------A-A-T---------TNG---     546
    ||   |  |      | ||  | |      | | |         | |
501 AACTCAGCAGCGGGGTCTCGGAGATCCGGCACACTGCGGACCGCTGGCGC     550

547 -T-T------AT-T-AA----TTC-C-----A--AG--GA-GG-CA---C     596
     | |      || | ||    ||| |     |  ||  || || ||   |
551 GTGTCCCTGGATGTCAACCACTTCGCCCCGGACGAGCTGACGGTCAAGAC     600

597 ---GG--GG--------A--T-A--GGTG-GT-----TA-T--T-CATGG     646
       ||  ||        |  | |  |||| |      |  |  | || ||
601 CAAGGATGGCGTGGTGGAGATCACCGGTGAGCCCCCCTGCTCCTGCAGGG     650
```





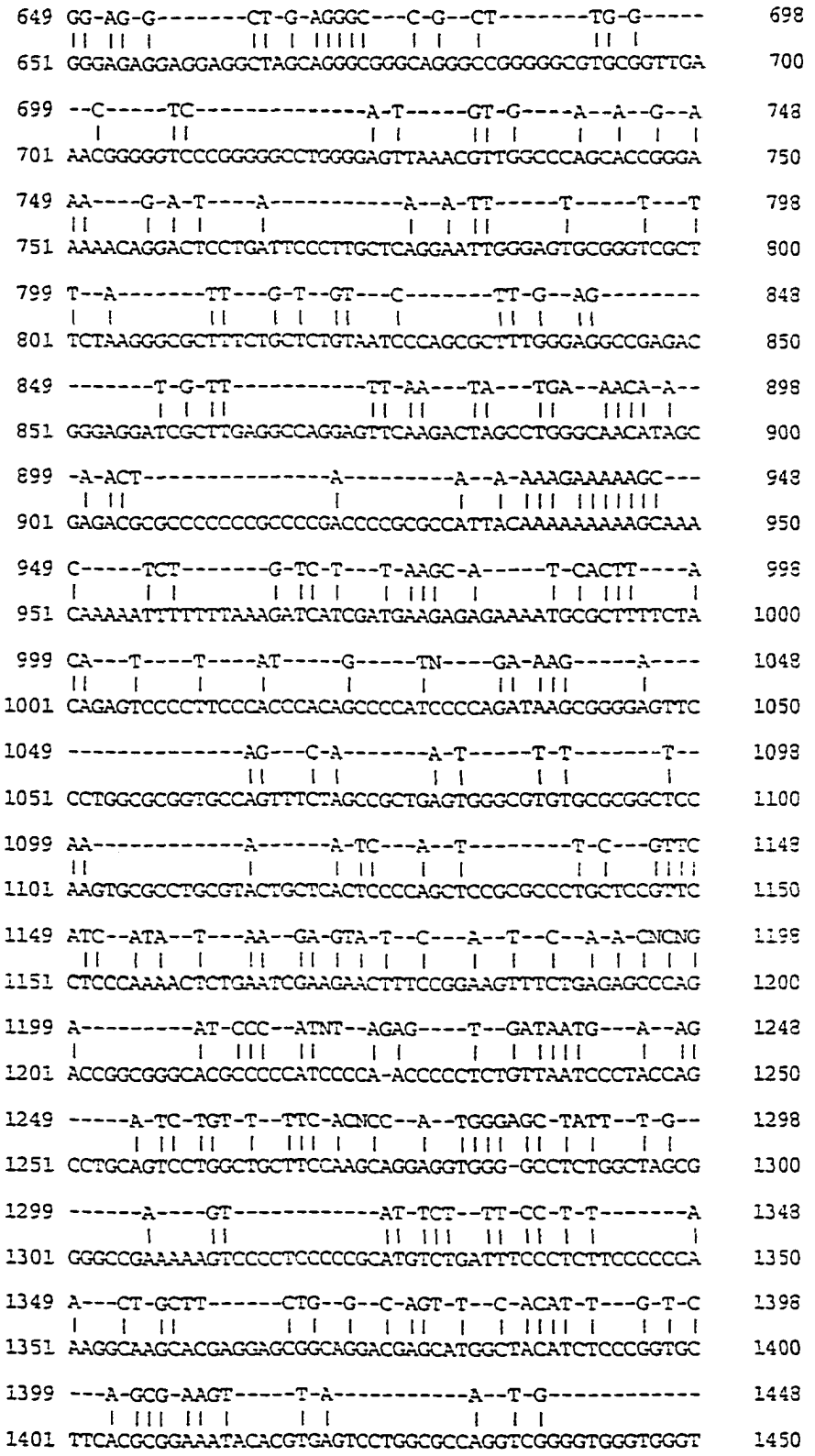


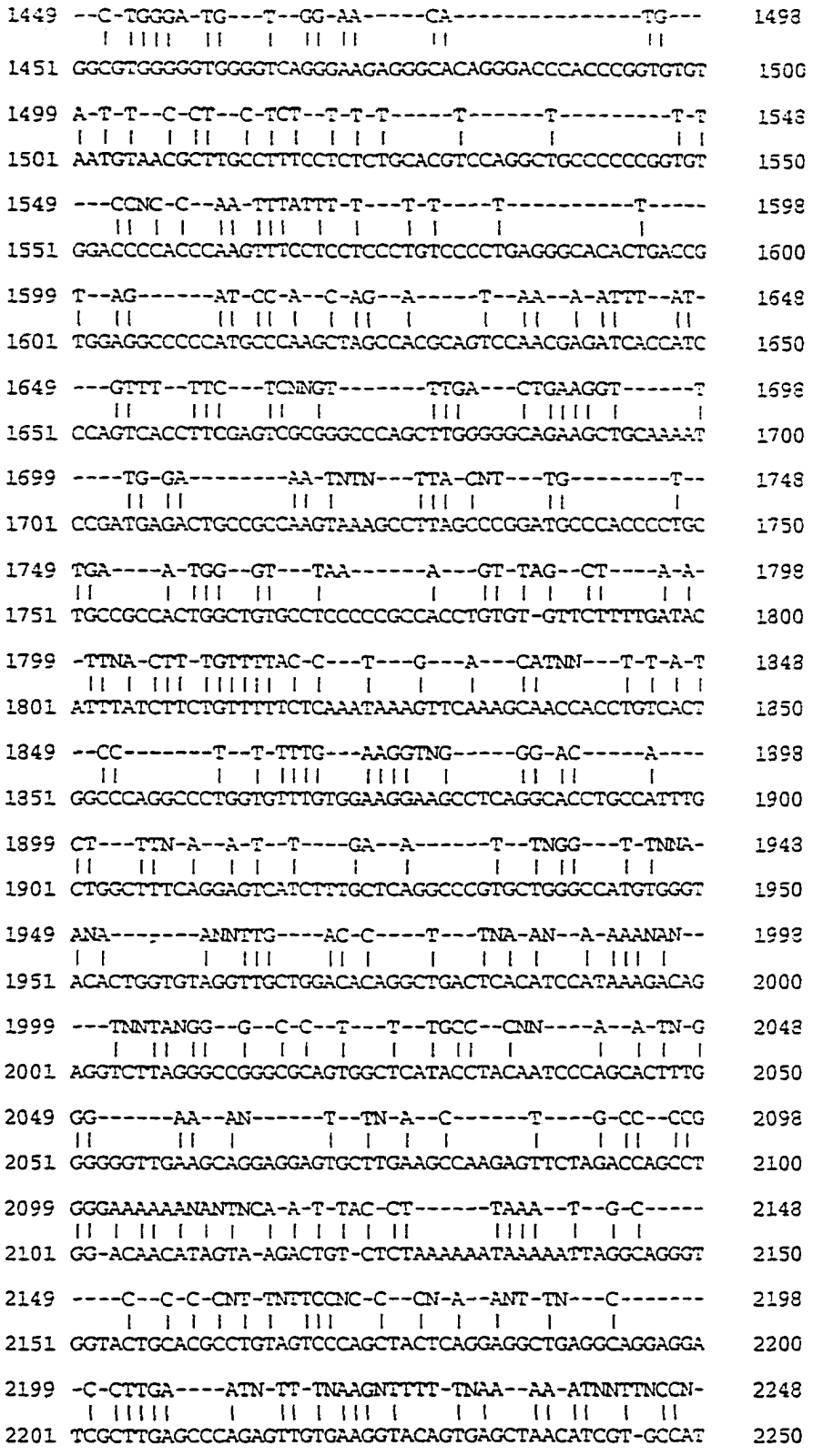
Alignment Of DNA 1/Region 3 (Normal Strand), And The Human hsp 27 Gene

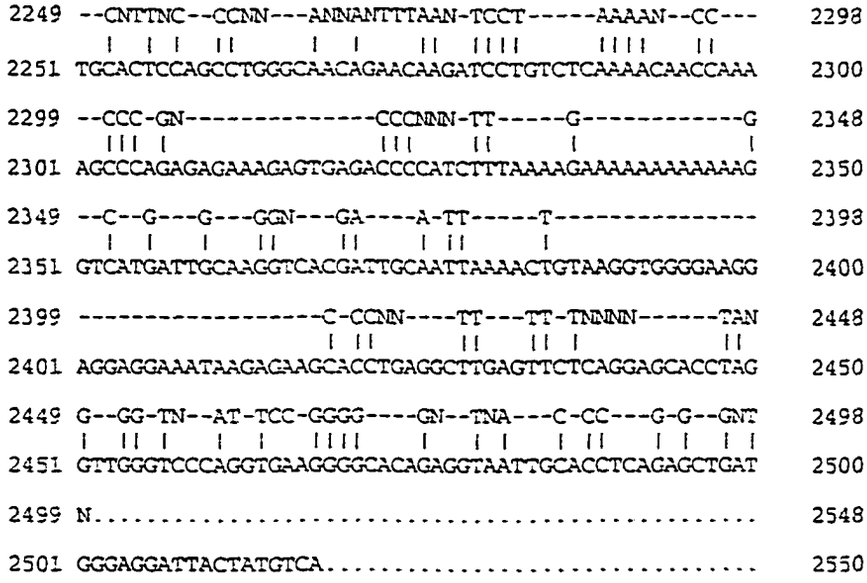
(The top row shows the DNA 1/Region 3 nucleotide sequence).

(The bottom row shows the human hsp 27 gene sequence).

```
-1 ..AT--A-------------AA--AGCA-AAG-TT-----T-----CT-A      48
     ||  |             ||  || | ||| ||     |     || |
 1 GAATTCATTTGCTTTTCCTTAACGAG-AGAAGGTTCCAGATGAGGGCTGA      50

49 ----T-TA----------A-G-----TGA-C-C------AGT-GT--ATT      98
       | |           | |     ||| | |      ||  ||  ||
51 ACCCTCTTCGCCCCGCCCACGGCCCCTGAACGCTGGGGGAGGAGTGCATG     100

99 --------------T-A------T-AT-GCGAT-AAGA-AGA--T-A--C     148
                 | |      | || || || || | |||  | |  |
101 GGGAGGGGCGGCCCTCAAACGGGTCATTGCCATTAATAGAGACCTCAAAC    150

149 TACT---T--T-----TA----A-TG-A--A--ATAAATA-CATA----A    198
     ||    |  |     ||    | || |  |  ||||| | |  |    |
151 -ACCGCCTGCTAAAAATACCCGACTGGAGGAGCATAAA-AGCGCAGCCGA    200

199 ----A-C-------A-TTTT--GA--A--CTGTCTACAGAA-A-A---A-    248
       | |       | ||||  ||  |  | |||  ||||  | |   |
201 GCCCAGCGCCCCGCACTTTTCTGAGCAGAC-GTC--CAGAGCAGAGTCAG    250

249 CTCAG-A-GA--GA--------T---------G-----G-----------    299
    | ||| | ||  ||        |         |     |
251 C-CAGCATGACCGAGCGCCGCGTCCCCTTCTCGCTCCTGCGGGGCCCCAG    300

299 -T---A----------A-TG--A-----A-A-----------AC-A--    348
     |   |          | ||  |     | |            || |
301 CTGGGACCCCTTCCGCGACTGGTACCCGCATAGCCGCCTCTTCGACCAGG    350

349 C-T-C----TG-------T----------T--T---A-TA-TAAAA-G--    398
    | | |    ||       |          |  |   | |  | |   |
351 CCTTCGGGCTGCCCCGGCTGCCGGAGGAGTGGTCGCAGTGGTTAGGCGGC    400

399 A--A-CTT-CTTCA--CTA-GT-C-CC----------------AT--AG    448
    |  | ||  |  ||  ||| || | ||                ||  ||
401 AGCAGCTGGC--CAGGCTACGTGCGCCCCCTGCCCCCCGCCGCCATCGAG    450

449 A-------A-TAG-----------A--G--------T--G-----AA-T    498
    |       | | |           |  |        |  |     || |
451 AGCCCCGCAGTGGCCGCGCCCGCCTACAGCCGCGCGCTCAGCCGGCAACT    500

499 -AG-AGTCTTG-TCTT--A-ATACAGCAT--TG---AT---T------TA    548
     || || |  | |||   | || | |||   ||   |    |      |
501 CAGCAG-CGGGGTCTCGGAGATCCGGCACACTGCGGACCGCTGGCGCGTG    550

549 T---T--ATGTNATATATCCATATTC-CC--G-A-----TGA---TCAAG    598
    |   |  |||| | |   |||  ||| ||  | |     |||   |||||
551 TCCCTGGATGTCA-A---CCAC-TTCGCCCCGGACGAGCTGACGGTCAAG    600

599 A-------T--C-T--TG-----C-----TG-G------TG-TN-T--AT    648
    |       |  | |  ||     |     || |      || |  |  |
601 ACCAAGGATGGCGTGGTGGAGATCACCGGTGAGCCCCCCTGCTCCTGCAG    650
```







Alignment Of DNA 1/Region 3

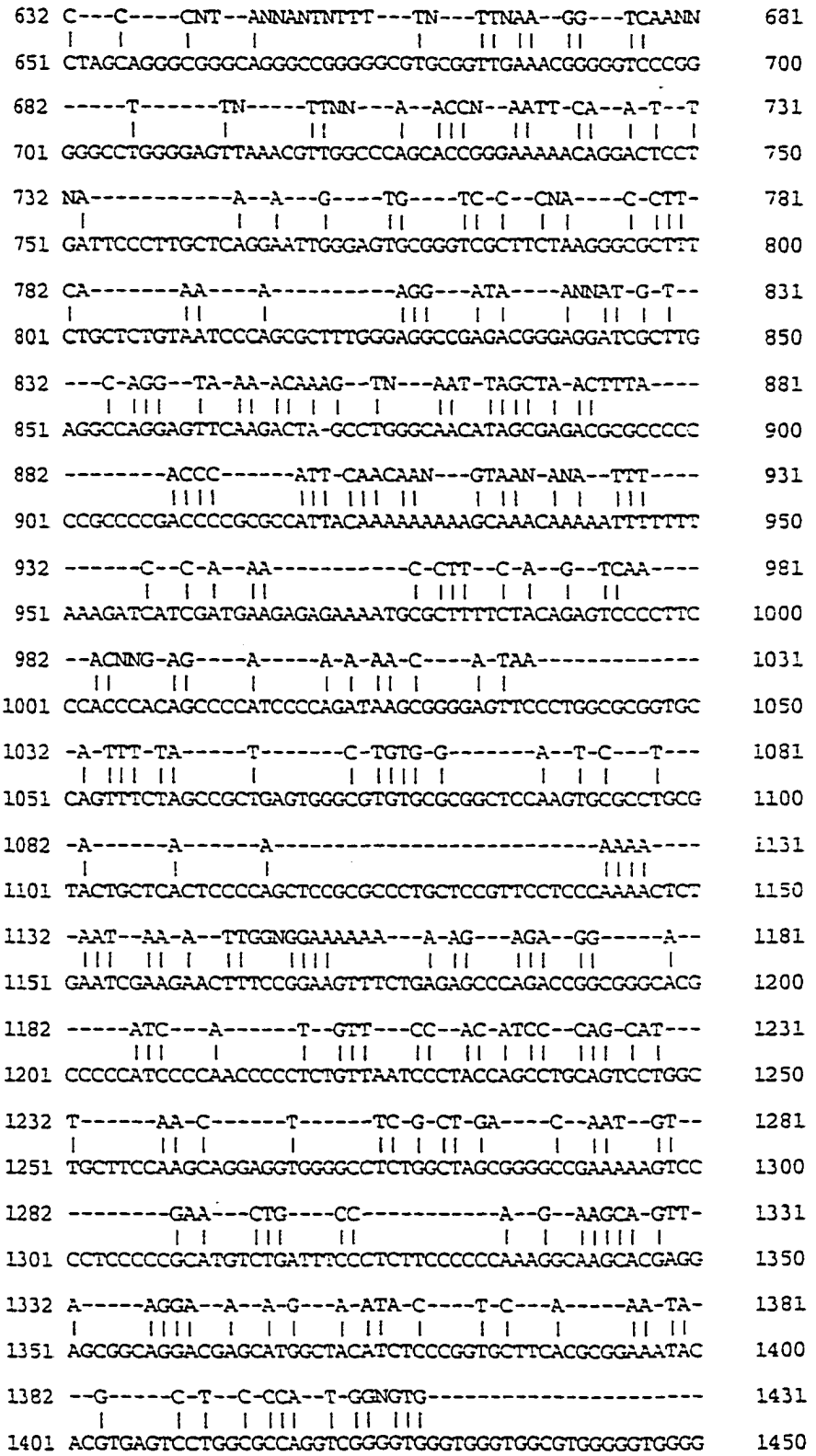
(Complementary Strand), And

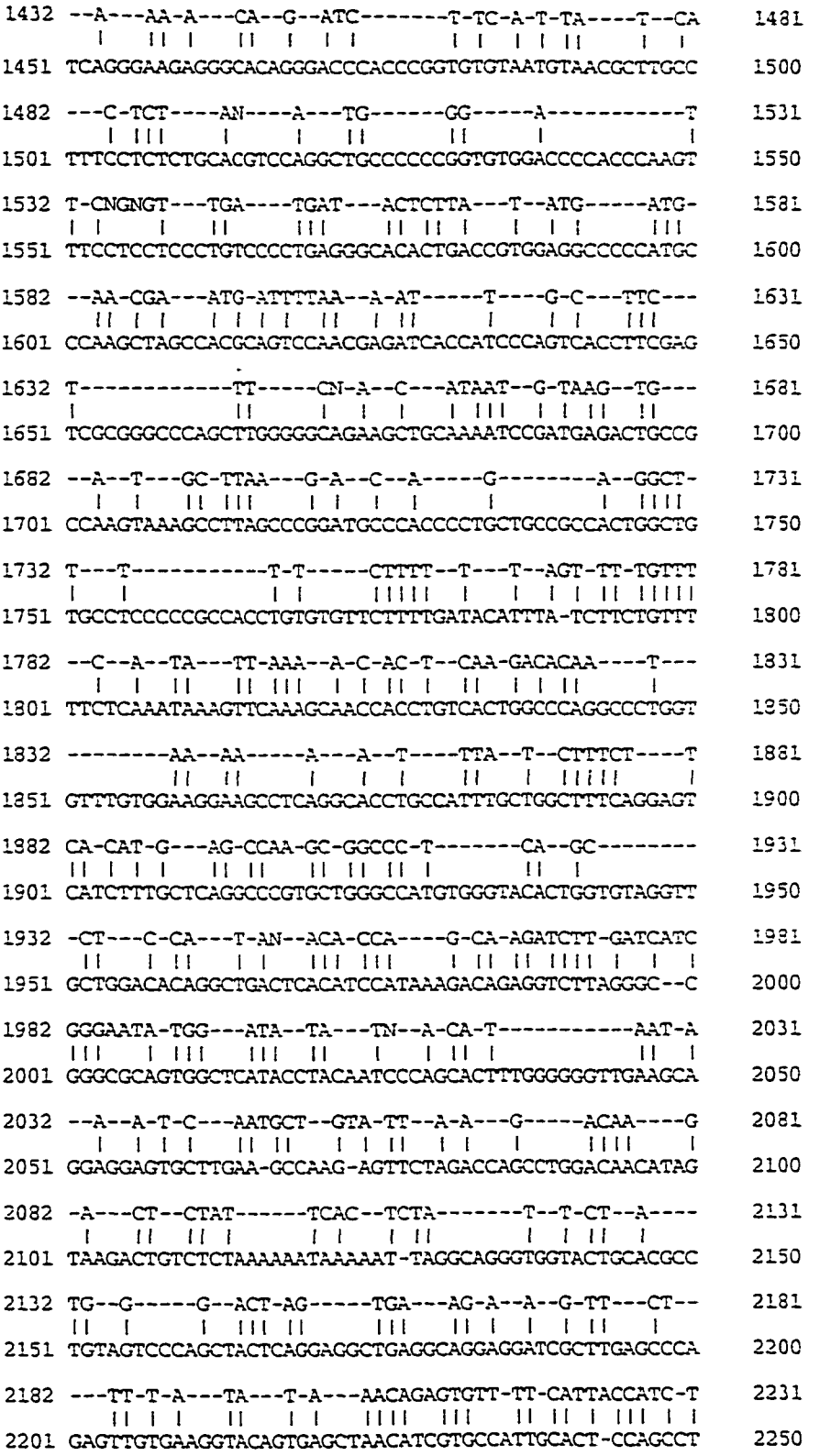
The Human hsp 27 Gene

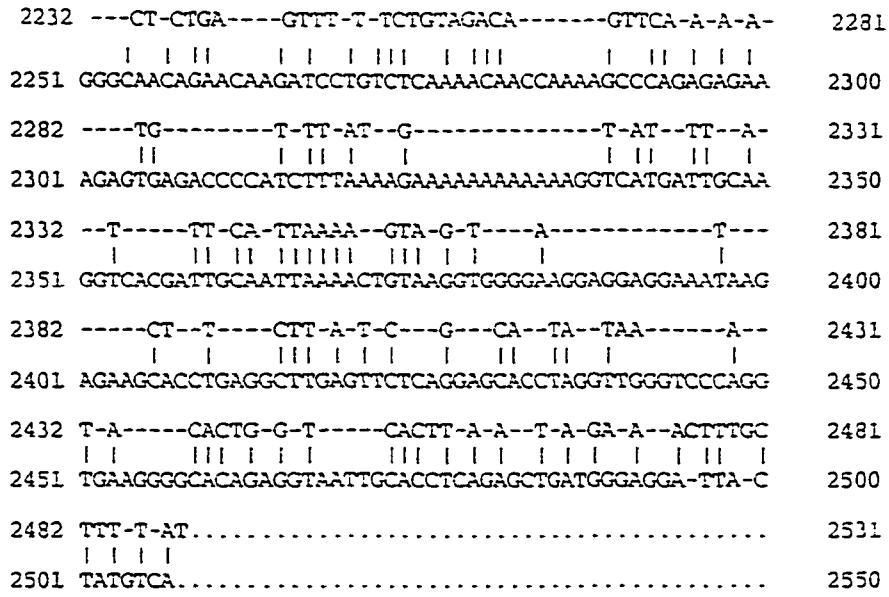
(The top row shows the DNA 1/Region 3 nucleotide sequence).

(The bottom row shows the human hsp 27 gene sequence).

```
-18 ..................NANC-----------CC-G--G-G---TNAN      31
                        | |           || |  | |   | |
  1 GAATTCATTTGCTTTTCCTTAACGAGAGAAGGTTCCAGATGAGGGCTGAA      50

 32 CCC-C--CG----GA--ATNACCCNT-ANNNN------A--A----A---      81
    ||| |  ||    |   |   ||| | |          |  |    |
 51 CCCTCTTCGCCCCGCCCACGGCCCCTGAACGCTGGGGGGAGGAGTGCATGG     100

 82 --A-----------ANN-GGG--A-----A--A-T-------TCN--C-      131
      |           |  |||  |     |  | |        ||   |
101 GGAGGGGCGGCCCTCAAACGGGTCATTGCCATTAATAGAGACCTCAAACA     150

132 CC-C--GC--------C---A--  .NNNG-------G-GN--C-G-G--      181
    || |  ||        |   |    |   |       | |   | | |
151 CCGCCTGCTAAAAATACCCGACTGGAGGAGCATAAAAGCGCAGCCGAGCC     200

182 --G-G----GN--TTTT---AG--GANTT--A-A--ANTNNTNNGG--GN     231
      | |    |  ||||   ||  ||  | | |  |    |  |   |
201 CAGCGCCCCGCACTTTTCTGAGCAGACGTCCAGAGCAG-AGTCAGCCAGC     250

232 A--AN-GNG-GNAANNA----TT-T---T--TN---------A------A     281
    |  |  | | |          || |   |  |          |      |
251 ATGACCGAGCGCCGCGTCCCCTTCTCGCTCCTGCGGGGCCCCAGCTGGGA     300

282 ----------A-----A-----AN--C----T-TN--A--A--------      331
              |     |     |   |    | |   |  |
301 CCCCTTCCGCGACTGGTACCCGCATAGCCGCCTCTTCGACCAGGCCTTCG     350

332 -----------------AN-A-T--TC--A------AGG-GN-A--AN-      381
                    |  | |  ||  |      ||| |  |  |
351 GGCTGCCCCGGCTGCCGGAGGAGTGGTCGCAGTGGTTAGGCGGCAGCAGC     400

382 T--------TN-G-G-GN-----G------G-----A---ANA-------      431
    |        |  | | |     |      |     |   | |
401 TGGCCAGGCTACGTGCGCCCCCTGCCCCCCGCCGCCATCGAGAGCCCCGC     450

432 AN-GG--G-G---GCAT-----------T----------T-A--AG-G      481
    |  ||  | |   || |           |          | |  || |
451 AGTGGCCGCGCCCGCCTACAGCCGCGCGCTCAGCCGGCAACTCAGCAGCG     500

482 ---T-----A-AT---------TGN--AN---TN-----T-T---T---T     531
       |     | ||         ||   |    |      | |   |   |
501 GGGTCTCGGAGATCCGGCACACTGCGGACCGCTGGCGCGTGTCCCTGGAT     550

532 -T-------T-C-CCC-GG--G-GC--A-G-TNAAN--------T----T     581
     |       | | ||| ||  | ||  | | | ||         |    |
551 GTCAACCACTTCGCCCCGGACGAGCTGACGGTCAAGACCAAGGATGGCGT     600

582 --T------C-CCN---A-------T--TNN-G--GGC-A-A--AGG---     631
      |      | ||    |       |  |   |  ||  | |  |||
601 GGTGGAGATCACCGGTGAGCCCCCCTGCTCCTGCAGGGGAGAGGAGGAGG     650
```





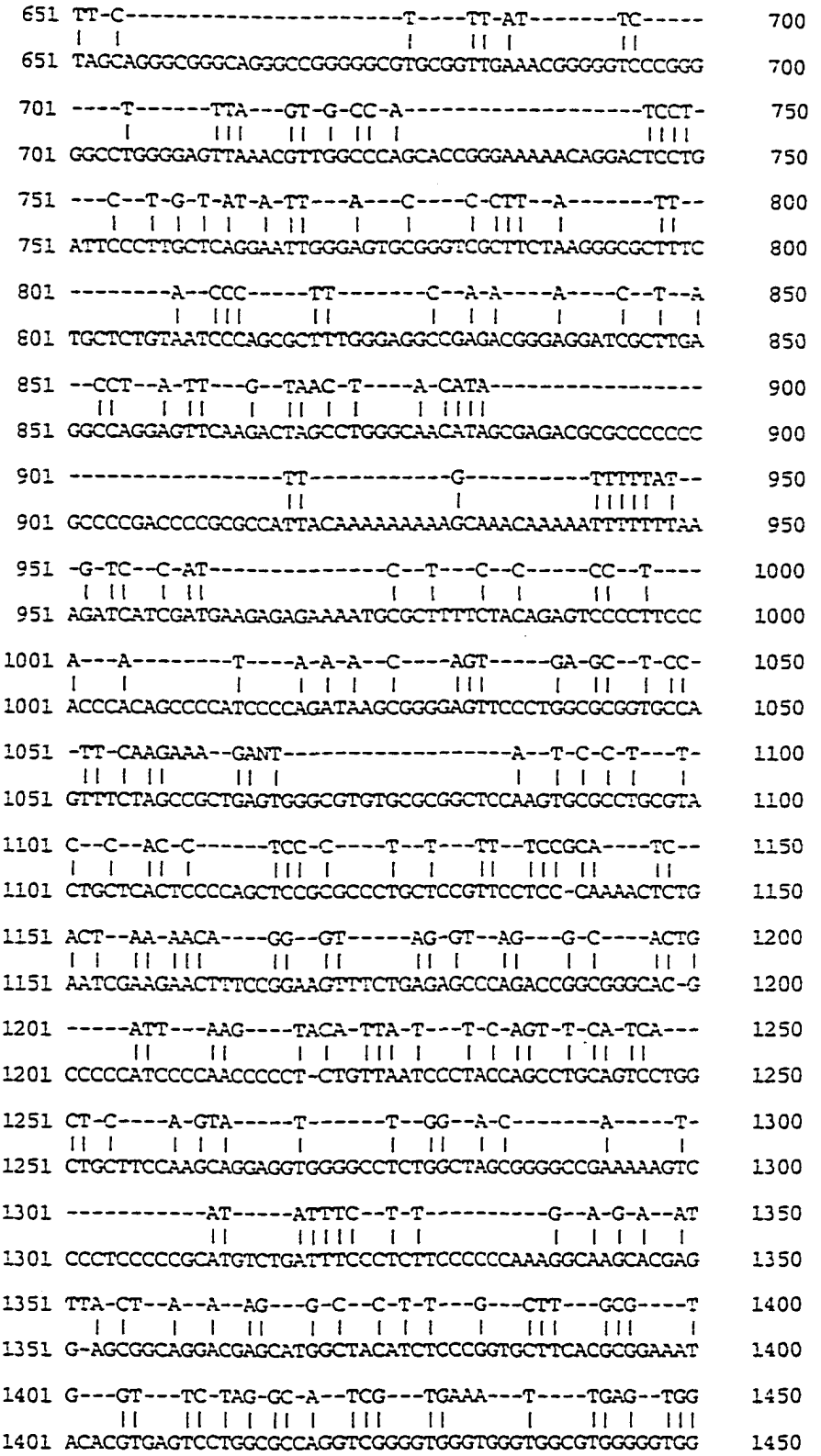


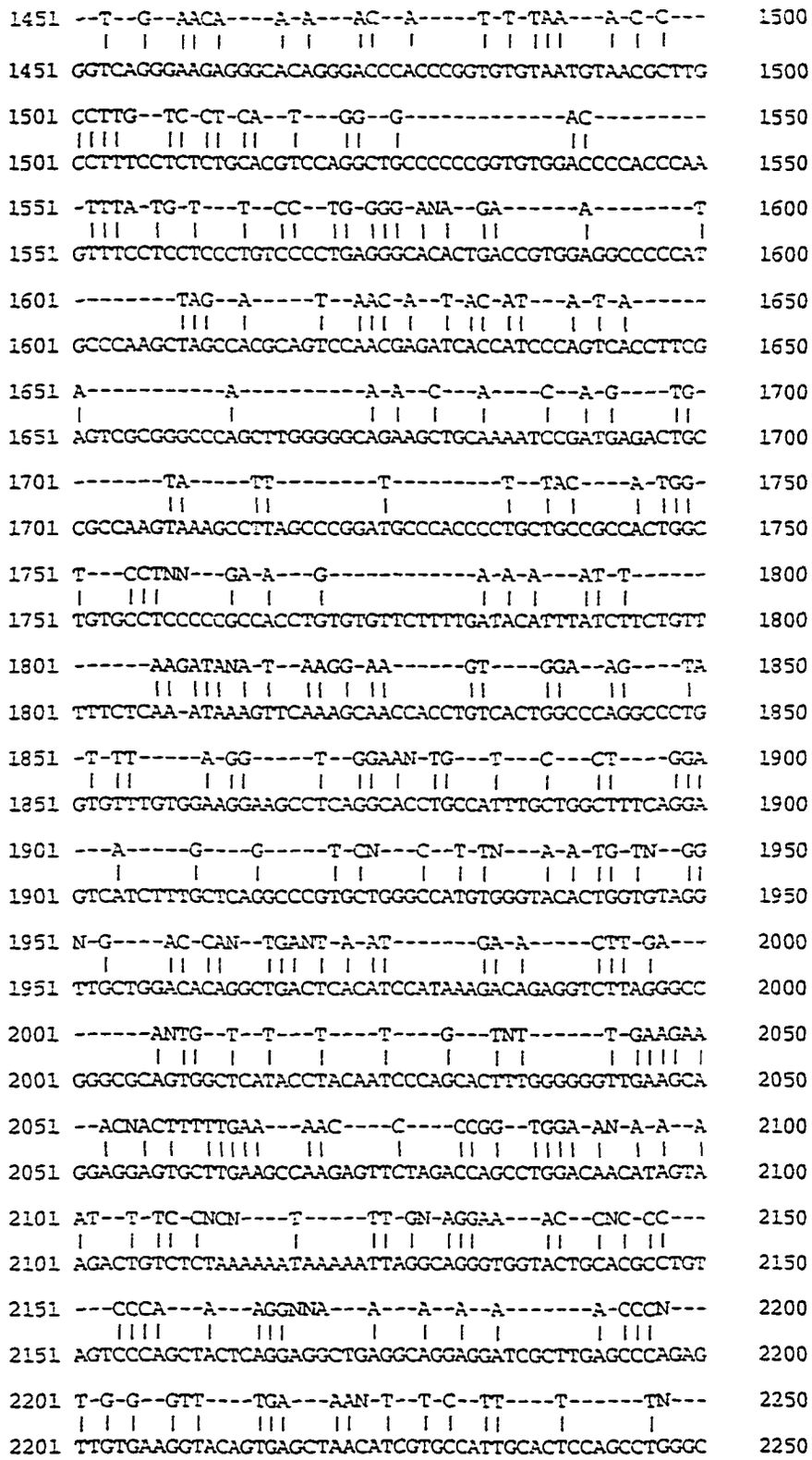
Alignment Of DNA 2/Region 1 (Normal Strand), And The Human hsp 27 Gene

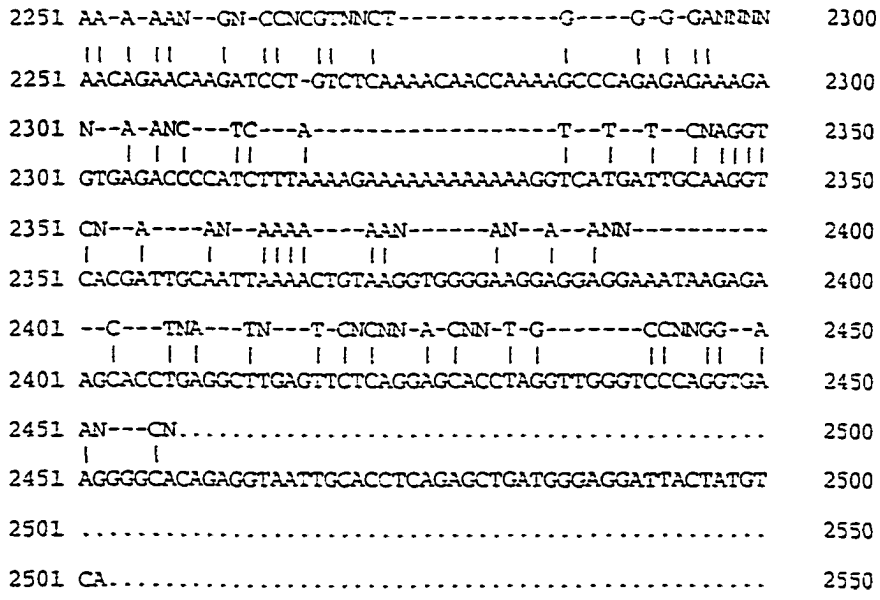
(The top row shows the DNA 2/Region 1 nucleotide sequence).

(The bottom row shows the human hsp 27 gene sequence).

```
  1 GACT--AT---CT----------G-G------T---A-A-G-----T-A-      50
    || |  ||   ||          | |      |   | | |     | |
  1 GAATTCATTTGCTTTTCCTTAACGAGAGAAGGTTCCAGATGAGGGCTGAA      50

 51 C--TCT----------A-------T-A-C-CT-----------C-T--      100
    |  |||           |       | | | ||           | |
 51 CCCTCTTCGCCCCGCCCACGGCCCCTGAACGCTGGGGGAGGAGTGCATGG      100

101 --A----C-------A--C-----A---CC-TT--T-------T-AAA--      150
      |    |       |  |     |   || ||  |       | |||
101 GGAGGGGCGGCCCTCAAACGGGTCATTGCCATTAATAGAGACCTCAAACA      150

151 --G--TG--AA---T-------TG-A----C-------C-----------      200
      |  ||  ||   |       || |    |       |
151 CCGCCTGCTAAAAATACCCGACTGGAGGAGCATAAAAGCGCAGCCGAGCC      200

201 ------------T------A-------T----------T---------      250
                |      |       |          |
201 CAGCGCCCCGCACTTTTCTGAGCAGACGTCCAGAGCAGAGTCAGCCAGCA      250

251 TG-C----------T----TT-T---TC-TA-GA------A--T-----      300
    || |          |    || |   || |  |      |  |
251 TGACCGAGCGCCGCGTCCCCTTCTCGCTCCTGCGGGGCCCCAGCTGGGAC      300

301 ---TT-------T--TAA----A-A------T-TT----------TT-G-      350
       ||       |  ||     | |      | ||          || |
301 CCCTTCCGCGACTGGTACCCGCATAGCCGCCTCTTCGACCAGGCCTTCGG      350

351 --T--------T-CC--A--------C-C-GT--T-A------A--A-C-      400
      |        | ||  |        | | ||  | |      |  | |
351 GCTGCCCCGGCTGCCGGAGGAGTGGTCGCAGTGGTTAGGCGGCAGCAGCT      400

401 --CC---C-A-GT--------TG----------AT----------          450
      ||   | | ||        ||          ||
401 GGCCAGGCTACGTGCGCCCCCTGCCCCCCGCCGCCATCGAGAGCCCCGCA      450

451 -T--C----------T-C-G----------A-C-------TCA--A-C--      500
     |  |          | | |          | |       |||  | |
451 GTGGCCGCGCCCGCCTACAGCCGCGCGCTCAGCCGGCAACTCAGCAGCGG      500

501 --T---G-A-A------A-ACT----A----T------T--C--T---T-      550
      |   | | |      | |||    |    |      |  |  |   |
501 GGTCTCGGAGATCCGGCACACTGCGGACCGCTGGCGCGTGTCCCTGGATG      550

551 -C------TT--------------T--CGG--AA-ACCT----T--CAT-      600
     |      ||              |  |||  || |||     |  | |
551 TCAACCACTTCGCCCCGGACGAGCTGACGGTCAAGACCAAGGATGGCGTG      600

601 -TG-A---C-----T-A-------T--TC----A----A-AGTT-----C      650
     || |   |     | |       |  ||    |    | ||       |
601 GTGGAGATCACCGGTGAGCCCCCCTGCTCCTGCAGGGGAGAGGAGGAGGC      650
```







Alignment Of DNA 2/Region 1

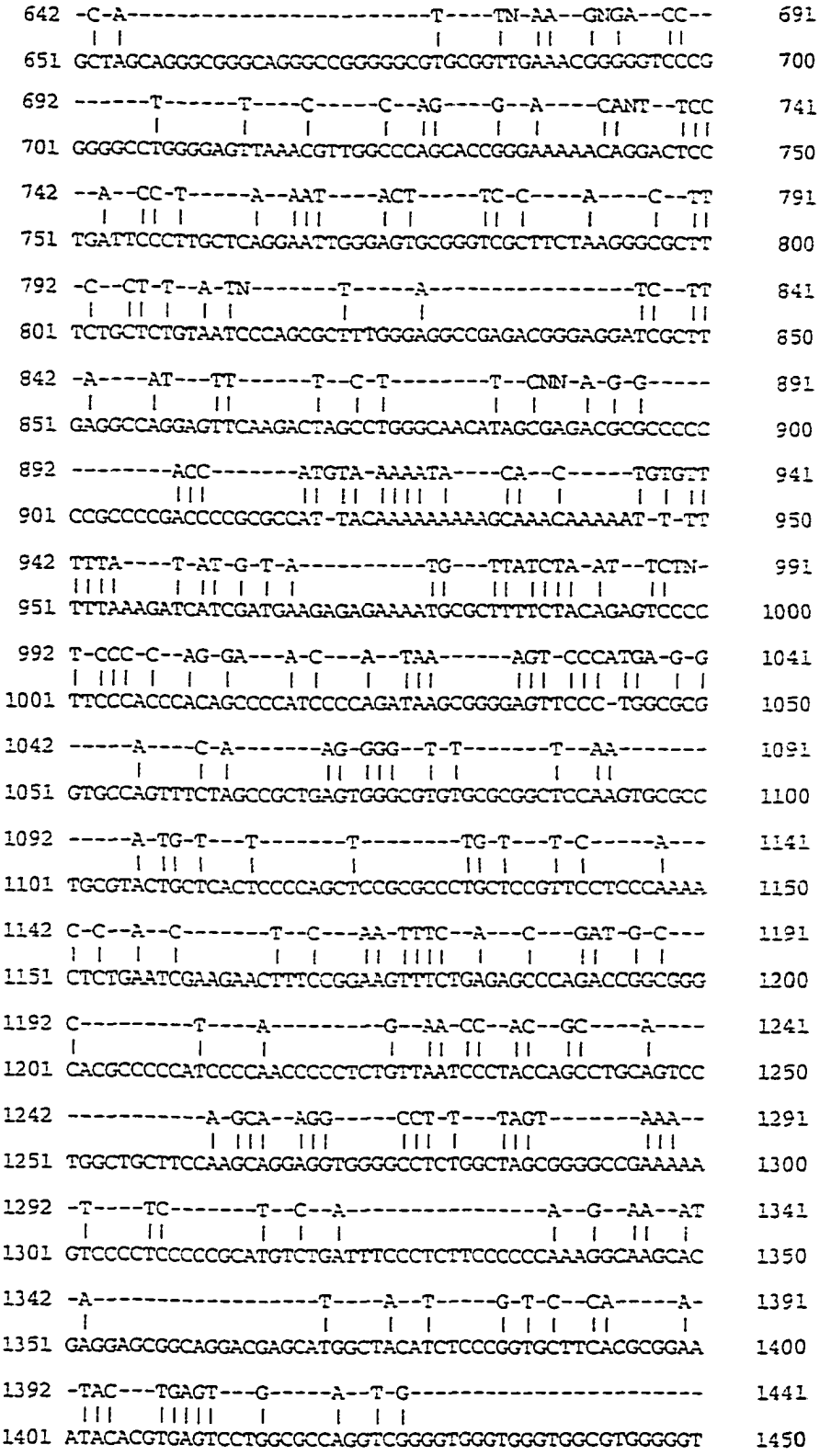
(Complementary Strand), And

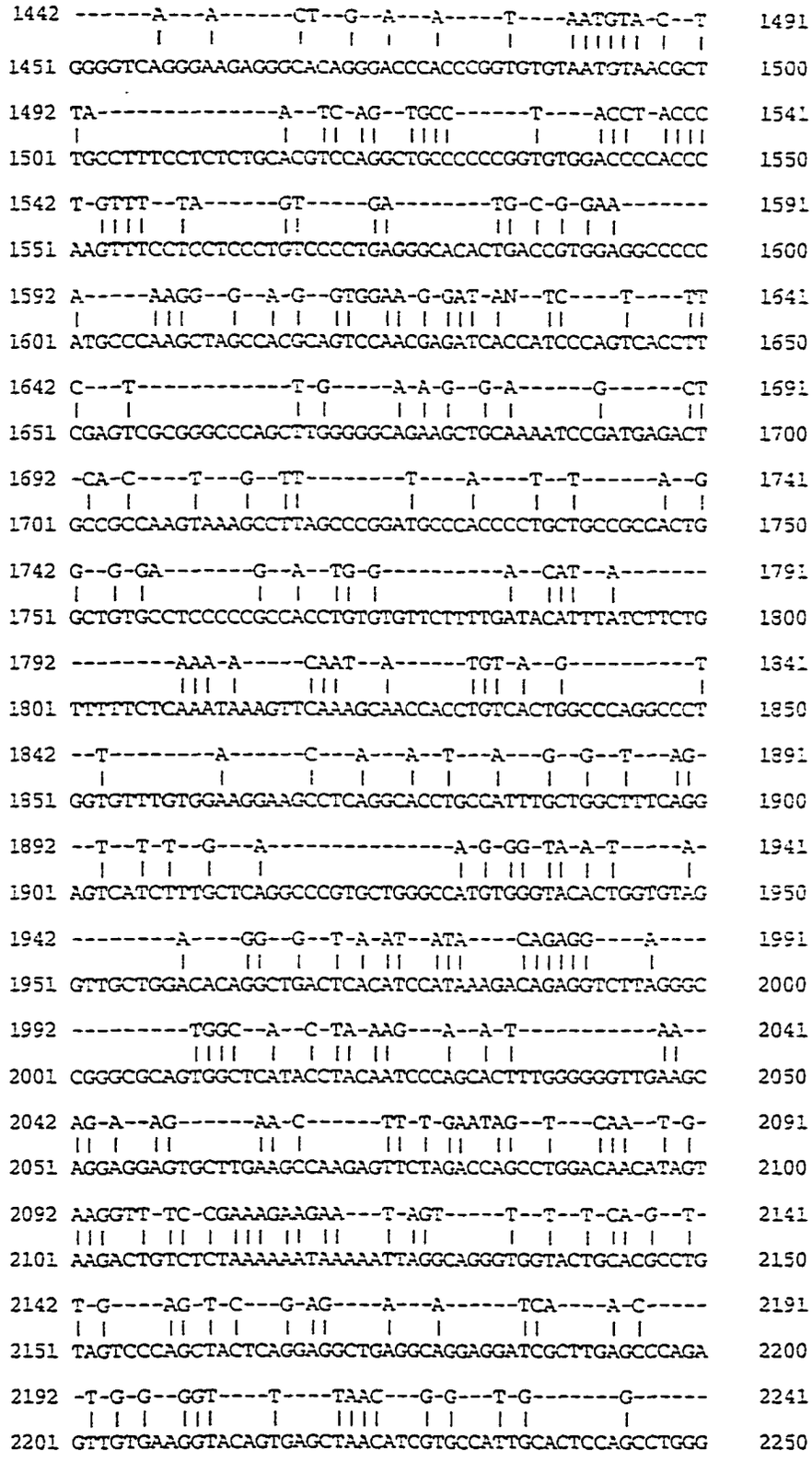
The Human hsp 27 Gene

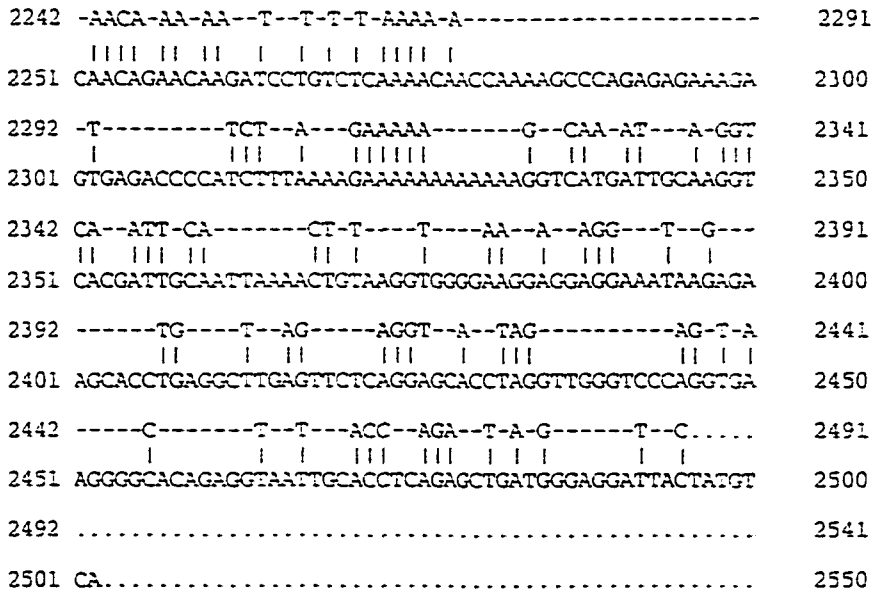
(The top row shows the DNA 2/Region 1 nucleotide sequence).

(The bottom row shows the human hsp 27 gene sequence).

```
 -8 .........NGNTT--CCNN---G-GC-ANNGTNN--GN-GAN-----A-      41
 1  GAATTCATTTGCTTTTCCTTAACGAGAGAAGGTTCCAGATGAGGGCTGAA      50

 42 ---TN-----------A-GNNT--TN-----TN---------T---T--      91
 51 CCCTCTTCGCCCCGCCCACGGCCCCTGAACGCTGGGGGAGGAGTGCATGG     100

 92 ----------T--------T--TTN---TTN---GA--CCTNGAAAT       141
101 GGAGGGGCGGCCCTCAAACGGGTCATTGCCATTAATAGAGACCTC-AAAC     150

142 GA--GN-T--TNNNNNT-CCC--C---AGNN--A------CGN-G--GN-     191
151 -ACCGCCTGCTAAAAATACCCGACTGGAGGAGCATAAAAGCGCAGCCGAG     200

192 CN----------TTTT-TNA--A-A-----AGA--ANTTTCAA--A-      241
201 CCCAGCGCCCCGCACTTTTCTGAGCAGACGTCCAGAGCAGAGTCAGCCAG     250

242 C----CC-AN-G--G-GT----TT-T---T--TNN----CC-----T---    291
251 CATGACCGAGCGCCGCGTCCCCTTCTCGCTCCTGCGGGGCCCCAGCTGGG     300

292 -----TTG-G-G---GG-----GN---G--G--T-TTC--C------TNC    341
301 ACCCCTTCCGCGACTGGTACCCGCATAGCCGCCTCTTCGACCAGGCCTTC     350

342 AAAN-GN---GGA------A--ATT--T-----TN-TT---C--CA-C--    391
351 GGGCTGCCCCGGCTGCCGGAGGAGTGGTCGCAGTGGTTAGGCGGCAGCAG     400

392 C-GG---GGTT---T--------T-CA----------A---A-A------    441
401 CTGGCCAGGCTACGTGCGCCCCCTGCCCCCCGCCGCCATCGAGAGCCCCG     450

442 -AGTNG----------T----------T----------TC------       491
451 CAGTGGCCGCGCCCGCCTACAGCCGCGCGCTCAGCCGGCAACTCAGCAGC     500

492 ----T-TC--A-AN-----ACA------A--------------------A    541
501 GGGGTCTCGGAGATCCGGCACACTGCGGACCGCTGGCGCGTGTCCCTGGA     550

542 ----A-C-ANTTC-------A--AG-T-----TCA----------T----    591
551 TGTCAACCACTTCGCCCCGGACGAGCTGACGGTCAAGACCAAGGATGGCG     600

592 T-----AN-TCAN---TG-G------T-CNCCN--A--------------    641
601 TGGTGGAGATCACCGGTGAGCCCCCCTGCTCCTGCAGGGGAGAGGAGGAG     650
```





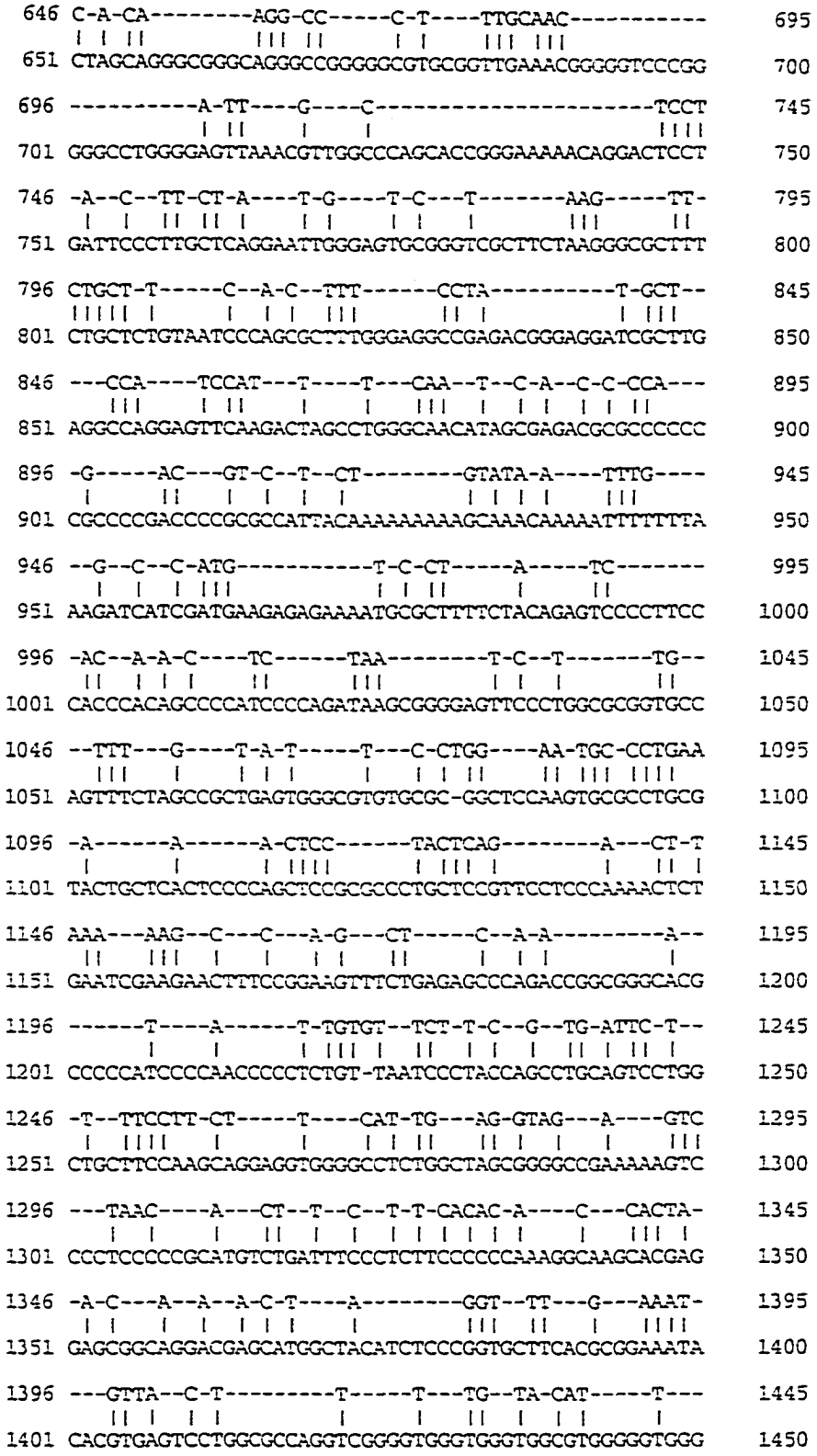


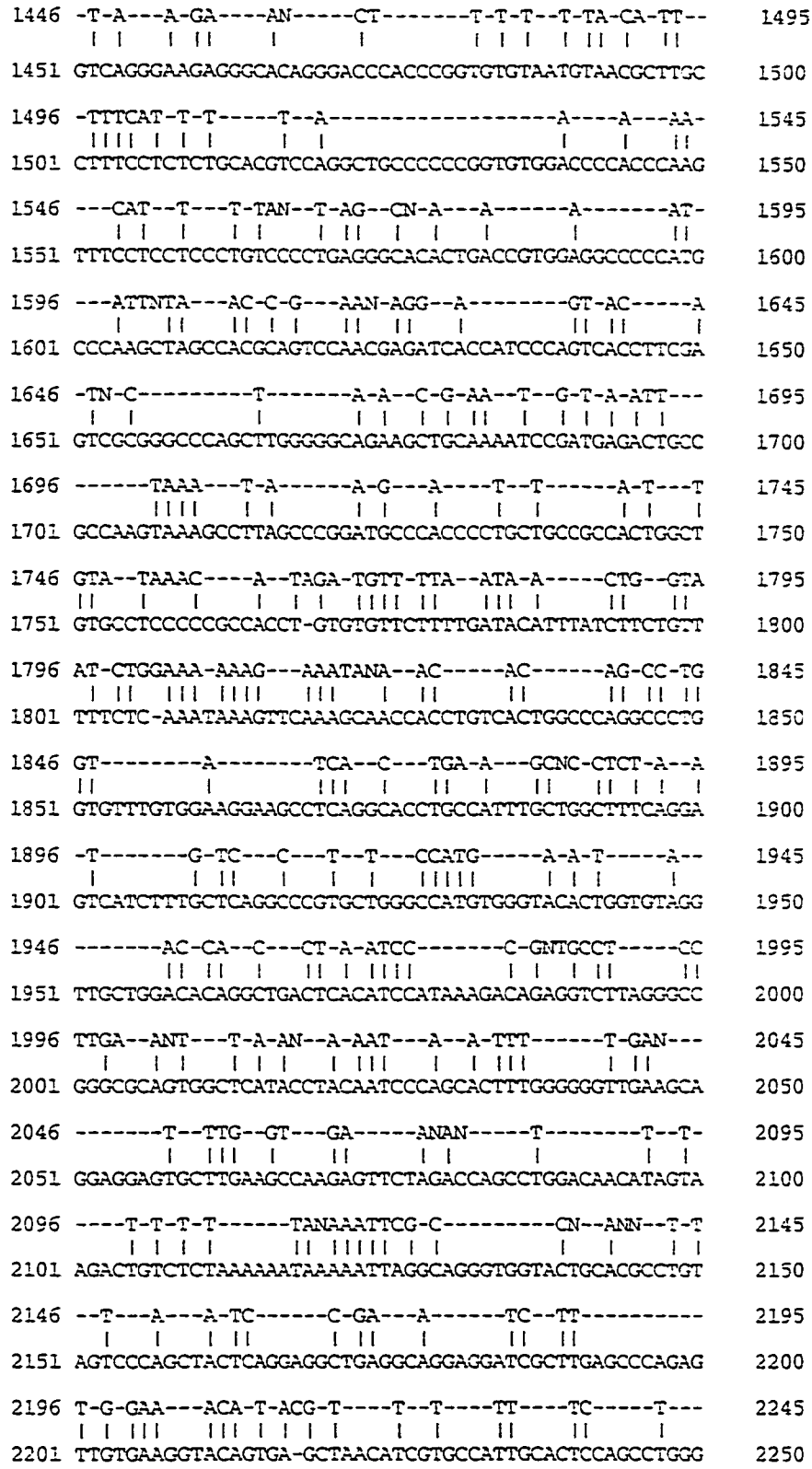
Alignment Of DNA 2/Region 2 (Normal Strand), And The Human hsp 27 Gene

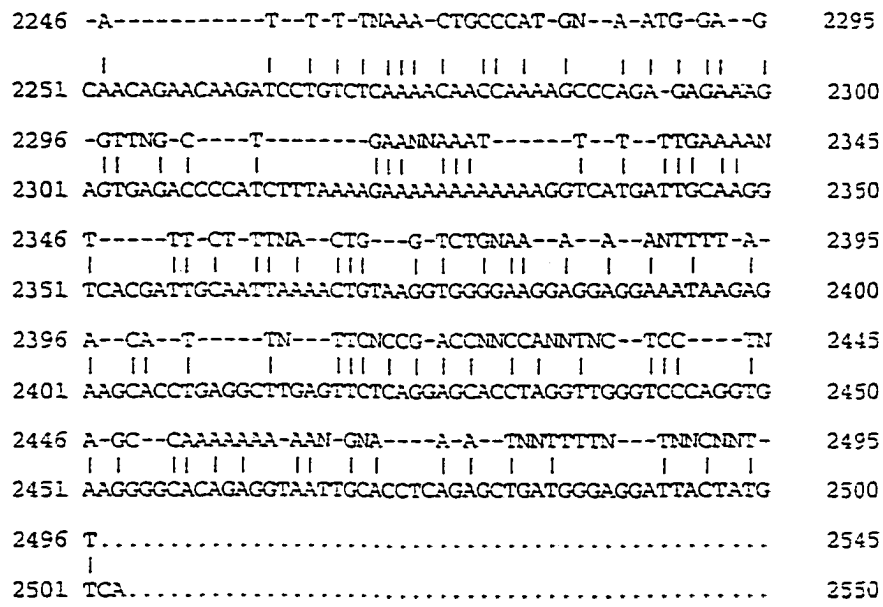
(The top row shows the DNA 2/Region 2 nucleotide sequence).

(The bottom row shows the human hsp 27 gene sequence).

```
-4 .....C-T--GC--------A--G-G------T-C-A-ACG-GGA-T---      45
       | |  ||        |  | |      | | | | | ||  |
 1 GAATTCATTTGCTTTTCCTTAACGAGAGAAGGTTCCAGATGAGGGCTGAA      50

46 C--T-TT-G----GC--A-------TGA----T-----ATT--T--AT--      95
   |  | || |    ||  |       |||    |     |    |  ||
51 CCCTCTTCGCCCCGCCCACGGCCCCTGAACGCTGGGGGAGGAGTGCATGG     100

96 --A----CA-C--T-----G--TC-T--C-ATTT-T-----CTTC-----     145
     |    | |  |     |  || |  | |||  |     | ||
101 GGAGGGGCGGCCCTCAAACGGGTCATTGCCATTAATAGAGACCTCAAACA     150

146 --GTTTG-TATA--T-------TGG-GGA---T----GTTTAT---A---     195
      |  || || |  |       ||| |||   |    |   |   |
151 CCGCCTGCTAAAAATACCCGACTGGAGGAGCATAAAAGCGCAGCCGAGCC     200

196 -ATT-------A-TTTT--G--CAG--G---A--------T---------     245
     |         | ||||  |  |||  |   |        |
201 CAGCGCCCCGCACTTTTCTGAGCAGACGTCCAGAGCAGAGTCAGCCAGCA     250

246 T-A-----------T----TT-T-GA--------------A--TG----     295
    | |           |    || | |                |  ||
251 TGACCGAGCGCCGCGTCCCCTTCTCGCTCCTGCGGGGCCCCAGCTGGGAC     300

296 ---T---G--A--G--A-----A-A------T-----A--A----TATCT     345
       |   |  |  |  |     | |      |     |  |    | ||
301 CCCTTCCGCGACTGGTACCCGCATAGCCGCCTCTTCGACCAGGCCT-TCG     350

346 ---T--------T-----A--ATT--T-G-A-T--TT-------------     395
       |        |     |  | |  | | | |  ||
351 GGCTGCCCCGGCTGCCGGAGGAGTGGTCGCAGTGGTTAGGCGGCAGCAGC     400

396 T--C-----T---T--------T-------------T-G----------     445
    |  |     |   |        |             | |
401 TGGCCAGGCTACGTGCGCCCCCTGCCCCCCGCCGCCATCGAGAGCCCCGC     450

446 --T-------------TA--GC-------T-A------AT-TAA--A---     495
      |             ||  ||       | |      |  | |  |
451 AGTGGCCGCGCCCGCCTACAGCCGCGCGCTCAGCCGGCAACTCAGCAGCG     500

496 ---T---GG---T-----A-ACTA---AC---T------T-T---T--AT     545
       |   ||   |     | |||    ||   |      | |   |  ||
501 GGGTCTCGGAGATCCGGCACACTGCGGACCGCTGGCGCGTGTCCCTGGAT     550

546 -TCA-----T---------A-----T-A-----A----------T----T     595
     |||     |         |     | |     |          |    |
551 GTCAACCACTTCGCCCCGGACGAGCTGACGGTCAAGACCAAGGATGGCGT     600

596 -----AG-T-A----TG-G------T-CTCAT-CA--------------     645
        || | |    || |       | ||| | ||
601 GGTGGAGATCACCGGTGAGCCCCCCTGCTCCTGCAGGGGAGAGGAGGAGG     650
```





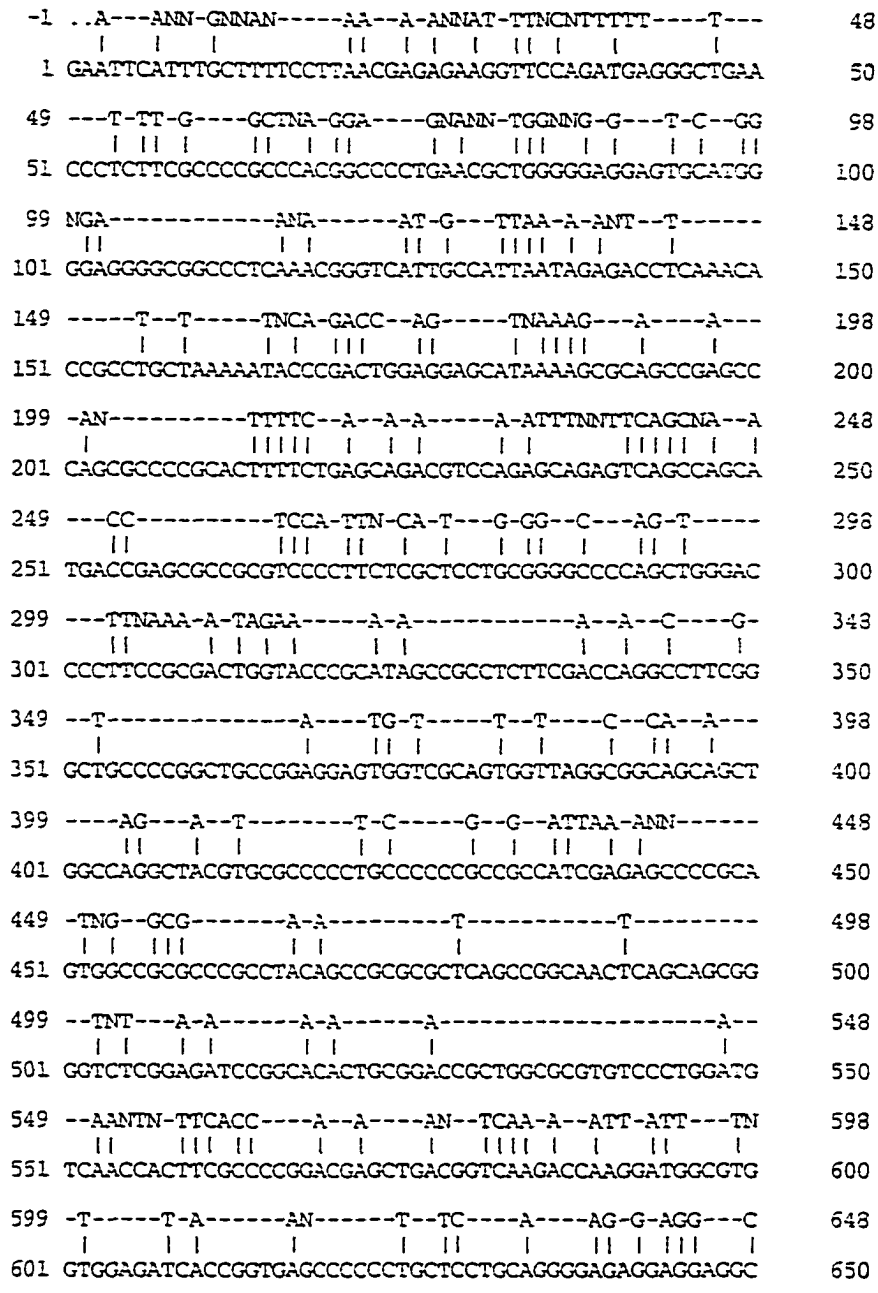


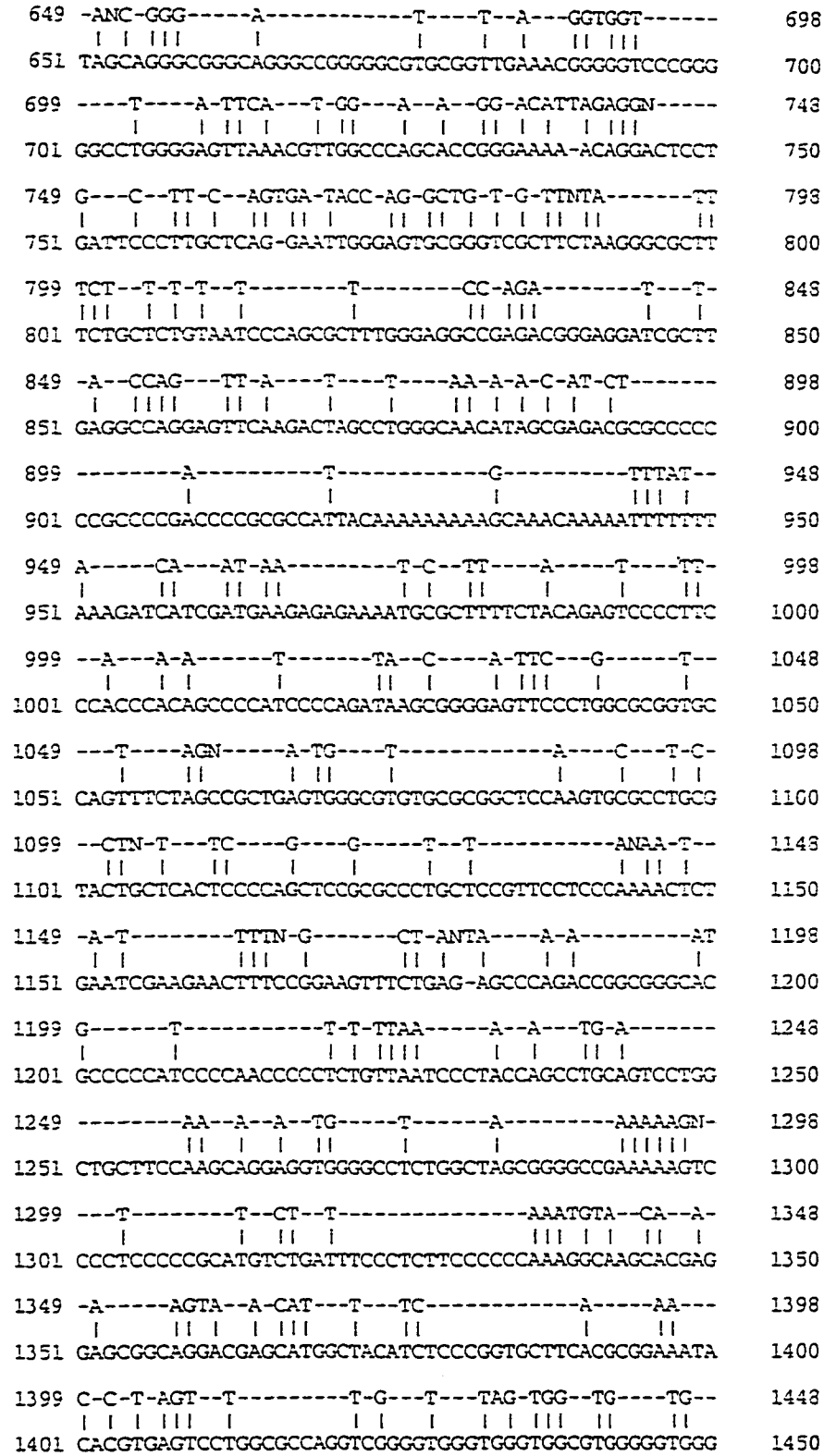
Alignment Of DNA 2/Region 2

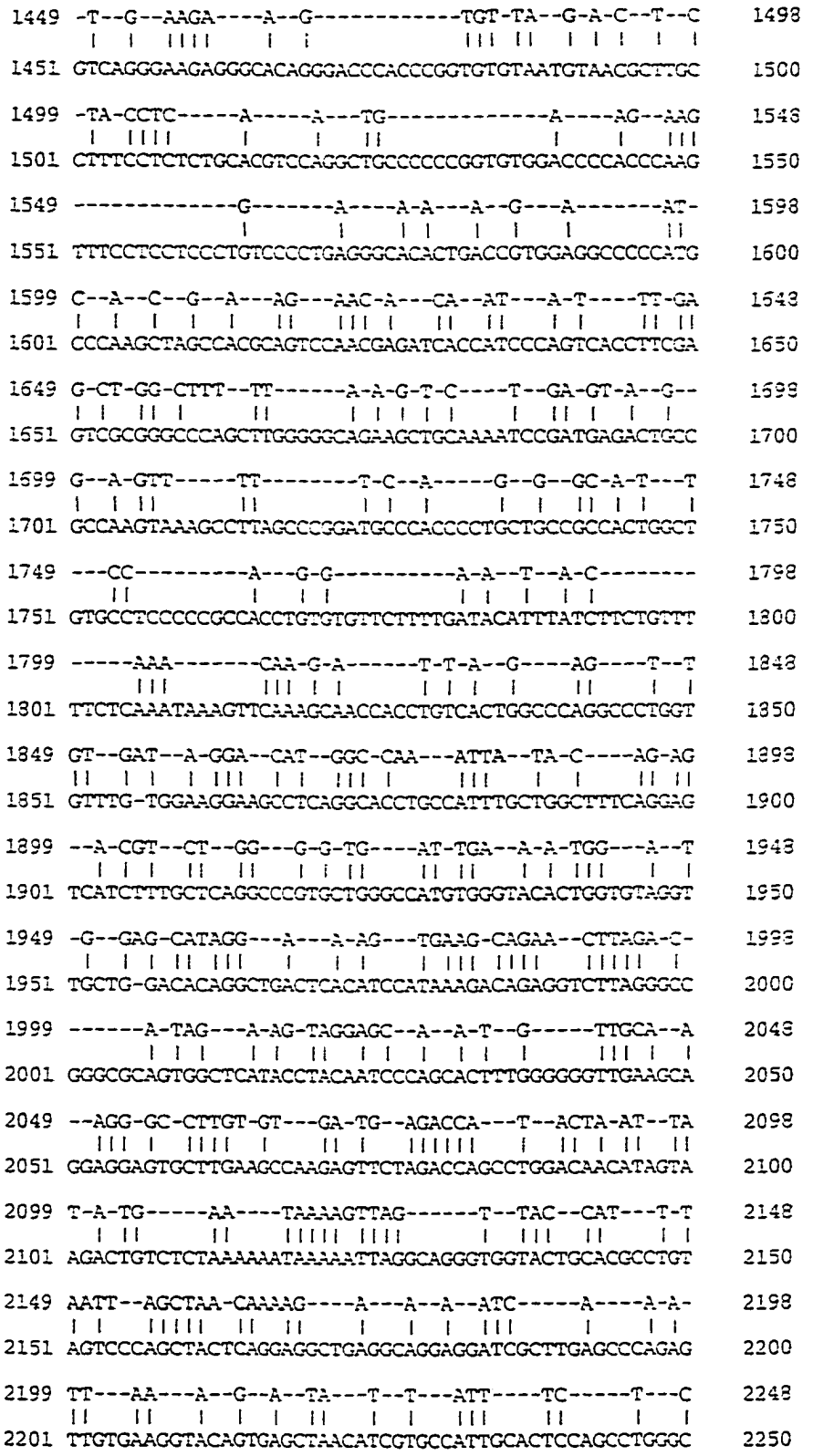
(Complementary Strand), And

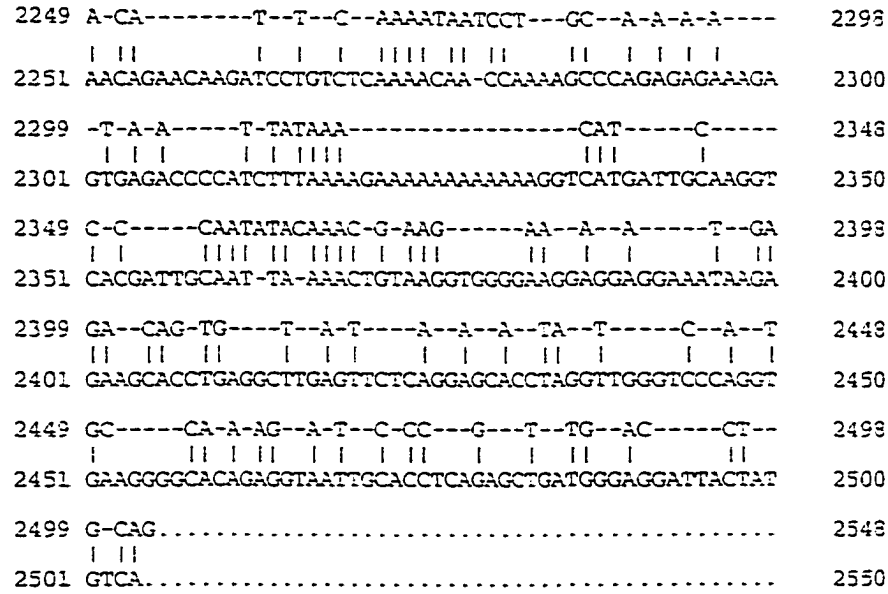
The Human hsp 27 Gene

(The top row shows the DNA 2/Region 2 nucleotide sequence).









APPENDIX V

Alignment Of DNA 4 (Normal Strand),

And The Most Similar-sized,

Homologous Area Of

Human hsp 27

(The top row shows the DNA 4 nucleotide sequence).

(The bottom row shows the human hsp 27 gene sequence).

```
  -3 ....TC-G-AC---CTTG---A-ATA-C--GGAGCTTCAAGCCAGGC--T      46
1349 AGCGGCAGGACGAGCATGGCTACATCTCCCGGTGCTTCACGC--GGAAAT    1398

  47 -C-C---AG-C-TGGCGTTTTAGGT-G---TGG-TGCAAGCAAGGC-TGG     96
1399 ACACGTGAGTCCTGGCGCC--AGGTCGGGGTGGGTGG--GT--GGCGTGG    1448

  97 GTATGTATTTCCAGG-AAGCAGT--AGANG-AGCTCGCC-G-TGCAATA-     146
1449 GGGTGGGGT-C-AGGGAAG-AGGGCACAGGGACC-CACCCGGTGTG-TAA    1498

 147 TG-AA---TTA--T--CAT-TA-GCA-G-CAAGGAAATTCGTCAACGATT     196
1499 TGTAACGCTTGCCTTTCCTCTCTGCACGTCCAGGC--TGCCCCC-CGGT-    1548

 197 TGTTGGTAA---ATG-AA-TTTCCTGTCAGATATCCAGTTGCTTG-GGAG     246
1549 -GT-GG-ACCCCACCCAAGTTTCCTC-C---TC-CCTGTCCCCTGAGG-G    1598

 247 AACGCTG-CCTTTGCG-CCTCCTGCGGACGGCTCAATCTGGCTTAAGT-A     296
1599 CACACTGACCGTGGAGGCC-CC--C--ATGCC-CAA---G-CT-A-GCCA    1648

 297 CG-ACTACACA-GAAGTTGAC-A-C--AGTAACATACGGGTTAAGCAGGA     346
1649 CGCAGTCCA-ACGA-GATCACCATCCCAGTCACCTTCGAGTC--GC-GG-    1698

 347 AGTGCAAATACTACG-TTGGCATGGTGCAGATTTCAG-TGTTCTTCAGTC     396
1699 -GC-C-----C-A-GCTTGG---GG-GCAGA----AGCTG--CAAAA-TC    1748

 397 CGGGGACTGGGATTGA-TAAGCCGAGACAAATAGCTAATCAATTGGCAGA     446
1749 CG---A-TG--A--GACT--GCCGC--CAAGTA---AAGCC-TTAGC---    1798

 447 ATCTATCGTTGATGGT-ACAA-TGCTTGAACAGCGGGACCATTTATGAGT     496
1799 --C---CG--GATGCCCACCCCTGCT-G--C--CG---CCACT---G-G-    1848

 497 CTGGAGTTGT-TAACCCGGTTATCAAATCCAAGTCTGGGTGGT-TTATTC     546
1849 CTG---T-GCCTCCCCCGC----CA---CCT-GTGTG--T--TCTT-TT-    1898

 547 CCGGTTCGTTTTTATGTTC-GTCTAGACT-AACA-AAAGG--AAA--AAT     596
1899 --GATACATTT--ATCTTCTGTTTTT-CTCAA-ATAAAGTTCAAAGCAAC    1948
```

```
 597 -ACATGGCCCATCTCAGCA-ATGGCACGCAGGTCTTCGTAGAA--A-GCT      646
1949 CACCTGTC--A-CTG-GCCCA-GGC-C-CTGGTGTTTGTGGAAGGAAGC-     1998

 647 CTCGCGGAAC-TGCAAT---C-GACGT--A--ACTGCAAT-TTCTAACGC      696
1999 CTCA-GGCACCTGCCATTTGCTGGCTTTCAGGAGT-CA-TCTT-TG-CTC     2048

 697 AGTTACCCCTGTTTTAACTGT---ATCTGACGCGT-C--TGGTTTGGTCG      746
2049 AGG--CCCGTG------CTGGGCCATGTG--G-GTACACTGGTGTAG--G     2098

 747 TTGGTGAATAC---CTG-CTTTTCACCTCTTC-T---GCTTCA-ACGCTT      796
2099 TTGCTGGACACAGGCTGACT---CACATC--CATAAAGA--CAGAGG-T-     2148

 797 CTAA---CTGA-CAANCAAGT---TCGAATGAC-TGCAATCCCTGGTACG      846
2149 CTTAGGGCCGGGC--GCA-GTGGCTC-A-T-ACCTACAATCCCAGC-AC-     2198

 847 TCAGTAACTG-----TTGAAG--GTTTGA-TACCTCCAGCACCACAAAGT      896
2199 T---T---TGGGGGGTTGAAGCAGGAGGAGTGCTTGAAGC-C-A-AGAGT     2248

 897 TC----CCGGC-TGG-CTTAAC-TGGTGAAGT-TGTC---AAGA--TCAC      946
2249 TCTAGACCAGCCTGGAC--AACATAGT-AAGACTGTCTCTAAAAAATAAA     2298

 947 C-TCCT-GGTTCGAAGTTCCGTGCGT--T-CAGGAC-GTA-TCT-A-CTG      996
2299 AAT--TAGG--C-AGG----GTG-GTACTGCACGCCTGTAGTCCCAGCT-     2348

 997 AC---GGTGGC-GAA-CAGCA--ATTCG-TTAA-CTTCCAGTGCTTGTCC     1046
2349 ACTCAGGAGGCTGAGGCAGGAGGAT-CGCTTGAGC--CCAGAG-TTGT--     2398

1047 GATG--ACCGTGA---A-CA--G--C-A--G-ATTCCGA-CCT----AT-     1096
2399 GAAGGTACAGTGAGCTAACATCGTGCCATTGCACTCC-AGCCTGGGCAAC     2448

1097 A-AA-----TC-TG-CGGT-AACTAACA-CCTTCA-CCTTC-GCTCACGA     1146
2449 AGAACAAGATCCTGTC--TCAA--AACAACCAAAAGCC--CAG---A-GA     2498

1147 GTACACCA-ACCCTGTATATCCGGTTCTGCGTAACTACGATGA-------     1196
2499 G-A-A--AGAG--TG-AGACCCCAT-CTT--TAAA-A-GAAAAAAAAAAA     2548

1197 -G-TC-TG---GCCAGGTT--G-TTGCGATTCGT-CTGTTCG-T----AC     1246
2549 AGGTCATGATTGCAAGGTCACGATTGCAATTAAAACTGTAAGGTGGGGAA     2598

1247 CTCGAGCTAGCGAAAT--GCGCTTGCAGTCCGGTA--CTATCGC-TT-TC     1296
2599 G--GAGG-AG-GAAATAAGAGAA-GCA--CCTG-AGGCT-T-GAGTTCTC     2648

1297 AACGA-CACCCCTACCATTGGTGTTA-ACGA-AATCG--A-A-ACGGTA-     1346
2649 AG-GAGCACC--TAGG-TTGG-GTCCCAGGTGAAGGGGCACAGA-GGTAA     2698

1347 TC-CATCGC-G-G-T-ATCC-A---TT-CG-TG.................     1396
2699 TTGCACCTCAGAGCTGATGGGAGGATTACTATGTCA.............     2748
```

Alignment Of DNA 4 (Complementary Strand), And The Most Similar-sized, Homologous Area Of Human hsp 27

(The top row shows the DNA 4 nucleotide sequence).

(The bottom row shows the human hsp 27 gene sequence).

```
   1 CA-CGA-ATGGATACCGCG-ATGGATACCGTTTCGATTTCGTTAACACCA     50
     | ||  | |||  | | | |||| |||       ||  | |   | ||
1348 .AGCGGCA-GGA--C-GAGCATGGCTAC-------AT--C-T---C-CC-  1397

  51 ATGGTAGGGGTGTC--GTTGAAAG-CGA--T-AGTACC-GGACTGCAAG-    100
     ||| |   | ||  |  ||||  | |  | ||| || || | || ||
1398 --GGT-GC--T-TCACGCGGAAATAC-ACGTGAGT-CCTGG-C-GCCAGG  1447

 101 -CGCA-T---T---TCGC-TAGC--TCGAGGT-ACGAACAGACGAATCGC    150
      ||   |   |   | || | |   | | ||| | | | ||| |    ||
1448 TCGGGGTGGGTGGGTGGCGTGGGGGTGG-GGTCAGGGA-AGAGG----GC  1497

 151 AACAACCTGGCCAGACTCATC-G-TAGT-TACGCA-G-AACCGGATATAC    200
     | ||    ||   ||| || | | | || ||   | | ||| |  | | |
1498 A-CA----GG---GACCCACCCGGT-GTGTA---ATGTAAC-GC-T-TGC  1547

 201 AGGGTTGGTG-TACTCGTG-A-G-CGAAGG-TGAA----GGTGTT-AGTT    250
         ||  |  | ||| || | | | | || ||      |||||  |
1548 C---TT--TCCT-CTC-TGCACGTCCA-GGCTGCCCCCCGGTGTGGACCC  1597

 251 -ACCGCAGATTTATAGGTCGGAATCTGCTGTTCACGGTCATCGG-ACAA-    300
      ||| || | || |   ||    ||  |||| | | |  |  || |||
1598 CACC-CA-AGTT-TCC-TCC---TCC-CTGTCCCCTG--AG-GGCACACT  1647

 301 GCAC--TGGAAGTTAACGAATTGCT---GTTCGCCAC-C-GTC-A--GTA    350
     | ||  |||| |    | | | ||    | | |||||  | ||| |  | |
1648 G-ACCGTGGAGGCCCCC-A-T-GCCCAAGCTAGCCACGCAGTCCAACG-A  1697

 351 GAT-AC-------GTC-C-T--GAA-CGCACGGA--A-CTTCGAACCAGG    400
     ||| ||       ||| | |  ||  |||  ||   | ||| |   |||
1698 GATCACCATCCCAGTCACCTTCGAGTCGCG-GGCCCAGCTTGGGGGCAGA  1747

 401 AGGTG-A---TCT--TGACAACTTCA-CCA-GTTAAGCC--AGCCGGGA-    450
     || || |   ||   ||| | || |  ||| || |||||  |||| |||
1748 AGCTGCAAAATCCGATGAGA-CTGCCGCCAAGTAAAGCCTTAGCCCGGAT  1797

 451 ----ACTT-TG-TGGTGC---TGGAGGTA--TCAAACCTTCAACAGTTAC    500
         ||   || ||  ||   |||  ||   ||   ||  |  | |  |
1798 GCCCACCCCTGCTGCCGCCACTGGCTGTGCCTC---CC--C--C-G---C  1847

 501 TGACGTACCAG-G-GATTGCAGTCATTCGA-AC-TTGNT-TGTCAGTTAG    550
        |  ||| | | | || |  |  || || || ||  | | || |||
1848 ---C--ACCTGTGTG-TT-CT-T--TT-GATACATTTATCT-TCTGTTTT  1897

 551 AAGCGTTGAAGCAGAAGAGGTGAAAAGCAGGTATTCACCA---AC-GACC    600
        | |  ||  | || || | ||| |||   |  ||||    || | ||
1898 T--C-TC-AA--ATAA-AGTTCAAA-GCA---AC-CACCTGTCACTGGCC  1947

 601 AAACCAGACGC-G-TCAGATACAGTTAAAACAGGGGTAACTGCGTTAGAA    650
        | || | | | |  | |   ||  ||   ||   ||  || | ||
1948 ---C-AGGCCCTGGT--GTTT--GTGGAA---GG---AA--GCCTCAGGC  1997
```

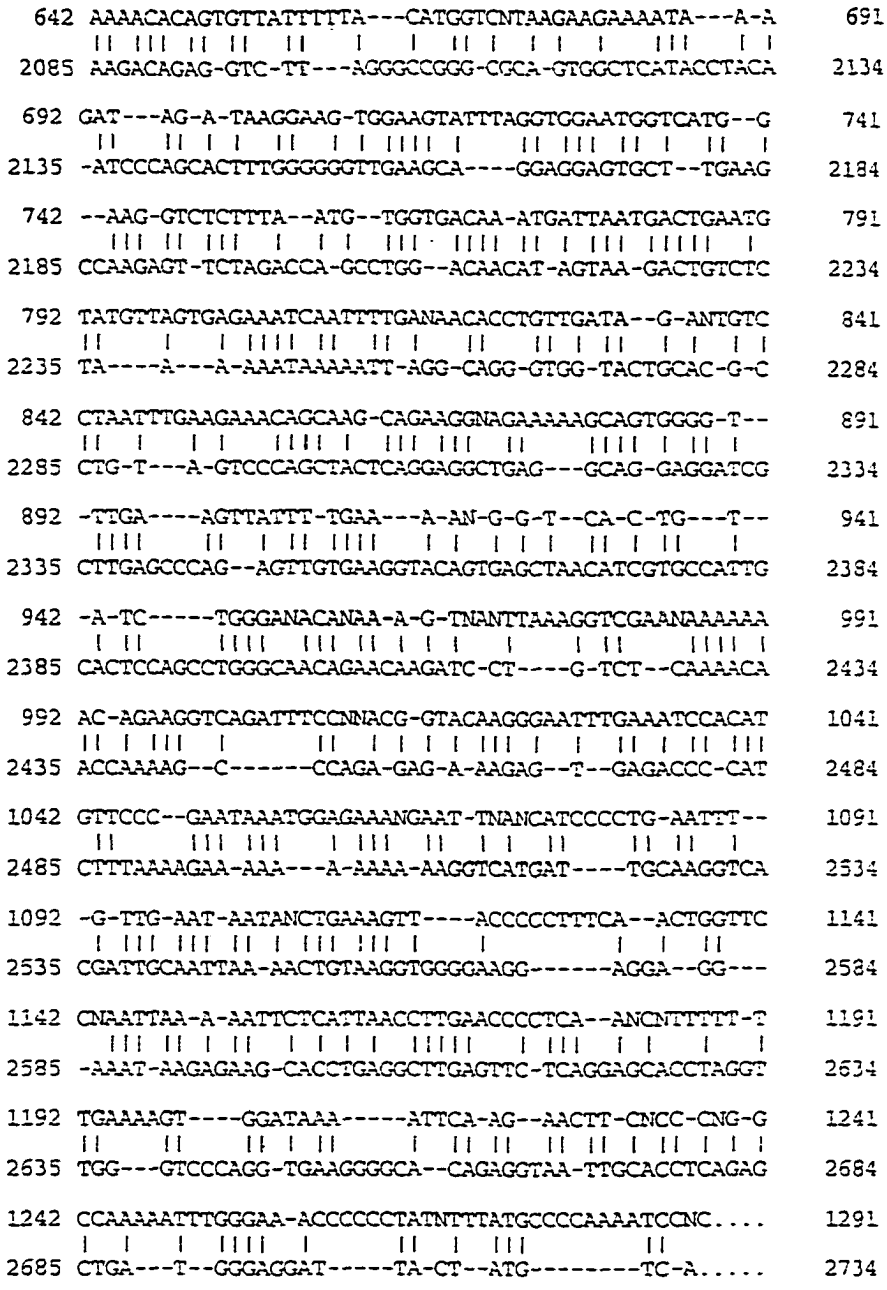
```
 651 AT-TGC-AGTTACGTCGATTGCAGTTCCGCGAG--AGCTTT-CTACGAAG     700
1998 ACCTGCCATTTGC-T-G---GCT-TTCAG-GAGTCATCTTTGCT-C-AGG    2047

 701 ACCTGCGTGCCATTGCTGAG--ATG-GGCCATGTATTTTTC-CTTTTGTT     750
2048 -CC--CGTGC--T-G--G-GCCATGTGG----GTA-----CACTGGTGT-    2097

 751 AG-T--CTAGACGAACA---TAAAA-ACGAACCGGGAATAAACCACCCAG     800
2098 AGGTTGCTGGAC-A-CAGGCTGACTCAC-ATCC---A-TAAAG-AC--AG    2147

 801 ACTTGGAT-TTGATAACCGGGTTAACAA---CTCCAGACTCATA-AATGG     850
2148 A---GG-TCTT-AGGGCCGGGCG--CAGTGGCTC-ATAC-C-TACAAT--    2197

 851 TCCC-GC--TGTTCAAGCATTGTACCATCA--ACGA-TAGATTCT-GCCA     900
2198 -CCCAGCACT-TTGGGGGGTTG-A--AGCAGGAGGAGT-GCTTGAAGCCA    2247

 901 ATTGA-TTAGCTATTTGTCTCGGCTT--ATCAATCCCAGTCCCCGGACTG     950
2248 A--GAGTT--CTA---GAC-CAGCCTGGA-CAA-CATAGTAA--G-ACTG    2297

 951 -----AAGAACACTGAAA-TCTGCACCATGCCAACGTAGTATTTGCACTT    1000
2298 TCTCTAAAAA-A-TAAAAAT-T--AG---GC-AGGGTGGTACT-GCACG-    2347

1001 CCTGCTTAA-CCCGTATGTTACT--GTGTCAA-CTT---CTGTGTAG--T    1050
2348 CCTG-T-AGTCCC--A-GCTACTCAG-G--AGGCTGAGGCAG-G-AGGAT    2397

1051 CGTACTTAAGCC-AGA-TTG--A-GCCGTCCGCAG-GAGGCGCAA-A--G    1100
2398 CG--CTTGAGCCCAGAGTTGTGAAG--GTAC--AGTGAG-CT-AACATCG    2447

1101 -GCAGCGTT-CTC-CCAAGCAACTGG--ATATCTGA-CAGGAAATTCAT-    1150
2448 TGC--CATTGCACTCCA-GC--CTGGGCA-A-CAGAACAAGA--TCC-TG    2497

1151 T-T-AC--CAAC-AAAT-C---GTTGACGAATT--T----CCT--TGCTG    1200
2498 TCTCAAAACAACCAAAAGCCCAGA-GA-GAAAGAGTGAGACCCCAT-CTT    2547

1201 CTAAT-GATAAT---------TCAT-ATTGCACGG-CGA-GCT--CN-TC    1250
2548 -TAAAAGAAAAAAAAAAAAAGGTCATGATTGCAAGGTC-ACGATTGCAAT-    2597

1251 TA---CTGCTTCC-TGG--AA--AT-AC-A--TA-------C-CC--AGC    1300
2598 TAAAACTG-TAAGGTGGGGAAGGAGGAGGAAATAAGAGAAGCACCTGAGG    2647

1301 CTTGC-TTG-CACCA-CACCTA----------A----AACG--C-CAGC    1350
2648 CTTGAGTTCTCAGGAGCACCTAGGTTGGGTCCCAGGTGAAGGGGCACAGA    2697

1351 TGG-A---GC-C-T--G-GCT--TGA-AGCTCCGTATT-CAAGGTCGA..    1400
2698 -GGTAATTGCACCTCAGAGCTGATGGGAG----G-ATTACTATGTC-A..    2747
```

Alignment Of DNA 1/Region 1 (Normal Strand), And The Most Similar-sized, Homologous Area Of Human hsp 27

(The top row shows the DNA 1/Region 1 nucleotide sequence).

(The bottom row shows the human hsp 27 gene sequence).

```
  -8 .........TA---TCTGGTA------CTCTAT--ACCTCTACACACCTT      41
1435 GCGTGGGGGTGGGGTCAGGGAAGAGGGCAC-AGGGACC-C-AC-C-CGGT   1484

  42 -T-TAAAGTGAA---TTGACCTATTTGCTTTTTCTAGAATTTTTAAAATT      91
1485 GTGTAATGT-AACGCTTG-CCT-TTC-CTCT--CT-GCACGTCCAGGCT-   1534

  92 TTTGTTCCACCG-T-TAAACCCCA-----GTTGAT--TC-TCGACT--CA     141
1535 ---GCCCC-CCGGTGTGGACCCCACCCAAGTT--TCCTCCTCC-CTGTCC   1584

 142 ACTGAA---A-ACT-ATTCTTCTTTCGGAAACCTTCATTG---A--CTAT     191
1585 CCTGAGGGCACACTGA--C--CGT--GGAGGCCCCCAT-GCCCAAGCTAG   1634

 192 TCAA--AGTTCTTCTTTATTCTTTAG-TG-CCATCCTCTGTATATTACCC     241
1635 CCACGCAGT-C--CA--A--CG--AGATCACCATCC-CAGTC-A---CC-   1684

 242 TT--ATTA-C---CCTTCAAACTAC----CTATTGTAA-CTACATATT--     291
1685 TTCGAGTCGCGGGCC--CAG-CTTGGGGGC-A--G-AAGCTGCAAAATCC   1734

 292 GTTTTTATGTCCATCT-CC-CCTAA-TAAACAGTGAGCTCCTTCAAGAAA     341
1735 GATG--A-G---A-CTGCCGCC-AAGTAAA--G----C-C-TT-A-GCCC   1784

 342 GAGTATCCTTCCACCTCCCTTTTTCCGCATCACTAGAACAGGGTAGG-TA     391
1785 G-G-ATGC--CCACC-CC-TGCTGCCGC--CACTGG--CTG--T-GCCTC   1834

 392 G---GC-AC-TGAT-TAAGTACATTATTCAGTTCATC-ACTCAGTAT-TG     441
1835 CCCCGCCACCTG-TGT--GTTC-TT-TTGA-TACATTTA-TC--T-TCTG   1884

 442 GACATATATTTCTTGAGAAT----TTACTAAAGG--CCTTGCTTG-CG-T     491
1885 ----T-T-TTTCTC-A-AATAAAGTT-C-AAAGCAACCAC-CT-GTCACT   1934

 492 GGTTCTANGCA-TCGTGAATTGAGTGGTGAACAAAACATTTAA--ACC--     541
1935 GGCCC-AGGCCCTGGTGT-TTG--TGG--AAGGAAGCCTC-AGGCACCTG   1984

 542 CC-TT-G-TC-CT--CATG-G--A-CTTTA-T--G-----T--T---CC-     591
1985 CCATTTGCTGGCTTTCAGGAGTCATCTTTGCTCAGGCCCGTGCTGGGCCA   2034

 592 TG-GGG-ACA--GA---A--TTA--G-ATA-A--C--A-T-ACAT--ATA     641
2035 TGTGGGTACACTGGTGTAGGTTGCTGGACACAGGCTGACTCACATCCATA   2084
```



Alignment Of DNA 1/Region 1

(Complementary Strand), And

The Most Similar-sized,

Homologous Area Of

Human hsp 27

(The top row shows the DNA 1/Region 1 nucleotide sequence).

(The bottom row shows the human hsp 27 gene sequence).

```
  -1 ..GNGGATTTTGGGG-CATA-AANATAGGG----GGGTTTCCCAAATTTT      48
       |  ||    | ||||  ||    ||  |  |||     |||    ||||
1435 GCGTGGGGGT-GGGGTCAGGGAAGA--GGGCACAGGGA--CCCA------    1484

  49 TGGCCNGG-GNG-AA-GTT-C--TTGAATTT--TATC--CACTTTTCAAA      98
        ||  ||  |  |  ||  ||   |  |||  |||  |  ||  |||  |  ||
1485 ---CCCGGTGTGTAATGTAACGCTTGCCTTTCCTCTCTGCACGTC-CA--    1534

  99 AAAANGNTTGAGG---GGT-TCAAGGTTAATG-AG--AA-TTT--T--TA     148
           |  |  |       |||  |    ||   |    |   ||  |||   |   |
1535 ----GGCT-GCCCCCCGGTGT---GG--ACCCCACCCAAGTTTCCTCCTC    1584

 149 A-T-TNG---GAA--C-CAGTTGAAAG-GG-GG------T----AA-CTT     198
       |  |      ||    |  ||  |  ||   |  ||  ||       |     ||  ||
1585 CCTGTCCCCTGAGGGCACACT-GACCGTGGAGGCCCCCATGCCCAAGCT-    1634

 199 TCAGNTATT-ATTCAACAAATTCAGGGGATGNTNAATTCNTTTCTCCATT     248
       ||   |    |  ||   |||    |  |    |  |  |  |    |  |   |  |||  |
1635 --AGCCACGCAGTC--CAA---C-G---A-GATCAC--CAT--C-CCAGT    1684

 249 TA--TTCG-G--GAACATGTGGATTTCAAATTCC--CTTGTAC-C-GTNN     298
       |   ||||  |   |   |   |  ||     |      ||  ||||     |  |
1685 CACCTTCGAGTCG--C--G-GG----C-----CCAGCTTGGGGGCAG-AA    1734

 299 G--G-AAA-TCTGACCTT---CTGTTTTTTTNTTCGACCTTTAANTNA--     348
       |   |  |||  ||  ||   |    |||           ||  ||    ||  |  |
1735 GCTGCAAAATCCGA--TGAGACTG-------C--CG-CC---AAGTAAAG    1784

 349 CTTTNTGTNTCCCAGATACAGTGACCNTTTTCAAAATAACTTCAAACCCC     398
       |  ||   |    ||  |||  |    |||      |     |   ||  |    |  ||
1785 CCTTA-G-C-CCG-GATGCCC--ACCC----C----TG-CTGC---CGCC    1834

 399 ACTG-CTTTTTCTNCCTTCTGCT---TGCTGT-TTCTTCAAATTAGG-AC     448
       ||||  ||  |   ||  ||   |  ||     ||  |||  |||||     ||  |   ||
1835 ACTGGCTGTGCCTCCCC-C-GCCACCTG-TGTGTTCTT----TT-GATAC    1884

 449 ANTCTATCAACAGGTGTTNTCAAAATTGATTTCTCACTAA-CATACATTC     498
       |  |  ||||   |  |  |  ||  |||||  |   |  |||   |   ||  ||  ||   |
1885 ATT-TATCTTCTGTTTTTCTCAAA-TAAAGTTC--A--AAGCA-AC---C    1934

 499 AGTCATTAATCATTTGTC-A--CCACA--T-TA-----AAGAGA--CCTT     548
       |  |  |   |||  |  |  |  |  ||  |   |  |       |||  ||  |||
1935 AC-C-TG--TCACTGGCCCAGGCC-CTGGTGTTTGTGGAAG-GAAGCCT-    1984
```

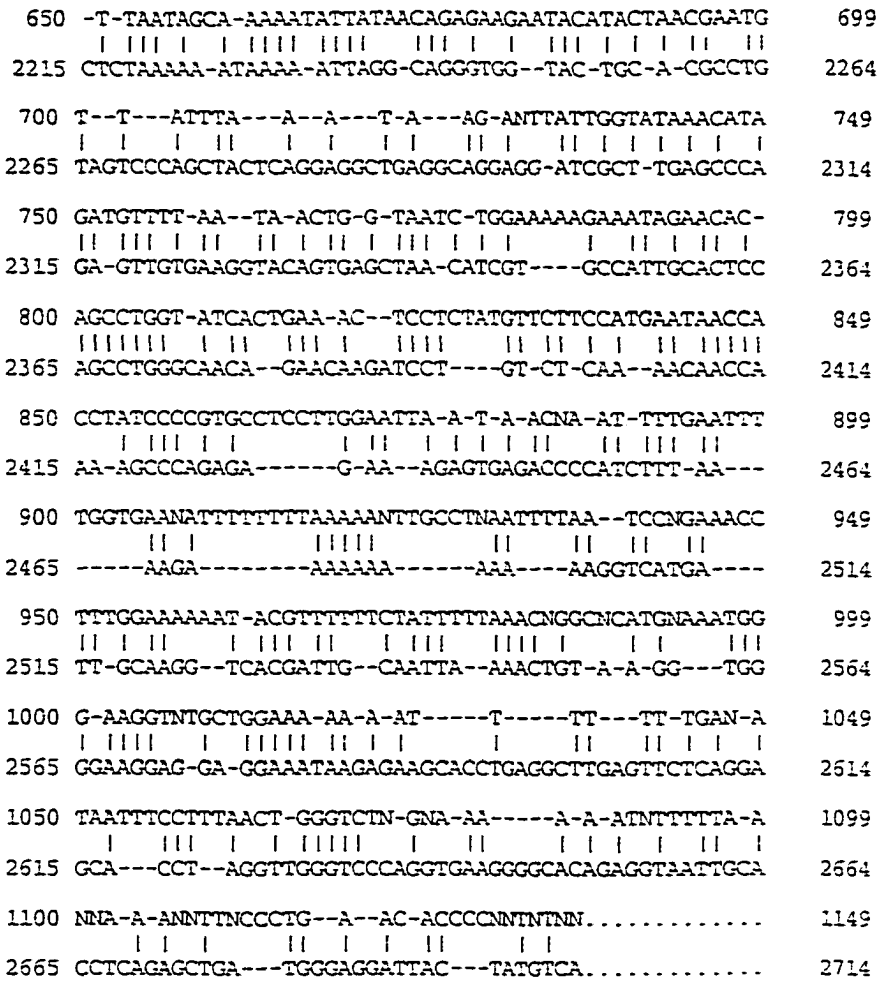
```
 549 CCATG-ACCATTCCACCTAAATACTTCCACTTCCTTATCT--A-TCTTTA     598
     | | | ||| | |||  |   |  | |   |  ||| ||   | ||   |
1985 C-AGGCACC-TGCCA--T---T--TGC---TGGCTT-TCAGGAGTC---A    2034

 599 T-TTT-CTT---C---T--TANGACCATGTAAAAATAACACTGTGTTTTT     648
     | ||| ||    |   |  |  | ||||||     || ||||| || |
2035 TCTTTGCTCAGGCCCGTGCTGGG-CCATGTGGG--TA-CACTG-GTGT--    2084

 649 ATATGTATGTT--ATCTA-ATTCTG--TC-C--CCAGGAACATA-A-AG-     698
     |   || || |  | | | |  |||  || |  |||  || | | | ||
2085 AG--GT-TGCTGGA-C-ACAGGCTGACTCACATCCAT-AA-AGACAGAGG    2134

 699 TCCATGAGGACAAGGG-GTTTAAATGTTTTGTTCACCAC-T-CAATTCAC     748
     ||  | ||| |  ||| |   |  ||    | |||  || | |||| | |
2135 TCT-T-AGGGCC-GGGCGC--AG-TG----GCTCAT-ACCTACAATCC-C    2184

 749 -G-A-T--GCN---TAGAACCACGCAAGCAAG-GCCTTTA-GT-AA-A-T     798
     | | |  |     | ||| || | | | | | || || | |  || | |
2185 AGCACTTTGGGGGGTTGAAGCAGG-AGG-A-GTGC-TTGAAGCCAAGAGT    2234

 799 TCTC-A--AG-----A-AATATATGTCCAATACTGAGTGATGAACTGAAT     848
     |||  |  ||     | || ||| ||  || |||| |  | ||   |||
2235 TCTAGACCAGCCTGGACAACATA-GT--AAGACTGTCTC-TAAAA--AAT    2284

 849 AATGTACTTAATCAG--TGCCTACCT--AC-CCTGT--TCT-AG-TGA-T     898
     ||   | |||  |||  ||  ||| |  || |||||  ||  || | | |
2285 AAAA-A-TTAGGCAGGGTGG-TAC-TGCACGCCTGTAGTCCCAGCT-ACT    2334

 899 GC-GGAAAAAGG--GAGGT-GGAAGGATACTCTTT--C------TTG--A     948
     | |||     ||  |||| ||| |||| | |||   |       |||  |
2335 -CAGGA----GGCTGAGGCAGGA-GGAT-CGCTTGAGCCCAGAGTTGTGA    2384

 949 AGG-A--G----CTCAC-T-GTTT-ATT--AGGGGAGA-TGGACATAAA-     998
     ||| |  |    || || | ||   |||  |    ||  ||| || | |
2385 AGGTACAGTGAGCTAACATCGTGCCATTGCACTCCAGCCTGGGCA-ACAG    2434

 999 AACAATATG-TAGT-TACAATAGGTAGTTTGAAGGGTAATAAGG---GTA    1048
     ||||| ||  | || | ||| |   |     ||    || |||    | |
2435 AACAAGATCCT-GTCT-CAA-A---AC----AACC--AA-AAGCCCAG-A    2484

1049 -ATATACAGAG-GATGGC---A-CT--AAA-GAATAAAGAAGAACTTTGA    1098
     | | | |||| || | |   | ||  ||| ||| ||| || ||     |
2485 GAGA-A-AGAGTGA-GACCCCATCTTTAAAAGAA-AAA-AA-AA-----A    2534

1099 ATAG-TCA--AT-G-AAGGTTTC-CGA----AAG-AAGAA-TAGTTTTCA    1148
     | || |||  || | |||||  | |||    ||  || || | ||    |
2535 A-AGGTCATGATTGCAAGGT--CACGATTGCAATTAA-AACT-GTA---A    2584

1149 GTTGAGTCGAGAATCAACTGG-GGTT-TAACGGTGGAA-CAAAA-A---T    1198
     | || |  | |||  |   || ||   ||| |  | || ||    |   |
2585 GGTG-G--G-GAAGGA---GGAGGAAATAA-GA-G-AAGCACCTGAGGCT    2634

1199 T---TTA--A--A--A---A--TTC--T---AGAAAAAG---CAAATAGG    1248
     |   ||    |  |  |   |  ||   |   ||   |||   || | |||
2635 TGAGTTCTCAGGAGCACCTAGGTTGGGTCCCAGGTGAAGGGGCACAGAGG    2684

1249 TCAATT-CACTTTAAAAGGTG-TGTAGAGGTATAGAGTACCA-GAT-A..    1298
     | |||| ||| | | | | || ||  |||| ||    ||| | | | |
2685 T-AATTGCACCTCAGA-GCTGATGG-GAGG-AT----TACTATG-TCA..    2734
```

Alignment Of DNA 1/Region 2 (Normal Strand), And The Most Similar-sized, Homologous Area Of Human hsp 27

(The top row shows the DNA 1/Region 2 nucleotide sequence).

(The bottom row shows the human hsp 27 gene sequence).

```
   0 .CTGCAGGTCA-ACGGGATCTTTG-G-C----ATG---ATATTTATACAC     49
     ||| ||| || || | | | | | |    |||   | |  ||  |||
1565 CCTG-AGGGCACACTG-ACCGTGGAGGCCCCCATGCCCA-AGCTAGCCAC  1614

  50 T--GTCT--C-ATTTT-CT-TC---GT----TT-GTATATTG-GGG---A     99
       |||   | |  |  |  ||   ||    || | |  | | |||   |
1615 GCAGTCCAACGAGATCACCATCCCAGTCACCTTCG-AG-TCGCGGGCCCA  1664

 100 TGTTTAT----A-ATTATTTTGCAGGATT--ATTTTGA-A-TGT-GAGAA    149
      | ||      | |  |  | |||  ||   ||   || | ||  |  ||
1665 -GCTTGGGGGCAGA--AGCT-GCAAAATCCGAT---GAGACTGCCGCCAA  1714

 150 ATAATATCTTTA------AT---------T--TG-----ATT---T-T-C    199
      ||| | | |||       ||         |  ||     | |   | | |
1715 GTAA-AGCCTTAGCCCGGATGCCCACCCCTGCTGCCGCCACTGGCTGTGC  1764

 200 -T----------T-T-TGTTAGCTAATT-A-A-ATGGTAACTAACT-TT    249
      |          | | ||||  ||  || | | ||  || ||  || ||
1765 CTCCCCCGCCACCTGTGTGTT--CTT-TTGATACATT-TATCTT-CTGTT  1814

 250 TATTCAT-ATAATTAGTATGGTCTCATCACA-CAAGGC-CCTTTGCAACA    299
     | ||| | | ||| |  | | | |||  | | |||  | ||| | || |
1815 T-TTC-TCA-AATAA--A-G-T-TCA--A-AGCAAC-CACCTGT-CA-C-  1864

 300 TTGCTCCTA--CTTCTATGTCTAAGTTCTGCTTCACTTT--CCT-ATGCT    349
     | || || |  |  ||  ||    ||| || |  |      ||| | ||
1865 TGGC-CC-AGGCC-CTG-GT----GTT-TG-TGGAAGGAAGCCTCAGGCA  1914

 350 CCAT-CCATTT-CAATCACCC-CAG-ACGTC-TCTGTATAATTTGGCCAT    399
     || | |||||| |  |  |   ||| | ||| ||| | |  |  ||||
1915 CC-TGCCATTTGC--TGGCTTTCAGGA-GTCATCT-T-TGCTCAGGCCC-  1964

 400 GTCCTAT-CACAACTC--TA-ATCTTGTTTGTA--TTCCTGGA-AT-GCC    449
     || ||   | ||  |   || | || ||  |||  || ||||| |  | |
1965 GTGCTGGGC-CATGTGGGTACA-CTGGT--GTAGGTTGCTGGACACAGGC  2014

 450 CTGAAAAACTCCTAC-TCAGACTTAAAAAGCCAGCTCAAATATTGTGTTC    499
      |||    |||  || ||  | | ||| | | ||   |       | || |
2015 -TGA----CTC--ACATCC-A-T-AAAGA-C-AG---A------G-GT-C  2064

 500 TTCGTGATTCTTTTCCTTCTTCATTGANGTAGANTCTAACACTTCTTCAC    549
     || | |   |    |   | ||  |  | |  || | ||  ||  ||
2065 TTAGGG---C----CGGGCG-CA--G-TG--G-CTC-AT-AC--CT--AC  2114

 550 ACA-CCACTAACAAACTA-GGT---TTGAA--ATGTT-ACTT-TTGTA-C    599
     | | || | | ||  ||  ||    |||||  | |   | |  ||| | |
2115 A-ATCC-C-AGCA--CTTTGGGGGGTTGAAGCAGGAGGAGTGCTTGAAGC  2164

 600 -A----TT-TAAGA--A-CCTTTTTTACATTTTTCATTTTAAAAACAT-T    649
      |    || || ||  | |||     |||     |||  ||| | | | |
2165 CAAGAGTTCTA-GACCAGCCTGG---ACAA----CATAGTAAGA-C-TGT  2214
```



Alignment Of DNA 1/Region 2

(Complementary Strand), And

The Most Similar-sized,

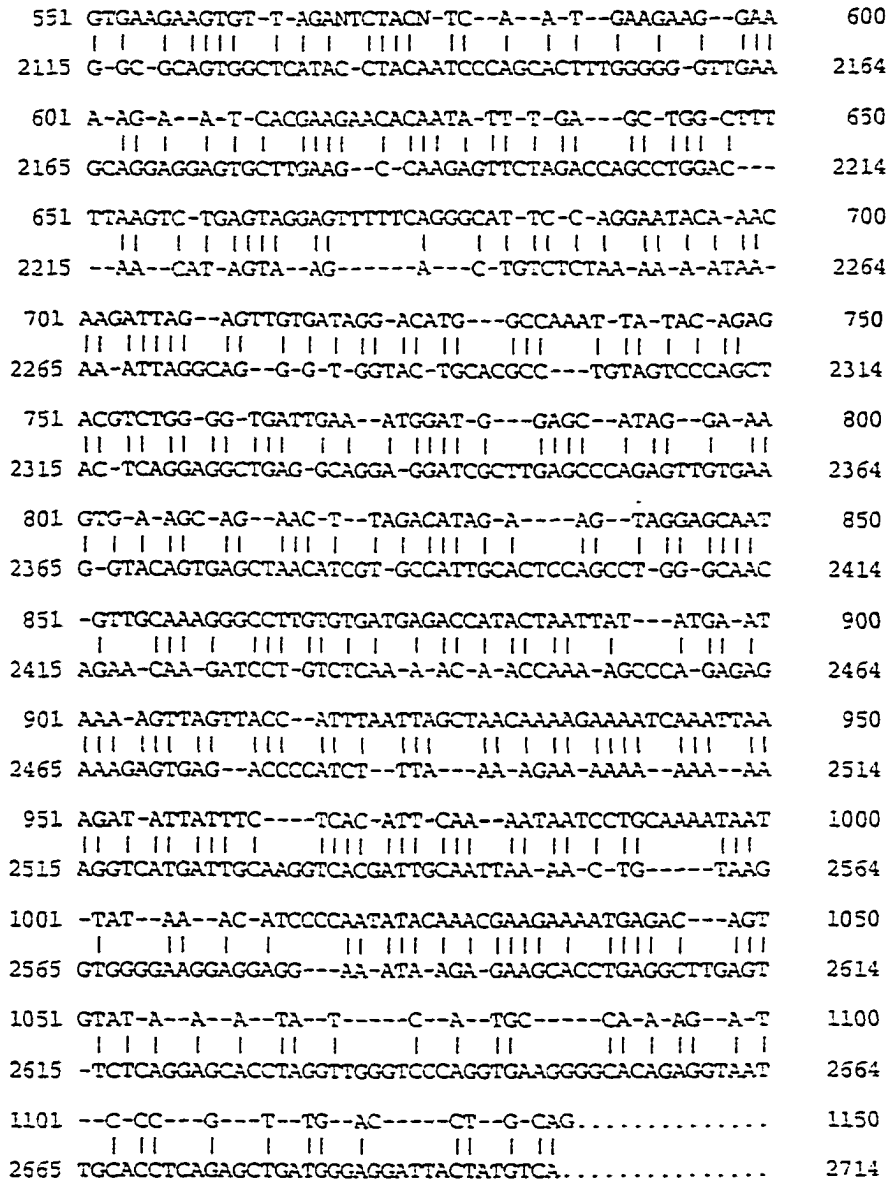
Homologous Area Of

Human hsp 27

(The top row shows the DNA 1/Region 2 nucleotide sequence).

(The bottom row shows the human hsp 27 gene sequence).

```
   1 NNANANNGGGG----TGTTCAGGGNAANNTTTNNTTAAAAANATTTTTTN      50
          |||      ||  | | | |                  ||
1565 CC-T-GAGGGCACACTGACC-GTGGAGGCC--CC--------CATGC---C  1614

  51 CNAG--A-CC-C--AGTTAAAGGAAATTATNTCAAAAAAATTTTTTTCCA     100
     | ||  | || |  |||  || || || |   ||       |      |||
1615 CAAGCTAGCCACGCAGTCCAACGAGATCA-C-CA------TC-----CCA  1664

 101 GCANACCTTCCCATTTNCATGNGCCN-GTTTAAAAATAGAAAAAACGTAT     150
     |  ||||||  | |  |  | |||  | ||       ||||    | |
1665 GT-CACCTTCG-AGTCGC--GGGCCCAGCTTGGGGGCAGAAG---C-TG-  1714

 151 TTTTTCCAAAGGTTTCNGGATTAAAATTNAGGCAANTTTTTAAAAAAAAT     200
       | |||  |  | | || | | |   | ||| |    |||
1715 -----CAAAA--T--CCG-ATGAGACTGCCGCCAAGT----AAAGC----  1764

 201 NTTCA-CCAAAATTCAAAATTNGTTATTA--ATTCCAAGGAG--GC-ACG     250
      || | ||      |       |  ||    |  ||  |  |  || ||
1765 CTT-AGCC------C------GG--ATGCCCACCCCT-GCTGCCGCCACT  1814

 251 GGGATAGGTGGTTATTCATGGAAGAACATAGAGGAGTTTCAGT-GATACC     300
     ||  | | ||  |   |  |  |   | | | |  ||| |  | |||||
1815 GGC-T-G-TGCCTCCCCC-GCCAC--C-T-GTGT-GTT-CTTTTGATAC-  1864

 301 AGGCTGTGT-TCTATTTCTTTTTC-CAGATTACCAGTT-ATTAA--AAC-     350
     |   | | | |||  | | ||||||  || ||  |  ||||  |  ||  |||
1865 A---T-T-TATCT-TCTGTTTTTCTCAAATAA--AGTTCA--AAGCAACC  1914

 351 ATCTATGTTTA-TA-CCAATAANTCT--TATTTA---AAT-AA-CATTCG     400
     | ||  ||  | |  || |     ||  | |||    ||  || | | |
1915 ACCT--GTC-ACTGGCCCAGGCC-CTGGTGTTTGTGGAAGGAAGCCT-C-  1964

 401 TTAGT-A--TGT-ATTCTTCT--CTGTT-ATAA-T-ATTTTTGCT-A---     450
       ||  |  ||  |||  | ||  || || |  | | || ||||||  |
1965 --AGGCACCTGCCATT-TGCTGGCT-TTCAGGAGTCATCTTTGCTCAGGC  2014

 451 ---T--TAAA--ATGTTTTTAAAATGAAAAATGTAAAAAAGGTT-CTT-A     500
        |  |     ||||   || | ||     ||||     |||| ||  |
2015 CCGTGCTGGGCCATGTGGGTACACTGG----TGTA-----GGTTGCTGGA  2064

 501 -A-ATG-T-ACAAA-AGTAACATTTCAA-ACCTAGTTTGTTAGTGGT-GT     550
      | | | | ||  | | |  |||   || ||  || |  |||| ||  |
2065 CACAGGCTGACTCACA-TC-CATA--AAGACAGAGGTC-TTAG-GGCCG-  2114
```

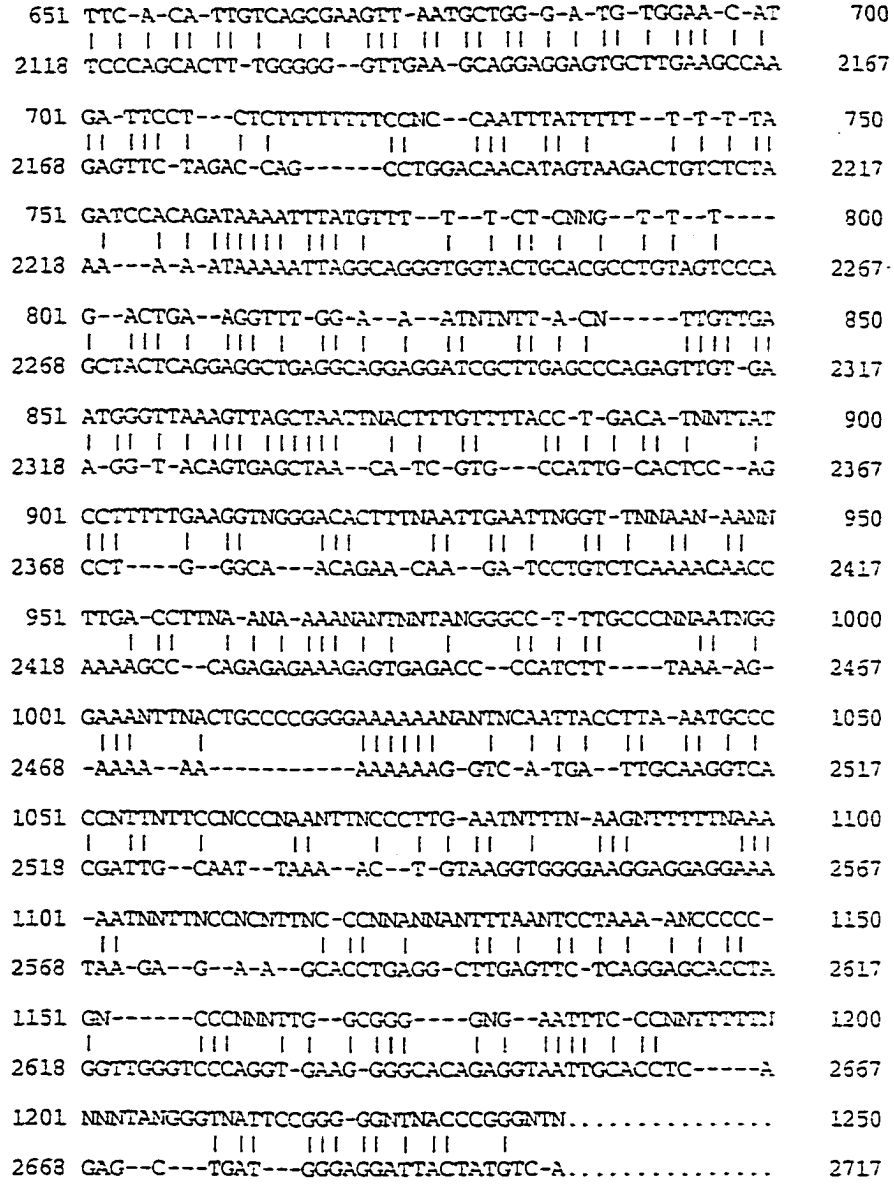


Alignment Of DNA 1/Region 3 (Normal Strand), And The Most Similar-sized, Homologous Area Of Human hsp 27

(The top row shows the DNA 1/Region 3 nucleotide sequence).

(The bottom row shows the human hsp 27 gene sequence).

```
   1 ATAAAAGCAAAGTTTCTA-T-TAA-G-T-GACCAGTGTATT-TATATGCG     50
1468 .....ACCCG-GTGTGTAATGTAACGCTTG-CC--T-T-TCCTCTCTGC-  1517

  51 A--TA-AGAAGATACTACTTT-TAATGAAAT--A---AA-TA-CATAAAC    100
1518 ACGTCCAG--GCTGCCCCCCGGTG-TGGACCCCACCCAAGTTTCCTC--C  1567

 101 AT---T-T----TGAACTGTCTACA--GAAAAACTCAGAGAGATGGTA--   150
1568 -TCCCTGTCCCCTGAGG-G-C-ACACTGA----C-C-GTG-GA-GGCCCC  1617

 151 -ATGA--AA---A-C-ACTCTGTTTATATAAAAGA--ACTTCTTCACTAG   200
1618 CATGCCCAAGCTAGCCACGCAGTCCA-ACGA--GATCAC--CATCCC-AG  1667

 201 TC-CCATAGAATAGAGTGAATAG---AGTCTTGT---CTTA-ATA-CAGC   250
1668 TCACCTTCGAGTCGCG-G----GCCCAG-CTTGGGGGC--AGA-AGCTGC  1717

 251 AT--T--GATTTATTA-TGTN---A--TATATCCATATTCCCG-ATGATC   300
1718 AAAATCCGATG-AG-ACTGCCGCCAAGTAAAGCCTTAG-CCCGGATGCCC  1767

 301 AAGATCTTGCTGGTGTNTA-TGGAGGCTGAGGG-CCGCTTGGCTCA--TG   350
1768 ACCC-CT-GCTGCCGCC-ACTGG---CTGTGCCTCCCCC-G-C-CACCTG  1817

 351 TGAAGAAAGATAAAT-TTTT--TA--TTGTGTCTTGAGTGTTTTAATATG   400
1818 TGT-G-----T---TCTTTTGATACATT-TATCTTC--TGTTTT--TCT-  1867

 401 AAACAAAACTAAA----AAAGAAAA-AGCCT--C--TGTCTTAAGCATCA   450
1868 ---CAAA--TAAAGTTCAAAGCAACCA-CCTGTCACTGGCCCAGGC--C-  1917

 451 CT--TACATTATGTNGAAAGAAGCAATTTTAAA-A--T-C-ATTCGTT--   500
1918 CTGGTG--TT-TGTGGAAGGAAGCC-TC--AGGCACCTGCCATTTGCTGG  1967

 501 CATCATATAAG-AGT-ATCAT--CA-A--CNCN-GAATC--CCATNTAGA   550
1968 C-T--T-TCAGGAGTCATCTTTGCTCAGGCCCGTGC-TGGGCCATGT-G-  2017

 551 GTGATA-A-TGAAGATCT-GTTTTC---ACNCCATGG--GAGCT-AT-T-   600
2018 G-G-TACACTG--G-TGTAGGTTGCTGGACAC-A-GGCTGA-CTCACATC  2067

 601 --T---GA--GTAT-TCTTT---CCTTAA-C--TG-CT--T-C-TGGCAG   650
2068 CATAAAGACAG-AGGTCTTAGGGCCGGGCGCAGTGGCTCATACCTA-CAA  2117
```



Alignment Of DNA 1/Region 3

(Complementary Strand), And

The Most Similar-sized,

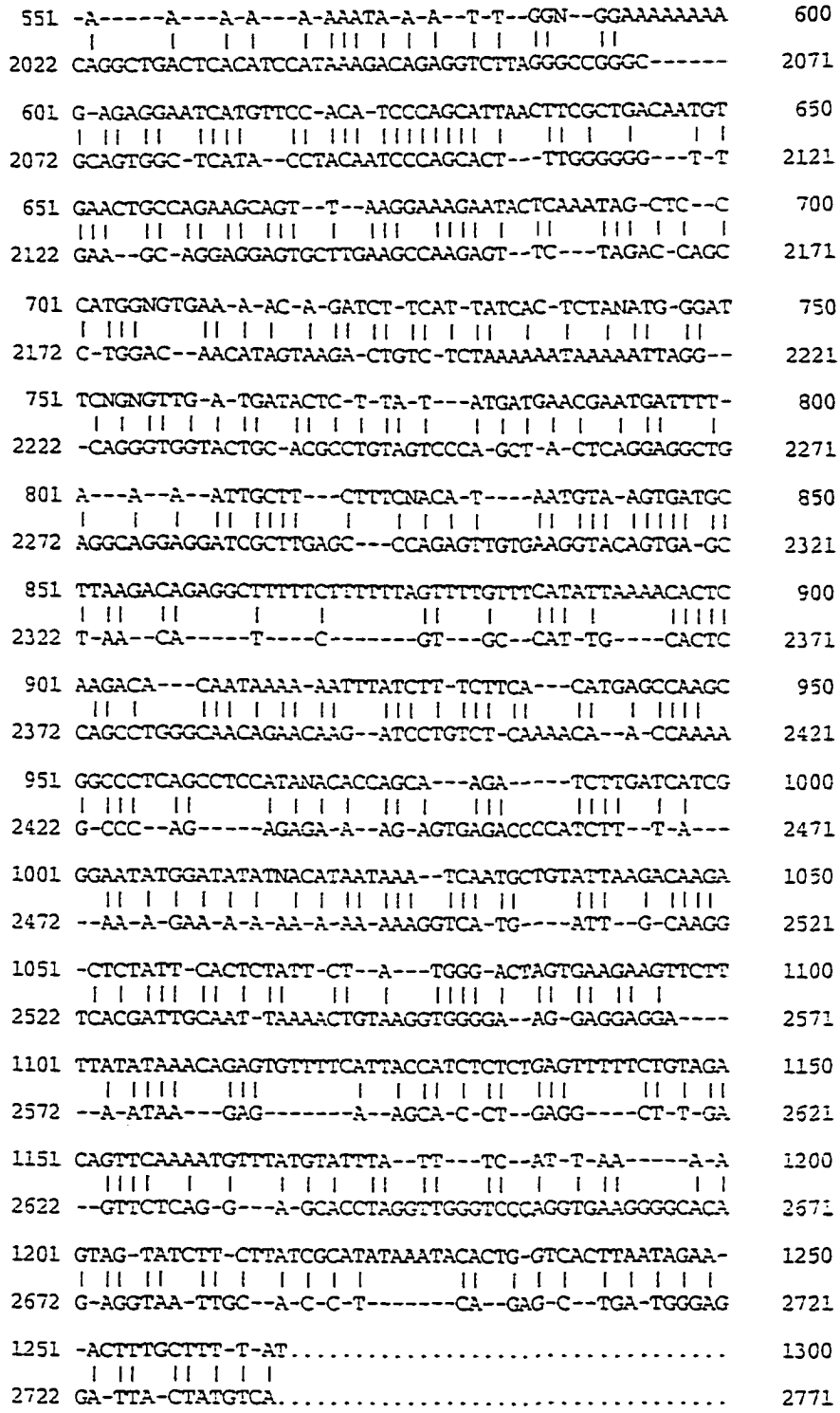
Homologous Area Of

Human hsp 27

(The top row shows the DNA 1/Region 3 nucleotide sequence).

(The bottom row shows the human hsp 27 gene sequence).

```
   1 NANCCCG-G-GTNANC---C-C---CCGGAATNACC-CN-T--ANNNNAA      50
     | || | | || |     | |   ||    | || |  |  |
1472 .ACCCGGTGTGTAATGTAACGCTTGCCT---TT-CCTCTCTGCACGTC--   1521

  51 AAANNGGGAAATTCNCCCCGCCAANNNG-GGNCGGGG-----GNTT--T-    100
       ||    !   |||| ||     | || |          | ||  |
1522 ---CAGGC---T--GCCCC-CC--GGTGTGGACCCCACCCAAGTTTCCTC   1571

 101 -TAGGANT-TAAANTNNTNNGGGNA-AN-GN--G-GNAANN----ATTTT    150
      |     | |        |  ||| | |  |   | | |       ||
1572 CTCC--CTGTC--C-CCTGAGGGCACACTGACCGTGGAGGCCCCCATGCC   1621

 151 TNAAA--A---AN-CT-TN-AA--ANATTCA--AGGGNA---AN-TTNG-    200
       ||   |   |  |  |  ||  | || ||  |    |   |  || |
1622 -CAAGCTAGCCACGCAGTCCAACGAGAT-CACCATCCCAGTCACCTTCGA   1671

 201 G--GNGGAANAAN---GGGGGCATTTAAGGT--AA--T----TGN-ANTN    250
     |  | ||     |    |||||||   ||| |  ||  |    ||  | |
1672 GTCGCGGGCCCAGCTTGGGGGCAG--AAGCTGCAAAATCCGATGAGACTG   1721

 251 TTTTTTCCCCGGGGCA-GTNAANT-TT--CCCN-ATTNN-------G--G    300
           || |    || || ||   ||  |||  ||           |  |
1722 ------CCGC----CAAGTAAAGCCTTAGCCCGGATGCCCACCCCTGCTG   1771

 301 --GCAAA-GGCC---CNTANN-----AN-TNT-T-TTNTTNAAGGT-CA-    350
       || |  |||     | |         |  | | | | || ||    | | ||
1772 CCGCCACTGGCTGTGCCTCCCCCGCCACCTGTGTGTTCTTTT-GATACAT   1821

 351 --ANNTTNT-TNNAAC-CNAATTCAA-TTNAAAGTGTCC-CNACCT-TCA    400
       |  || | |    | | |||  || || ||||   |  | |||| |||
1822 TTATCTTCTGTTTTTCTCAAATA-AAGTTCAAAG---CAACCACCTGTCA   1871

 401 AAA----AGGA--TAANNAT-GTCAGGTAA--AA-CAA-AGTNAATTAGC    450
          |||    |      | ||  || ||  || |   ||  |  | ||
1872 CTGGCCCAGGCCCTGGTGTTTGT--GG-AAGGAAGCCTCAGGCACCT-GC   1921

 451 TAACTTTAACC--CATTCAACAANGTAANANATTT-C-CAAACCTTCAGT    500
      |  |||  |   | ||||  |   | |    ||| | ||  ||  | ||
1922 CA--TTTG-CTGGCTTTCAGGA-G-TCAT-C-TTTGCTCAGGCC--C-GT   1971

 501 CAAACNNGAGAAAAACATA-----A-ATT--T-TA--TCTG-TGGATCTA    550
       |  | |     |||      | | |  | ||  | ||  |||| | |
1972 G---CTGG-GC----CATGTGGGTACACTGGTGTAGGT-TGCTGGA-C-A   2021
```

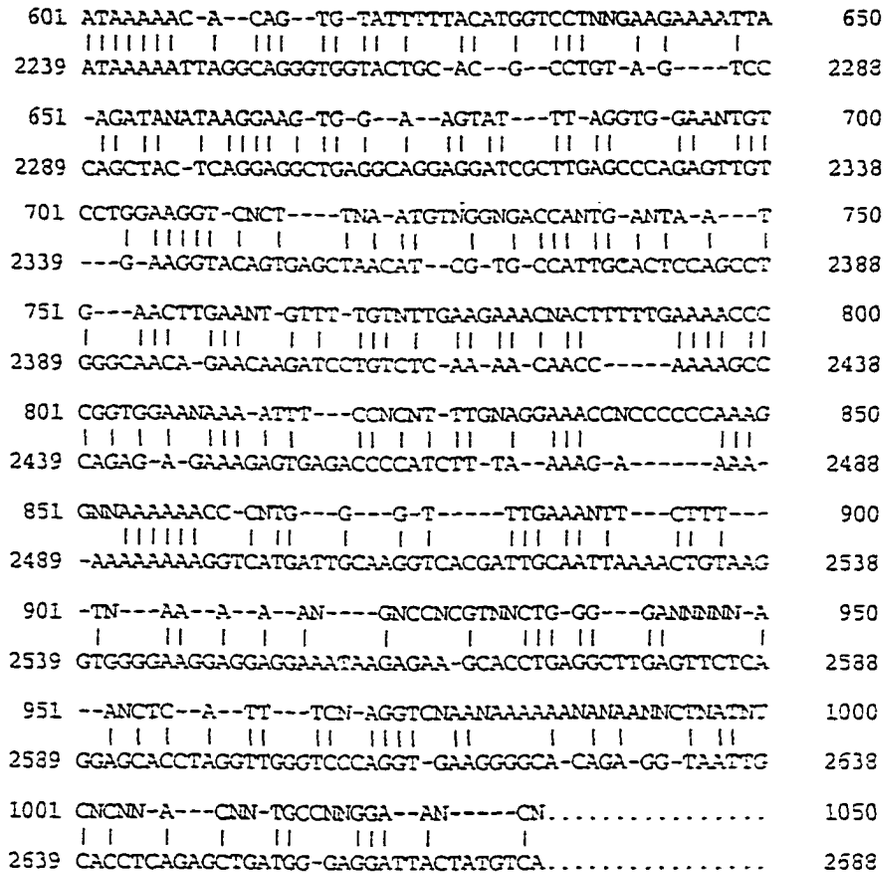


Alignment Of DNA 2/Region 1 (Normal Strand), And The Most Similar-sized, Homologous Area Of Human hsp 27

(The top row shows the DNA 2/Region 1 nucleotide sequence).

(The bottom row shows the human hsp 27 gene sequence).

```
   1 GAC-TATCTGGTAAGTACTCTATACCTCTACACACCTTTT----A-AAG-     50
     || | || |  | || | |    || | |  |  ||       | |||
1639 CACCT-TC-G--A-GT-CGCGGG-CC-C-AG-C---TTGGGGGCAGAAGC  1688

  51 TG-AATTGA-CCTATTTG-CTTTTTCTAG--AATTTTTAAAATTTTTGTT    100
     || ||   | || ||  | ||    |  |  || |   |||   || |
1689 TGCAAA--ATCCGATGAGACTGC--C--GCCAAGT---AAAGCCTTAG--  1738

 101 CCACCGTTAA--ACCCCAGTTGATTCTCGACTCAACTGAAAACTATTCTT    150
     || | | |    ||||| | ||   | || | || |||    || | | |
1739 CC-CGGATGCCCACCCCTGCTG---C-CG-C-CA-CTGG---CTGTGCCT  1788

 151 CTTTCGGAAACCTTCATTG--AC-TAT-TCAAAGTTCTTCTTTATTC-TT    200
     |   |     ||  |   | || | | |    |||||| || || | ||
1789 C---C-----CC--C---GCCACCTGTGT----GTTCTT-TTGATACATT  1838

 201 TAGTGCCATCCTCTGTATATTACCCTTATTACCCTTCAAA-CTACCTAT-    250
     || | |  | ||||| | || | |  || |   |||||| | ||| |
1839 TA-T-C--T--TCTGTTT-TT-CTCAAATAAAG-TTCAAAGCAACC-ACC  1888

 251 TGTAACTA-CATAT----TG-T-TTT-T--ATGTCCATC-TC---C-CCT    300
     ||| |||  |  |      || | ||| |  | |   | | ||   | |||
1889 TGTCACTGGCCCAGGCCCTGGTGTTTGTGGAAGGA-AGCCTCAGGCACCT  1938

 301 A--ATAAACAGTGAGCTCCTTCAAGAAAGANT-ATCCTT-C-CA--CC--    350
       ||   |  || |||  |||| |   || | ||| || | ||  ||
1939 GCCATTTGC--TG-GCT--TTCA-G---GAGTCATCTTTGCTCAGGCCCG  1988

 351 T-C----CCTTTTTCCGCATCACTAAAACAGG-GTAGGTAGGCACTGATT    400
     | |    || | |   | | ||||       || |||||| | |  ||
1989 TGCTGGGCCATGTGG-GTA-CACT------GGTGTAGGTTG-C--TG---  2033

 401 AAGTACATTATTCAGTTCATCACTCAGTATTGGACATATATTTCTTGAGA    450
       | |||     |||  | | |||||  ||    |||| |   |  |||
2039 --G-ACA-----CAGG-C-TGACTCAC-ATC---CATAAAGA-CA-GAGG  2088

 451 ATTTACTAAAGGCCTTGCTTGC-GTGG-T--T-C-TAGGCA-TCGT-GAA    500
     | |  ||  |||| ||  || |||| |  | | ||  || ||   | |
2089 -TCT--TAG-GGCCGGGC--GCAGTGGCTCATACCTA--CAATCCCAGCA  2138

 501 ATT-GAGTGGT-GAA-CAAAAC-A-T--TTAAACCCCTT-GTCCTCATGG    550
     || | | ||| ||| ||  |  | |  || || ||    || || | |
2139 CTTTGGGGGGTTGAAGCAGGAGGAGTGCTTGAAGCCAAGAGTTCT-A-G-  2188

 551 GACTTTATGTTCCTGGGGANAGA-AT--TA-GA-T------AACATACAT    600
     ||   | |  |||||  | | | ||  || || |       || | | |
2189 -ACC--A-G--CCTGG--ACA-ACATAGTAAGACTGTCTCTAA-A-A-A-  2238
```



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Alignment Of DNA 2/Region 1

(Complementary Strand), And

The Most Similar-sized,

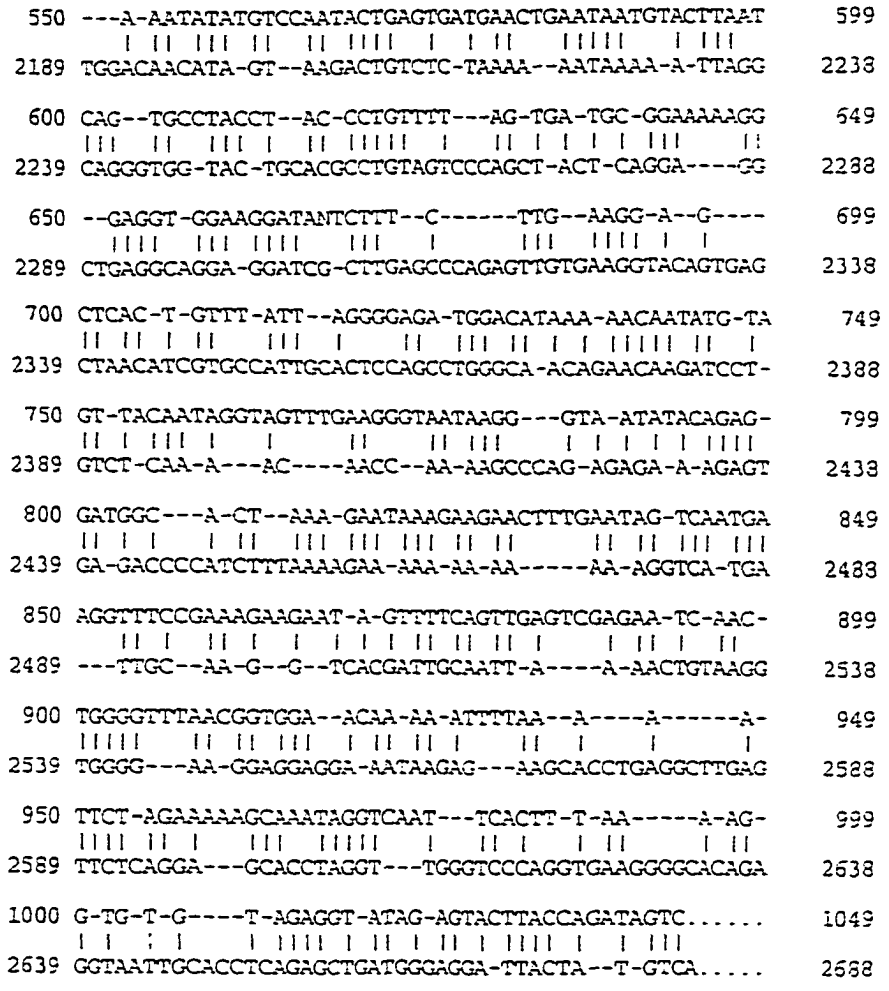
Homologous Area Of

Human hsp 27

(The top row shows the DNA 2/Region 1 nucleotide sequence).

(The bottom row shows the human hsp 27 gene sequence).

```
   0 .NGNTTC----CNNGG-CANNG-TNNGNG--ANA---TN-AGNNTTNTNT     49
1639 CACCTTCGAGTCGCGGGCCCAGCTTGGGGGCAGAAGCTGCAAAATCCGAT  1688

  50 TTTTTNTTNGACCTN--GA-AA-TGA-GN-TTNNNNN---TCCCCAGNNA     99
1689 -----G--AGAC-TGCCGCCAAGTAAAGCCTTAGCCCGGATGCCCACCC-  1738

 100 C-GN-G--GNCNTTTTTNAAAAGAANT-T---TCAAACCCANGGGTTTTT    149
1739 CTGCTGCCGCCACT---G----G--CTGTGCCTC---CCC-CGC------  1788

 150 TNNCCTTTGGGGGGNGGTTTCCTNCAAANGNGGA-A-ATTT-TNTTCCAC    199
1789 -CACCTGTG-----TG-TT-C-TT----T-TG-ATACATTTATCTTCT--  1838

 200 CGGGGTTT-TCAAA-AAGTNGTTTCTTCAANACAAAACAN-T-TCAA-GT    249
1839 -GTT-TTTCTCAAATAA--AGTT-CA--AAG-CAAC-CACCTGTCACTGG  1888

 250 TCATTANTCANTGGTCNCCNACAT-TN-AAGNGA-CCTTCCAGGACANTT    299
1889 CCC--AGGCCCTGGTGT--T---TGTGGAAGGAAGCCT-C-AGG-CACCT  1938

 300 -CCACCTAAATACTTCCACTTCCTTATNTA-TCTTAAT-TTT-CT-----    349
1939 GCCAT-T---TGCTGGCT-TTC---AGG-AGTC--A-TCTTTGCTCAGGC  1988

 350 ---T-CNNAGGACCATGTAAAAATACACTGTGTTTTTATATGTATGTT--    399
1989 CCGTGCTG-GG-CCATGTGGG--TACACTG-GTGT--AG--GT-TGCTGG  2038

 400 ATCTA-ATTCTN--TC-C--CCAGGAACATA-A-AG-TCCCATGAGGACA    449
2039 A-C-ACAGGCTGACTCACATCCAT-AA-AGACAGAGGTCTTA-G-GGCC-  2088

 450 AGGG-GTTTAAATGTTTTGTTCACCAC-T-CAATTTCACG-A-T--GCC-    499
2089 -GGGCGC--AG-TG----GCTCAT-ACCTACAATCCCA-GCACTTTGGGG  2138

 500 --TAGAACCACGCAAGCAAG-GCCTTTA-GT-AA-A-TTCTC-A--AG--    549
2139 GGTTGAAGCAGG-AGG-A-GTGC-TTGAAGCCAAGAGTTCTAGACCAGCC  2188
```

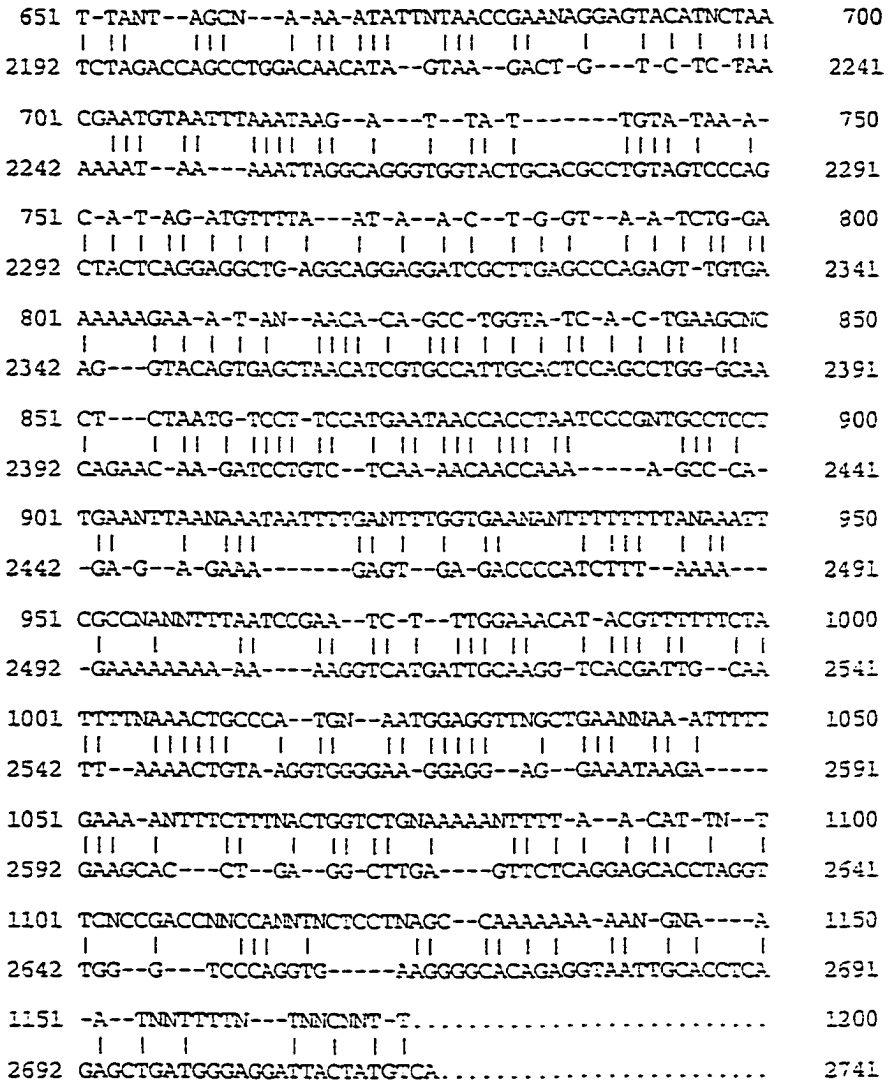


Alignment Of DNA 2/Region 2 (Normal Strand), And The Most Similar-sized, Homologous Area Of Human hsp 27

(The top row shows the DNA 2/Region 2 nucleotide sequence).

(The bottom row shows the human hsp 27 gene sequence).

```
   1 CTGCAGGT-CAACGGGATCTT---TGGCA--TGATATTTATACACTGTCT      50
       || | |  |    || |   || |   |||     |  |||||
1542 ....AGTTTC--C----TCCTCCCTGTCCCCTGAGGGC-A--CACTG---  1591

  51 CATTTTCTTCGTTTGTATATTGGGG---ATGTTTATAATTATTTTGC-AG     100
      |    | |||  | |    ||     |||   | |  ||     || |
1592 -A----C--CGTG-G-A----GGCCCCCATGCCCA-AGCTA----GCCAC  1641

 101 GATTATTTTGAATGTGAGAAAT-A--AT---A-TCTTTAATTT-GATTTT     150
      |  | |   || |  |||  | |  ||   | ||   |  || || |
1642 GC--AGTCC-AACG--AGA--TCACCATCCCAGTC---ACCTTCGAGTCG  1691

 151 CTTTTGTT-AGCTAATTAAATGGTAACTAA-CTTTTATTCATATAATTAG     200
      |    |  ||||  |     || |  || ||      || | |||  |
1692 CGG--GCCCAGCT--TGGG--GGCAG--AAGCTG-----CA-A-AATCCG  1741

 201 TATG-GTCT-CATCACACAAG----GCCCTTTGCAAC--ATTGCTCCTAC     250
      ||| | || |  | | ||||    ||| || ||  |  || || || ||
1742 -ATGAGACTGC--CGC-CAAGTAAAGCC-TTAGCC-CGGAT-GC-CC-AC  1791

 251 TTCTATGTCTAAGTTCTGCTTCACTTTCCTATGCTCCATCCATTTCAATC     300
       ||  | ||  |  | ||  ||||  | | ||| |  |||    |  |
1792 CCCT--G-CT--GC-C-GC--CACTGGC-TGTGC-C--TCC----C---C  1841

 301 ACCCCAGACGTCTCTGTATAATTTGGCCATGTCCTATCACAACTCTAATC     350
      | |||  | | | ||| |  |||     ||    || |||  | || ||
1842 -CGCCAC-C-TGTGTGT-TC-TTT-----TG----AT-ACAT-T-TA-TC  1891

 351 TTGT-TTGTATTCCTGGAAT---GCCCTGAAA--AACT-CCT---ACT--     400
      || | || | ||| |  |||   |  | |||  |||  |||   |||
1892 TTCTGTT-T-TTC-TCAAATAAAGTTC--AAAGCAACCACCTGTCACTGG  1941

 401 C--AGAC--T--TA-------AA--AAGCC--AG-C---T-CAAATATTG     450
      |  || |  |  |        ||  |||||  || |   | | | | |||
1942 CCCAGGCCCTGGTGTTTGTGGAAGGAAGCCTCAGGCACCTGCCA-T-TTG  1991

 451 -TGTTCTT-C-GTGA-T--TCTTT--TC---CT--T-CTT--C-AT-TGA     500
      ||  ||| | | || |  |||||  ||   |   | ||   | || ||
1992 CTGG-CTTTCAG-GAGTCATCTTTGCTCAGGCCCGTGCTGGGCCATGTG-  2041

 501 GGTA----G----AG-T--CT--A-ACA--CTT-CTTCACACACCACTAA     550
      ||||    |    || |  ||  | |||  ||  || ||||  ||| |||
2042 GGTACACTGGTGTAGGTTGCTGGACACAGGCTGACT-CACAT-CCA-TAA  2091

 551 CAAACT-AGGT-TT-GAA---------A-TG--T--TACTTTTGTACATTT     600
      | ||  |||| || |           | ||  |  ||| |   ||| |
2092 -AGACAGAGGTCTTAGGGCCGGGCGCAGTGGCTCATACCT----ACAATC  2141

 601 A-AGAANCTTTT-----TT-A--CATTTTTCATT--TTAAA---AACATT     650
       || |  |||       || |  ||       | |  || ||   || | |
2142 CCAGCAC-TTTGGGGGGTTGAAGCAGGAGG-AGTGCTTGAAGCCAAGAGT  2191
```



Alignment Of DNA 2/Region 2

(Complementary Strand), And

The Most Similar-sized,

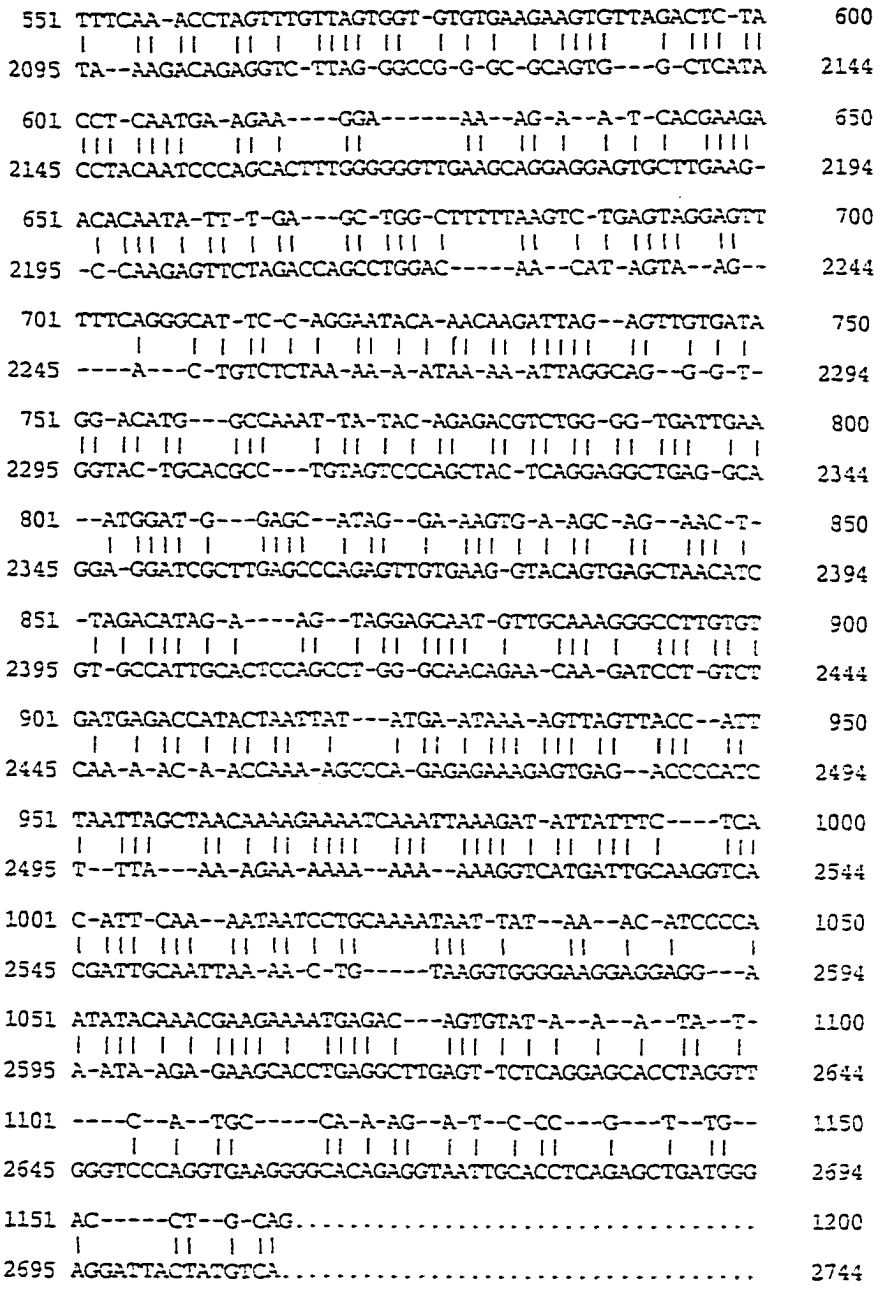
Homologous Area Of

Human hsp 27

(The top row shows the DNA 2/Region 2 nucleotide sequence).

(The bottom row shows the human hsp 27 gene sequence).

```
   1 AANNGNNANAAAAANNATTTNCNTTTTTTTTTGGCTNAGGAGNANN-TGGN      50
     |                 |   | | |         || ||| | |   ||
1545 .AGTTTCCT----CC-T--CCCTGTCCC-----CTGAGG-GCACACTGAC    1594

  51 NGGTCGGNGAANA--ATGTTAAAANTTTTTNCA-GACCAGTNAAA-GAAA     100
     | | || |       |||   ||  |     || | | |||  || || |
1595 CG-T-GGAGGCCCCCATGCCCAAGCTAG--CCACG-C-AGTCCAACGAGA    1644

 101 NTTTTCAAAAATTTNNTTCAGCNAACCTCC-ATTN-CATGGGC--AGTTT     150
          ||  | |       |||  | ||| | | |  |  ||||  || ||
1645 T----CACCA-T--CC--CAGTCA-CCTTCGAGTCGC--GGGCCCAGCTT    1694

 151 NAAAATAGAAAAAACGTATGTTTCCA-AAGATTCGGATTAAANNTNGGCG     200
         |       |   |   | | ||| | |  |  |||    |
1695 GG-----G-------G---G---C-AGAAGCTGC--A--AAATC-CG---    1744

 201 AATTTNTAAAAAAAAANTNTTCACCAAANTCAAAATTATTTNTTAANTTCA     250
     | |    | |       |   | ||||  | |||         |||    |
1745 A-T--G-AGA-----CTGC-CGCCAA-GT-AAAGC-----CTTAGCC-C-    1794

 251 AGGAGGC--ANC---G--G--G--ATTAGG-TGGTTATTCATGGAAGG-A     300
     |||  ||  | |   |  |  |  | | ||  || |   ||       | |
1795 -GGATGCCCACCCCTGCTGCCGCCACT-GGCTG-TGCCTCCCCC--GCCA    1844

 301 CATTAGAGGNGCTTCAGT-GATACCAGGCTGTGTTNTATTTCT-TTTTTC     350
     | |  | |  | |||  | |||||  |   | | |  | || || ||||||
1845 CCT--GTG-TG-TTCTTTTGATAC-A---T-T-T-ATCTT-CTGTTTTTC    1894

 351 -CAGATTACCAGTT-ATTAA--AAC-ATCTATGTTTA-TA-CA-ATAATC     400
      || || |  |||| |  ||  ||| | ||  ||  | |  |  |    |
1895 TCAAATAA--AGTTCA--AAGCAACCACCT--GTC-ACTGGCCCAGGCCC    1944

 401 T--TATTT----AA--ATTA-CATTCGTTAGNATGTAC-T-CCTNTT-C-     450
     |  | |||    ||  |  | | | |   ||     || | ||  || |
1945 TGGTGTTTGTGGAAGGA--AGCCT-C---AGGC---ACCTGCCATTTGCT    1994

 451 GG-TTANA--A-T-ATTTTNGCT-ANTAAAATGTT------T-T---TAA     500
     || ||  |  | | || || ||| |      || |       | |   ||
1995 GGCTTTCAGGAGTCATCTTTGCTCAGGCCCGTGCTGGGCCATGTGGGTAC    2044

 501 AATGAAAAATGTAAAAAAGNTT-CTT-A-A-ATG-T-ACAAA-AGTAACA     550
     | ||     ||||      | || ||  | | | | | ||  | | |  ||
2045 ACTGG----TGTA-----GGTTGCTGGACACAGGCTGACTCACA-TC-CA    2094
```



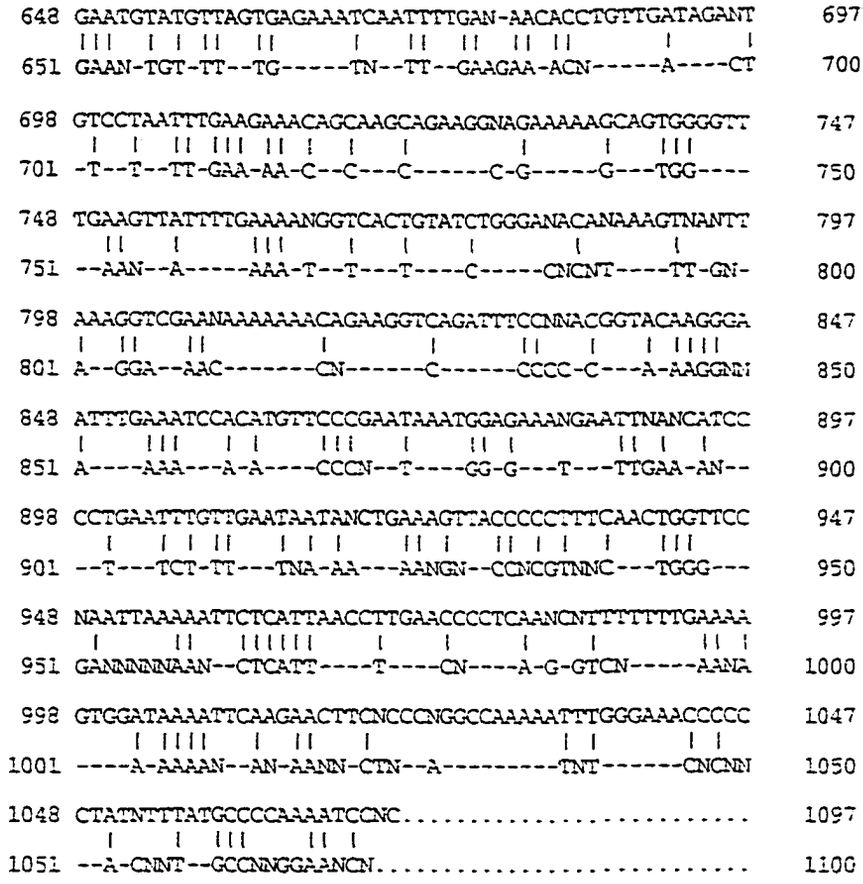
APPENDIX VI

Alignment Of DNA 1/Region 1 (Normal Strand), And DNA 2/Region 1 (Normal Strand)

(The top row shows the DNA 1/Region 1 nucleotide sequence).

(The bottom row shows the DNA 2/Region 1 nucleotide sequence).

```
  -2 ...TATCTGGTA----CTCTATACCTCTACACACCTTTTAAAGTGAATTG    47
     |||||||||    ||||||||||||||||||||||||||||||||||
   1 GACTATCTGGTAAGTACTCTATACCTCTACACACCTTTTAAAGTGAATTG    50

  48 ACCTATTTGCTTTTTCTAGAATTTTTAAAATTTTTGTTCCACCGTTAAAC    97
     |||||||||||||||||||||||||||||||||||||||||||||||||
  51 ACCTATTTGCTTTTTCTAGAATTTTTAAAATTTTTGTTCCACCGTTAAAC   100

  98 CCCAGTTGATTCTCGACTCAACTGAAAACTATTCTTCTTTCGGAAACCTT   147
     |||||||||||||||||||||||||||||||||||||||||||||||||
 101 CCCAGTTGATTCTCGACTCAACTGAAAACTATTCTTCTTTCGGAAACCTT   150

 148 CATTGACTATTCAAAGTTCTTCTTTATTCTTTAGTGCCATCCTCTGTATA   197
     |||||||||||||||||||||||||||||||||||||||||||||||||
 151 CATTGACTATTCAAAGTTCTTCTTTATTCTTTAGTGCCATCCTCTGTATA   200

 198 TTACCCTTATTACCCTTCAAACTACCTATTGTAACTACATATTGTTTTTA   247
     |||||||||||||||||||||||||||||||||||||||||||||||||
 201 TTACCCTTATTACCCTTCAAACTACCTATTGTAACTACATATTGTTTTTA   250

 248 TGTCCATCTCCCCTAATAAACAGTGAGCTCCTTCAAGAAAGAGTATCCTT   297
     |||||||||||||||||||||||||||||||||||||||||| |||||||
 251 TGTCCATCTCCCCTAATAAACAGTGAGCTCCTTCAAGAAAGANTATCCTT   300

 298 CCACCTCCCTTTTTCCGCATCACTAGAACAGGGTAGGTAGGCACTGATTA   347
     |||||||||||||||||||||||| ||||||||||||||||||||||||
 301 CCACCTCCCTTTTTCCGCATCACTAAAACAGGGTAGGTAGGCACTGATTA   350

 348 AGTACATTATTCAGTTCATCACTCAGTATTGGACATATATTTCTTGAGAA   397
     |||||||||||||||||||||||||||||||||||||||||||||||||
 351 AGTACATTATTCAGTTCATCACTCAGTATTGGACATATATTTCTTGAGAA   400

 398 TTTACTAAAGGCCTTGCTTGCGTGGTTCTANGCATCGTGAA-TTGAGTGG   447
     |||||||||||||||||||||||||||| |||||||||| ||||||||
 401 TTTACTAAAGGCCTTGCTTGCGTGGTTCTAGGCATCGTGAAATTGAGTGG   450

 448 TGAACAAAACATTTAAACCCCTTGTCCTCATGG-ACTTTATGTTCCTGGG   497
     |||||||||||||||||||||||||||||||| ||||||||||||||||
 451 TGAACAAAACATTTAAACCCCTTGTCCTCATGGGACTTTATGTTCCTGGG   500

 498 GACAGAATTAGATAACATACATATAAAAACACAGTGTTATTTTTACATGG   547
     || ||||||||||||||||||||||||||||| ||||||||||||
 501 GANAGAATTAGATAACATACATATAAAAACACAGTGT-ATTTTTACATGG   550

 548 TCNTAAGAAGAAAATAAAGATAGATAAGGAAGTGGAAGTATTTAGGTGGA   597
     || |  ||||||||| |||||| ||||||||||||||||||||||||||
 551 TCCTNNGAAGAAAATTAAGATANATAAGGAAGTGGAAGTATTTAGGTGGA   600

 598 ATGGTCATGGAAGGTCTCTTTAATGTGG-TGACAAATGATTAATGA-CT-   647
     |  ||| |||||||| ||| ||||| |  ||| | ||| |||||| ||
 601 ANTGTCCTGGAAGGTCNCTTNAATGTNGGNGACCANTGANTAATGAACTT   650
```

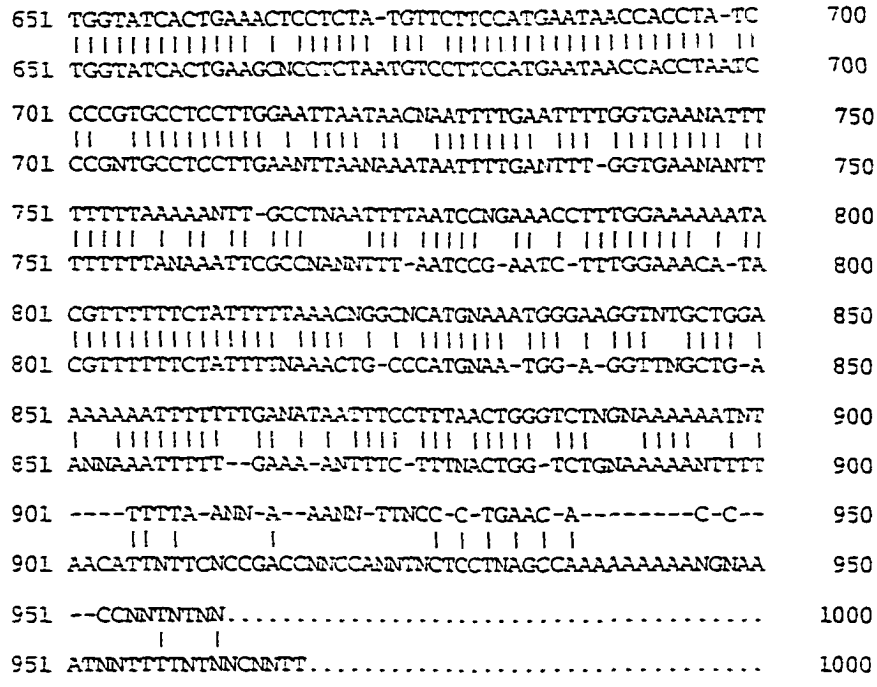


Alignment Of DNA 1/Region 2 (Normal Strand), And DNA 2/Region 2 (Normal Strand)

(The top row shows the DNA 1/Region 2 nucleotide sequence).

(The bottom row shows the DNA 2/Region 2 nucleotide sequence).

```
   1 CTGCAGGTCAACGGGATCTTTGGCATGATATTTATACACTGTCTCATTTT   50
     ||||||||||||||||||||||||||||||||||||||||||||||||||
   1 CTGCAGGTCAACGGGATCTTTGGCATGATATTTATACACTGTCTCATTTT   50

  51 CTTCGTTTGTATATTGGGGATGTTTATAATTATTTTGCAGGATTATTTTG  100
     ||||||||||||||||||||||||||||||||||||||||||||||||||
  51 CTTCGTTTGTATATTGGGGATGTTTATAATTATTTTGCAGGATTATTTTG  100

 101 AATGTGAGAAATAATATCTTTAATTTGATTTTCTTTTGTTAGCTAATTAA  150
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 101 AATGTGAGAAATAATATCTTTAATTTGATTTTCTTTTGTTAGCTAATTAA  150

 151 ATGGTAACTAACTTTTATTCATATAATTAGTATGGTCTCATCACACAAGG  200
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 151 ATGGTAACTAACTTTTATTCATATAATTAGTATGGTCTCATCACACAAGG  200

 201 CCCTTTGCAACATTGCTCCTACTTCTATGTCTAAGTTCTGCTTCACTTTC  250
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 201 CCCTTTGCAACATTGCTCCTACTTCTATGTCTAAGTTCTGCTTCACTTTC  250

 251 CTATGCTCCATCCATTTCAATCACCCCAGACGTCTCTGTATAATTTGGCC  300
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 251 CTATGCTCCATCCATTTCAATCACCCCAGACGTCTCTGTATAATTTGGCC  300

 301 ATGTCCTATCACAACTCTAATCTTGTTTGTATTCCTGGAATGCCCTGAAA  350
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 301 ATGTCCTATCACAACTCTAATCTTGTTTGTATTCCTGGAATGCCCTGAAA  350

 351 AACTCCTACTCAGACTTAAAAAGCCAGCTCAAATATTGTGTTCTTCGTGA  400
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 351 AACTCCTACTCAGACTTAAAAAGCCAGCTCAAATATTGTGTTCTTCGTGA  400

 401 TTCTTTTCCTTCTTCATTGANGTAGANTCTAACACTTCTTCACACACCAC  450
     |||||||||||||||||||| ||||| ||||||||||||||||||||||||
 401 TTCTTTTCCTTCTTCATTGAGGTAGAGTCTAACACTTCTTCACACACCAC  450

 451 TAACAAACTAGGTTTGAAATGTTACTTTTGTACATTTAAGAACCTTTTTT  500
     |||||||||||||||||||||||||||||||||||||||| |||||||
 451 TAACAAACTAGGTTTGAAATGTTACTTTTGTACATTTAAGAANCTTTTTT  500

 501 ACATTTTTCATTTTAAAAACATTTTAATAGCAAAAATATTATAACAGAGA  550
     |||||||||||||||||||||||| |||| |||||||| |||| ||
 501 ACATTTTTCATTTTAAAAACATTTTANTAGCNAAAATATTNTAACCGAAN  550

 551 AGAA-TACATACTAACGAATGTTATTTAAATAAGANTTATTGGTATAAAC  600
     || | ||||| |||||||||| |||||||||| | |||| |||||||
 551 AGGAGTACATNCTAACGAATGTAATTTAAATAAGATT-ATTG-TATAAAC  600

 601 ATAGATGTTTTAATAACTGGTAATCTGGAAAAA-GAAATAGAACACAGCC  650
     |||||||||||||||||||||||||||||||| |||||| |||||||||
 601 ATAGATGTTTTAATAACTGGTAATCTGGAAAAAGAAATANAACACAGCC  650
```



Alignment Of DNA 4 (Normal Strand),

And DNA 2/Region 1 (Normal Strand)

(The top row shows the DNA 4 nucleotide sequence).

(The bottom row shows the DNA 2/Region 1 nucleotide sequence).

```
  1 TCGACCTTGAATAC-GG-A-G--CTTC-A-AGCCAGGCT-C-CAGCTGGC      50
    ||| |  | | | || | |  || | | | ||   || | || |   |
 -1 ..GAC-T--A-T-CTGGTAAGTACT-CTATA-CCT--CTACACA-C---C      48

 51 GTTTTAG-GTG--TGGTGCAAGCAAGGCTGGGTATG-TATTTCCAGGAAG     100
     |||||  |||  |  || |  | |   |   | || | |||| || ||
 49 -TTTTAAAGTGAAT--TG-AC-CTA---T---T-TGCTTTTTCTAG-AAT      98

101 CAGTAGACGA-----GCTCG-CCGTGCAAT---A-T-GAATTATC-ATTA     150
      || |  |     | ||  |||| ||    | | || || || | |
 99 TTTTAAA--ATTTTTGTTCCACCGTTAAACCCCAGTTGA-TTCTCGACT-     148

151 GCAGCAAGGAAATTCGTCAA-CGATT-TGTT-GGTAAATGAATTTCCTGT     200
     || |  | |||  | | |  |  || | || || |||        ||| |
149 -CAACT-GAAAA--C-T-ATTC--TTCT-TTCGG-AAA-------CCT-T     198

201 CA--GA-TATCCA--GTTGCTTGGGAGAACGCTGCCTTT-GCGCC-TCCT     250
    ||  || ||| ||  ||| |||     |   | | ||| | ||| ||||
199 CATTGACTATTCAAAGTT-CTTCTTT-AT---T-C-TTTAGTGCCATCCT     248

251 GCGG-ACGGCTCAATCTGGC-TTAAGTACGACTACACAGAAGTTGACACA     300
     | | |    | | | |  | |||  |||  || || | |  |  || |
249 -CTGTA----T-A-T-TACCCTTAT-TACC-CTTCA-A-AC-T--AC-C-     298

301 GTAACATACGGGTTAAGCAGGAAGTGCAAATACTACGTTGGCATGGTGCA     350
     ||   |  | | || |   |    ||  || |   | |   |  |
299 -TAT--T--G--T-AA-CT--A----CA--TA-T---T-G---T--T---     348

351 GATTTCAGTGTTCTTCAGTCCGGGGACTGGGATTGATAAGCCGAGACAAA     400
      ||| | ||| | ||  |||     ||   ||  | || |  ||
349 --TTT-A-TGTCCATC--TCCC----CTA--AT--A-AA-C--AG-----     398

401 T-AGCTAATCAATTGGCA-GAATCTATCGTTGATGGTACAATGCTTGAAC     450
    | ||||   |  ||  || |||   |  |   |   ||   | |||   |
399 TGAGCTC--C--TT--CAAGAA---A--G---AN--TA---TCCTT---C     448

451 AGCGGGACCATTTATGAGTCTGGAGTTGTTAACCCGGTTATCA--AATCC     500
      |  ||| |         ||    || ||  || |   ||||  ||
449 --C---ACC-TCC------CT----TT-TT--CC-GC--ATCACTAA---     498

501 AAGTCTGGGT-GGTTTATTCCCGG-TTC-GT---TT-TT-ATGTTCGTCT     550
    ||  | |||| |||  |  | | | ||  ||    || || | |||| ||
499 AA--CAGGGTAGGT--AGGCACTGATTAAGTACATTATTCA-GTTCATC-     548

551 AGACTAACAAAA--GGAAAAATA---CATG-GCCCATCTCAGC-AATGGC     600
    |  ||  ||  |  ||| | |||   | || |   || | | | || |||
549 A--CT--CAGTATTGGACATATATTTCTTGAGA--ATTT-A-CTAAAGGC     598

601 AC--GCA-G-GTC-TTCGTAGAAAGCTCTCGCGGAACTGCAATCGACGT-     650
     |  ||  | ||  ||| |||   ||  ||| | || || | | |  ||
599 -CTTGCTTGCGTGGTTC-TAG---GCA-TCGTGAAATTG-AGT-G--GTG     648
```

```
651 AACTGCAATTTCTAACGCAGTTA--CCCCTGTTTTAACTGTATCTGACGC  700
    |||   ||     |||  | |||  ||||| | |   ||  ||  |  |
649 AAC---AA-----AAC--ATTTAAACCCCT-TGTC--CTC-AT--G--G-  698

701 GTCTGGTT-TGGTCGTTGGTGAATACCTGCTTTTCACCTCTTCTGCTTCA  750
    | ||  || || ||  | || ||  |   |  ||  |         | |  |
699 GACT--TTATGTTCCTGGG-GAN-A---GAATT--A--------GAT--A  748

751 ACGCTTCTAACTGACAATCAAGTTCGAATGACTGCAATCCCTGGTACGTC  800
    ||      |  | ||| | |  |   ||  |    || | | | |  ||
749 AC------A--T-ACA-T-A--T---AA--A----AA-CACAG-T--GT-  798

801 AGTAACTGTTGA-AGGTTTGATACCTCCAGCACCACAAAGTTCCCGGCTG  850
    | |   | || | | |   | | |||   | |  | ||| ||   |  |
799 ATT---T-TT-ACATG---G-T-CCTNN-GAAG-A-AAA-TTAA-GA-TA  848

851 GCTTAACTGGTGAAGTTGTCAAGATCACCTCCT-GGTTCGAAGT-TCCGT  900
      | ||  ||  |||| |  ||| | |  |  | ||| ||| | ||| |
849 NAT-AA--GG--AAGTGG--AAG-T-AT-T--TAGGTG-GAANTGTCC-T  898

901 GCGTTCAGGACG-TATCTACTGACGGTGGCGAAC-AGCAATTCGTTAACT  950
    | |   ||| |  | |  | ||  || | | |   |  |||  |  ||||
899 G-GA--AGGTCNCT-TNAA-TGTNGGNGACCANTGANTAAT--G--AACT  948

951 TCCAGTGCTTGTCCGATGACCGTGAACAGCAGATTCCGACCTAT--AAAT  1000
    |  | || || |  | |     |||| |  | |  |  || | |  |||
949 TGAANTGTTT-T--G-TNT---TGAAGA--A-A--CN-ACTTTTTGAAAA  998

1001 CTGCGGT--AACTAACACCTTC-ACCTTCGCTCACGAGTACACCAACCCT  1050
     |  ||||  ||  || |  ||| | || |   | ||  | |||  |||
999  CCCCGGTGGAANAAA-AT-TTCCNCNTTTGN--AGGA--A-ACCNCCCCC  1048

1051 GTATATCCGGTTCTGCGTAACTACGATGAGTCTGGCCA-GGTTGTTGCGA  1100
       | |   ||         ||  |  |  |  |   ||  || | |||  |
1049 CAA-A---GGNNA-----AA--A--A--A--C---CCNTGGGT-TTGAAA  1098

1101 TTCGTCTGTTCG----T-ACCTCG-AGCTAGCGAAATGC--GCTTGCAGT  1150
       |  ||| ||          || ||   || | ||        ||  || |
1099 NT--TCTTTTNAAAAANGNCCNCGTNNCTGGGGANNNNNAANCT--CATT  1148

1151 -CC-GGTACTATCGCTTTCAACGACACCCCTACCATTGGTGTTAACGAAA  1200
     |  ||| | |         ||  | |   ||   ||   |     |
1149 TCNAGGT-CNAANAAAA--AAN-ANAANNCTN--ATN--T-----CN---  1198

1201 TCGAAACGGTATCCATCGCGGTATCCATTCGTG.................  1250
       |  | |  |  ||   | |  |   |  |
1199 -CNNA-CNNTG-CCNN-G-G--A---AN-CN...................  1248
```

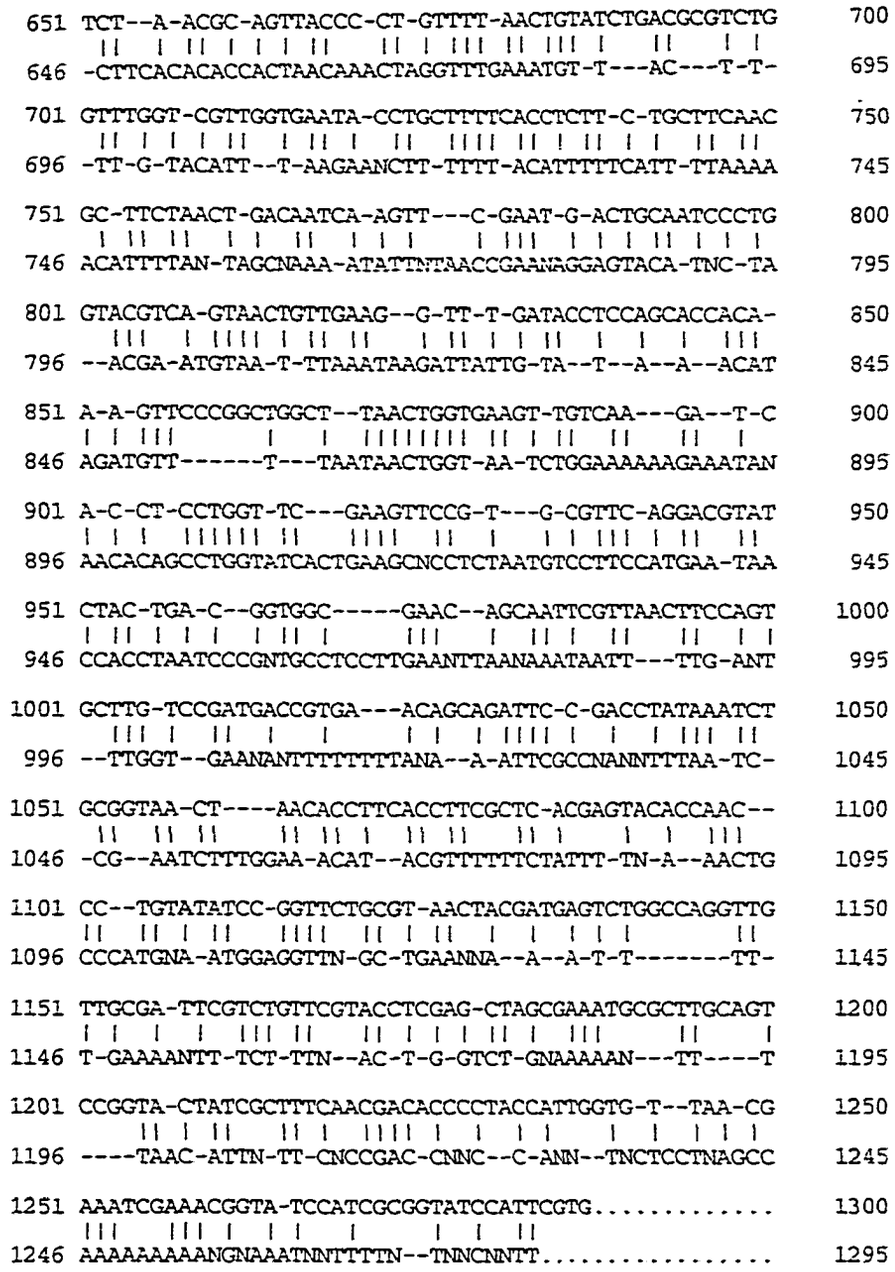
Alignment Of DNA 4 (Normal Strand),

And DNA 2/Region 2 (Normal Strand)

(The top row shows the DNA 4 nucleotide sequence).

(The bottom row shows the DNA 2/Region 2 nucleotide sequence).

```
  1 TCGACCTTGAATACGGAGCTTCAAGCCAGGCTCCAGCTGGCGTTTTAGGT      50
-4 .....CT-GCA---GG---T-CAA-CG-GGATCTT--TGGCATGATA--T      45

 51 GTGGTGCA-AGCAAGGCTGGGT-ATGTAT-TTCCAG---G-A-A--GCAG     100
46 -T--T--ATA-CA---CTGTCTCATTT-TCTTC--GTTTGTATATTG--G      95

101 TAGACGAGCTCGCCGTGCAATATGAATTATCATTAGCAGCAAGGAA-ATT     150
 96 --G--GA--T-GT--T----TAT-AATTATT-TT-GCAG---G-ATTATT     145

151 CGTCAACGATTTGTTG-GTAAATGAATTTCCTGTCAGATATCCAGTTGCT     200
146 T-TGAA---T--GT-GAG-AAAT-AATATCTT-T-A-ATTTG-A-TT--T     195

201 TGGGAGAACGCTGCCTTTGCGCCTCCTGCGGACGGCTCAATC---TGGCT     250
196 T-------C--T---TTTG----T--T----A-G-CT-AATTAAATGG-T     245

251 TAAGTACGACT---A--CACAGAA---GT-TGA-CACAGT-A-ACATACG     300
246 -AACTA--ACTTTTATTCATATAATTAGTATGGTCTCA-TCACACA-A-G     295

301 GG--TTAAGCAGGA-AGTGCAAATACTACGTTGGC-ATGGTGC-AGATTT     350
296 GCCCTTT-GCA--ACATTGC---TCCTAC-TT--CTATG-T-CTA-AGTT     345

351 CAGTGTTCTTCAGTC-CGGGGACTGGGATTGATAAGCCGAGACAAATAGC     400
346 C--TG--CTTCACTTTC-----CT---AT-GCT---CC-ATCCA--T---     395

401 TAATCAATTGGCAGAATCTATCGTTGATGGTACAATGCTTGAACAGCGGG     450
396 T--TCAAT---CACCC-C-A--G---ACG-T-C--T-CT-GTATA-----     445

451 ACCATTTATGAGTCTGGAGTTGTTAACCCGGTTATCA-AA-TCCAAGTCT     500
446 A---TTT--G-GCC---A-T-GT---CC---T-ATCACAACTCTAA-TCT     495

501 GGGTGGTT-TATTCCCGGTTCGTTTTTATGTTCGTCTAGACTAACAA---     550
496 TG-T--TTGTATTCCTGGA--------ATGC-C--CT-GA--AA-AACTC     545

551 --A---AGGAA--AAATACATGGCCCATCTCAGCAATGGCAC-GCAGGTC     600
546 CTACTCAG-ACTTAAA-A-A-G-CC-AGCTCA--AAT---ATTGT-GTTC     595

601 TTCGTAGAAAGCTCT-CGCGGAACTGCAATCGACGTA-A--CTG-CAATT     650
596 TTCGT-GATT-CTTTTC-CTT--CTTCATT-GAGGTAGAGTCTAACACTT     645
```



REFERENCES

1. Adams, D. J., and McGuire, L. Quantitative enzyme-linked immunosorbent assay for the estrogen-regulated Mr 24,000 protein in human breast tumors: Correlation with estrogen and progestrone receptors. Cancer Res., 45: 2445-2449, 1985.

2. Love, S., and King, R.J.B. A 27 kDA heat shock protein that has anomolous prognostic powers in early and advanced breast cancer. Br. J. Cancer, 69: 743-748, 1994.

3. Carper, S. W., Duffy, J. J., Gerner, E.W. Heat shock proteins in thermotolerance and other cellular processes. Cancer Res., 47: 5249-5255, 1987.

4. Becker, J., and Graig, E.A. Review - Heat shock proteins as molecular chaperones. Eur. J. Biochem., 219: 11-23, 1994.

5. Leonhardt, S.A., Fearon, K., Danese, P.N., and Mason, T.L. HSP78 encodes a yeast mitochondrial heat shock protein in the Clp family of ATP-dependent proteases. Mol. Cell. Biol., 13: 6304-6313, 1993.

6. Pratt, W.B., Schemer, L.C., Hutchinson, K.A., and Daiman, F.C. A model of glucocorticoid receptor unfolding and stabilization by a heat-shock protein complex. J. Steroid Biochem., 267: 13728-13734, 1992.

7. Cheng, M.Y., Hartl, F.U., Martin, J., Pollock, R.A., Kalousek, F., Neupert,W., Halberg, E.M., Harlberg, R.L., and Horwich, A.L. Mitochondrial heat- shock protein hsp60 is essential for assembly of proteins imported into yeast mitochondria. Nature, 337: 620-625, 1989.

8. Chiang. H.L., Terlecky, S.R., Plant C.P., and Dice, J.F. A Role for a 70-kilodalton Heat Shock Protein in Lysosomal Degradation of Intracellular Proteins. Science, 246: 382-384, 1989.

9. Hemmingsen, S.M., Woolford, C., van der Vies, S.M., Tilly, K.,Dennis, D.T., Georgopoulos, C.P., Hendris, R.W., and Ellis, R.J. Homologous plant and bacterial

proteins chaperone oligomeric protein assembly. Nature, 333: 330-334, 1988.

10. Mizzen, L., and Welch., W.J. Characterization of the thermotolerant cell. I. Effects on protein synthesis activity and the regulation of HSP70 expression. J. Cell. Biol., 106: 1105-1116, 1988.

11. Chretien, P., and Landry, J. Enhanced constitutive expression of the 27-kDA heat shock proteins in heat-resistant variants from Chinese hamster cells. J. Cell Phys., 137: 157-166, 1988.

12. Lavoie, J.N., Gingras-Breton, G., Tanguay, R.M., and Landry, J. Induction of Chinese hamster HSP27 gene expression in mouse cells confers resistance to heat shock. HSP27 stabilization of the microfilament organization. J. Biol. Chem., 268: 3420-3429, 1993.

13. Jakob, U., Gaestel, M., and Engel, K., Buchner, J. Small heat shock proteins are molecular chaperones. J. Biol. Chem., 268: 1517-1520,1993.

14. Oesterreich, S., Weng, C.N., Qium, G., Hilsenbeck, S.G., Osborn, C.K. and Fuqua, S.A.W. The small heat shock protein hsp27 is correlated with growth and drug-resistance in human breast cancer cell lines. Cancer Res., 53: 4443-4448, 1993.

15. Hickey, E., Brandon, S.E., Potter, R., Stein, G., Stein., J. and Weber, L.A. Sequence and organization of genes encoding the human hsp 27 kDA heat shock protein. Nucleic Acids Res.,14: 4127-4145, 1986.

16. McGuire, S.E., Fuqua, S.A.W., Naylor, S.L., Helin-Davis, D.A., and McGuire, W.L. Chromosomal assignment of human 27-kDa heat shock protein gene family. Somat. Cell Mol. Genet., 15: 167-171, 1989.

17. Carper, S.W., Rocheleau, T.A., Storm, F.K. cDNA sequence of a human heat shock protein HSP27. Nucleic Acids Res., 18: 6457, 1990.

18. Oesterreich, S., Hickey, E., Weber, L.A., and Fuqua, S.A.W. Basal Regulatory Promoter Elements of the hsp27 Gene in Human Breast Cancer Cells. Biochemical and Biophy. Res. Comm., 22: 155-163, 1996.

19. de Jong , W.W., Leunissen, J.A.M., Voorter, C.E.M. Evolution of the alpha-crystallin/small heat-shock protein family. Mol. Biol. Evol., 10: 103-126, 1993.

20. de Jong, W.W., Leunissen, J.A.M., Leenen, P.J.M., Zweers, A., and Versteeg, M. J. Dogfish and alpha-crystallin sequences: comparison with small heat shock proteins and Schistosoma egg antigen. J. Biol. Chem., 263: 5141-5149, 1988.

21. Gaestel, M.R., Gotthardt, and Muller, T. Structure and organization of a murine gene

encoding small heat-shock protein Hsp25. Gene, 128: 279-283, 1993.

22. Frohli, E., Aoyama, A., and Klemenz, R. Cloning of the mouse hsp 25 gene and an extremely conserved hsp 25 pseudogene. Gene, 128: 273-277, 1993.

23. Uoshima, K., Handelman, B. , and Cooper, L. F. Isolation and characterization of a rat HSP27 gene. Biochem. Biophys. Res. Commun., 197: (3) 1388-1395, 1993.

24. Sambrook, J., Fritsch, E.F., and Maniatis, T. Molecular Cloning: A laboratory manual. 2nd edition. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, N.Y.

25. Needleman, S.B. and Wunsch, C.D. A General Method Applicable to the Search for Similarities in the Amino Acid Sequencre of Two Proteins. J. Mol. Bio., 48: 443, 1970.

26. Stryer, L. Biochemistry, 3rd Edition. W.H. Freeman and Company, New York, NY.

27. Fickett, J.W. Inferring genes from open reading frames. Comput. Chem., 18: 203-205, 1994.

28. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, J. Basic local alignment search tool. J. Mol. Biol., 215: 403-10, 1990.

29. Ninio, J. Molecular approaches to evolution. 1st edition. Princeton University Press, Princeton, N.J.

VITA

Graduate College

University of Nevada, Las Vegas

Ann A. Ohiaeri

Home Address:
962 Carnival Avenue
Las Vegas, NV 89123

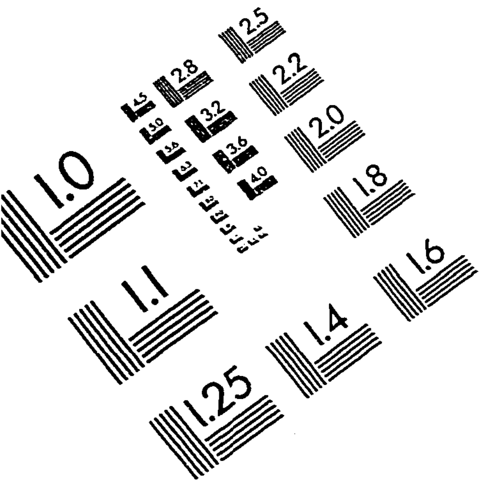
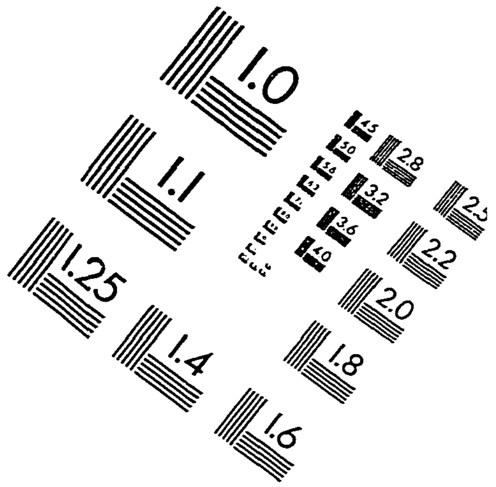
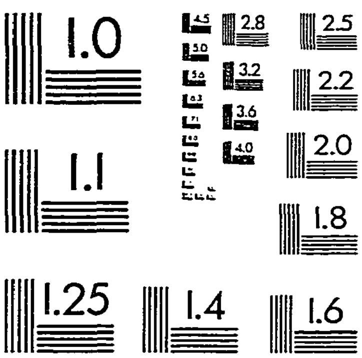
Degrees:
Bachelor of Science, Biology, 1995
University of Nevada, Las Vegas

Special Honors and Awards:
UNLV Graduate College Grant
UNLV Graduate College Scholarship

Thesis Title: Towards the Cloning of the Human Heat Shock Protein Twenty-Seven Multigene Family

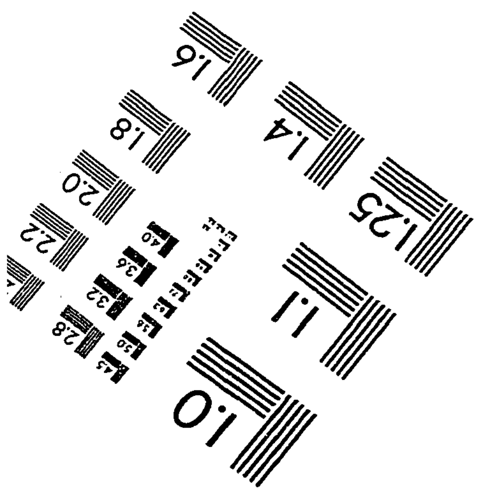
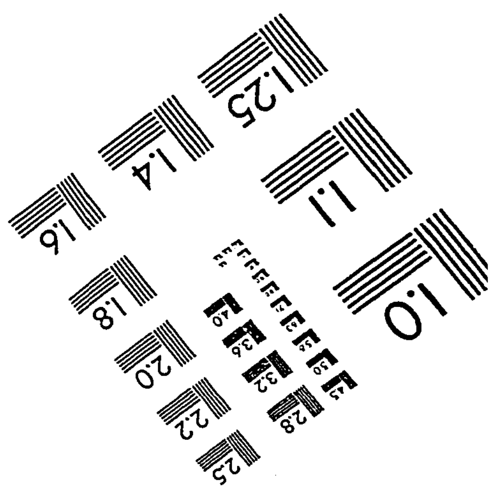
Thesis Examination Committee:
Chairperson, Dr. Stephen Carper, Ph.D.
Committee Member, Dr. Bryan Spangelo, Ph.D.
Committee Member, Dr. Vernon Hodge, Ph.D.
Graduate Faculty Representative, Roberta Williams, M.S.

IMAGE EVALUATION TEST TARGET (QA-3)

150mm

6″

APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989