PREDICTION PROGRAM OF SECONDARY STRUCTURE FROM SEQUENCE OF PROTEINS ACCORDING TO THE METHOD OF CHOU AND FASMAN

by

ANNE-MARIE PHAM

B.Sc., The University of Montpellier, 1978

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE SUDIES

(Department of Food Science)

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

July 1981

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Food Science

The University of British Columbia
2075 Wesbrook Place
Vancouver, Canada
V6T 1W5

Date July 7, 1981

ABSTRACT

Several methods have been proposed for predicting the secondary structure of proteins. The method of Chou and Fasman (1974a, 1974b, 1978a, 1978b) is relatively simple in theory and reasonably accurate. Unfortunately, the rules of Chou and Fasman are sometimes ambiguous and can be interpreted differently by researchers.

Several attempts have been made for computerization of the rules of Chou and Fasman (Argos et al., 1976; Chou and Fasman, 1978b; Dzionara et al., 1977). However, they are for computation of only a portion of the protein secondary structure. The final assignment of the entire structure has to rely on the individual's manipulation.

In addition to three separate computer programs for prediction of the α-helix, β-sheet and β-turn structures, a fourth program was written for clarifying overlapping areas between α-helix and β-sheet. Although the predicted structures of 24 proteins with known conformation were in general satisfactory, there were a number of missing areas and boundary values different from X-ray diffraction patterns.

In an attempt to improve the accuracy of the prediction, the nucleation rules were modified to emphasize

importance of the type and positions of amino acid residues in the region.

Furthermore, an extra step for boundary adjustment was added to the search for α-helix and β-sheet regions. This step compared the importance of the boundary conformational parameters and the possible interference of the different conformations at the boundaries of the predicted regions. These modifications produced predicted secondary structures which were in good agreement with the X-ray diffraction patterns and the predicted patterns of Chou and Fasman (1974b, 1978b).

The Matthews' coefficient (C) calculated for α-helix and β-sheet were 0.39 or above, meaning that the prediction would be quite useful although there might be one or two helical regions missed or overpredicted. The paired-sample t-test revealed that the values of C_α ($P < 0.01$) and C_β ($P \leq 0.05$) calculated for the present prediction were significantly improved from the values of Chou and Fasman. The computer-assisted technique described in this thesis, therefore, would decrease the discrepancy between the predicted data from different researchers due to the ambiguous interpretations of the rules of Chou and Fasman.

The second part of this study involved the application of the program to several food related proteins

(bovine serum albumin, α_{s1}-casein, β-casein, κ-casein, chymosin, α-lactalbumin, α-lactoglobulin, ovalbumin, pepsin and trypsinogen). Although references could not be found for all proteins tested, the results obtained for κ-casein and α-lactalbumin were comparable to those reported by other researchers.

Since conformational data have long been recognized as contributing to the information on protein and enzyme functionality, the computerization of the predictive method of Chou and Fasman will definitely be a tool for explaining the protein functionality in food processing.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES

(cont'd)

ACKNOWLEDGEMENTS

The author wishes to express her sincere appreciation to her supervisor, Dr. S. Nakai, Professor, Department of Food Science, for his constant advise, help and encouragement throughout the course of this study, and in the preparation of the thesis.

She is also thankful to the members of her graduate committee:

Dr. R. C. Fitzsimmons, Associate Professor,
Department of Poultry Science

Dr. W.D. Powrie, Professor and Head,
Department of Food Science

Dr. B.J. Skura, Assistant Professor,
Department of Food Science

for their interest in this research and for the review of this thesis.

INTRODUCTION

Broadly, the functional properties of proteins denote any physico-chemical property that affects the processing and behavior of proteins in food systems, as judged by the quality attributes of the final product. The functional properties are influenced by and vary according to: a) the source of proteins, b) the method of isolation and purification, c) the concentration of proteins, d) the type of modifications (enzymatic, acid, or alkaline hydrolysis), and e) environmental conditions (pH, temperature, and ionic strength).

An extensive review of the various studies on protein functionality was published by Kinsella (1976). In general most of the changes in protein functionality have been found to be related to the degree of denaturation that proteins undergo. For example, in gelation a heat treatment is usually required to cause at least partial denaturation or unfolding of the polypeptide chains. Those unfolded chains will then gradually associate to form a gel matrix if attractive forces and thermodynamic condtions are suitable.

It is necessary to consider the hydrophobic, electronic and steric parameters of molecules to understand the mechanism of folding and unfolding of proteins, their biological activity, and to predict their behavior upon certain treatments

(i.e., possible areas of denaturation, extent of unfolding, exposing of the hydrophobic core). The electronic parameters can be evaluated by electrophoresis , while a fluorometric method has been developed by Kato and Nakai (1980) for evaluation of hydrophobic parameters.

X-ray diffraction has been used to study the steric parameters of more than sixty proteins. In addition, X-ray diffraction has contributed a great deal to our knowledge of protein-protein, protein-metal, protein-solvent interactions (Liljas and Rossman, 1974 ; Matthews, 1975b). However , thèse crystallographic analyses cannot be applied to most food proteins, or to many membrane and ribosomal proteins due to problems of crystallization. Furthermore, the X-ray technique is quite laborious, expensive and time-consuming.

Anfinsen et al. (1961), studying the reformation of reduced bovine ribonuclease, observed that the native structure of a protein is controlled by its amino acid sequence. This finding has become the motivation for many attempts to obtain patterns of protein structure from sequence data.

Although protein functionality depends on its unique three-dimensional topology, one can still learn much from the prediction of its secondary structure. Nishikawa and Ooi (1972, 1973), in a study of protein tertiary structure, based their energy calculations on the conformation derived from the

computation of sets of dihedral angles ϕ and ψ. To fit the polypeptide chain of the tobacco mosaic virus protein (TMV) to a low resolution Fourier map, some information on the secondary structure was found to be desirable (Leberman, 1971).

Secondary structure predictions will provide useful information on areas of protein molecules where the X-ray pattern is not yet clearly resolved, especially at the N-terminal. For instance, areas predicted as helical by Chou and Fasman (1974b) for cytochrome b_5 and ferricytochrome c had not been detected by X-ray diffraction at 2.8Å resolution. Their results were later confirmed by X-ray at 2.45Å and 2.0Å resolution (Dickerson et al., 1971; Mathews et al., 1972; Takano et al., 1973).

Conformational information may also be used to design experimental models for checking the effects of conformational changes on hormonal or enzymatic activity (Dunn and Chaiken, 1975; Fink and Bodanszky, 1976; Peña et al., 1975). Some researchers (Deber et al., 1976; Kopple et al., 1975) considered study of the β-turn a good starting point for elucidating the influence of sequence and surroundings on protein conformation. The β-turn structure is potentially identifiable and is simple enough to be characterized by experimental and predictive techniques (^{13}C NMR, circular dichroism, confor-

mational energy calculations). It also helps to explain the mode of activation of biologically active peptides (Bradbury et al.,1976).

Another application of the secondary structure prediction is the comparison of proteins of the same family which may maintain some conformational homology despite variations in sequence data,such as the case of proinsulins and proteinase inhibitors (Chou and Fasman,1978b).

The method of Chou and Fasman (Chou and Fasman, 1978a,1978b) has been frequently considered the least complicated in use for the prediction of the secondary structure of proteins. Yet,it possesses an overall accuracy higher than random guessing for a three-state model (α-helix,β-sheet and coil state). The percent of total residues correctly identified in a protein is 75 for this method versus 33 for random guessing. Furthermore,Chou and Fasman's work has improved on the earlier studies (Davies,1964; Havsteen,1966; Goldsack, 1969) since it takes into account combinations of residues that are α-helix,β-sheet and β-turn formers and breakers. The computed percentage of secondary structure obtained by their method agrees quite well with estimates based on CD studies (Kawauchi and Li, 1974; Garel et al.,1975; Garnier et al.,1975; Green,1975; Matthews,1975a; Scanu et al.,1975; Holladay and Puett,1976; Muñoz et al.,1976; Wallace,1976).

With the exception of Argos et al. (1976), most of the laboratories that have applied the method of Chou and Fasman for specific investigations on protein structure have not yet reported a common computerized technique which can be used for other proteins. The objectives of this study were as follows: a) design a program which would provide similar results to those published by Chou and Fasman (1974b, 1978b) and b) if successful, extend this program to food related proteins so that possible correlation between protein functionality and conformational changes may be better understood.

LITERATURE REVIEW

Definition of the Different Conformational Regions

According to the IUPAC - IUB Commission on biochemical nomenclature (1970) the secondary structure of a segment of a polypeptide chain is the local spatial arrangement of its main chain atoms without regard to the conformation of its side chain or its relationship with other segments. The four typical conformations encountered in the secondary structure are the α-helix, the β-sheet, the β-turn (bend), and the random coil.

A. Alpha-Helix

The α-helix contains 3.6 amino acid residues per turn of the protein backbone, with the R groups of the amino acids extending outward from the axis of the helical structure. Hydrogen bonding can occur between the hydrogen of the NH group of one peptide bond and the oxygen of the CO group of another peptide bond four residues along the protein chain. The hydrogen bonds are nearly parallel to the axis of the helix, lending strength to the helical structure. Since natural amino acids exist in L configuration, a right-handed helix is more stable than a

left-handed helix. Therefore, if helical structures exist in proteins they are invariably right-handed helices (Anglemier and Montgomery, 1976). Since the α-helix has the lowest feasible free energy, formation of this structure is spontaneous, provided there are no interactions between charged R groups or steric hindrance by residues on the larger amino acids. Examples of protein types in which the α-helix predominates are enzymes and respiratory proteins.

Taking into account the structural requirements that are specific to globular proteins, Lim (1974a) proposed a number of conditions necessary for a helix to exist along the peptide chain. Each separate helical region must have a hydrophobic side group or a group which would permit the helix to attach itself to the hydrophobic core of the globule. From the analysis of immersion of the hydrophobic side chain situated on the α-helix surface, Lim (1974a) emphasized the role of hydrophobic pairs, (1-5), and hydrophobic triplets, (1-2-5) or (1-4-5), in the attachment of the α-helix to the hydrophobic core. Hydrophobic-hydrophilic triplets, (1-2-5) and (1-4-5), are also important for helix stabilization.

Another way to describe the conformation of a protein chain is to measure the dihedral angles ϕ and ψ which correspond to rotations about the N-C^{α} and C^{α}-C bonds.

The ϕ, ψ angles for residues in a regular right-handed helix are given by (-57, -47°) (IUPAC-IUB, 1970). Since 3.6 residues are required to form a hydrogen bond in a single turn of the α-helix, all consecutive sequences of four or more residues having ϕ, ψ angles within 40° of (-60°, - 50°) are considered to be helical. Some residues at the helical ends may have dihedral angles that fall outside the range $-100^{\circ} \leq \phi \leq -20^{\circ}$ and $-90^{\circ} \leq \psi \leq -10^{\circ}$ but are included as helical if they show hydrogen bonding. Based on the above criteria a total of 152 helical regions were identified in 29 proteins (Chou and Fasman, 1978a, 1978b).

B. Beta-Sheet

In this conformation, the peptide backbone forms a zig-zag pattern with the R groups of the amino acids extending above and below the peptide chain. Since all peptide bonds are available for hydrogen bonding, this conformation allows maximum cross-linking between adjacent peptide chains and, thus, good stability. Both parallel and anti-parallel pleated sheets are possible. This conformation predominates in many fibrous proteins such as silk and insect fibres.

According to Lim (1974a), β-structural regions can be divided into three types by their relative position

to the surface of the globule: the internal,the surface,and the semi-surface type. In order to exist without violating structural requirements for globular proteins,each type should be formed from a certain number of hydrophobic/hydrophilic residues . For instance,entirely hydrophobic regions or hydrophobic regions with one or two hydrophilic residues in the first two and/or last two positions on the N- and C-terminal will favor the internal type.

The condition for a β-chain to be located on the surface of the globule requires that one side of the band have only hydrophobic groups and the other side only hydrophilic groups. The semi-surface type may exist in peripheral regions of the β-sheets. These regions must have only hydrophilic side groups or mainly residues of Gly.

The position of certain amino acid residues can also be very critical. Pro cannot be included in the β-structure because of the stereochemistry of its side group. The surface type must not have Gly on the hydrophobic side or Gly and Ala on the hydrophilic side. This is stipulated by the fact that the presence of Gly on the hydrophobic side will impede the tight packing formation in the hydrophobic core. Water molecules can loosen hydrogen bonds of the peptide groups neighbouring with the C^{α} atoms of Gly or Ala when these two amino acids occur on the hydrophilic side.

The ϕ,ψ angles for residues in a parallel-chain β-sheet and an antiparallel β-sheet have values of $(-119^{\circ},113^{\circ})$ and $(-139^{\circ}, 135^{\circ})$,respectively (IUPAC-IUB,1970). A consecutive sequence of three or more residues having ϕ,ψ angles within 40° of $(-120^{\circ},110^{\circ})$ or $(-140^{\circ},135^{\circ})$ are considered to be in the β-conformation,even if these residues are not involved in hydrogen bonding. However,residues at the β-ends that have dihedral angles outside the range $-180^{\circ} \leq \phi < -80^{\circ}$ and $175^{\circ} \leq \psi \leq 70^{\circ}$ are included in the β-region if they participate in at least one hydrogen bond. The two end residues that are not hydrogen bonded in antiparallel β-sheets are not counted as β-residues but instead are assigned to the coil conformation and /or the β-turn conformation.

Chou and Fasman (1978a,1978b),analyzing 137 β-regions,observed 3 two-residue β-segments (papain 111-112, 130-131, and ferrodoxin 50-51), 10 three-residue β-segments, and 9 four-residue β-segments. This number increases to 28 and 24 for the five-residue and six-residue β-segments,respectively. The three longest β-regions contain 17 residues (thermolysin 16-32), 16 residues (ribonuclease 96-111), and 15 residues (lactate dehydrogenase 280-294). In contrast, Chou and Fasman (1978a,1978b) identified 24 helical segments longer than 17 residues in 29 proteins. The reason that helices are longer

than β-sheets may be because of the greater ease of helical intrachain hydrogen bond formation compared to β-sheet interchain hydrogen bond formation.

C. Coil Regions

Residues in the protein that are not classified to be in the helix or β-regions are assigned to the coil conformation, irrespectively of the ϕ, ψ angles of the residue. Hence, three consecutive residues having the α-conformation or two consecutive residues having the β-conformation but without hydrogen bonding are considered to be in the coil state (Chou and Fasman, 1978b). The four longest coils regions found among 29 proteins contained 54 residues (thermolysin 181-234), 51 residues (carboxypeptidase 123-173), 46 residues (ferrodoxin 4-49) and 41 residues (rubredoxin 14-54). These coil regions cannot be considered completely structureless since they may contain many β-turns. In the case of ferredoxin and rubredoxin, the flexibility of these coil regions is severely restricted by the iron-sulfur coordinations.

D. β-Turn Regions

The β-turn involves four consecutive residues in

a protein where the polypeptide chain folds back on itself by nearly 180^{o}. It is these regions of chain reversal that give a protein its globularity rather than linearity. Lewis et al. (1971) proposed that chain reversals play the important role of bringing distant parts of the peptide chain together, enabling interactions between helix-helix, antiparallel-parallel β-pleated sheet, or helix-β-sheet.

Venkatachalam (1968) was the first to characterize three types of turns in a tetrapeptide where there is a hydrogen bond between the CO group of residue i and the NH group of residue (i+3). Most bends (80%) from 8 proteins contain at least one or more of the following residues: Ser, Thr, Asp, Asn and Pro. This supports the idea that these residues are responsible for bend stability and perhaps for bend formation. With the exception of Pro which can occupy only a few backbone conformations, these residues have been shown to be capable of forming side chain-backbone hydrogen bonds with their own backbone (Lewis et al., 1973).

Using the X-ray atomic coordinates from 29 proteins, Chou and Fasman (1977) computed the C_i - C_{i+3} distances of 4651 tetrapeptides. Those whose distances were below 7Å and not in a helical region were considered as β-turns. Of the 457 β-turns elucidated, 243 of them also have O_i - N_{i+3}

distances $\leq$ 3.5Å and were considered to have hydrogen bonding. Chou and Fasman (1977) also assigned β-turns to 11 types similar to those of Lewis et al. (1973) based on the ϕ, ψ dihedral angles of the second and third residues of the bend.

Review of the Various Predictive Methods

Several researchers have attempted to predict the secondary structure of proteins from their sequence data.

Szent-Gyorgyi and Cohen (1957) through their study with the KMEF proteins (keratin, myosin, epidermin and fibrinogen) and with collagen, demonstrated that the helix content determined by optical rotatory dispersion (ORD) is inversely proportional to the percentage of Pro residues distributed throughout the sequence. They concluded that less than 3 percent Pro distributed randomly in a chain permits more than 50 percent α-helix. About 8 percent Pro deforms the backbone into a rancom coil. Very high Pro content may favor a poly-L-proline helix type.

Davies (1964) using the b_o value of ORD, found a strong correlation between the helix content of fifteen proteins and the mole percentage of (Ser + Thr + Val + Ile + Cys), residues classified as "nonhelical-formers" by Blout

et al. (1960) and Blout (1962). No strong correlation between the helix content and the mole percentage of any particular amino acid was observed (Davies, 1964). Furthermore, the correlation reported by Szent-Gyorgyi and Cohen (1957) was not sustained when the number of proteins was increased to that used by Davies (1964). Therefore, the correlations previously mentioned should be applied with caution until they are supported by additional data.

Havsteen (1966) carried out a statistical analysis of the correlation between the content of certain amino acids in 40 proteins and their ORD parameter b_o. A linear relationship was observed between $-1/b_o$ and the percentage content of (Ser + Thr + Pro); thus, supporting the previous findings on the interactions between hydroxyl groups of Ser, Thr and peptide linkages which may interfere with the formation of α-helices. Pro residues tend to destabilize helices by requiring a 90^o bend of the peptide chain. The presence of a β-form of left-handed helices also seems to markedly influence b_o. The influence of the amino acid side chains on b_o justify their classification as helix-favoring, helix-indifferent, and helix-inhibiting groups.

Goldsack (1969), using the data of 107 proteins, demonstrated that the parameter b_o can be correlated to the total content of the so-called helix-forming amino acids

(Ala + Arg + Asp + Cys + Glu+ Leu + Lys), as well as, to that of the nonhelix-forming group of amino acids (Gly + Phe + Pro + Ser + Thr + Trp + Tyr). On the other hand, using the a_o parameter, it seemed that no particular amino acid side chain grossly controls the amount of β-structure in a protein. Nevertheless. further ORD characterization of the different β-structures (intramolecular parallel and antiparallel, as well as, intermolecular cross-β structure) will be useful to elucidate the relationship between a_o and the amino acid composition.

These preliminary efforts in predicting protein conformation relied heavily on ORD data and amino acid composition. The X-ray analysis of protein structure was at an early stage of development and the amino acid sequence was still unknown for many proteins.

Scheraga (1960) attempted to construct a three-dimensional model of ribonuclease on the basis of available data on the primary, secondary and tertiary structures. Its importance lies in the fact that it provides a basis to plan experiments for the investigation of side chain group interactions and it may also be of help in Fourier analysis of X-ray data on ribonuclease crystals.

On the basis of known sequence and structure of myoglobin, alpha- and beta-hemoglobin, Guzzo (1965)

suggested that the presence of the four critical groups; Pro, Asp, Glu and His may be a necessary condition for a section of proteins to be non-helical. Analyses of Pro replacement by Asp and Glu in mutant and variant proteins supported his theory. This was applied in an effort to predict the secondary structure of lysozyme and tobacco mosaic virus. Absence of hydrophobic bonding and weakening of interpeptide hydrogen bonding as a result of water competition in the vicinity of those polar residues might be the reason for the unfavorable effect of those residues on helix formation.

Prothero (1966) compared his results to that of Guzzo (1965) on six proteins and proposed a rule which seems to achieve a reasonable degree of fit with the known protein structures. The rule states: any region of five residues will be α-helical if at least three of its residues are comprised of Ala, Val, Leu, or Glu. Alternatively, any region of seven residues will be α-helical if at least three residues are comprised of Ala, Val, Leu, Glu and an additional residue includes Gln, Ile, or Thr. Using this rule, goodness of fit between 65 and 68% was obtained for α-, β- and γ-hemoglobin, lysozyme and myoglobin.

Periti et al.(1967) carried out a systematic statistical analysis of the available data for horse

hemoglobin, and sperm whale myoglobin. This led them to the consideration of helical and anti-helical pairs of amino acid residues (1 2, 1 3, 1 4, ... , 1 7; 2 3, 2 4, ... , 2 8, 3 4, ...). Histograms for the recognition of helical segments of egg white lysozyme were constructed according to their method.

Finding that it was undesirable to represent the helical segments by the usual linear way, Schiffer and Edmunson (1967) proposed a two-dimensional representation called the "helical wheel". The wheels are projections of the amino acid side chains onto a plane perpendicular to the axis of the helix. Side chains interactions and general characteristics of the helices can be better visualized. Using data form four proteins, it was observed that areas with hydrophobic residues located in the $n\pm 3$, n, and $n\pm 4$ positions have the greatest potential for helicity. Such hydrophobic arcs are absent in nonhelical wheels. Hence, the wheel representation may be of help to identify areas with helical potential. Among the six proteins chosen for testing the wheel method, the prediction of helical segments in insulin is the most accurate and closest to X-ray data later proposed by Blundell *et al*. (1972)

Low et al. (1968) looked for sequence identities of length varying between that of di- and hexapeptides. The theory behind their method is based on the assumption that if helix-forming sequences in which local interactions predominate can be recognized then their position along the polypeptide chain may be irrelevant. A computer program was written to locate sequence identities from available data. Although this method gives less over-prediction of helical regions compared to other methods, it results in more omissions. The authors recognized that the procedure needs to be improved by taking into account the effects of long-range interactions and that of non-helical sequences.

Kotelchuck and Scheraga (1969), from earlier energy computations, formulated a set of rules in which various single peptide units were assigned as helix-forming (Ala, Val, Leu, Ile, Met, Thr, Gln, Glu, Phe, Cys, His, and Arg) or helix-breaking (Ser, Asn, Asp, Trp, Tyr, and Lys). Their designations were quite similar to those of previous studies (Prothero, 1966; Schiffer and Edmunson, 1967). This allowed correct identification of 61% of the helices and 78% of the total residues in four proteins; myoglobin, lysozyme, tosyl-α-chymotrypsin and ribonuclease A. They did attempt to define conditions for helix nucleation and termination, ruling that five or more peptide

units constitute the minimum length for any helical area and that a sequence of two helix-breakers will stop the helix propagation. They agreed, however, that their model was not very accurate for smaller protein systems where long-range interactions may play an important role in helix nucleation and stabilization.

Using a combination of the Kotelchuck and Scheraga (1969) and the Schiffer and Edmunson (1967) schemes, Leberman (1971) succeeded in correctly assigning 82% of all residues in seven proteins as helical and nonhelical regions. The omission of observed regions was explained as an effect of the tertiary or even the quaternary structure, or the binding of a prosthetic group (e.g., human hemoglobin, myoglobin).

Lewis et al. (1970) based their method on the Zimm and Bragg (1959) σ and s parameters for helix initiation and elongation. The parameters were obtained from melting curves of random copolymers of amino acids. Helix probability profiles constructed for eleven proteins yield 68% accuracy. Correlation between the propensity of a residue to be a helical former in the denatured protein and its occurrence in a helical area in the corresponding native protein was suggested. The correlation supports the hypothesis that residues in the α_R conformation may be involved in the

nucleation of protein folding. A comparison was made of the conformational structure of denatured cytochrome c from various species (Lewis and Scheraga, 1971). They showed that, even though there were amino acid replacements in cytochrome c throughout evolution, there remains a conservation of the nature of the helix-forming power at each position in the chain.

Despite the progress in protein prediction, there was still a lack of information on β-sheet structure. This was because the earliest proteins elucidated by X-ray diffraction were hemoglobin and myoglobin which are devoid of β-sheet conformation. Hence, most of the researchers at that time often chose to ignore the β-sheet conformation in their calculations. Furthermore, it was difficult to obtain β-sheet in solution for spectrophotometric analysis. However, as more protein structures were elucidated by X-ray diffraction, it became increasingly apparent that the presence of β-sheet was as important as that of α-helix. Interpretation of an electron density map at 2Å resolution indicates that the predominant conformation in concanavalin A is formed by two antiparallel β-sheets. Residues not included in the β structures are arranged in regions of random coil. One of the pleated sheets contributes extensively to the interactions among the monomers to form

both dimers and tetramers (Edelman et al., 1972). X-ray analysis of tosyl-α-chymotrypsin revealed only a small fraction of α-helix but several adjacent, anti-parallel pleated sheets stabilized by hydrogen bonds (Birktoft and Blow, 1972).

Ptitsyn and Finkelshtein (1970) classified the various amino acid residues as helical or antihelical, and tentatively as β-breaker or β-former according to their tendency of stabilizing the various β-structures. Nonpolar amino acids (Leu, Ala, Met) except Cys and Tyr, have a greater tendency to enter α helical zones than the polar ones. The amino acid residues with compact hydrocarbon sidegroups are assigned with positive β-potential whereas those with charged side groups and Pro are considered as β-breakers. Although their classification takes into account only the interactions of the side groups with the main chain backbone and not with each other, they obtained good agreement between their predictive method and X-ray data ($Q\alpha$ = 79% and $Q\beta$ = 79%) for nine proteins. This supports the suggestion that instead of competing with local interactions and dictating the secondary structure, distant interactions work in harmony with the local ones and help to stabilize the conformation which mainly results from local interactions.

Nagano (1973) developed a computer method to predict helices, loops and β-structures from the primary structure. The basis of his method lies on the assumption that short-range interactions are due to amino acid residue pairs separated by m residues (m = 0, 1, 2, 5, ... 6). Four prediction functions (helix, loop, random coil, and β-structure) were estimated by a linear combination of statistical quantities of different m values as a measure of the statistical constraint. The coefficients used in the combination were determined to make the number of correct assignments as large as possible. Very successful results were obtained (85.3% for helix prediction, 64.4% for loop, and 90.1% for β-structures).

On the basis of the influence of nearest neighbouring pairs of amino acids (n-1) and (n+1) on the conformation of amino acid (n), Kabat and Wu (1973a, 1973b) designed, then later revised their 20x20 table of frequency of occurrences of various conformations tabulating three values: α-helix, β-sheet and neither. The frequencies were then used to locate helix-breaking positions in various proteins. Due to limited data on proteins with extensive β-sheet fragments, recognition of the β-sheet breaking regions was made on papain only. The regions between two β-sheet breaking residues would be permissively β-sheet regions.

Application of the method on concanavalin A, which has many β-sheet regions, allows location of 10 out of the 13 β-sheet areas. Although no guidelines were given to prevent overprediction of α- and β-regions, the conjunction of this method with the helical wheel method or other schemes may lead to a higher degree of accuracy (Chou and Fasman, 1978b).

Lim (1974b) proposed another method that takes into account both quantitative evaluation of energy and qualitative stereochemical considerations. Based on the most characteristic features of globular proteins (compactness of form; presence of a tightly packed hydrophobic core; a polar shell) and the role of the different types of long-range interactions, different requirements were set up to find the most energetically advantageous conformations for the protein chain. Lim (1974a) also elaborated on the structural role of the different hydrophilic side groups in the stabilization of the proteins tertiary structure. α-Helix and β-sheet are classified into various types according to their specific orientation relative to the globule surface. Regions which do not belong to helix or β-sheet type are classified as irregular regions. Through the use of helical and antihelical pairs and triplets at positions [1-2], [1-3], [1-4], [1-5], [1-2-5] and [1-4-5], Lim (1974b) developed a predictive

algorithm for helices. The search of β-structural areas is only done on fragments of the chain not attributed to α-helical regions, because it is energetically more advantageous to have one long helix than several shorter β-regions. The accuracy of the predictive method applied to 25 proteins of known structure was 81% for α-helix and 85% for β-sheet. The conformation of 25 unknown proteins was also tested with the method.

Chou et al. (1972) through CD conformational studies of poly(N(3-hydroxypropyl)-L-glutamine) and of copolymers of hydroxypropyl-L-glutamine with L-leucine reached the following conclusions. The helical content of the homopolymer and copolymers was found to increase with: a) decreasing temperature, b) increasing methanol concentration, and c) increasing molar ratios of Leu in the copolymers. A survey of the conformation of eleven proteins reveals that of all the amino acids occuring in the inner helical regions Leu occurs most frequently. This suggests that Leu may be the strongest helical-forming amino acid residue in polypeptides, as well as, in proteins (Chou and Fasman, 1973). For the first time, the helix and β-sheet conformational potential of all 20 amino acids were established in their hierarchical order. Following this study more complete investigation (Chou and Fasman,1974a)

on the conformational parameters P_α, P_β and P_t of each amino acid residue in 15 proteins served as the basis for a new predictive method (Chou and Fasman, 1974b). The major advantages of their method are its simplicity and its accuracy. Without recourse to complicated computer analysis, one can expediently locate the helix, β-sheet and coil regions of proteins with 70-80% accuracy (Chou and Fasman, 1978a, 1978b) by simply averaging the P_α, P_β and P_t values of the residues in the segment under consideration. Another way of locating the various conformations is to assign each residue as a former, an indifferent, or a breaker based on its helix and β-sheet potential. The β-turn conformational parameter P_t was also computed, enabling the prediction of chain reversals and tertiary folding in proteins. The simplicity and effectiveness of the method are the main reasons for its wide use (Chou and Fasman, 1978b). Indeed, according to Argos et al. (1976), the complexity of some proposed algorithms is such that their computerization has not been developed. This problem may be the reason why these methods have limited popularity compared to the method of Chou and Fasman or other popular methods (Lim, 1974b; Kabat and Wu, 1973a).

MATERIALS AND METHODS

The Predictive Method of Chou and Fasman

Using the criteria of dihedral angles and hydrogen bond formation, Chou and Fasman (1978a, 1978b) first determined the different conformational states in 29 proteins. The frequency of all 20 amino acids in each conformation was then calculated by dividing their occurrence in the conformation under consideration by their total occurrence in the 29 proteins. The percentages of residues in the 29 proteins found in the helical, sheet, coil, and β-turn regions are respectively represented by their average fractions $\langle f_\alpha \rangle$ = 0.38, $\langle f_\beta \rangle$ = 0.20, $\langle f_c \rangle$ = 0.42, and $\langle f_t \rangle$ = 0.20. Each residue is assigned to the α, β, or coil state so that $\langle f_\alpha \rangle + \langle f_\beta \rangle + \langle f_c \rangle = 1.00$. The β-turn residue assignment is made independently. Each amino acid is then assigned as former, indifferent, or breaker according to its conformational parameters P_α, P_β which are obtained by dividing the frequency of its occurrence in a conformation by the respective average frequency (e.g., $P_\alpha = f_\alpha/\langle f_\alpha \rangle$, $P_\beta = f_\beta/\langle f_\beta \rangle$, $P_t = f_t/\langle f_t \rangle$). The conformational parameters P_α and P_β for the 20 amino acids are listed in Table 1 in hierarchical order along with their assignment as former, indifferent, or breaker.

Table 1. Conformational parameters for α-helical and β-sheet residues based on 29 proteins.[a]

α-Residues	P_α	Helical Assignment[b]	β-Residues	P_β	β-Sheet Assignment[c]
Glu^-	1.51	H_α	Val	1.70	H_β
Met	1.45	H_α	Ile	1.60	H_β
Ala	1.42	H_α	Tyr	1.17	H_β
Leu	1.21	H_α	Phe	1.38	h_β
Lys^+	1.16	h_α	Trp	1.37	h_β
Phe	1.13	h_α	Leu	1.30	h_β
Gln	1.11	h_α	Cys	1.19	h_β
Trp	1.08	h_α	Thr	1.19	h_β
Ile	1.08	h_α	Gln	1.10	h_β
Val	1.06	h_α	Met	1.05	h_β
Asp^-	1.01	I_α	Arg	0.93	i_β
His	1.00	I_α	Asn	0.89	i_β
Arg	0.98	i_α	His	0.87	i_β
Thr	0.83	i_α	Ala	0.83	i_β
Ser	0.77	i_α	Ser	0.75	b_β
Cys	0.70	i_α	Gly	0.75	b_β
Tyr	0.69	b_α	Lys^+	0.74	b_β
Asn	0.67	b_α	Pro	0.55	B_β
Pro	0.57	B_α	Asp^-	0.54	B_β
Gly	0.57	B_α	Glu^-	0.37	B_β

[a] Chou and Fasman (1978b)

[b] Helical assignments: H_α, strong α-former; h_α, α-former; I_α, weak α-former; i_α, α-indifferent; b_α, α-breaker; B_α, strong α-breaker.

[c] β-sheet assignements: H_β, strong β-former; h_β, β-former; I_β, weak β-former; i_β, β-indifferent; b_β, β-breaker; B_β, strong β-breaker.

The symbols H and h may be thought of as strong and moderate hydrogen bonding, respectively with the subscripts α,β denoting helical or β-sheet conformation. Each amino acid residue can also be characterized by its boundary conformational parameters ($P_{\alpha N}$, $P_{\alpha C}$, $P_{n\alpha N}$, $P_{n\alpha C}$, $P_{\beta N}$, $P_{\beta C}$, $P_{n\beta N}$, $P_{n\beta C}$) as listed in Tables 2 and 3.

When all the residues in a protein sequence have been classified, one can use the empirical rules discussed below to predict its secondary structure (Chou and Fasman, 1978a, 1978b).

A. Search for Helical Regions

The search was carried out according to the method of Chou and Fasman (1978a, 1978b), which can be described as follows:

1. Helix nucleation. A cluster of four helical residues (h_α or H_α) out of six residues along the protein sequence will initiate a helix. A weak helical residue (I_α) counts as 1/2 h_α (i.e., three h_α and two I_α residues out of sic may also cause helix nucleation).

2. Helix propagation. Extend the helical segment in both directions as long as adjacent tetrapeptides are not helix breakers (see below). When overlapping segments all satisfy the helix nucleation rule, they are linked together into a long helix. The nucleated helix of six

Table 2. Conformational Parameters of Helical Boundary Residues[a] in 29 Proteins.

	$P_{\alpha N}$		$P_{\alpha C}$		$P_{n\alpha N}$		$P_{n\alpha C}$
Glu(-)	2.44	Lys(+)	1.83	Ser	1.55	His(+)	1.86
Asp(-)	2.02	His(+)	1.77	Asn	1.42	Asn	1.64
Pro	2.01	Met	1.57	Gly	1.41	Gly	1.64
Trp	1.47	Val	1.25	His(+)	1.22	Pro	1.58
Ala	1.29	Arg(+)	1.20	Pro	1.10	Lys(+)	1.49
Gln	1.22	Glu(-)	1.24	Thr	1.09	Agr(+)	1.24
Thr	1.08	Gln	1.22	Glu(-)	1.04	Asp(-)	1.06
Asn	0.81	Ala	1.20	Lys(+)	1.01	Phe	1.04
Gly	0.76	Leu	1.13	Tyr	0.99	Tyr	0.96
Ser	0.74	Cys	1.11	Asp(-)	0.98	Cys	0.94
His(+)	0.73	Phe	1.10	Phe	0.93	Ser	0.93
Met	0.71	Ile	0.98	Leu	0.85	Ile	0.87
Tyr	0.68	Ser	0.96	Met	0.83	Thr	0.86
Ile	0.67	Thr	0.75	Ile	0.78	Leu	0.84
Cys	0.66	Tyr	0.73	Gln	0.75	Gln	0.70
Lys(+)	0.66	Asp(-)	0.61	Val	0.75	Glu(-)	0.59
Phe	0.61	Asn	0.59	Ala	0.70	Ala	0.52
Val	0.61	Gly	0.42	Cys	0.65	Met	0.52
Leu	0.58	Trp	0.40	Trp	0.62	Val	0.32
Arg(+)	0.44	Pro	0.00	Arg(+)	0.34	Trp	0.16

[a]Helix boundary residues include the three helical residues on both ends of a helical region and the three nonhelical residues adjacent to the helical end residues, a total of six residues on each end of the helix. $P_{\alpha N}$ = normalized frequency of residues in the N-terminal helix region; $P_{\alpha C}$ = normalized frequency of residues in the C-terminal helix region; $P_{n\alpha N}$ = normalized frequency of residues in the N-terminal nonhelical region; $P_{n\alpha C}$ = normalized frequency of residues in the C-terminal nonhelical region.

Table 3. Conformational Parameters of β-Sheet Boundary Residues[a] in 29 Proteins.

	$P_{\beta N}$	$P_{\beta C}$	$P_{n\beta N}$	$P_{n\beta C}$
Ile	1.94 Tyr	1.96 Asn	1.86 Pro	1.69
Val	1.69 Val	1.79 Pro	1.58 Gly	1.68
Gln	1.65 Phe	1.50 Gly	1.46 Trp	1.59
Phe	1.40 Ile	1.35 Ser	1.41 Ser	1.49
Trp	1.49 Leu	1.27 Asp(-)	1.39 Asp(-)	1.32
Met	1.43 Asn	1.21 Cys	1.34 Thr	1.16
Leu	1.30 Trp	1.19 Tyr	1.23 Asn	1.13
Thr	1.17 Cys	1.11 Lys(+)	1.09 Arg(+)	1.05
Tyr	1.07 Met	0.95 Gln	1.09 Tyr	1.01
Lys(+)	1.00 His(+)	0.90 Thr	1.09 His(+)	0.96
Arg(+)	0.90 Arg(+)	0.90 Glu(-)	0.92 Met	0.85
Cys	0.87 Asp(-)	0.85 Arg(+)	0.89 Glu(-)	0.85
Als	0.86 Ser	0.79 His(+)	0.78 Lys(+)	0.82
Pro	0.66 Thr	0.75 Ala	0.67 Gln	0.77
Asn	0.66 Ala	0.75 Ile	0.59 Ala	0.74
Gly	0.63 Gly	0.74 Met	0.52 Val	0.59
Ser	0.63 Lys(+)	0.74 Trp	0.48 Leu	0.59
His(+)	0.54 Gln	0.65 Leu	0.46 Ile	0.53
Asp(-)	0.38 Glu(-)	0.55 Val	0.42 Cys	0.53
Glu(-)	0.35 Pro	0.40 Phe	0.30 Phe	0.44

[a] β-sheet boundary residues include the three residues on both ends of a β regions and the three non-β residues adjacent to the β-sheet end residues, a total of six residues on each end of the β-sheet region. $P_{\beta N}$ = normalized frequency of residues in the N-terminal β region; $P_{\beta C}$ = normalized frequency of residues in the C-terminal region; $P_{n\beta N}$ = normalized frequency of residues in the N-terminal non-β region; $P_{n\beta C}$ = normalized frequency of residues in the C-terminal non-β region.

residues should contain at least two thirds h's, while the propagated helix should be comprised of one half or more helix formers. It is important to utilize the rule that a weak helical former (I_α) counts as 1/2h in the segment. Both the helix nucleation segments and the entire helix should have fewer than one third helix breakers (b_α or B_α).

3. Helix Termination. The propagated helix is terminated on both sides by the following tetrapeptide breakers with $\langle P_\alpha \rangle < 1.00$: b_4, b_3i, b_3h, b_2i_2, b_2ih, b_2h_2, bi_3, bi_2h, bih_2, and i_4. Some tetrapeptides, such as hi_3 and h_2i_2, may have $\langle P_\alpha \rangle < 1.00$ but are not listed as breakers since they allow helix propagation to continue. Once the helix is defined, some of the residues (h or i) in the above tetrapeptide breakers may be incorporated at the helical ends. For example, the hi of the breaker bbhi may be added to the predicted helix only at the N-terminal side, but the bb may not be included at either the N- or C-terminal helix. The notations i, b, h in the tetrapeptide breakers also include I, B, and H, respectively. Adjacent β-regions that have higher β- than α-potential (i.e., $\langle P_\beta \rangle > \langle P_\alpha \rangle$) can also terminate helix propagation.

4. Proline as Helix Breaker. Pro cannot occur in the inner helix or at the C-terminal helical end but

can occupy the first turn (i.e., third residues) in the N-terminal helix.

5. Helix boundaries. Pro, $Asp^{(-)}$, $Glu^{(-)}$ are incorporated into the N-terminal helical end, while $His^{(+)}$, $Lys^{(+)}$ and $Arg^{(+)}$ are incorporated into the C-terminal helical end. I_{α} assignments are given to Pro and Asp (near the N-terminal helix), as well as, Arg (near the C-terminal helix) if necessary to satisfy condition A.1. Glu is still assigned as H at the N-terminal helix while His and Lys are still h and I, respectively, at the C-terminal helix.

Rule 1. Any segment of six residues or longer in a native protein with $\langle P_{\bar{\alpha}} \rangle \geq 1.03$ and $\langle P_{\alpha} \rangle > \langle P_{\beta} \rangle$, and satisfying conditions A.1 through A.5, is predicted as helical.

B. Search for β-sheet Regions

The search for β-pleated sheet regions was carried out by applying the set of rules outlined by Chou and Fasman (1978a, 1978b) as follows:

1. β-sheet Nucleation. A sequence of three β-formers (h_{β} or H_{β}) or a cluster of three β-formers out of four or five residues along the protein sequence will initiate a β-sheet

2. β-Sheet Propagation. Extend the β-sheet segment in both directions as long as adjacent tetrapeptides are not β-sheet breakers (see below). β-Sheet formation is unfavorable if the entire segment contains one third or more β-sheet breakers (b_β or B_β) or less than one half β-sheet formers.

3. β-Sheet Termination. Apply conditions A.3 outlined for helix termination by utilizing the same tetrapeptide breakers with $\langle P_\beta \rangle < 1$ for stopping β-sheet propagation. Adjacent α-regions that have higher α- than β-potential (i.e., $\langle P_\alpha \rangle > \langle P_\beta \rangle$) can also terminate β-propagation.

4. Strong β-Sheet Breakers. Glu and Pro are the strongest β-sheet breakers and should not be incorporated into β-sheets unless they occur in tetrapeptides with $\langle P_\alpha \rangle < \langle P_\beta \rangle > 1$.

5. β-Sheet Boundaries. Charged residues and Pro are unfavorable to β-sheet formation and should not be incorporated into β-sheets unless they occur in tetrapeptides with $\langle P_\alpha \rangle < \langle P_\beta \rangle > 1$.

Rule 2. Any segment of three residues or longer in a native protein with $\langle P_\beta \rangle > 1.05$ and $\langle P_\beta \rangle > \langle P_\alpha \rangle$, and satisfying conditions B.1 through B.5 is predicted as β-sheet.

C. Overlapping α- and β-Regions

In most cases, utilization of the set of rules described above was adequate to locate the secondary structures of proteins. However there were regions in proteins containing both α- and β residues where ambiguities arose, so that additional measures were required to resolve the dilemma. Chou and Fasman (1978a, 1978b) followed the procedure described below to determine whether the overlapping region was predominately α or β.

1. Calculate the $\langle P_\alpha \rangle$ and $\langle P_\beta \rangle$ for the overlapping region; if $\langle P_\alpha \rangle > \langle P_\beta \rangle$, the region is helical, if $\langle P_\beta \rangle > \langle P_\alpha \rangle$, it is β-sheet. The α- and β-potential of the overlapping residues can also be compared by grouping the α- and β-assignments. Thus a region of six residues with $(H_2h_2ib)_\alpha$ and $(Hh_3iB)_\beta$ assignments should be helical, since there are two strong α-formers (H_α) and one α-breaker (b_α) compared to one strong β-former (H_β) and one strong β-breaker (H_β).

2. Use Tables 2 and 3 on the frequency of helix and α-sheet boundaries to delineate whether the region is α or β.

3. Since helices are longer than β-sheets, a long segment containing both α- and β-potential is predicted as helical if $\langle P_\alpha \rangle > \langle P_\beta \rangle$, even though there may be a smaller fragment, that is, five residues within the segment whose $\langle P_\beta \rangle > \langle P_\alpha \rangle$. Hence, in the example given above for carboxypeptidase, 173-186 is predicted as one long helix instead of a short helix 173-178 and a β-region 179-183.

4. Regions with both α- and β-potential adjacent to a predicted β-turn (see below) are predicted to be β-sheet as long as there are at least three β-formers on each side of the β-turn; that is, the minimum β length is reduced from five to three, with the middle two residues of the β-turn counting as coil residues. For example, regions 105-110 and 115-124 in ribonuclease have both α- and β-potential. However, the high probability of a β-turn at 113-115 easily allows the assignment of 105-110 $[(H_3\dot{h}Ib)_\beta > (Hh_2I_2i)_\alpha$ and $\langle P_\beta \rangle = 1.31 > \langle P_\alpha \rangle = 1.10]$ and 115-124 $[(H_3h_2Iib_3)_\beta > (Hh_5i_2bB)_\alpha$ and $\langle P_\beta \rangle = 1.13 > \langle P_\alpha \rangle = 1.02]$ as β-sheets rather than α-helices.

Rule 3. Any segment containing overlapping α- and β-residues is resolved through conformational boundary analysis (C.2) with $\langle P_\alpha \rangle > \langle P_\beta \rangle$ for the predicted

α-region (C.1). β-Formers may be incorporated into a long helix if they are not helical tetrapeptide breakers (C.3). Helix propagation may be terminated by α residues if these same residues favor the formation of antiparallel β-sheets.

In summary, according to Chou and Fasman (1978b), there are only three basic rules for predicting protein secondary structure. While the α and β search conditions elaborated above seem to be quite extensive they are given so that incorrect predictions will be minimized.

D. Search for β-Turns

At present, 408 β-turns have been elucidated from 29 proteins and the frequency of occurrence for the 20 amino acids in those 408 turns, at positions i to i+3, as well as their P_t values ($P_t = f_t/\langle f_t \rangle$) are given in Table 4. (Chou and Fasman, 1977, 1979). The probability of β-turn occurrence at residue i is computed from $p_t = (f_i)\ (f_{i+1})\ (f_{i+2})\ (f_{i+3})$ with the aid of Table 4. The average probability of β-turn occurrence is $\langle p_t \rangle = 0.55 \times 10^{-4}$. Two cut-off values were selected: $p_t = 1.0 \times 10^{-4}$ (a value approximately double that of the average) and $p_t = 0.75 \times 10^{-4}$ (a value that is 1 1/2 times that of the average). According to Chou and Fasman (1979),

Table 4. Frequency Hierarchies of Amino Acids in the β-Turns of 29 Proteins.[a]

	i	i + 1	i + 2	i + 3	P_t	P_{t2}
Asn	0.161 Pro	0.301 Asn	0.101 Trp	0.167 Asn	1.56 Pro	2.04
Cys	0.149 Ser	0.139 Gly	0.190 Gly	0.152 Gly	1.56 Gly	1.63
Asp	0.147 Lys	0.115 Asp	0.179 Cys	0.128 Pro	1.52 Asp	1.61
His	0.140 Asp	0.110 Ser	0.125 Tyr	0.125 Asp	1.46 Asp	1.56
Ser	0.120 Thr	0.108 Cys	0.117 Ser	0.106 Ser	1.43 Ser	1.52
Pro	0.102 Arg	0.106 Tyr	0.114 Gln	0.098 Cys	1.19 Lys	1.13
Gly	0.102 Gln	0.098 Arg	0.099 Lys	0.095 Tyr	1.14 Tyr	1.08
Thr	0.086 Gly	0.085 His	0.093 Asn	0.091 Lys	1.01 Arg	1.05
Tyr	0.082 Asn	0.083 Glu	0.077 Arg	0.085 Gln	0.98 Thr	0.98
Trp	0.077 Met	0.082 Lys	0.072 Asp	0.081 Thr	0.96 Cys	0.92
Gln	0.074 Ala	0.076 Thr	0.065 Thr	0.079 Trp	0.96 Gln	0.84
Arg	0.070 Tyr	0.065 Phe	.065 Leu	0.070 Arg	0.95 Glu	0.80
Met	0.068 Glu	0.060 Trp	0.064 Pro	0.068 His	0.95 His	0.77
Val	0.062 Cys	0.053 Gln	0.037 Phe	0.065 Glu	0.74 Ala	0.64
Leu	0.061 Val	0.048 Leu	0.036 Glu	0.064 Ala	0.66 Phe	0.62
Ala	0.060 His	0.047 Ala	0.035 Ala	0.058 Met	0.60 Met	0.51
Phe	0.059 Phe	0.041 Pro	0.034 Ile	0.056 Phe	0.60 Trp	0.48
Glu	0.056 Ile	0.034 Val	0.028 Met	0.055 Leu	0.59 Val	0.43
Lys	0.055 Leu	0.025 Met	0.014 His	0.054 Val	0.50 Leu	0.36
Ile	0.043 Trp	0.013 Ile	0.013 Val	0.053 Ile	0.47 Ile	0.29

Table 4. Frequency Hierarchies of Amino Acids in the β-Turns of 29 Proteins.[a] (cont'd)

---

[a] i, i+1, i+2, and i+3 represent the frequencies of the first, second, third, and fourth residues, respectively, in a reverse β-turn. P_t is the conformational potential of a residue in a β-turn based on all four positions of a reverse turn. P_{t2} is the conformational potential of a residue in a β-turn based on the second and third positions of a reverse turn. This frequency table was based on 408 β-turns in 29 proteins.

the lower cut-off value predicts more bend residues correctly while the higher cut-off value predicts more non-bend residues correctly. However it appears that the predictive accuracy is similar for the two values. The cut-off value of 0.75×10^{-4} has been used by Chou and Fasman (1978b, 1979) in their search for β-turns in 29 proteins.

Rule 4. Tetrapeptides with $p_t > 0.75 \times 10^{-4}$ as well as $\langle P_t \rangle > 1.00$ and $\langle P_\alpha \rangle < \langle P_t \rangle > \langle P_\beta \rangle$ are selected as probable bends. Adjacent probable bends (i.e., 11-14, 12-15, 13-16) are compared pairwise, and the tetrapeptide with the highest p_t value is predicted as a β-turn.

E. Evaluation of the Predictive Accuracy

To evaluate the success of any predictive scheme it is necessary to compare the predicted conformational state for each residue of a protein with the observed assignment based on X-ray diffraction. The percentage of residues n_k predicted in the conformational state k is given by:

$$\%k = \frac{100\ (n_k - n_x)}{n_k} \qquad (1)$$

where k represents the α-, β- or coil regions in the native protein structure as determined by x-ray crystallography and n_x is the number of incorrectly predicted residues in the state k.

The percentage of overprediction is given by the criteria:

$$\%\, nk = \frac{100(n_{nk} - n_{nx})}{n_{nk}} \qquad (2)$$

where %nk represents the percentage of correctly predicted residues not in the conformational state k, $n_{nk} = N - nk$, and n_{nx} is the number of k residues overpredicted. Hence the quality of prediction for a given type of conformational k can be expressed as the mean of %k (eq. 1) and %nk (eq. 2).

$$Q_k = \frac{\%k + \%nk}{2}$$

A value of 100% for %k, %nk, and Q_k indicates total agreement between observation and prediction, while 0% indicates total disagreement (Chou and Fasman, 1978a, 1978b).

Recently, Matthews (1975) introduced a correlation coefficient that indicates how much better a given prediction is than a random one.

$$C_\alpha = \frac{[(n_\alpha - \alpha_m)/N] - [(n_\alpha - \alpha_m + \alpha_o)/N](n_\alpha/N)}{\{[(n_\alpha - \alpha_m + \alpha_o)/N](n_\alpha/N)(1 - n_\alpha/N)[1 - (n_\alpha - \alpha_m + \alpha_o)/N]\}^{1/2}}$$

(3)

The correlation coefficient for β-sheet and β-turn may be obtained by substituting β and t, respectively, for α in equation 3. A correlation of C=1 indicates perfect agreement between prediction and observation, C=0 indicates that a prediction is no better than random, and C=-1 indicates total disagreement or 0% accuracy. If $C_\alpha \geq 0.6$, the predicted structure is near that of the observed structure with no helical regions generally missed but with N- and C-terminal points off by a few residues. If $C_\alpha \geq 0.4$, generally one or two helical regions might be missed or overpredicted, however the prediction would still be quite useful. Similar statements can be made regarding sheet and turn (Argos et al., 1976).

Amino Acid Sequence of Proteins

The amino acid sequence of the various proteins used for testing our program comes from the Atlas of protein sequence and structure (Dayhoff, 1972, 1973, 1976, 1978).

Programming

The program was written in Fortran language and tested at the UBC computing centre. The amino acid sequence of each protein was converted into a sequence of integers. The 20 amino acid residues were sorted alphabetically and each of them assigned a fixed number between 1 and 20. For instance: 1 → Ala, 2 → Arg, ..., 19 → Tyr, and 20 → Val. Hence in order to use the program, one must convert the protein sequence into a corresponding series of integers. All the necessary details concerning the use of the program are given in the appendix.

RESULTS AND DISCUSSION

Programming of the method

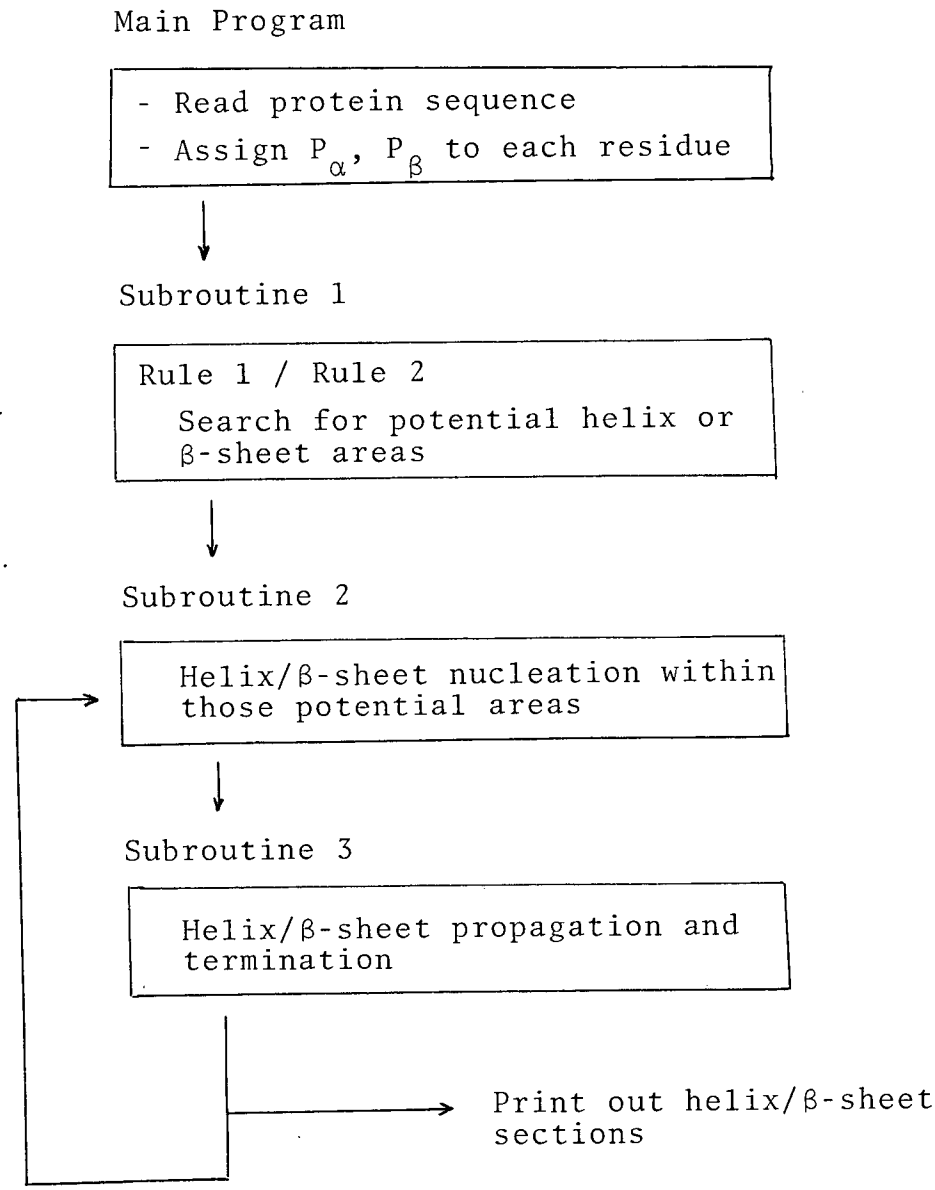
Following the rules outlined by Chou and Fasman (1978a, 1978b) four different programs were written to predict α-helix, β-sheet, and β-turn, and to solve the overlapping areas between α-helix and β-sheet.

Each program consisted of the main program and several subroutines. In every case, the purpose of the main program was to read in the sequence of the protein under consideration, and then to assign to each amino acid residue the corresponding values of the conformational parameters (P_α, P_β, P_t) and the boundary conformational parameters ($P_{\alpha N}$, $P_{\alpha C}$, $P_{n\alpha N}$, $P_{n\alpha C}$, $P_{\beta N}$, $P_{\beta C}$, $P_{n\beta N}$, $P_{n\beta C}$). The subroutines were then called on to search for α-helix, β-sheet, and β-turn regions, and to solve overlapping areas.

A. Scheme for the search of α-helix and β-sheet regions

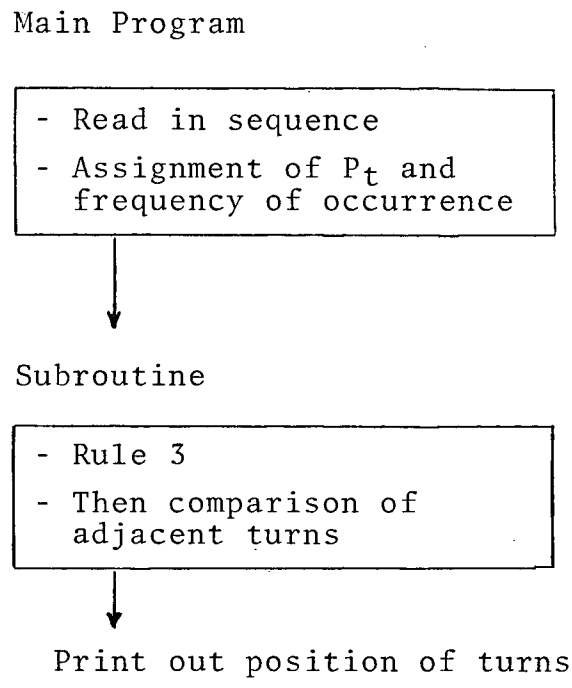
In the case of α-helix and β-sheet prediction, once the whole sequence has been recorded, the first subroutine is called to detect the areas in the sequence with helix or β-sheet potential according to rule 1 or rule 2, respectively. Then within the limits of those potential areas, the rules for nucleation, propagation and termination are

applied to locate the different sections more accurately. Those rules were elaborated in the second and third subroutines. The various important factors such as strong helix or β-sheet breakers, and helix or β-sheet boundaries were also taken into account.



B. Scheme for the β-Turn Search

For the β-turn search, only one subroutine was needed to locate the different turns according to rule 3 and to compare the adjacent predicted turns so as to consider only the one with the highest probability of occurrence (p_t).



C. Scheme for Solving Overlapping α- and β-Areas

In this case, the purpose of the main program was to record the whole sequence of each protein as well as the consecutive pairs of overlapping areas which were formatted

in the following manner: H1 S1 H2 S2 H3 S3 H4 S4 ...

H1 - H2 : boundary values of the helical fragment

S1 - S2 : boundary values of the β-sheet fragment

The first subroutine carried out the comparison of the average P_α and P_β of each fragment itself and that of the overlapping area. In case the β-sheet was contained within the α-helix, the overlapping area was the β-sheet itself. The results obtained at this step could already suggest whether the entire fragment (β-sheet/α-helix) and the overlapping area had a higher propensity to exist in one of the conformations than the other (i.e., in helical state if $\langle P_\alpha \rangle > \langle P_\beta \rangle$, or in β-sheet conformation if $\langle P_\alpha \rangle < \langle P_\beta \rangle$).

In the second subroutine, instead of assigning to each amino acid residue in the fragments under consideration the alphabetic representation of H_α, H_β, I_α, I_β, B_α, B_β, the alphabetic representations were converted to numerical ones (i.e., H_α,$H_\beta \rightarrow 2.00$; h_α, $h_\beta \rightarrow 1.00$; I_α, $I_\beta \rightarrow 0.50$; i_α, $i_\beta \rightarrow 0.25$; b_α, $b_\beta \rightarrow -0.50$; B_α, $B_\beta \rightarrow -1.00$). Hence, instead of comparing sets of characters (H_u h_v I_x i_y b_z B_w) as did Chou and Fasman (1978a, 1978b), numerical values were used to represent the conformational potential of the regions under consideration. This "character analysis" was performed

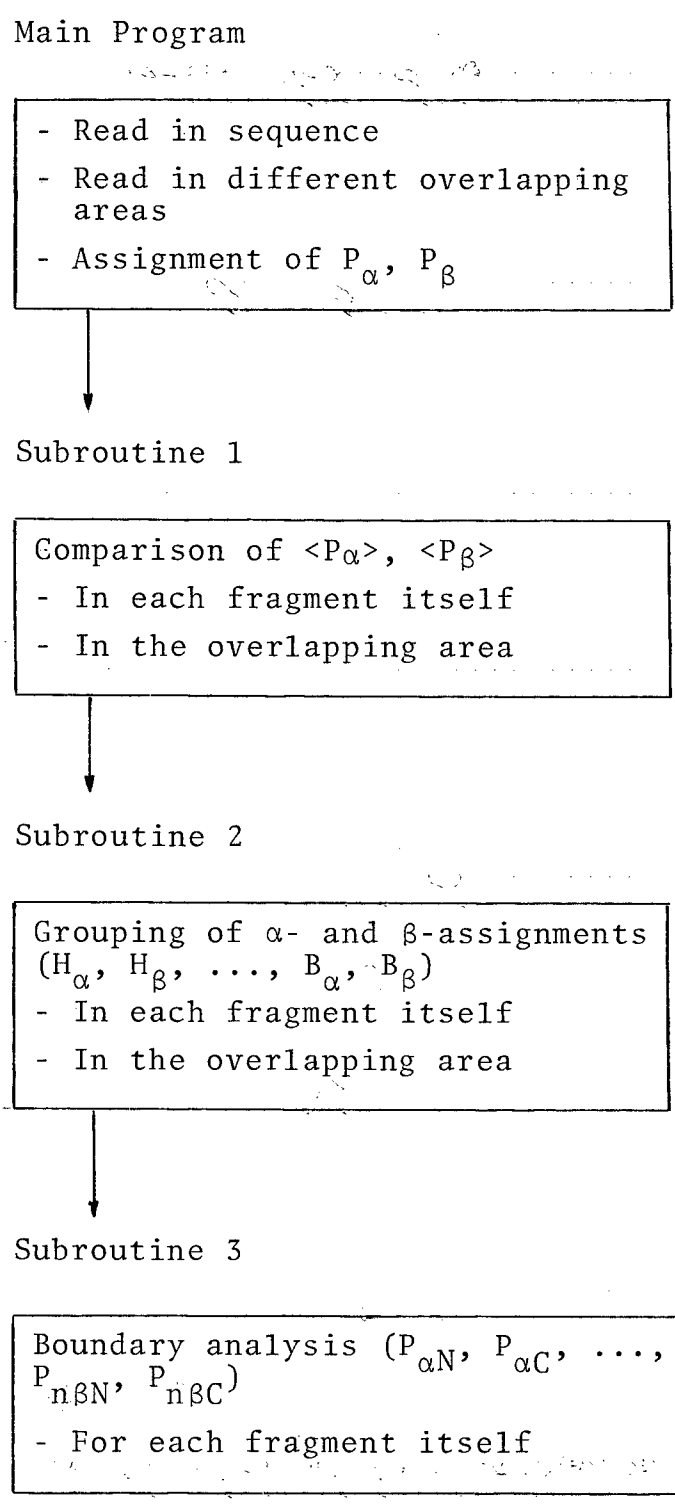
on the α-helix and β-sheet fragments, as well as, on the overlapping area.

The "boundary analysis" was carried out in the third subroutine. This consisted of summing up the boundary conformational parameters of the three residues belonging to the fragment and those of the three residues adjacent to the fragment ends, using the values from Tables 2 and 3.

e.g. $P_{\alpha N}(H1) + P_{\alpha N}(H1+1) + P_{\alpha N}(H1+2)$
$P_{\alpha C}(H2) + P_{\alpha C}(H2-1) + P_{\alpha C}(H2-2)$

$P_{n\alpha N}(H1-1) + P_{n\alpha N}(H1-2) + P_{n\alpha N}(H1-3)$
$P_{n\alpha C}(H2+1) + P_{n\alpha C}(H2+2) + P_{n\alpha C}(H2+3)$

Similar procedures were applied to the β-sheet fragment. The "boundary analysis" thus took into consideration the influence of the neighbouring residues at the boundaries. Hence, if a fragment has very high potential for the helical state, the neighbouring residues at the boundaries may participate in stabilization of the helix if they are favorable to its presence.


Main Program
- Read in sequence
- Read in different overlapping areas
- Assignment of P_α, P_β
Subroutine 1
Comparison of $<P_\alpha>$, $<P_\beta>$
- In each fragment itself
- In the overlapping area
Subroutine 2
Grouping of α- and β-assignments (H_α, H_β, ..., B_α, B_β)
- In each fragment itself
- In the overlapping area
Subroutine 3
Boundary analysis ($P_{\alpha N}$, $P_{\alpha C}$, ..., $P_{n\beta N}$, $P_{n\beta C}$)
- For each fragment itself

Efficiency of the α-helix prediction

When the rules established by Chou and Fasman (1978a, 1978b) were strictly followed, several areas were missed in the present prediction and the boundaries of the predicted regions were quite different from those of Chou and Fasman or from X-ray analysis (Table 5, p. 189)

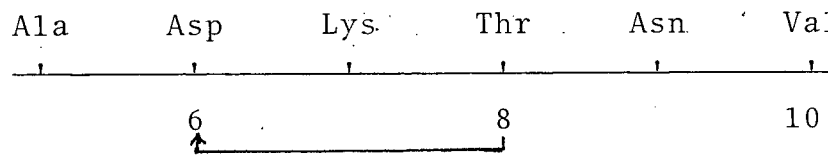
However, when the results obtained for the various proteins were analyzed, the difference between the boundary values of the present study and those of Chou and Fasman or from X-ray analysis could be reduced by taking into account factors such as: a) the boundary conformational parameters ($P_{\alpha N}$, $P_{\alpha C}$, $P_{n\alpha N}$, $P_{n\alpha C}$) and b) the β-turn or β-sheet potential in the vicinity of the helical boundaries. In other words, after going through the entire procedure of helical search, if a region delineated by two values J1 and J2 (J1: N-terminal of the predicted region and J2: C-terminal of the same region) was predicted as α-helix then the values $P_{\alpha N}$ of J1 and $P_{\alpha C}$ of J2 would be compared to those of their neighbouring residues so that the new boundaries J1 $\pm$ n, J2 $\pm$ n' (n and n': integers) would have the most favorable $P_{\alpha N}$ and $P_{\alpha C}$ for helix stabilization. The parameters $P_{n\alpha N}$ and $P_{n\alpha C}$ of the nonhelical residues adjacent to the helical bound-

aries were also important in this "move of the boundaries". When considering the possibility of β-turn presence at the helix boundaries or the overlapping of the end residues with a fragment possessing high β-sheet potential, it may also be necessary to move J1 and J2 to new positions which are dictated by $P_{\alpha N}$, $P_{\alpha C}$ respectively (cf. subroutines MOJ1, MOJ2, RMJ1, and RMJ2 for α-helix prediction).

The following are some examples of α-helix boundaries adjustment to illustrate the concept of "move of boundaries".

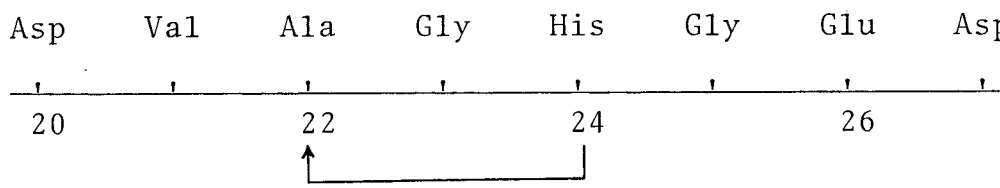
(1) J1 = J1 - 2

a. α-Hemoglobin: 8-17



Asp (6) has the second highest $P_{\alpha N}$ value. As there is neither β-turn nor β-sheet potential in this region, by moving back to Asp (6), the helical area 8-17 gains 1 h_{α}, Lys (7), and 1 I_{α}, Asp (6). There is no need to move further to the N-terminal because the program has already predicted the area 1-8 as helical.

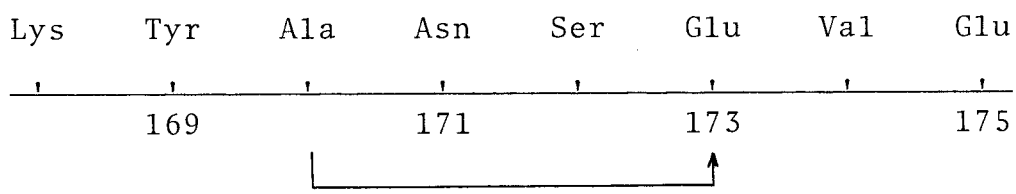
b. Myoglobin: 24-36



Ala (22) has a higher $P_{\alpha N}$ value than His (24). The incorporation of a breaker, Gly (23) is balanced by that of Ala, a strong helix former. This boundary adjustment helps to link the helical regions 13-22 and 24-36.

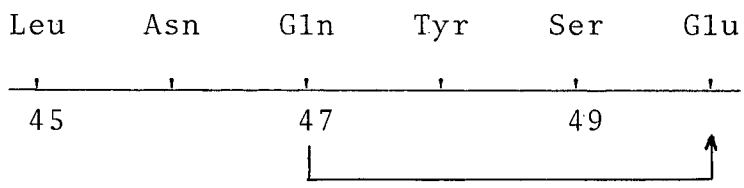
(2) J1 = J1 + 3

a. Carboxypeptidase: 170-182



The boundary adjustment is justified by the strong potential β-turn of the tetrapeptide 169-172 and by the very high $P_{\alpha N}$ of Glu (173). Furthermore, Asn (171) and Ser (172) have the highest $P_{n\alpha N}$ values, hence their presence is favorable to Glu (173) if this residue is chosen as the N-boundary.

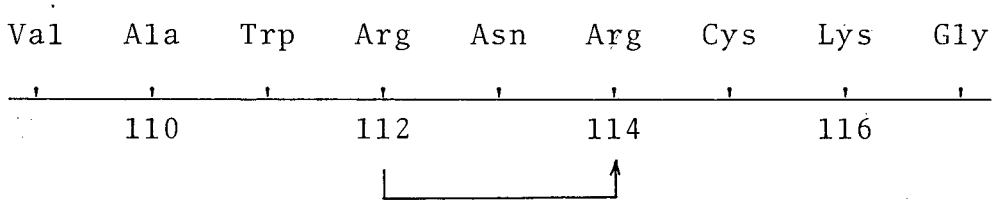
b. Papain: 47-60



By moving the N-boundary to position 50 (Glu), advantage is taken of the good $P_{\alpha N}$ of Glu and at the same time the helix breaker Tyr (48) and the helix indifferent Ser (49) are removed. Ser (48) has high propensity to be found at the nonhelical N-boundary.

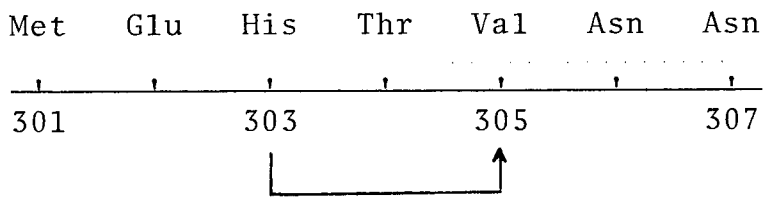
(3) J2 = J2 + 2

a. Lysozyme: 105-112



The region 105-112 has enough helix formers to balance the incorporation of an extra helix breaker, Asn (113). And Arg (114) has good $P_{\alpha C}$ value which justifies its consideration as the new C-boundary. The residues Cys (115), Lys (116) and Gly (117) are favorable to the new position of J2.

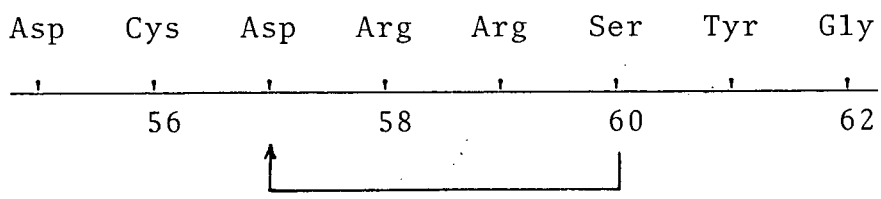
b. Carboxypeptidase: 297-303



Despite the lower $P_{\alpha C}$ of Val (305) compared to that of His (303), the move of J2 to J2+2 allows the addition of an extra i_{α} (Thr) and H_{α} (Val) to the region. In fact $P_{\alpha C}$ of Val is higher than the average value and the two residues Asn (306, 307) are listed second for their $P_{n\alpha C}$.

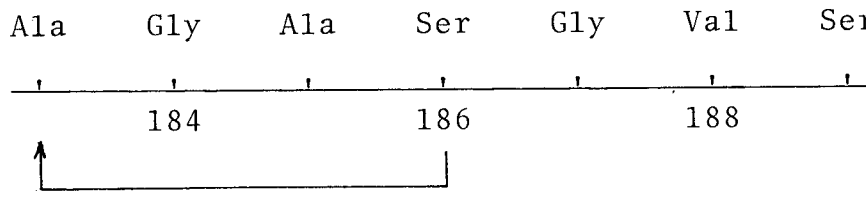
(4) J2 = J2 - 3:

a. Papain: 47-60



The tetrapeptide 57-60 has β-turn potential and although $P_{\alpha C}$ of Asp (57) is lower than that of Cys (56), the helical conformational parameter of Asp (I_{α}) is higher than that of Cys (i_{α}). The residues Arg (58, 59) exhibit good $P_{n\alpha C}$.

b. α-Chymotrypsin: 172-186



The tetrapeptide 185-188 exhibits β-turn potential and by moving J2 to J2-3, there are two advantages. First, the breaker, Gly (184), is avoided and second, a new boundary with relatively good $P_{\alpha C}$, Ala (183), is obtained.

The boundary adjustment procedures were elaborated in the subroutines MOJ1, RMJ1 and MOJ2, RMJ2 for the residues J1 and J2, respectively. Although such considerations are not always simple and may give unexpected results because the boundary conditions vary from fragment to fragment as well as from protein to protein, they still help to save part of the analysis of overlapping areas between helix and β-sheet and to avoid conflicts between β-turns and helices.

A certain number of segments were missed in the prediction by the present program due to one of the following reasons:

It was ruled that the nucleation segment should contain fewer than one third helix breakers (B_α or b_α). This condition strictly eliminated some segments which have two breakers out of six residues although they met

the requirement of having at least two thirds h's (e.g. α-hemoglobin 23-28 and papain 120-126).

- The nucleation segment did not have at least two thirds h's (e.g. carboxypeptidase 116-121 and lysozyme 80-85) , although by incorporating an extra residue, the final segment (carboxypeptidase 116-122) would respect the rule, or by shifting the whole fragment by one position to the left (lysozyme 79-84) the new boundaries (J1-1, J2-1) would have better $P_{\alpha N}$ and $P_{\alpha C}$ respectively.

- The position of the residue Pro at the N-terminal end is also critical. For example, in concanavalin A 83-88, although this fragment had the rigth number of h's, the presence of Pro as the fourth residue (position 86) impeded the fragment from being taken into accoung. In this case it was not possible to shift the fragment to the right to position 84-89 because from position 89 the area had higher β-sheet than α-helix potential. Therefore we attempted a preliminary consideration of the fragment 83-88 as a possible helix then tried to shift it to avoid having Pro as the fourth residue (e.g. concanavalin A 80-85).

- For Russell's Viper Venom 47-55, as there were more I_α, i_α and b_α than H_α and h_α in this part of the polypeptide chain, the nucleation search skipped the entire area because none of the combinations of six consecutive amino acid

residues had at least two thirds h's. Nevertheless the "one half h's " requirement was met by the entire fragment 47-55 (5 h's out of 9 amino acid residues). Hence this requirement could not be applied strictly in all cases, and it may be modified to such an extent that the final predicted segment will not be considered as having deviated from the normal criteria.

- The gathering of two or three Pro residues in the same area was another possible reason for missing a fragment (e.g. bovine colostrum inhibitor 5-10). The fragment 1-10 of this protein contained three Pro at positions 4, 5, and 11, hence the program could only detect the segment 9-14 as the most suitable to avoid having Pro in the inner helix. However, if the residue Pro 5 was considered acceptable as an N-terminal end residue (good $P_{\alpha N}$, although not at the third position), then the fragment 5-10 would be a potential α-helix since it met the requirement for two thirds h's.

In summary, in order to obtain results which are closer to those of Chou and Fasman or to X-ray data, the nucleation rule was slightly modified (i.e., under specific circumstances the nucleation segment may have one third breakers, which then must be compensated by the addition of h's during the α-helix propagation, and the presence of Pro at the 2 first positions of the N-terminal end instead of

at the third one does not always constitute an obstacle to α-helix formation). In addition, before asserting a segment of the polypeptide chain as α-helix, the search for its most suitable boundaries was first carried out. This extra analysis was developed in two extra subroutines, the first one dealing with the N-terminal residue (cf. subroutines MOJ1, RMJ1) and the second one with the C-terminal residue (cf. subroutines MOJ2, RMJ2).

The following program, for the α-helix search, was eventually adopted.

```
C
C
C     ..........................................................
C     .                                                        .
C     .          MAIN PROGRAM OF HELIX PREDICTION              .
C     .                                                        .
C     ..........................................................
C
C
C     ..........................................................................
C
C     PURPOSE
C        TO READ IN THE SEQUENCE OF THE PROTEIN AND TO ASSIGN TO EACH AMI
C        NO ACID RESIDUE ITS CONFORMATIONAL PARAMETERS(PA,PB,PT) AND ITS
C        BOUNDARY CONFORMATIONAL PARAMETERS(PAN,PAC,PNAN,PNAC,PBN,PBC,
C        PNBN,PBNC)
C
C     ..........................................................................
C
C
      REAL S,T1,T2,A1,A2 ,T3,T4,T5,TT,P
      INTEGER G,F,H,U,D,V1,V2,W, V3,V4,V5,V6,V7,V8,Q
      LOGICAL HELLO,BYE   ,BALL,MOVE
      DIMENSION S(1000,20),M(1000),H(1000),D(1000,16),P(1000,10)
      COMMON S,T1,T2,T3,T4,T5,TT,A1,A2,P,F,H,U,D,W,M,M1,M2,M3,M4,M5,M6,
     1L,I,K,L1,L2,NZ,NY,JA,JB,JC,JD,J1,J2,KM,N1,N2,NN,J,G,K3,V1,V2,V3,V4
     2,V5,V6,V7,Q,HELLO,BYE,BALL,MOVE
C
C
C
C     DESCRIPTION OF PARAMETERS
C        S  - ARRAY RECORDING THE DIFFERENT CONFORMATIONAL PARAMETERS FOR
C             EACH AMINO ACID RESIDUE (K)
C             S(K,1)  - PA
C             S(K,2)  - PB
C             S(K,5)  - PT
C             S(K,6)  - PNAN
C             S(K,7)  - PNAC
C             S(K,8)  - PAN
C             S(K,9)  - PAC
C        T1  - SUM OF PA OF N AMINO ACID RESIDUES
C        T2  - SUM OF PB OF N AMINO ACID RESIDUES
C        T5  - SUM OF PT OF N AMINO ACID RESIDUES
C        T3  - SUM OF THE ASSIGNMENTS AS FORMER,BREAKER,INDIFFERENT IN
C             THE NUCLEATION FRAGMENT
C        T4  - SUM OF THE ASSIGNMENTS AS FORMER,BREAKER,INDIFFERENT IN
C             THE ENTIRE PREDICTED HELICAL AREA
C        TT  - ALLOWED NUMBER OF BREAKERS IN THE ENTIRE PREDICTED AREA
C             ( EQUAL TO ONE THIRD OF THE LENGTH)
C        P   - FREQUENCY OF THE RESIDUES IN A REVERSE B-TURN
```

```
C               P(K,1)  - FREQUENCY OF THE FIRST RESIDUE
C               P(K,2)  - FREQUENCY OF THE SECOND RESIDUE
C               P(K,3)  - FREQUENCY OF THE THIRD RESIDUE
C               P(K,4)  - FREQUENCY OF THE FOURTH RESIDUE
C         M   - ARRAY RECORDING THE NUMERICAL ASSIGNMENT OF EACH AMINO
C               ACID RESIDUE
C         NN  - TOTAL NUMBER OF RESIDUES OF THE PROTEIN
C         N   - NUMBER OF LINES USED TO ENTER THE WHOLE SEQUENCE (16 RESI
C               DUES PER LINE)
C         D   - ARRAY RECORDING THE POSITION OF EACH AMINO ACID RESIDUE
C               ON THE NTH LINE
C               D(K,L)  - AMINO ACID RESIDUE AT POSITION K ON LINE L
C
C     REMARK
C     SOME OF THE PARAMETERS WILL BE DESCRIBED IN THE SUBSEQUENT SUBROU
C     TINES SINCE THEIR DEFINITION MAY CHANGE FROM ONE SUBROUTINE TO ANO
C     THER
C
          PRINT 100
      100 FORMAT('1',35X,'*********************************')
          PRINT 102
      102 FORMAT(' ',35X,'*',31X,'*')
          PRINT 103
      103 FORMAT(' ',35X,'*',4X,'ALPHA-HELIX PREDICTION',5X,'*')
          PRINT 102
          PRINT 104
      104 FORMAT(' ',35X,'*********************************'//)
          READ (5,106)  NN ,N
      106 FORMAT(6X,I4,6X,I4)
          WRITE (6,107)  NN
      107 FORMAT('O','TOTAL NUMBER OF AA:',I7)
          WRITE (6,108)  N
      108 FORMAT(' ','NUMBER OF DATA LINES:',I5,/)
          PRINT 109
      109 FORMAT('O','PROTEIN SEQUENCE')
          PRINT 110
      110 FORMAT(' ','................'/)
          READ(5,111)  ((D(J,K),K=1,16),J=1,N)
      111 FORMAT(16I5)
          WRITE (6,112) ((D(J,K),K=1,16),J=1,N)
      112 FORMAT(' ',16I5)
C
C     TO CHECK THE NUMERICAL ASSIGNMENT OF EACH AMINO ACID RESIDUE IN
C     THE SEQUENCE SO TO ASSIGN ITS CORRESPONDING CONFORMATIONAL PARAMETERS
C
          I=1
          DO 21  J=1,N
          DO 22  K=1,16
          M(I)=D(J,K)
          IF  (M(I).EQ.0)  GO TO 999
```

```
                I=I+1
          22   CONTINUE
          21   CONTINUE
          999   DO 32  K=1,NN
                IF (M(K).EQ.1)   GO TO 1
                IF (M(K).EQ.2)   GO TO 2
                IF (M(K).EQ.3)   GO TO 3
                IF (M(K).EQ.4)   GO TO 4
                IF (M(K).EQ.5)   GO TO 5
                IF (M(K).EQ.6)   GO TO 6
                IF (M(K).EQ.7)   GO TO 7
                IF (M(K).EQ.8)   GO TO 8
                IF (M(K).EQ.9)   GO TO 9
                IF (M(K).EQ.10)  GO TO 10
                IF (M(K).EQ.11)  GO TO 11
                IF (M(K).EQ.12)  GO TO 12
                IF (M(K).EQ.13)  GO TO 13
                IF (M(K).EQ.14)  GO TO 14
                IF (M(K).EQ.15)  GO TO 15
                IF (M(K).EQ.16)  GO TO 16
                IF (M(K).EQ.17)  GO TO 17
                IF (M(K).EQ.18)  GO TO 18
                IF (M(K).EQ.19)  GO TO 19
                IF (M(K).EQ.20)  GO TO 20
                IF (M(K).EQ.25)  GO TO 25
      C
      C
          1     S(K,1)=1.42
                S(K,2)=0.83
                S(K,5)=0.66
                S(K,6)=0.70
                S(K,7)=0.52
                S(K,8)=1.29
                S(K,9)=1.20
                P(K,1)=0.060
                P(K,2)=0.076
                P(K,3)=0.035
                P(K,4)=0.058
                GO TO 32
          2     S(K,1)=0.98
                S(K,2)=0.93
                S(K,5)=0.95
                S(K,6)=0.34
                S(K,7)=1.24
                S(K,8)=0.44
                S(K,9)=1.25
                P(K,1)=0.070
                P(K,2)=0.106
                P(K,3)=0.099
                P(K,4)=0.085
```

```
151             GO TO 32
152         3   S(K,1)=0.67
153             S(K,2)=0.89
154             S(K,5)=1.56
155             S(K,6)=1.42
156             S(K,7)=1.64
157             S(K,8)=0.81
158             S(K,9)=0.59
159             P(K,1)=0.161
160             P(K,2)=0.083
161             P(K,3)=0.191
162             P(K,4)=0.091
163             GO TO 32
164         4   S(K,1)=1.01
165             S(K,2)=0.54
166             S(K,5)=1.46
167             S(K,6)=0.98
168             S(K,7)=1.06
169             S(K,8)=2.02
170             S(K,9)=0.61
171             P(K,1)=0.147
172             P(K,2)=0.110
173             P(K,3)=0.179
174             P(K,4)=0.081
175             GO TO 32
176         5   S(K,1)=0.70
177             S(K,2)=1.19
178             S(K,5)=1.19
179             S(K,6)=0.65
180             S(K,7)=0.94
181             S(K,8)=0.66
182             S(K,9)=1.11
183             P(K,1)=0.149
184             P(K,2)=0.053
185             P(K,3)=0.117
186             P(K,4)=0.128
187             GO TO 32
188         6   S(K,1)=1.11
189             S(K,2)=1.10
190             S(K,5)=0.98
191             S(K,6)=0.75
192             S(K,7)=0.70
193             S(K,8)=1.22
194             S(K,9)=1.22
195             P(K,1)=0.074
196             P(K,2)=0.098
197             P(K,3)=0.037
198             P(K,4)=0.098
199             GO TO 32
200         7   S(K,1)=1.51
```

```
201                 S(K,2)=0.37
202                 S(K,5)=0.74
203                 S(K,6)=1.04
204                 S(K,7)=0.59
205                 S(K,8)=2.44
206                 S(K,9)=1.24
207                 P(K,1)=0.056
208                 P(K,2)=0.060
209                 P(K,3)=0.077
210                 P(K,4)=0.064
211                 GO TO 32
212           8     S(K,1)=0.57
213                 S(K,2)=0.75
214                 S(K,5)=1.56
215                 S(K,6)=1.41
216                 S(K,7)=1.64
217                 S(K,8)=0.76
218                 S(K,9)=0.42
219                 P(K,1)=0.102
220                 P(K,2)=0.085
221                 P(K,3)=0.190
222                 P(K,4)=0.152
223                 GO TO 32
224           9     S(K,1)=1.00
225                 S(K,2)=0.87
226                 S(K,5)=0.95
227                 S(K,6)=1.22
228                 S(K,7)=1.86
229                 S(K,8)=0.73
230                 S(K,9)=1.77
231                 P(K,1)=0.140
232                 P(K,2)=0.047
233                 P(K,3)=0.093
234                 P(K,4)=0.054
235                 GO TO 32
236          10     S(K,1)=1.08
237                 S(K,2)=1.60
238                 S(K,5)=0.47
239                 S(K,6)=0.78
240                 S(K,7)=0.87
241                 S(K,8)=0.67
242                 S(K,9)=0.98
243                 P(K,1)=0.043
244                 P(K,2)=0.034
245                 P(K,3)=0.013
246                 P(K,4)=0.056
247                 GO TO 32
248          11     S(K,1)=1.21
249                 S(K,2)=1.30
250                 S(K,5)=0.59
```

```
251                 S(K,6)=0.85
252                 S(K,7)=0.084
253                 S(K,8)=0.58
254                 S(K,9)=1.13
255                 P(K,1)=0.061
256                 P(K,2)=0.025
257                 P(K,3)=0.036
258                 P(K,4)=0.070
259                 GO TO 32
260           12    S(K,1)=1.16
261                 S(K,2)=0.74
262                 S(K,5)=1.01
263                 S(K,6)=1.01
264                 S(K,7)=1.49
265                 S(K,8)=0.66
266                 S(K,9)=1.83
267                 P(K,1)=0.055
268                 P(K,2)=0.115
269                 P(K,3)=0.072
270                 P(K,4)=0.095
271                 GO TO 32
272           13    S(K,1)=1.45
273                 S(K,2)=1.05
274                 S(K,5)=0.60
275                 S(K,6)=0.83
276                 S(K,7)=0.52
277                 S(K,8)=0.71
278                 S(K,9)=1.57
279                 P(K,1)=0.068
280                 P(K,2)=0.082
281                 P(K,3)=0.014
282                 P(K,4)=0.055
283                 GO TO 32
284           14    S(K,1)=1.13
285                 S(K,2)=1.38
286                 S(K,5)=0.60
287                 S(K,6)=0.93
288                 S(K,7)=1.04
289                 S(K,8)=0.61
290                 S(K,9)=1.10
291                 P(K,1)=0.059
292                 P(K,2)=0.041
293                 P(K,3)=0.065
294                 P(K,4)=0.065
295                 GO TO 32
296           15    S(K,1)=0.57
297                 S(K,2)=0.55
298                 S(K,5)=1.52
299                 S(K,6)=1.10
300                 S(K,7)=1.58
```

```
301             S(K,8)=2.01
302             S(K,9)=0.00
303             P(K,1)=0.102
304             P(K,2)=0.301
305             P(K,3)=0.034
306             P(K,4)=0.068
307             GO TO 32
308         16  S(K,1)=0.77
309             S(K,2)=0.75
310             S(K,5)=1.43
311             S(K,6)=1.55
312             S(K,7)=0.93
313             S(K,8)=0.74
314             S(K,9)=0.96
315             P(K,1)=0.120
316             P(K,2)=0.139
317             P(K,3)=0.125
318             P(K,4)=0.106
319             GO TO 32
320         17  S(K,1)=0.83
321             S(K,2)=1.19
322             S(K,5)=0.96
323             S(K,6)=1.09
324             S(K,7)=0.86
325             S(K,8)=1.08
326             S(K,9)=0.75
327             P(K,1)=0.086
328             P(K,2)=0.108
329             P(K,3)=0.065
330             P(K,4)=0.079
331             GO TO 32
332         18  S(K,1)=1.08
333             S(K,2)=1.37
334             S(K,5)=0.96
335             S(K,6)=0.62
336             S(K,7)=0.16
337             S(K,8)=1.47
338             S(K,9)=0.40
339             P(K,1)=0.077
340             P(K,2)=0.013
341             P(K,3)=0.064
342             P(K,4)=0.167
343             GO TO 32
344         19  S(K,1)=0.69
345             S(K,2)=1.47
346             S(K,5)=1.14
347             S(K,6)=0.99
348             S(K,7)=0.96
349             S(K,8)=0.68
350             S(K,9)=0.73
```

```
      P(K,1)=0.082
      P(K,2)=0.065
      P(K,3)=0.114
      P(K,4)=0.125
      GO TO 32
20    S(K,1)=1.06
      S(K,2)=1.70
      S(K,5)=0.50
      S(K,6)=0.75
      S(K,7)=0.32
      S(K,8)=0.61
      S(K,9)=1.25
      P(K,1)=0.062
      P(K,2)=0.048
      P(K,3)=0.028
      P(K,4)=0.053
      GO TO 32
25    S(K,1)=0.00
      S(K,2)=0.00
      S(K,5)=0.00
      S(K,6)=0.00
      S(K,7)=0.00
      S(K,8)=0.00
      S(K,9)=0.00
      P(K,1)=0.00
      P(K,2)=0.00
      P(K,3)=0.00
      P(K,4)=0.00
32    CONTINUE
C
C
      PRINT 40
40    FORMAT('-',12X,'PRELIMINARY SEARCH FOR REGIONS WITH HELIX POTENTIA
     1L - RULE 1')
      PRINT 41
41    FORMAT(' ',12X,'.....
     1.........')
C
C     TO CALL SUBROUTINE ONE TO CARRY OUT THE PRELIMINARY SEARCH OF HELI
C     CAL REGIONS
C
      CALL ONE
      STOP
      END
End of File
```

```
C
C
      SUBROUTINE ONE
C     ..............
C
C
C     ..........................................................
C     .                                                        .
C     .           PRELIMINARY SEARCH FOR HELICAL REGIONS       .
C     .                                                        .
C     ..........................................................
C
C
C     ..................................................................
C
C     PURPOSE
C        PRELIMINARY SEARCH FOR HELICAL REGIONS BY APPLYING RULE 1 :
C        <PA> > 1.03 AND <PA> > <PB>
C
C     ..................................................................
C
C
      REAL S,T1,T2,A1,A2 ,T3,T4,T5,TT,P
      INTEGER G,F,H,U,D,V1,V2,W, V3,V4,V5,V6,V7,V8,Q
      LOGICAL HELLO,BYE  ,BALL,MOVE
      DIMENSION S(1000,20),M(1000),H(1000),D(1000,16),P(1000,10)
      COMMON S,T1,T2,T3,T4,T5,TT,A1,A2,P,F,H,U,D,W,M,M1,M2,M3,M4,M5,M6,
     1L,I,K,L1,L2,NZ,NY,JA,JB,JC,JD,J1,J2,KM,N1,N2,NN,J,G,K3,V1,V2,V3,V4
     2,V5,V6,V7,Q,HELLO,BYE,BALL,MOVE
C
C
C     DESCRIPTION OF PARAMETERS
C
C        H    - BOUNDARY RESIDUES OF A PREDICTED REGION
C               H(K)  - N-TERMINAL RESIDUE
C               H(K+1)- C-TERMINAL RESIDUE
C        J    - FIRST RESIDUE OF A SECTION TO BE CONSIDERED FOR THE PRELI
C               MINARY SEARCH BUT WILL CHANGE DURING N-PROPAGATION (J-1)
C        JA   - FIRST RESIDUE OF A SECTION TO BE CONSIDERED FOR THE PRELI
C               MINARY SEARCH BUT WILL CHANGE DURING C-PROPAGATION (JA+1)
C        N1   - FIRST RESIDUE OF A SECTION TO BE CONSIDERED FOR THE PRELI
C               MINARY SEARCH
C        N2   - LAST RESIDUE OF A SECTION  TO BE CONSIDERED FOR THE PRELI
C               MINARY SEARCH
C        A1   - AVERAGE <PA>  OF A SECTION
C        A2   - AVERAGE <PB>  OF A SECTION
C        I    - SWITCHING VALUE FOR DECISION MAKING
C               I=1   N-PROPAGATION
C               I=2   C-PROPAGATION
C        K    - COUNTER USED WITH THE ARRAY H TO STORE THE BOUNDARY RESI
```

```
C               DUES OF PREDICTED REGIONS
C
C     THE SEARCH WILL STOP WHEN THE LAST SEGMENT AT THE C-TERMINAL HAS
C     ONLY 5 AMINO ACID RESIDUES. IT IS NOT LONG ENOUGH FOR THE HELICAL
C     STATE
C
   10     K=2
          H(K)=0
          H(K-1)=0
          NZ=NN-5
          J=1
          JA=1
   15     I=0
   20     N2=JA+5
          HELLO=.FALSE.
          IF (J.EQ.H(K))  HELLO=.TRUE.
          IF (HELLO)      N1=H(K)+1
          IF (.NOT.HELLO)  N1=J
C
C     IF ARG OR CYS IS AT THE C-TERMINAL THEY CAN BE ADDED TO THE POTEN
C     TIAL FRAGMENT BECAUSE OF THEIR GOOD PAC VALUE
C
C
C     TO CALCULATE THE AVERAGE <PA>,<PB> AND TO COUNT THE NUMBER OF BREA
C     KERS IN THE SECTION N1-N2
C
          T1 = 0
          T2 = 0
          L  = 0
          LB = 0
          L  = N1+1+(N2-N1)/2
          DO  25  LN=L,N2
          IF (M(LN).EQ.2.OR.M(LN).EQ.5)  S(LN,1)=1.00
   25     CONTINUE
          DO 30  L=N1,N2
          T1 = T1 + S(L,1)
          T2 = T2 + S(L,2)
          IF (S(L,1).LE.0.69)  LB=LB+1
   30     CONTINUE
          A1 = T1/(N2-N1+1)
          A2 = T2/(N2-N1+1)
C
C
C     IF <PA> < 1.03 TO START THE SEARCH AGAIN FROM NEXT POSITION J+1
C
          IF (A1.LT.1.03000000)  GO TO 45
C
C     SPECIAL SITUATION WHERE THE SECTION MAY HAVE HELICAL POTENTIAL EVEN
C     THOUGH <PA> < <PB>
C
```

```
101              IF (A1.GT.1.1100.AND.(A2-A1).LT.0.0640.AND.M(N1).EQ.4.AND.M(N2+1)
102             1.EQ.2.AND.M(N1+3).EQ.1.AND.M(N1-1).EQ.17.AND.M(N1-2).EQ.8.AND.
103             2S(N2,1).GT.1.01.AND.LB.EQ.0)  GO TO 60
104      C
105      C     IF <PA> < <PB> EVEN IF <PA> > 1.03 TO START SEARCH AGAIN FROM NEXT
106      C     POSITION J+1 UNLESS THE LAST AMINO ACID RESIDUE HAS BEEN REACHED
107      C
108              IF (A1.LT.A2 .AND. N2.EQ.NN .AND.(N2+1-N1).EQ.6)  GO TO 80
109              IF (A1.LT.A2 .AND. N2.EQ.NN .AND.(N2+1-N1).GT.6)  GO TO 70
110              IF (A1.LT.A2 .AND. N2.NE.NN)  GO TO 45
111      C
112      C     TO PROPAGATE AT THE C-TERMINAL SIDE WHEN THE SEARCH HAS NOT REACHED
113      C     THE LAST AMINO ACID RESIDUE YET (NN)
114      C
115              IF (I.EQ.2 .AND. N2.EQ.NN)  GO TO 55
116              IF (I.EQ.2 .AND. N2.NE.NN)  GO TO 50
117      C
118      C     TO START N-PROPAGATION WHEN <PA> > 1.03 AND <PA> > <PB> UNLESS THE
119      C     HELICAL SEGMENT STARTS FROM POSITION 1
120      C
121              J=J-1
122              I = 1
123              BYE=.FALSE.
124              IF (J.EQ.H(K))   BYE=.TRUE.
125              IF (BYE)   I = 2
126              IF (BYE)   GO TO 50
127              IF (.NOT. BYE)  GO TO 20
128      C
129      C     TO SWITCH FROM N-PROPAGATION TO C-PROPAGATION WHEN THE REMAINING
130      C     SECTION OF THE SEQUENCE HAS MORE THAN 5 RESIDUES
131      C
132          45  J=J+1
133              IF (I.EQ.2)  GO TO 70
134              IF (I.EQ.1)  I = 2
135          50  JA=JA+1
136              IF (JA.LE.NZ)  GO TO 20
137              IF (JA.GT.NZ)  GO TO 80
138      C
139      C     TO PRINT OUT THE LAST HELIX POTENTIAL AREA H(K),H(K+1) AND THE MA
140      C     XIMUM VALUE OF THE COUNTER K WHICH WILL BE USED INT THE NEXT SUBROU
141      C     TINE
142      C
143          55  K=K+1
144              H(K)=N1
145              K=K+1
146              H(K)=N2
147              PRINT  58,H(K-1),H(K)
148          58  FORMAT('0',30X,I6,10X,I6)
149              KM=K
150              GO TO 80
```

```
151     C
152     C
153     C     TO PRINT OUT THE HELIX POTENTIAL AREAS H(K),H(K+1),THEN THE PRELI
154     C     MINARY SEARCH STARTS AGAIN FROM POSITION (H(K+1) +1)
155     C
156         60     K=K+1
157                H(K)=N1
158                K=K+1
159                H(K)=N2
160                GO TO 75
161         70     K=K+1
162                H(K)=N1
163                K=K+1
164                H(K)=N2-1
165         75     PRINT 78,H(K-1),H(K)
166         78     FORMAT('O',30X,I6,10X,I6)
167                J=H(K)
168                JA=H(K)
169                KM=K
170                IF (JA.LE.NZ)  GO TO 15
171         80     PRINT  85 ,KM
172         85     FORMAT('O',40X,'KM:',I4)
173                K=2
174                W=1
175                PRINT 90
176         90     FORMAT('-',12X,'SEARCH FOR ACTUAL HELICES FROM THE POTENTIAL REGIO
177               1NS')
178                PRINT 95
179         95     FORMAT(' ',12X,'.....................................................
180               1..'//)
181     C
182     C     TO CALL SUBROUTINE TWO TO CARRY OUT THE NUCLEATION SEARCH ON THOSE
183     C     POTENTIAL AREAS
184     C
185                CALL TWO
186                RETURN
187                END
End of File
```

```
C
C
      SUBROUTINE TWO
C     .............
C
C
C     ..................................................
C     .                                                .
C     .          SEARCH FOR HELIX NUCLEATION           .
C     .                                                .
C     ..................................................
C
C
C
C     ..................................................................
C
C     PURPOSE
C        SEARCH FOR NUCLEATING HELICAL REGIONS WHICH SHOULD CONTAIN AT
C        LEAST 4 FORMERS OUT OF 6 RESIDUES
C
C     ..................................................................
C
      REAL S,T1,T2,A1,A2 ,T3,T4,T5,TT,P
      INTEGER G,F,H,U,D,V1,V2,W, V3,V4,V5,V6,V7,V8,Q
      LOGICAL,HELLO,BYE  ,BALL,MOVE
      DIMENSION S(1000,20),M(1000),H(1000),D(1000,16),P(1000,10)
      COMMON S,T1,T2,T3,T4,T5,TT,A1,A2,P,F,H,U,D,W,M,M1,M2,M3,M4,M5,M6,
     1L,I,K,L1,L2,NZ,NY,JA,JB,JC,JD,J1,J2,KM,N1,N2,NN,J,G,K3,V1,V2,V3,V4
     2,V5,V6,V7,Q,HELLO,BYE,BALL,MOVE
C
C
C     DESCRIPTION OF PARAMETERS
C        J    - FIRST RESIDUE OF THE 6 RESIDUE PEPTIDE SUBJECTED TO THE
C               NUCLEATION SEARCH
C        JA   - SIXTH RESIDUE OF THE 6 RESIDUE PEPTIDE SUBJECTED TO THE
C               NUCLEATION SEARCH
C        W    - SWITCHING VALUE FOR DECISION MAKING
C               W=1    THE CURRENT POTENTIAL AREA IS STILL LONG ENOUGH (>
C                      6 RESIDUES) TO BE SUBJECTED TO THE NUCLEATION SEARCH
C               W=2    THE CURRENT POTENTIAL AREA IS TOO SHORT FOR ANOTHER
C                      HELIX SO TO START WITH THE NEXT POTENTIAL AREA
C
C     REMARKS
C     UNLESS NOTIFIED THE OTHER PARAMETERS STILL HAVE THE SAME DEFINITION
C
C
C     IF W=2  THE NUCLEATION SEARCH WILL START ON A NEW POTENTIAL AREA
C     SINCE THE PREVIOUS ONE HAS BEEN THOUROUGHLY ANALYZED. EACH TIME K
C     INCREASES BY 1 THE NEXT POTENTIAL AREA IS SUBJECTED TO THE NUCLEA
C     TION PROCEDURE
```

```
      C
   10     IF (W.EQ.2)  GO TO 20
   15     K=K+1
          IF (K.GT.KM)  GO TO 170
          N1=H(K)
          K=K+1
          N2=H(K)
          IF (W.EQ.1)  J=N1
          NY=N2-5
   20     JA=J+5
      C
      C     TO COUNT THE DIFFERENT TYPES OF ASSIGNMENTS (T3) AND THE NUMBER OF
      C     BREAKERS (L) IN THE SEGMENT J-JA
      C        S(I,3) = 0.0 IF RESIDUE I IS A BREAKER OR AN INDIFFERENT
      C        S(I,3) = 0.5 IF RESIDUE I IS A WEAK FORMER
      C        S(I,3) = 1.0 IF RESIDUE I IS A FORMER
      C
          T3 = 0
          L  = 0
          DO  25 I=J,JA
          S(I,3)=0
          IF (S(I,1).GE.1.00) S(I,3)=0.5
          IF (S(I,1).GE.1.06) S(I,3)=1.0
          T3 = T3+S(I,3)
          IF (S(I,1).LE.0.69)  L=L+1
   25     CONTINUE
          PRINT 30,J,JA,T3,L
   30     FORMAT(' ',10X,'J :',I4,5X,'JA:',I4,5X,'T3:',F7.4,5X,'L:',I3,5X,
         1'HELIX NUCLEATION')
      C
      C     IF CASE ARG IS AT THE C-TERMINAL IT MAY SWITCH FROM INDIFFERENT TO
      C     FORMER SO THAT THE NUCLEATION RULE CAN BE SATISFIED
      C
          IF (T3.EQ.3.5.AND.M(JA).EQ.2) S(JA,1)=1.00
          IF (T3.EQ.3.5.AND.M(JA).EQ.2) T3=4.0
      C
      C
      C     LIST OF SPECIAL SITUATIONS WHERE THE NUCLEATION RULE AND THE TYPES
      C     OF RESIDUES IN THE SEGMENT SHOULD BE COMBINED TOGETHER SINCE THE
      C     NUCLEATION RULE BY ITSELF IS TOO DISCRIMINATIVE
      C
          IF ((JA+2).GT.NN.OR.(J-2).LE.0)  GO TO 35
          IF (T3.GE.4.0.AND.L.LE.2.AND.M(J).EQ.7.AND.M(JA).EQ.1.AND.M(JA-1)
         1.EQ.7.AND.M(J+3).EQ.1.AND.M(J-2).EQ.1.AND.S(JA+1,1).GT.1.16.AND.
         3S(JA+2,1).GT.1.16)  GO TO 90
      C
   35     IF ((JA+1).GT.NN.OR.(J-3).LE.0)  GO TO 40
          IF (T3.GE.4.5.AND.L.EQ.1.AND.M(J+3).EQ.15.AND.M(J).EQ.4.AND.M(J+2)
         1.EQ.11.AND.M(J+4).EQ.7.AND.M(JA).EQ.18.AND.S(J-3,8).GE.2.01.AND.
         2 S(J-2,1).GT.1.16.AND.S(JA+1,1).GT.1.01)  GO TO 110
```

```
      C
   40   IF ((JA+2).GT.NN.OR.(J-4).LE.0)  GO TO 45
        IF (T3.GE.3.5.AND.L.EQ.0.AND.M(J).EQ.14.AND.M(J+1).EQ.1.AND.M(JA+1
     1).EQ.7.AND.M(JA-1).EQ.9.AND.S(J-1,1).LT.0.67.AND.S(JA+2,1).LE.0.69
     2 .AND.S(J-2,8).LT.1.08.AND.S(J-3,8).GT.1.47.AND.S(J-4,8).LT.1.08)
     3  GO TO 120
      C
   45   IF ((J-2).LE.0)  GO TO 50
        IF (T3.GE.4.0.AND.L.LE.2.AND.M(J+1).EQ.6.AND.M(J+3).EQ.1.AND.M(JA
     1-1).EQ.11.AND.M(JA).EQ.11.AND.S(J,1).LE.0.69.AND.S(J-1,1).LE.0.69
     2.AND.S(J-2,8).GT.1.47)  GO TO 110
      C
   50.  IF ((JA+2).GT.NN)  GO TO 55
        IF (T3.GE.4.0.AND.L.EQ.1.AND.M(J).EQ.15.AND.M(J+1).EQ.4.AND.M(JA)
     1.EQ.7.AND.M(J+3).EQ.5.AND.M(JA+1).EQ.15.AND.M(JA+2).EQ.15)  GO TO
     2  130
      C
   55   IF ((JA+1).GT.NN.OR.(J-2).LE.0)  GO TO 60
        IF (T3.GE.3.5.AND.L.EQ.0.AND.M(J-1).EQ.15.AND.S(J+2,1).GT.1.21.AND
     1 .S(J+3,1).GT.1.16.AND.S(J+4,1).GT.1.16.AND.S(JA,9).GT.0.75.AND.S(
     2 JA+1,9).GT.0.75.AND.S(J-2,8).LT.1.08)  GO TO 140
      C
   60   IF ((JA+2).GT.NN.OR.(J-1).LE.0)  GO TO 65
        IF (T3.GE.3.0.AND.L.EQ.3.AND.M(J).EQ.15.AND.M(J+2).EQ.1.AND.M(J+3)
     1.EQ.20.AND.M(J+4).EQ.13.AND.M(JA).EQ.8.AND.S(JA+1,2).LT.0.93.AND.
     2S(JA+2,2).LT.0.74.AND.S(J-1,2).LT.0.93)  GO TO 130
      C
   65   IF ((JA+7).GT.NN)  GO TO 70
        IF (T3.EQ.2.5.AND.L.EQ.2.AND.M(J+2).EQ.14.AND.S(J+3,8).GT.2.02.AND
     1 .S(J+4,1).GT.0.77.AND.S(JA,1).GT.0.83.AND.S(JA+1,8).GT.2.01.AND.
     2 S(JA+5,9).GT.1.10.AND.S(JA+5,1).GT.1.16.AND.M(JA+2).NE.15.AND.S(
     3 JA+3,9).GT.1.10.AND.S(JA+4,9).GT.1.24.AND.S(JA+6,9).LT.1.10.AND.S
     4 (JA+7,1).LT.1.06)  GO TO 150
      C
   70   IF ((JA+1).GT.NN.OR.(J-3).LE.0)  GO TO 75
        IF (T3.GE.4.0.AND.L.LE.1.AND.M(JA).EQ.15.AND.S(JA+1,1).LT.S(JA-1,1
     1 ).AND.S(JA-1,1).GT.1.16.AND.S(JA-1,9).GT.1.10.AND.S(J,8).GT.2.01
     2 .AND.S(J-1,8).GT.1.29.AND.S(J-2,6).GT.1.09.AND.S(J-3,8).LT.S(J-1
     3 ,8).AND.S(J+1,1).GT.1.16.AND.S(J+2,9).GT.1.10.AND.S(J+3,9).GT.1.
     4 20)  GO TO 140
      C
   75   IF ((JA+3).GT.NN.OR.(J-3).LE.0)  GO TO 80
        IF (T3.GE.3.50.AND.L.LE.1.AND.M(J).EQ.15.AND.S(JA,9).GE.1.10.AND.S
     1 (JA,1).GT.1.08.AND.S(JA+1,1).LE.0.69.AND.S(JA+2,1).LE.0.69.AND.S(
     2 JA+3,9).LT.0.98.AND.S(J+1,1).GT.1.16.AND.S(J+2,1).GT.1.16.AND.S(J
     3 +3,9).GE.1.57.AND.S(J+4,9).GT.0.75.AND.S(J-1,1).GT.1.16.AND.S(J-2
     4 ,1).GT.1.13.AND.S(J-3,1).GT.1.01)  GO TO 130
      C
      C     THE NUCLEATION RULE BY ITSELF IS THE CRITERIA FOR SELECTION IF NO
      C     NE OF THE ABOVE CONDITIONS IS SATISFIED
```

```
      C
   80     IF (T3.GE.4.0.AND.L.LT.2)  GO TO 90
      C
      C
      C
      C     THE NUCLEATION SEARCH FAILED FOR THE SEGMENT J-JA. TO START AGAIN
      C     FROM NEXT POSITION J+1
            J=J+1
            IF (J.LE.NY)  GO TO 20
            GO TO 15
      C
      C
      C     A VALID NUCLEATION SEGMENT ACCORDING TO RULE HELIX-4 SHOULD NOT HA
      C     VE PRO RESIDUE IN THE INNER HELIX
      C
   90     DO 95  I=J,JA
            IF (M(I).EQ.15 .AND. I.EQ.(J+2))  J1=J
            IF (M(I).EQ.15 .AND. I.EQ.(J+2))  GO TO 100
            IF (M(I).EQ.15 .AND. I.NE.(J+2))  GO TO 105
   95     CONTINUE
      C
      C     TO CALL SUBROUTINE THREE FOR THE PROPAGATION OF THE VALID NUCLEATI
      C     NG SEGMENT
      C
  100     CALL THRE
            GO TO 10
      C
      C     THE PRESENCE OF PRO IN THE INNER HELIX IS UNFAVORABLE TO THE NUCLE
      C     TION SO TO START THE SEARCH AGAIN FROM NEXT POSITION J+1
      C
  105     J=J+1
            IF (J.LE.NY)  GO TO 20
            GO TO 15
      C
      C
      C     TO PRINT OUT THE POSSIBLE HELICAL REGIONS WHICH ARE THEN SUBJECTED
      C     TO THE BOUNDARY ADJUSTMENT(SUBROUTINE MOJ1). THOSE ARE ALSO SPECIAL
      C     CASES BECAUSE THE PROPAGATION PROCEDURE IS OMITTED
      C
  110     PRINT 115,J,JA
  115     FORMAT('0',10X,'PSEUDO HELIX FROM',5X,'J :',I5,3X,'TO JA:',I5,10X,
           1'SPECIAL CASE'/)
            J1=J
            J2=JA
            GO TO 160
      C
  120     J1=J - 3
            J2=JA+1
            PRINT 125,J1,J2
  125     FORMAT('0',10X,'PSEUDO-HELIX FROM',5X,'J1:',I5,3X,'TO J2:',I5,10X,
```

```
     1'SPECIAL CASE'/)
      GO TO 165
C
130   J1=J
      J2=JA
      PRINT 125,J1,J2
      GO TO 165
C
140   J1=J-1
      J2=JA-1
      PRINT 125,J1,J2
      GO TO 165
C
150   J1= J+2
      J2=JA+5
      PRINT 125,J1,J2
      GO TO 165
C
C
C     TO CALL SUBROUTINE MOJ1 FOR THE BOUNDARY ADJUSTMENT OF THE PREDIC
C     TED AREA. WHEN RETURNING FROM THAT PROCEDURE IF THE POTENTIAL AREA
C     IS NOT LONG ENOUGH FOR ANOTHER HELIX THEN TO START ANALYZING THE
C     NEXT POTENTIAL AREA
C
160   CALL MOJ1
165   IF (J2.LT.NY)  J=J2+1
      IF (J2.LT.NY)  W=2
      IF (J2.GE.NY)  W=1
      GO TO 10
C
170   PRINT 175
175   FORMAT('-','END OF PROGRAM')
      RETURN
      END
End of File
```

```
      C
      C
            SUBROUTINE THRE
      C     .............
      C
      C
      C           ..................................................
      C           .                                                .
      C           .       PROPAGATION OF THE ALPHA-HELIX           .
      C           .                                                .
      C           ..................................................
      C
      C
      C     ....................................................................
      C
      C     PURPOSE
      C        TO ADD TO THE NUCLEATING FRAGMENT TETREPEPTIDES WHICH HAVE
      C        <PA> > 1.00 AND WHICH SATISFY THE PROPAGATION SET OF RULES
      C
      C     ....................................................................
      C
      C
            REAL S,T1,T2,A1,A2 ,T3,T4,T5,TT,P
            INTEGER G,F,H,U,D,V1,V2,W, V3,V4,V5,V6,V7,V8,Q
            LOGICAL HELLO,BYE  ,BALL,MOVE
            DIMENSION S(1000,20),M(1000),H(1000),D(1000,16),P(1000,10)
          COMMON S,T1,T2,T3,T4,T5,TT,A1,A2,P,F,H,U,D,W,M,M1,M2,M3,M4,M5,M6,
         1L,I,K,L1,L2,NZ,NY,JA,JB,JC,JD,J1,J2,KM,N1,N2,NN,J,G,K3,V1,V2,V3,V4
         2,V5,V6,V7,Q,HELLO,BYE,BALL,MOVE
      C
      C
      C     DESCRIPTION OF PARAMETERS
      C        JB  - WHETHER IT IS N- OR C-PROPAGATION JB WILL ALWAYS BE THE
      C              FIRST LEFT RESIDUE OF THE ADJACENT TETRAPEPTIDE
      C        JC  - WHETHER IT IS N- OR C-PROPAGATION JC WILL ALWAYS BE THE
      C              FOURTH RESIDUE OF THE ADJACENT TETRAPEPTIDE
      C        N1  - N-TERMINAL RESIDUE OF THE CURRENT POTENTIAL AREA
      C        U   - SWITCHING VALUE FOR DECISION MAKING
      C              U=1   N-PROPAGATION
      C              U=2   C-PROPAGATION
      C
      C
      C     IF PRO OCCUPY THE FIRST TURN OF THE NUCLEATING SEGMENT TO START C-
      C     PROPAGATION IMMEDIATELY BECAUSE N-PROPAGATION IS NOT POSSIBLE ACCOR
      C     DING TO RULE HELIX-4
      C
         10     M1=0
                M2=0
                M3=0
                M4=1
```

```
      M6=0
      IF (M(I).EQ.15 .AND. I.EQ.(J+2))  GO TO 25
      U=1
C
C     AS LONG AS JB BELONGS TO THE CURRENT POTENTIAL AREA THE N-PROPAGA
C     TION CAN BE CARRIED OUT
20    M1=M1+1
      JB=J-(4*M1)
      IF (JB.GT.0.AND.JB.GE.N1)  GO TO 30
      IF (JB.LT.N1.AND.M1.EQ.1)  J1=J
      IF (JB.LT.N1.AND.M1.NE.1)  J1=J-4*(M1-1)
25    U=2
      M2=0
30    T3=0
      IF (U.EQ.1)  GO TO 35
C
C     TO START C-PROPAGATION WHEN N-PROPAGATION HAS BEEN STOPPED AND AS
C     LONG AS THE ADJACENT TETRAPEPTIDE IS WITHIN THE LIMITS OF THE POT
C     ENTIAL AREA
C
      IF (M2.NE.0)  JB=JA+1+(4*M2)
      IF (M2.EQ.0)  JB=JA+1
      M2=M2+1
      IF (JB.GT.N2)  GO TO 70
C
C     TO CALCULATE THE <PA> OF THE ADJACENT TETRAPEPTIDE (JB-JC)
C
35    JC=JB+3
      IF (JC.GT.N2 .AND.JB.LE.N2)  GO TO 70
      DO 40  I=JB,JC
      T3=T3+S(I,1)
40CONTINUE
C
      PRINT  45,JB,JC,T3
45    FORMAT(' ',10X,'JB:',I4,5X,'JC:',I4,5X,'T3:',F7.4,15X,'HELIX PROPA
     1GATION')
C
C
C     IF <PA> > 1.00 TO CHECK THE NUMBER OF BREAKERS AND FORMERS IN THE
C     SECTION FORMED BY THE TETRAPEPTIDE AND THE TWO ADJACENT RESIDUES
C     OF THE NUCLEATING FRAGMENT OR OF THE PROPAGATING ONE
C
      IF (T3.GE.4.0)  GO TO 190
C
C     TETRAPEPTIDES WITH <PA> <1.00 SHOULD NOT CONTAIN ANY BREAKER NOR
C     ONLY 4 IA IN ORDER TO ALLOW HELIX PROPAGATION TO CONTINUE
C
      DO 50  I=JB,JC
      IF (S(I,1).LE.0.69)  GO TO 60
50    CONTINUE
```

```
      L=0
      DO 55  I=JB,JC
      IF (S(I,1).LE.1.01 .AND.S(I,1).GE.0.70)  L=L+1
55    CONTINUE
      IF (L.EQ.4)  GO TO 60
      IF (L.NE.4)  GO TO 190
C
C     TO SWITCH TO C-PROPAGATION WHEN N-PROPAGATION HAS BEEN STOPPED
C
60    BALL=.FALSE.
      IF (U.EQ.1)  BALL=.TRUE.
      IF (BALL)  J1=JB+4
      IF (BALL)  U=2
      IF (BALL)  GO TO 30
C
C
C     BOTH N- AND C-PROPAGATIONS BY TETRAPEPTIDE ADDITION HAVE BEEN STOP
C     PED. TO START ADDING ONE RESIDUE AT A TIME TO N-TERMINAL FIRST
C     THEN TO C-TERMINAL OF THE PROPAGATING SECTION.
C     WHEN ADDING IA TO EACH END TO CHECK IMMEDIATELY WHETHER THE RULE
C     OF AT LEAST HALF OF FORMERS IS STILL SATISFIED OR NOT
C
70    IF (M(J1+2).EQ.15)  GO TO 80
75    L1=J1-1
      IF (L1.LT.(N1)  .OR.L1.EQ.0)  GO TO 80
      IF (M(L1).EQ.4 .OR.M(L1).EQ.17)  S(L1,1)=1.00
      IF (S(L1,1).GT.1.00)  J1=L1
      IF (S(L1,1).LE.1.00.AND.S(L1,1).GE.0.70)  J1=L1
      IF (S(L1,1).GT.1.00)  GO TO 75
80    J2=JB-1
85    L2=J2+1
      IF (L2.GT.NN)  GO TO 90
      IF (L2.GT.(N2)  )  GO TO 90
      IF (M(L2).EQ.2.OR.M(L2).EQ.5)  S(L2,1)=1.00
      IF (S(L2,1).GT.1.00)  J2=L2
      IF (S(L2,1).LE.1.00 .AND. S(L2,1).GE.0.70)  J2=L2
      IF (S(L2,1) .GT.1.00)  GO TO 85
C
C
C     ...CHECK FOR THE # OF HELIX FORMERS IN THE ENTIRE HELIX...
C
C     TO COMPARE THE ACTUAL NUMBER OF FORMERS (T4) TO ITS THEORITICAL
C     ONE (TT: EQUAL TO AT LEAST HALF OF THE SECTION)
C
90    T4=0
      DO 95  I=J1,J2
      S(I,4)=0
      IF (S(I,1).GE.1.00) S(I,4)=0.5
      IF (S(I,1).GE.1.06) S(I,4)=1.0
      T4=T4+S(I,4)
```

```
      95    CONTINUE
            TT=(J2-J1+1)/2.0
            PRINT 100,J1,J2,T4,TT
      100   FORMAT(' ',10X,'J1:',I4,5X,'J2:',I4,5X,'T4:',F7.4,5X,'TT:',F7.4,
     1 4X,'ACTUAL AND THEORIT. # FORMERS FROM J1 TO J2')
C
C     TO CONTINUE ADDING ONE RESIDUE AT A TIME TO BOTH ENDS IF T4 > TT
C
            IF (T4.GE.TT .AND. S(J1-1,1) .LE.0.69)  GO TO 110
            IF (T4.GE.TT .AND. S(J1-1,1) .GT.0.69.AND.L1.GT.(N1))  GO TO 70
      110   IF (L2.GT.NN)  GO TO 170
            IF (T4.GE.TT .AND. S(J2+1,1) .GT.0.69.AND.L2.LT.(N2))  GO TO 85
            IF((T4.GE.TT.AND.S(J2+1,1).LE.0.69) .OR.(T4.GE.TT.AND.S(J2+1,1).GT
     1.0.69.AND.L2.GE. N2))  GO TO 170
C
C
C     IF T4 < TT THEN TO WITHDRAW SOME BOUNDARY RESIDUES (ESPECIALLY
C     BA,IA) SO THAT T4 > TT
C
      120   IF (S(J2,1) .LT.1.00)  GO TO 125
            IF (S(J1,1).LT.1.00 .AND.M(J1+2).NE.15)  GO TO 130
            IF (S(J2,1) .LT.1.06)  GO TO 135
            IF (S(J1,1).LT.1.06 .AND.M(J1+2).NE.15)  GO TO 140
      125   J2=J2-1
            IF (S(J2+1,1).LT.1.00)  GO TO 150
      130   J1=J1+1
            IF (S(J1-1,1).LT.1.00)  GO TO 150
      135   J2=J2-1
            IF (S(J2+1,1).LT.1.06)  GO TO 150
      140   J1=J1+1
            IF (S(J1-1,1).LT.1.06)  GO TO 150
C
C
C     TO CHECK T4 AND TT EVERY TIME A BOUNDARY RESIDUE IS WITHDRAWN
C
      150   T4=0
            DO 155  I=J1,J2
            S(I,4)=0
            IF (S(I,1).GE.1.00)  S(I,4)=0.5
            IF (S(I,1).GE.1.06)  S(I,4)=1.0
            T4=T4+S(I,4)
      155   CONTINUE
            TT=(J2-J1+1)/2.0
            PRINT 160,J1,J2,T4,TT
      160   FORMAT(' ',10X,'J1:',I4,5X,'J2:',I4,5X,'T4:',F7.4,5X,'TT:',F7.4,
     1 4X,'ACTUAL AND THEORIT. # FORMERS FROM J1 TO J2')
            IF (T4.GE.TT)  GO TO 170
            IF (T4.LT.TT)  GO TO 120
      170   PRINT 175,J1,J2
      175   FORMAT('0',10X,'PSEUDO-HELIX FROM J1:',I5,3X,'TO  J2:',I5,/)
```

```
C
C
C     TO CALL SUBROUTINE MOJ1 TO CARRY OUT THE BOUNDARY ADJUSTMENT
C
      CALL MOJ1
          IF (J2.LT.NY)  J=J2+1
          IF (J2.LT.NY)   W=2
          IF (J2.LT.NY)   N1=J2
          IF (J2.LT.NY)   RETURN
  180     W=1
          RETURN
C
C     ... CHECK FOR THE NUMBER OF FORMERS IN THE 6 RESIDUE UNIT ...
C
C     PRO CAN ONLY  EXIST AT THE FIRST TURN OF N-TERMINAL SIDE. ANY OTH
C     ER POSITION ESPECIALLY AT THE C-TERMINAL WILL IMPEDE THE PROPAGA
C     TION
C
  190     DO 200  I=JB,JC
          IF (M(I).EQ.15.AND.I.EQ.(JB+2).AND.U.EQ.1)  GO TO 210
          IF (M(I).EQ.15.AND.I.NE.(JB+2).AND.U.EQ.1)  GO TO 220
          IF (M(I).EQ.15.AND.U.EQ.2)  GO TO 70
  200     CONTINUE
          IF (U.EQ.1)  GO TO 210
          IF (JB.EQ.(JA+1))  JB=JA-1
          IF (JB.NE.(JA-1)) JB=JB-2
C
C     IF PRO IS NOT FOUND IN THE TETRAPEPTIDE THEN TO CHECK THE NUMBER
C     OF FORMERS OF THE 6 RESIDUE UNIT (= TETRAPEPTIDE + 2 ADJACENT RESI
C     DUES)
C
  210     JC=JB+5
          T4=0
          DO 215 I=JB,JC
          S(I,4)=0
          IF (S(I,1).GE.1.00)  S(I,4)=0.5
          IF (S(I,1).GE.1.06)  S(I,4)=1.0
          T4=T4+S(I,4)
  215     CONTINUE
          PRINT 218,JB,JC,T4
  218     FORMAT(' ',10X,'JB:',I4,5X,'JC:',I4,5X,'T4:',F7.4,14X,' HELIX FORM
         1IN 6 OVERL. RESIDUES')
          IF (T4.GE.4.0)  GO TO 240
C
C     IF THE 6 RESIDUE UNIT DOES NOT HAVE AT LEAST TWO THIRDS FORMERS
C     THEN EITHER TO SWITCH FROM N-PROPAGATION TO C-PROPAGATION OR TO
C     START ADDING ONE RESIDUE AT A TIME TO BOTH ENDS
C
          IF (U.EQ.2)  GO TO 230
  220     U=2
```

```
      J1=JB+4
      GO TO 30
230   JB=JC-3
      GO TO 70
C
C     ... TO CHECK THE NUMBER OF BREAKERS IN THE ENTIRE POLYPEPTIDE ...
C
C     DESCRIPTION OF PARAMETERS
C        JB  - N-TERMINAL RESIDUE OF THE HELICAL POLYPEPTIDE
C        JD  - C-TERMINAL RESIDUE OF THE HELICAL POLYPEPTIDE
C        M3  - COUNTER
C        M4  - COUNTER
C
C
C     IF THE ACTUAL NUMBER OF BREAKERS (L) IS LESS THAN THE THEORITICAL
C     ONE (M5: ONE THIRD OF THE SECTION) THEN THE REGION CAN KEEP ON PRO
C     PAGATING. OTHERWISE EITHER TO SWITCH FROM N-PROPAGATION TO C-PROPA
C     GATION OR TO START ADDING ONE RESIDUE AT A TIME
C
240   M5=0
      IF (U.EQ.1)  GO TO 250
      JB=JB-(4*M4)
250   JD=JB+9+(4*M3)
      M3=M3+1
      M4=M4+1
      M5= (JD-JB+1)/3
      L=0
      DO 255 I=JB,JD
      IF (S(I,1).LE.0.69)  L=L+1
255   CONTINUE
      PRINT 258,JB,JD,M5 ,L
258   FORMAT(' ',10X,'JB:',I4,5X,'JD:',I4,5X,'M5:',I7,5X,'L:',I3,5X,
     1 'THEORIT. AND ACTUAL # BREAKERS FROM JB TO JD')
      IF (L.LT.M5.AND.U.EQ.1.AND.M(JB+2).EQ.15)  GO TO 260
      IF (L.LT.M5.AND.U.EQ.1)  GO TO 20
      IF (L.LT.M5.AND.U.EQ.2)  GO TO 30
      M6=M2
      IF (U.EQ.2.AND.M6.EQ.0)  JB=JB+6
      IF (U.EQ.2.AND.M6.NE.0)  JB=JB+6+(4*M6)
      IF (U.EQ.2)  GO TO 70
      U=2
      J1=JB+4
      GO TO 30
260   J1=JB
      U=2
      GO TO 30
      END
End of File
```

```
C
C
      SUBROUTINE  MOJ1
C     ................
C
C
C     ..........................................................
C     .                                                        .
C     .            BOUNDARY MOVE OF THE N-TERMINAL             .
C     .                                                        .
C     ..........................................................
C
C     ..................................................................
C     PURPOSE
C        TO FIND OUT THE MOST FAVORABLE N-BOUNDARY RESIDUE FOR THE PREDIC
C        TED HELIX BASED ON THE BOUNDARY CONFORMATIONAL PARAMETERS OF THE
C        ADJACENT RESIDUES
C     ..................................................................
C
C
      REAL S,T1,T2,A1,A2 ,T3,T4,T5,TT,P
      INTEGER G,F,H,U,D,V1,V2,W, V3,V4,V5,V6,V7,V8,Q
      LOGICAL HELLO,BYE  ,BALL,MOVE
      DIMENSION S(1000,20),M(1000),H(1000),D(1000,16),P(1000,10)
      COMMON S,T1,T2,T3,T4,T5,TT,A1,A2,P,F,H,U,D,W,M,M1,M2,M3,M4,M5,M6,
     1L,I,K,L1,L2,NZ,NY,JA,JB,JC,JD,J1,J2,KM,N1,N2,NN,J,G,K3,V1,V2,V3,V4
     2,V5,V6,V7,Q,HELLO,BYE,BALL,MOVE
C
C
C     DESCRIPTION OF PARAMETERS
C        V1  - ACTUAL NUMBER OF BREAKERS IN THE PREDICTED HELIX (=L)
C        V2  - COUNTER INDICATING THE POSITION OF THE ADJESTMENT BECAUSE
C              THE PROCEDURE CONTAINS SEVERAL DIFFERENT POSSIBILITIES OF AD
C              JUSTMENT (COUNTER USED FOR N-TERMINAL ADJUSTMENT)
C        J1  - N-TERMINAL RESIDUE OF THE PREDICTED HELIX
C        J2  - C-TERMINAL RESIDUE OF THE PREDICTED HELIX
C        K3  - C-TERMINAL RESIDUE OF THE PREVIOUS PREDICTED HELIX
C
C
C     ..... SITUATION WITH J1 CLOSE TO ZERO .....
C
C     TO TAKE INTO ACCOUNT THE POSITION OF J1 WHEN IT IS CLOSE TO THE N-
C     TERMINAL OF THE PROTEIN SINCE THERE IS LESS FREEDOM FOR MOVING IT
C     TOWARDS THIS SIDE
C
C
C
      V1=L
      V2=0
      V3=0
```

```
      C
            PRINT 5
          5 FORMAT('0',30X,'BOUNDARY ANALYSIS OF THE N-TERMINAL')
      C
      C     *** 1 ***
            IF ((J1-1).LE.0)  GO TO 10
            BALL=.FALSE.
            IF (J1.EQ.2.AND.S(J1,8).GT.1.47.AND.S(J1-1,1).LE.0.69.AND.S(J1+1,
           2 1).GT.1.01.AND.S(J1+2,8).GE.S(J1,8).AND.M(J1-1).NE.15)BALL=.TRUE.
            IF (BALL)  J1=J1
            IF (BALL)  V2=1
            IF (BALL)  GO TO 300
      C
      C     *** 2 ***
         10 BALL=.FALSE.
            IF (J1.EQ.1.AND.S(J1,8).GT.1.08.AND.S(J1+1,8).LT.S(J1,8).AND.S(J1+
           1 2,8).LE.S(J1,8).AND.S(J1+3,8).LT.1.08)  BALL=.TRUE.
            IF (BALL)  J1=J1
            IF (BALL)  V2=2
            IF (BALL)  GO TO 300
      C
      C     *** 3 ***
            BALL=.FALSE.
            IF (J1.EQ.1.AND.S(J1,8).GT.1.08.AND.S(J1,1).GT.1.01.AND.S(J1+1,1).
           1 LT.1.06.AND.S(J1+2,8).LT.S(J1,8).AND.S(J1+3,8).LT.1.08)  BALL=.
           2  .TRUE.
            IF (BALL)  J1=J1
            IF (BALL)  V2=3
            IF (BALL)  GO TO 300
      C
      C     *** 4 ***
            BALL=.FALSE.
            IF (J1.EQ.1.AND.S(J1,1).GT.1.16.AND.S( J1+1,8).GT.2.02.AND.S(J1+2,
           1 8).GT.2.02.AND.S(J1+3,1).GT.1.11.AND.S(J1+4,1).GT.1.16)  BALL=
           2 .TRUE.
            IF (BALL)  J1=J1
            IF (BALL)  V2=4
            IF (BALL)  GO TO 300
      C
      C     *** 5 ***
            BALL=.FALSE.
            T1=0
            T2=0
            T5=0
            T1=S(J1+1,1)+S(J1+2,1)+S(J1+3,1)+S(J1+4,1)
            T2=S(J1+1,2)+S(J1+2,2)+S(J1+3,2)+S(J1+4,2)
            T5=S(J1+1,5)+S(J1+2,5)+S(J1+3,5)+S(J1+4,5)
            PRINT   2,T1,T2,T5
          2 FORMAT(' ',30X,'T1,T2,T5',3(F7.3),'    STEP 5,MOJ1 CLOSE TO 0')
            IF (T5.GT.T1.AND.T5.GT.T2.AND.S(J1,8).LT.1.08.AND.S(J1+4,8).LT.1.
```

```
     1 08.AND.S(J1+5,8).GT.1.47.AND.S(J1+6,8).GT.1.08.AND.S(J1,8).LT.1.0
     2 8.AND.S(J1+2,8).LT.0.66.AND.S(J1+2,6).LT.1.01.AND.S(J1+3,8).LT.1.
     3 08)  BALL=.TRUE.
      IF (BALL)  J1=J1+5
      IF (BALL)  V2=5
      IF (BALL)  GO TO 300
C
C     *** 6 ***
      BALL=.FALSE.
      IF (J1.EQ.1.AND.S(J1,8).LT.1.08.AND.S(J1+1,8).GE.1.08.AND.S(J1+2,
     1 8).LT.S(J1+1,8).AND.S(J1+3,8).LT.S(J1+1,8).AND.S(J1+4,8).LT.S(J1+
     2 1,8))  BALL=.TRUE.
      IF (BALL)  J1=J1+1
      IF (BALL)  V2=6
      IF (BALL)  GO TO 300
C
C     *** 7 ***
      BALL=.FALSE.
      IF (J1.EQ.1.AND.S(J1,8).LT.1.08.AND.S(J1+1,8).LT.1.08.AND.S(J1+2,8
     1 ).GE.1.08.AND.S(J1+3,8).LT.1.08.AND.S(J1+4,8).LT.1.08)  BALL=
     2 .TRUE.
      IF (BALL)  J1=J1+2
      IF (BALL)  V2=7
      IF (BALL)  GO TO 300
C
C     *** 8 ***
      BALL=.FALSE.
      IF ( J1.EQ.1.AND.S(J1,8).LT.1.08.AND.S(J1+1,8).LT.1.08.AND.S(J1+2,
     1 8).LT. 1.08.AND.S(J1+3,8).GE.1.08)  BALL=.TRUE.
      IF (BALL)  J1=J1+3
      IF (BALL)  V2=8
      IF (BALL)  GO TO 300
C
C     *** 9 ***
      BALL=.FALSE.
      IF (K.EQ.3)  K3=N2
      IF (S(J1,8).LT.1.08.AND.(J1-2).LT.(K3-1).AND.S(J1+1,8).LT.1.08.AND
     1 .S(J1+2,1).LE.0.69.AND.M(J1+2).NE.15.AND.S(J1+3,8).GE.1.08) BALL
     2  =.TRUE.
      IF (BALL)  J1=J1+3
      IF (BALL)  V2=9
      IF (BALL)  GO TO 300
C
C     *** 10 ***
      IF ((J1-3).LE.0)  GO TO 6
      BALL=.FALSE.
      IF (S(J1,2).GE.1.47.AND.S(J1-1,2).GE.1.47.AND.S(J1-2,2).GT.0.93.AN
     1D.S(J1+2,2).GE.1.47.AND.S(J1+3,8).GT.1.47.AND.S(J1-3,1).LT.1.06.AN
     2D.(S(J1+1,2).GT.0.75  .OR.M(J1+2).EQ.1))  BALL=.TRUE.
      IF (BALL)  J1=J1+3
```

```
151               IF (BALL)  V2=10
152               IF (BALL)  GO TO 300
153     C
154     C      *** 11 ***
155           6   BALL=.FALSE.
156               IF ((J1-2 ).LE.0)  GO TO 15.
157               IF (J1.EQ.3.AND.S(J1,8).LT.1.08.AND.S(J1+1,8).LT.1.08.AND.S(J1+2,
158              1 8).LT.1.08.AND.S(J1+3,8).LE.1.08.AND.S(J1+4,8).GT. 1.08.AND.S(J1-
159              2 1,8).LT.1.08.AND.S(J1-2,8).LT.1.08)  BALL=.TRUE.
160               IF (BALL)  J1=J1+4
161               IF (BALL)  V2=11
162               IF (BALL)  GO TO 300
163     C
164     C
165     C      ..... TO REPEAT THE B-TURN CHECK .....
166     C
167     C      TO CHECK THE PRESENCE OF TURNS IN THE VICINITY OF THE HELIX BOUNDA
168     C      RIES WHICH MAY FORCE THE PREDICTED BOUNDARIES TO BE MOVED TO A NEW
169     C      POSITION. WE CHECK IT FROM POSITION J1-3 (I=0) TO J1+3 (I=6)
170     C
171           15  I=0
172               LE=J1-3
173               IF (LE.LE.0)  GO TO 200
174           20  LF=LE+3
175               IF ((LE+3).GT.NN)  GO TO 210
176     C
177     C      TO COMPARE PA (T1),PB (T2),AND PT (T5) AND TO CALCULATE THE PROBABI
178     C      LITY OF B-TURN OCCURRENCE (TT) OF THE TETRAPEPTIDE LE-LF
179     C
180               T1=0
181               T2=0
182               T5=0
183               TT=0
184               HELLO=.FALSE.
185               DO 25   L=LE,LF
186               T1=T1+S(L,1)
187               T2=T2+S(L,2)
188               T5=T5+S(L,5)
189           25  CONTINUE
190     C
191               TT=P(LE,1)*P(LE+1,2)*P(LE+2,3)*P(LE+3,4)
192               PRINT  30,LE,T1,T2,T5,TT,I
193           30  FORMAT(' ',10X,'LE,T1,T2,T5,TT,I',I5,3(F7.4,2X),F13.9,I4,3X,
194              1 'B-TURN SEARCH AT N-TERMINAL')
195     C
196               IF (T5.GT.T1.AND.T5.GT.T2.AND.TT.GT.0.000075000)  HELLO=.TRUE.
197     C
198     C      *** 1 ***
199               IF (HELLO.AND.LE.EQ.(J1-3).AND.S(J1+1,8).GE.1.08.AND.S( J1,8).LT.1
200              1 .08.AND.S(J1+2,8).GT.1.08.AND.S(J1+3,8).GT.1.08)  GO TO 101
```

```
C
C     *** 2 ***
          IF (HELLO.AND.LE.EQ.(J1-2).AND.S(J1-2,5).GT.1.52.AND.S(J1-1,5).GT.
     1 1.52.AND.S(J1,5).GT.1.43.AND.S(J1,1).LT.1.06.AND.S(J1+1,1).GT.1.0
     2 6.AND.S(J1+1,8).GT.1.08.AND.S(J1+2,1).LT.S(J1+1,1))  GO TO 101
C
C     *** 3 ***
          IF (HELLO.AND.LE.EQ.(J1-1).AND.S(J1+2,8).GT.1.47.AND.S(J1+4,8).LT.
     1 1.08.AND.S(J1+5,8).LT.1.08.AND.S(J1+6,8).LT.S(J1+2,8).AND.S(J1+1,
     2 8).LT.S(J1+2,8))  GO TO 102
C
C     *** 4 ***
          IF (HELLO.AND.S(J1+2,8).LT.1.08.AND.S(J1+3,8).LT.1.08.AND.LE.EQ.(J
     1 1-2).AND.S(J1+5,1).GT.1.13.AND.S(J1+6,8).LT.1.08.AND.S(J1+7,8).LT
     2 .1.08.AND.(S(J1+4,8).LT.1.08.OR.S(J1+4,1).LT.1.06))  GO TO 105
C
C     *** 5 ***
          IF (HELLO.AND.LE.EQ.J1.AND.S(J1+4,1).GT.1.11.AND.S(J1+5,1).GT.1.21
     1 .AND.S(J1+6,1).GT.1.21.AND.S(J1+3,1).LT.S(J1+4,1).AND.S(J1+3,6).
     2 GT.1.01.AND.S(J1+2,6).GT.1.22)  GO TO 104
C
C     *** 6 ***
          IF (HELLO.AND.S(J1+3,8).GT.1.47.AND.S(J1+4,8).LT.1.08.AND.LE.EQ.(
     1 J1-1).AND.S(J1+5,8).LE.S(J1+3,8).AND.S(J1-1,1).LT.1.16.AND.S(J1-2
     2 ,1).LE.0.69.AND.S(J1+2,1).LT.0.98.AND.S(J1+2,6).GT.1.41)  GO TO
     3  102
C
C     *** 7 ***
          IF (HELLO.AND.S(J1+5,8).GE.1.08.AND.S(J1+6,8).LT.1.08.AND.S(J1+7,8
     1 ).LT.1.08.AND.LE.EQ.(J1+2))  GO TO 105
C
C     *** 8 ***
          IF (HELLO.AND.LE.EQ.(J1-2).AND.S(J1+2,8).LT.1.08.AND.S(J1+3,8).GE.
     1 1.08.AND.S(J1+3,1).GT.1.01.AND.S(J1+4,8).LE.S(J1+3,8).AND.S(J1,1)
     2 .LT.0.83.AND.(S(J1+1,2).GT.1.47.OR.S(J1+1,1).LT.0.67).AND.(S(J1+2
     3 ,2).GT.1.47.OR.S(J1+2,1).LT.0.83))  GO TO 103
C
C     *** 9 ***
          IF (HELLO.AND.S(J1+4,8).GT.1.47.AND.S(J1+3,8).LE.S(J1+4,8).AND.LE.
     1 EQ.J1.AND.S(J1+5,8).LT.1.08.AND.S(J1+6,8).LT.S(J1+4,8)) GO TO 104
C
C     *** 10 ***
          IF (HELLO.AND.LE.EQ.J1.AND.S(J1+3,8).LT.1.08.AND.S(J1+4,8).LT.1.08
     1 .AND.S(J1+5,8).LT.1.08.AND.S(J1+6,8).LT.1.08.AND.S(J1+4,1).GT.0.
     2 69.AND.S(J1,8).LT.1.47.AND.S(J1+1,1).LT.1.16.AND.S(J1+2,1).LT.1.
     3 21)  GO TO 104
C
C     *** 11 ***
          IF ((J1-1).LE.0)  GO TO 40
          IF(HELLO.AND.S(J1+3,8).LT.1.08.AND.S(J1+4,8).GE.1.08.AND.LE.EQ.(J1
```

```
251               1 -1).AND.(S(J1+2,1).LE.0.69.OR.S(J1+1,1).LE.0.69).AND.S(J1-1,1).LE
252               2.0.69)  GO TO 104
253      C
254      C      *** 12 ***
255           40    IF (HELLO.AND.LE.EQ.(J1+1).AND.S(J1+4 ,8).LE.1.08.AND.S(J1+5,8).LT
256               1 .1.08.AND.S(J1+6,8).GT.1.47)  GO TO 106
257      C
258      C      *** 13 ***
259                IF (HELLO.AND.LE.EQ.(J1+2).AND.S(J1+6,8).GT.1.08.AND.S(J1+5,8).LT.
260               1 S(J1+6,8) )  GO TO 106
261      C
262      C      *** 14 ***
263                IF (HELLO.AND.LE.EQ.(J1+1).AND.S(J1+6,8).GE.S(J1+5,8).AND.S(J1+6,1
264               1 ).GT.S(J1+5,1).AND.S(J1+7,8).LT.S(J1+6,8).AND.S(J1+4,8).LT.1.08)
265               2  GO TO 106
266      C
267      C      *** 15 ***
268                IF ((J1-3).LE.0)  GO TO 50
269                IF (S(J1,8).LT.1.08.AND.S(J1,2).GE.1.47.AND.S(J1-1,2).GT.1.38.AND.
270               1S(J1-3,1).LE.0.69.AND.S(J1+1,8).GT.1.47.AND.S(J1+2,2).GT.1.19.AND.
271               2S(J1+3,8).LT.1.08.AND.HELLO.AND.LE.EQ.(J1+3))  GO TO 107
272      C
273      C      *** 16 ***
274           50    IF (HELLO.AND.LE.EQ.(J1-3).AND.S(J1+1,8).LT.1.08.AND.S(J1+2,8).LT.
275               1 1.08.AND.S(J1+3,8).LT.1.08.AND.S(J1+4,8).GT.1.08.AND.S(J1+4,1).GT
276               2 .1.08.AND.S(J1+5,1).LT.S(J1+4,1))  GO TO 104
277      C
278      C      *** 17 ***
279                IF (HELLO.AND.S(J1+3,8).LT.1.08.AND.S(J1+4,8).LT.1.08.AND.S(J1+5,8
280               1 ).GT.1.08.AND.S(J1+5,1).GT.1.08.AND.S(J1+6,8).LT.1.08.AND.S(J1+5,
281               2 1).GT.S( J1+4,1).AND.LE.EQ.(J1-1))  GO TO 105
282      C
283      C      *** 18 ***
284                IF ((J1+7).GT.NN)  GO TO 80
285                IF (HELLO.AND.S(J1+4,8).LT.1.08.AND.S(J1+5,8).LT.1.08.AND.S(J1+6,
286               1 8).LT.1.08.AND.S(J1+6,1).LE.0.69.AND.S(J1+7,8).GE.1.08.AND.LE.EQ.
287               2 (J1+1))  GO TO 107
288      C
289      C      *** 19 ***
290           80    IF (HELLO.AND.LE.EQ.J1.AND.S(J1+3,8).LT.1.08.AND.S(J1+4,8).LT.1.08
291               1 .AND.S(J1+5,1).GT.1.16.AND.S(J1+6,8).LT.1.08.AND.(J1+7).GE.J2)
292               2  GO TO 105
293      C
294      C      *** 20 ***
295                IF ((J1+8).GT.NN.OR.(J1-1).LE.0)  GO TO 90
296                IF (HELLO.AND.LE.EQ.J1.AND.S(J1+2,5).GT.0.74.AND.S(J1+3,5).GT.1.52
297               1 .AND.S(J1+1,5).GT.0.98.AND.S(J1+4,2).GT.1.47.AND.S(J1+5,2).GT.1.6
298               2 0.AND.S(J1+6,8).LT.0.58.AND.S(J1+7,1).GT.1.13.AND.S(J1+7,8).GT.S(
299               3J1+8,8).AND.S(J1+8,1).GT.1.01.AND.S(J1-1,8).LT.1.08.AND.S(J1-1,1).
300               4 LE.0.69)  GO TO 107
```

```
C
C     *** 21 ***
90      IF ((J1+13).GT.NN)  GO TO 100
        IF (HELLO.AND.S(J1+2,8).LT.1.08.AND.LE.EQ.(J1-2).AND.S(J1+3,2).GT.
       1 0.93.AND.S(J1+4,2).GT.1.38.AND.S(J1+5,2).GT.1.19.AND.M(J1+6).EQ.1
       2 .AND.S(J1+7,8).LT.1.08.AND.S(J1+8,2).GT.1.38.AND.S(J1+9,8).LT.0.6
       3 8.AND.M(J1+9).EQ.M(J1+10).AND.M(J1+11).EQ.M(J1+9).AND.S(J1+12,8).
       4 GT.0.81.AND.S(J1+13,8).GT.2.02)  GO TO 112
C
100     IF (I.EQ.0)  GO TO 200
        IF (I.EQ.1)  GO TO 200
        IF (I.EQ.2)  GO TO 200
        IF (I.EQ.3)  GO TO 200
        IF (I.EQ.4)  GO TO 200
        IF (I.EQ.5)  GO TO 200
        IF (I.EQ.6)  GO TO 210
C
C
C     MOVE OF N-BOUNDARY AS A CONSEQUENCE OF STRONG B-TURN POTENTIAL IN
C     THE VICINITY OF THE PREDICTED HELIX
C
101     J1=J1+1
        GO TO 110
102     J1=J1+2
        GO TO 110
103     J1=J1+3
        GO TO 110
104     J1=J1+4
        GO TO 110
105     J1=J1+5
        GO TO 110
106     J1=J1+6
        GO TO 110
107     J1=J1+7
        GO TO 110
112     J1=J1+12
        GO TO 110
C
110     V2=80
        GO TO 300
C
200     I=I+1
        LE=LE+1
        GO TO 20
C
C
C     ..... B-TURN PROBLEMS OR OTHER PROBLEMS .....
C
C     ADJUSTMENT OF N-BOUNDARY MAY ALSO BE CAUSED BY EITHER RANDOM COIL
C     OR B-SHEET POTENTIAL OR BY THE LOW BOUNDARY CONFORMATIONAL PARAME
```

```
C      TER OF THE CURRENT BOUNDARY RESIDUE
C
C
C      *** 12 ***
  210   BALL=.FALSE.
        IF ((J1+7).GT.NN)  GO TO 230
        LC=0
        JN=J1+6
        DO 215 L=J1,JN
        IF (S(L,2).LE.0.75)  LC=LC+1
  215   CONTINUE
        JN=J1+4
        JM=J1+1
        T1=0
        T2=0
        T5=0
        DO 218 L=JM,JN
        T1=T1+S(L,1)
        T2=T2+S(L,2)
        T5=T5+S(L,5)
  218   CONTINUE
        PRINT 220,T1,T2,T5,LC
  220   FORMAT(' ',30X,'T1,T2,T5',3(F7.3),2X,'LC:',I3,'   STEP 12 ,MOJ1,
       1 B-TURN PROBLEM')
        IF (LC.GT.2.AND.T5.GT.T1.AND.(T2-T5).LT.0.500.AND.S(J1+5,8).LT.1.0
       1 8.AND.S(J1+6,8).GT.1.47.AND.S(J1+7,8).LT.2.01.AND.S(J1,1).LT.1.21
       2 .AND.S(J1+5,1).LT.1.21.AND.S(J1-1,6).LT.1.01)  BALL=.TRUE.
        IF (BALL)  J1=J1+6
        IF (BALL)  V2=12
        IF (BALL)  GO TO 300
C
C      *** 13 ***
        BALL=.FALSE.
        IF (J1.LE.K3.AND.(P(J1+1,1)*P(J1+2,2)*P(J1+3,3)*P(J1+4,4)).GT.0.00
       1 0100.AND.P(J1+1,1).GT.0.120.AND.P(J1+2,2).GT.0.139.AND.S(J1,5).
       2 GT.0.96.AND.S(J1+4,2).GT.1.19.AND.S(J1+4,8).LT.1.08.AND.S(J1+5,2)
       3 .GT.1.47.AND.S(J1+5,8).LT.S(J1+6,8).AND.S(J1+7,8).LT.S(J1+6,8).AN
       4 D.S(J1+6,1).GT.0.69)  BALL=.TRUE.
        IF (BALL)  J1=J1+6
        IF (BALL)  V2=13
        IF (BALL)  GO TO 300
C
C      *** 14 ***
  230   BALL=.FALSE.
        IF ((J1-2).LE.0)  GO TO 250
        T1=0
        T2=0
        T5=0
        T1=S(J1-2,1)+S(J1-1,1)+S(J1,1)+S(J1+1,1)
        T2=S(J1-2,2)+S(J1-1,2)+S(J1,2)+S(J1+1,2)
```

```
401               T5=S(J1-2,5)+S(J1-1,5)+S(J1,5)+S(J1+1,5)
402               PRINT 235,T1,T2,T5
403          235  FORMAT(' ',30X,'T1,T2,T5',3(F7.3),'    STEP 14,MOJ1 B-T PROBL.')
404               IF (T5.GT.T1.AND.T5.GT.T2.AND.S(J1+1,8).GT.1.47.AND.S(J1+2,8).LT.
405              11.08.AND.S(J1+3,8).LT.S(J1+1,8))  BALL=.TRUE.
406               IF (BALL)  J1=J1+1
407               IF (BALL)  V2=14
408               IF (BALL)  GO TO 300
409       C
410       C      *** 15 ***
411               BALL=.FALSE.
412               IF (S(J1,2).GT.1.37.AND.S(J1+1,2).GT.1.47.AND.S(J1+2,2).GT.1.47.AN
413              1 D.S(J1+3,1).LE.0.69.AND.S(J1+3,8).LT.1.08.AND.S(J1+4,8).GT.1.08.A
414              2 ND.S(J1+4,1).GT.1.16.AND.S(J1-1,2).GT.1.10.AND.S(J1-2,8).LT.1.08.
415              3AND.S(J1+5,1).LT.S(J1+4,1))  BALL=.TRUE.
416               IF (BALL)  J1=J1+4
417               IF (BALL)  V2=15
418               IF (BALL)  GO TO 300
419       C
420       C      *** 16 ***
421               BALL=.FALSE.
422               IF ((J1+7).GT.NN)  GO TO 240
423               IF ( (P(J1-2,1)*P(J1-1,2)*P(J1,3)*P(J1+1,4)).GT.0.000075.AND.S(J1+
424              1 8,8).GE.1.08.AND.(P(J1+4,1)*P(J1+5,2)*P(J1+6,3)*P(J1+7,4)).GT.
425              2 0.000075)  BALL=.TRUE.
426               IF (BALL)  J1=J1+8
427               IF (BALL)  V2=16
428               IF (BALL)  GO TO 300
429       C
430       C      *** 17 ***
431          240  BALL=.FALSE.
432               IF ((J1-3).LE.0)  GO TO 250
433               IF(S(J1,1).LE.0.69.AND.M(J1).NE.15.AND.S(J1-1,1).LE.0.69.AND.M(J1-
434              1 1).NE.15.AND.(S(J1-2,1).LE.0.69.OR.S(J1-2,8).LT.1.08).AND.S(J1-3,
435              22).GT.0.93.AND.S(J1+1,8).GE.1.08.AND.S(J1+2,8).LT.1.08.AND.S(J1+3,
436              3 1).GT.1.01)  BALL=.TRUE.
437               IF (BALL)  J1=J1+1
438               IF (BALL)  V2=17
439               IF (BALL)  GO TO 300
440       C
441       C      *** 18 ***
442               BALL=.FALSE.
443               IF (S(J1,8).LT.1.08.AND.(P(J1-3,1)*P(J1-2,2)*P(J1-1,3)*P(J1,4)).
444              1GT.0.000075.AND.S(J1+1,1).GT.1.11.AND.S(J1+2,1).GE.1.13.AND.S(J1+3
445              2 ,1).GT.0.69.AND.S(J1+4,1).GT.1.21.AND.(S(J1+5,1).GT.1.21.OR.S(J1+
446              3 5,2).LE.0.75))  BALL=.TRUE.
447               IF (BALL)  J1=J1+1
448               IF (BALL)  V2=18
449               IF (BALL)  GO TO 300
450       C
```

```
      C     *** 19 ***
            BALL=.FALSE.
            IF (S(J1,8).LT.1.08.AND.S(J1+1,1).LE.0.69.AND.M(J1+1).NE.15.AND.S(
     1 J1+2,8).GE.1.08.AND.S(J1+3,8).LT.1.08.AND.S(J1+4,8).LE.S(J1+2,8)
     2 .AND.(P(J1-3,1)*P(J1-2,2)*P(J1-1,3)*P(J1,4)).GE.0.000075)  BALL
     3  =.TRUE.
            IF (BALL)  J1=J1+2
            IF (BALL)  V2=19
            IF (BALL)  GO TO 300
      C
      C     *** 20 ***
            BALL=.FALSE.
            IF ((J1-4).LE.0)  GO TO 250
            IF (S(J1,8).LT.1.08.AND.S(J1,1).LE.0.69.AND.S(J1+1,1).GT.1.08.AND.
     1 S(J1+2,8).LT.1.08.AND.S(J1+2,1).GT.0.69.AND.S(J1+3,8).LT.2.01.AND
     2 .(P(J1-4,1)*P(J1-3,2)*P(J1-2,3)*P(J1-1,4)).GT.0.000075)  BALL=
     3  .TRUE.
            IF (BALL)  J1=J1+1
            IF (BALL)  V2=20
            IF (BALL)  GO TO 300
      C
      C     *** 21 ***
            BALL=.FALSE.
            IF (S(J1,8).LT.1.08.AND.S(J1+1,8).LT.1.08.AND.(S(J1+2,8).GT.1.47.
     1 OR.S(J1+2,8).GT.S(J1+3,8)).AND.S(J1+3,8).LT.1.47.AND.S(J1+4,8).LT
     2 .1.47.AND.(P(J1-4,1)*P(J1-3,2)*P(J1-2,3)*P(J1-1,4)).GE.0.000075.A
     3 ND.S(J1+1,8).LT.0.73.AND.S(J1,2).GT.1.19.AND.S(J1+1,2).GT.1.47)
     4  BALL=.TRUE.
            IF (BALL)  J1=J1+2
            IF (BALL)  V2=21
            IF (BALL)  GO TO 300
      C
      C     *** 22 ***
            BALL=.FALSE.
            T1=0
            T2=0
            T5=0
            T1=S(J1-4,1)+S(J1-3,1)+S(J1-2,1)+S(J1-1,1)
            T2=S(J1-4,2)+S(J1-3,2)+S(J1-2,2)+S(J1-1,2)
            T5=S(J1-4,5)+S(J1-3,5)+S(J1-2,5)+S(J1-1,5)
            PRINT 245,T1,T2,T5
        245 FORMAT(' ',30X,'T1,T2,T5',3(F7.3),'    STEP 22,MOJ1 B-T PROBL.')
            IF (T5.GT.T1.AND.T5.GT.T2.AND.S(J1,8).LT.1.08.AND.S(J1,1).LT.1.01.
     1 AND.S(J1+1,8).LT.1.08.AND.S(J1+2,8).GT.1.08.AND.S(J1+3,8).LE.S(J1
     2 +2,8))  BALL=.TRUE.
            IF (BALL)  J1=J1+2
            IF (BALL)  V2=22
            IF (BALL)  GO TO 300
      C
      C     *** 23 ***
```

```
501               BALL=.FALSE.
502               IF (S(J1,1).LT.1.06.AND.S(J1-1,2).GT.1.38.AND.S(J1-2,2).GT.1.60.AN
503              1 D.S(J1-4,2).GT.1.38.AND.S(J1+1,2).GT.1.19.AND.S(J1+1,8).LT.0.66.A
504              2 ND.S(J1+2,8).GT.0.81.AND.S(J1+2,1).GT.0.77.AND.S(J1+3,8).LT.S(J1+
505              3 2,8).AND.S(J1+3,1).GT.1.13)  BALL=.TRUE.
506               IF (BALL)  J1=J1+2
507               IF (BALL)  V2=23
508               IF (BALL)  GO TO 300
509       C
510       C    *** 24 ***
511               BALL=.FALSE.
512               IF ((J1-5).LE.0)  GO TO 250
513               IF (S(J1,8).LT.1.08.AND.S(J1+2,8).GT.1.08.AND.S(J1+1,8).LT.S(J1+2,
514              1 8).AND.S(J1-1,8).LT.1.08.AND.(S(J1+3,8).LE.S(J1+2,8).OR.S(J1+3,8)
515              2 .LT.2.01) .AND.S(J1+2,1).GT.1.01.AND.(P(J1-5,1)*P(J1-4,2)*P(J1-3,
516              3 3)*P(J1-2,4)).GE.0.000075)  BALL=.TRUE.
517               IF (BALL)  J1=J1+2
518               IF (BALL)  V2=24
519               IF (BALL)  GO TO 300
520       C
521       C      *** 25 ***
522               BALL=.FALSE.
523               IF ((J1-6).LE.0)  GO TO 250
524               IF (S(J1,8).LT.1.08.AND.S(J1+4,8).GT.1.47.AND.S(J1-1,1).LE.0.69
525              1 .AND.M(J1-1).NE.15.AND.S(J1-2,8).LE.S(J1+4,8).AND.S(J1+1,8).LT.
526              2 1.08.AND.S(J1+2,8).LT.1.08.AND.S(J1+2,1).LE.0.69.AND.S(J1+3,8).LT
527              3 .1.08.AND.(P(J1-6,1)*P(J1-5,2)*P(J1-4,3)*P(J1-3,4)).GT.0.000100)
528              4  BALL=.TRUE.
529               IF (BALL)  J1=J1+4
530               IF (BALL)  V2=25
531               IF (BALL)  GO TO 300
532       C
533       C      *** 26 ***
534           250   IF ((J1-1).LE.0)  GO TO 260
535               BALL=.FALSE.
536               T1=0
537               T2=0
538               T5=0
539               T1=S(J1-1,1)+S(J1,1)+S(J1+1,1)+S(J1+2,1)
540               T2=S(J1-1,2)+S(J1,2)+S(J1+1,2)+S(J1+2,2)
541               T5=S(J1-1,5)+S(J1,5)+S(J1+1,5)+S(J1+2,5)
542               PRINT 255,T1,T2,T5
543           255   FORMAT(' ',30X,'T1,T2,T5',3(F7.3),'    STEP 26,MOJ1 B-T PROBL.')
544               IF (T5.GT.T1.AND.T5.GT.T2.AND.S(J1+3,8).GE.1.08.AND.S(J1+4,8).LT.
545              1 1.47.AND.S(J1+5,8).LT.1.47.AND.((J1+6).GE.J2.OR.S(J1+6,8).LT.S(J1
546              2+3,8)).AND.S(J1,5).GT.1.19.AND.S(J1+1,5).GT.0.74.AND.S(J1+1,1).LT.
547              3 1.21.AND.S(J1+3,1).GT.S(J1,1).AND.S(J1+2,2).GT.1.19.AND.S(J1+3,2)
548              4 .GT.1.05.AND.S(J1+4,2).GT.1.05.AND.S(J1+5,2).GT.0.75.AND.S(J1+1,
549              5 2).GT.0.89)  BALL=.TRUE.
550               IF (BALL)  J1=J1+3
```

```
      IF (BALL)  V2=26
      IF (BALL)  GO TO 300
C
C     *** 27 ***
260   BALL=.FALSE.
      IF (S(J1+4,8).GT.1.47.AND.S(J1+5,8).LE.S(J1+4,8).AND.S(J1+6,8).LE.
     1 S(J1+4,8).AND.(P(J1+1,1)*P(J1+2,2)*P(J1+3,3)*P(J1+4,4)).GT.0.0001
     2 00.AND.S(J1+3,1).LT.S(J1+4,1).AND.S(J1+2,8).LT.2.01.AND.S(J1+1,8)
     3 .LT.2.01)  BALL=.TRUE.
      IF (BALL)  J1=J1+4
      IF (BALL)  V2=27
      IF (BALL)  GO TO 300
C
C
C     IF NONE OF THE ABOVE CONDITIONS IS SATISFIED TO CALL THE NEXT SUB
C     ROUTINE RMJ1 TO KEEP ON CHECKING FOR POSSIBILITIES OF N-TERMINAL
C     ADJUSTMENT
C
          CALL RMJ1
          RETURN
C
C     N-TERMINAL OF THE PREDICTED HELIX HAS BEEN ADJUSTED TO CALL SUBROU
C     TINE MOJ2 FOR C-TERMINAL ANALYSIS
C
300   CALL MOJ2
      RETURN
      END
End of File
```

```
      C
      C
            SUBROUTINE  RMJ1
      C     .................
      C
      C
      C     ..........................................................
      C     .                                                        .
      C     .          RMJ1 = REMAINING OF MOVE OF J1                .
      C     .                                                        .
      C     ..........................................................
      C
      C
      C     ..................................................................
      C
      C     PURPOSE
      C        TO KEEP ON CHECKING FOR THE BEST POSITION FOR N-BOUNDARY,THIS
      C        SUBROUTINE IS A CONTINUATION OF MOJ1
      C
      C     ..................................................................
      C
      C
            REAL S,T1,T2,A1,A2 ,T3,T4,T5,TT,P
            INTEGER G,F,H,U,D,V1,V2,W, V3,V4,V5,V6,V7,V8,Q
            LOGICAL HELLO,BYE  ,BALL,MOVE
            DIMENSION S(1000,20),M(1000),H(1000),D(1000,16),P(1000,10)
            COMMON S,T1,T2,T3,T4,T5,TT,A1,A2,P,F,H,U,D,W,M,M1,M2,M3,M4,M5,M6,
           1L,I,K,L1,L2,NZ,NY,JA,JB,JC,JD,J1,J2,KM,N1,N2,NN,J,G,K3,V1,V2,V3,V4
           2,V5,V6,V7,Q,HELLO,BYE,BALL,MOVE
      C
      C
      C     REMARKS
      C        THE PARAMETERS DESCRIBED IN THE SUBROUTINE MOJ1 STILL KEEP THE SA
      C        ME DEFINITION IN THIS SUBROUTINE
      C
      C        THE DIFFERENT COMMENTS J1=J1+1,J1=J1+5,...,J1=J1-1 INDICATE THE EV
      C        ENTUAL POSITION OF J1 IF ITS ENVIRONMENT MEETS ONE OF THE CONDI
      C        TIONS DESCRIBED BELOW. IF NOT J1 WILL STILL REMAIN AT THE SAME PO
      C        SITION BECAUSE IT APPEARS TO BE THE MOST FAVORABLE ONE
      C
      C
      C ... J1 = J1+7 .....
      C .....................
      C
      C     *** 28 ***
            BALL=.FALSE.
            IF ((J1+8).GT.NN)  GO TO 20
            IF (J1.LE.K3.AND.(P(J1+1,1)*P(J1+2,2)*P(J1+3,3)*P(J1+4,4)).GT.0.00
           1 007500.AND.S(J1+3,5).GT.1.43.AND.S(J1+4,5).GT.1.19.AND.S(J1+5,8)
           2 .LT.0.66.AND.S(J1+6,8).LT.1.08.AND.S(J1+7,8).GT.0.71.AND.S(J1+7,1
```

```
     3 ).GT.0.69.AND.S(J1+8,8).LT.S(J1+7,8).AND.S(J1+8,1).GT.1.01.AND.
     4 S(J1+6,1).LE.0.69)  BALL=.TRUE.
      IF (BALL)  J1=J1+7
      IF (BALL)  V2=28
      IF (BALL)  GO TO 300
C
C ... J1 = J1+5 .....
C ....................
C
C     *** 29 ***
20    BALL=.FALSE.
      IF ((J1-2).LE.0)  GO TO 300
      T1=0
      T2=0
      T5=0
      T1=S(J1-2,1)+S(J1-1,1)+S(J1,1)+S(J1+1,1)
      T2=S(J1-2,2)+S(J1-1,2)+S(J1,2)+S(J1+1,2)
      T5=S(J1-2,5)+S(J1-1,5)+S(J1,5)+S(J1+1,5)
      PRINT  25,T1,T2,T5
25    FORMAT(' ',30X,'T1,T2,T5',3(F7.3),'    STEP 29, J1+5 ,RMJ1')

      IF (T5.GT.T1.AND.T5.GT.T2.AND.S(J1+1,8).LT.1.08.AND.S(J1+2,8).LT.
     11.08.AND.S(J1+3,8).LT.1.08.AND.S(J1+4,8).LT.1.08.AND.S(J1+5,8).GE.
     2 1.08.AND.S(J1+4,1).LT.1.06.AND.S(J1+3,1).LE.0.69.AND.S(J1+1,1).LT
     3 .1.06)  BALL=.TRUE.
      IF (BALL)  J1=J1+5
      IF (BALL)  V2=29
      IF (BALL)  GO TO 300
C
C
C ... J1 = J1+4 .....
C ....................
C
C     *** 30 ***
      BALL=.FALSE.
      IF ((J1-3).LE.0)  GO TO 300
      IF (S(J1,8).LE.1.08.AND.S(J1+1,8).LT.1.08.AND.S(J1+2,8).LT.1.08.AN
     1 D.S(J1+4,8).GT.1.08.AND.S(J1-1,8).LT.1.08.AND.S(J1-2,8).LT.1.08
     2 .AND.(S(J1-3,8).LT.1.08.OR.S(J1-3,1).LE.0.69).AND.S(J1+3,8).LE.
     3 1.08.AND.(S(J1+4,8).GT:1.29.OR.(S(J1+4,8)-S(J1,8)).GT.0.65)) BALL
     4  =.TRUE.
      IF (BALL)  J1=J1+4
      IF (BALL)  V2=30
      IF (BALL)  GO TO 300
C
C     *** 31 ***
      BALL=.FALSE.
      IF (S(J1,8).LE.1.08.AND.S(J1+4,8).GT.1.47.AND.S(J1+1,8).LT.1.08.AN
     1 D.S(J1+2,8).LT.2.01.AND.S(J1+3,8).LT.2.01.AND.S(J1-1,8).LT.1.08
     2 .AND.(S(J1-2,8).LT.1.08.OR.S(J1-2,8).LT.S(J1+4,8)).AND. S(J1-3,8)
```

```
              3 .LT.S(J1+4,8))  BALL=.TRUE.
               IF (BALL)  J1=J1+4
               IF (BALL)  V2=31
               IF (BALL)  GO TO 300
C
C
C ... J1 = J1-5 .....
C ......................
C
C      *** 32 ***
               IF ((J1-6).LE.0)  GO TO 30
               BALL=.FALSE.
               IF (S(J1,8).LT.1.08.AND.M(J1+1).EQ.4.AND.M(J1-2).EQ.7.AND.S(J1-1,8
              1 ).LT.1.08.AND.S(J1-1,1).GT.1.01.AND.M(J1-3).EQ.4.AND.S(J1-4,1).GT
              2.1.01.AND.S(J1-4,8).LT.S(J1-5,8).AND.S(J1-5,8).GE.1.08.AND.S(J1-6,
              3 6).GE.1.22)  BALL=.TRUE.
               IF (BALL)  J1=J1-5
               IF (BALL)  V2=32
               IF (BALL)  GO TO 300
C
C      *** 33 ***
               BALL=.FALSE.
               IF (S(J1,8).LT.1.08.AND.S(J1,1).GT.1.01.AND.M(J1-5).EQ.1.AND.S(J1+
              1 1,8).GE.1.08.AND.S(J1+2,8).GE.1.08.AND.S(J1+3,8).GE.1.08.AND.M(J1
              2 -1).EQ.18.AND.S(J1-2,8).GE.0.81.AND.S(J1-3,1).GT.1.01.AND.S(J1-4,
              3 1).GT.1.01.AND.S(J1-3,8).LT.1.08.AND.S(J1-4,8).LT.1.08.AND.S(J1-6
              4 ,6).GT.1.04)  BALL=.TRUE.
               IF (BALL)  J1=J1-5
               IF (BALL)  V2=33
               IF (BALL)  GO TO 300
C
C ... J1 = J1-4 .....
C ......................
C
C      *** 34 ***
   30      IF ((J1-5).LE.0)  GO TO 40
               BALL=.FALSE.
               IF (S(J1,8).LT.1.08.AND.S(J1,1).GT.1.01.AND.S(J1-4,8).GT.1.47.AND.
              1 S(J1-3,8).GE.2.01.AND.S(J1-1,8).LT.1.08.AND.S(J1-2,8).LT.2.01.AND
              2 .S(J1-2,1).GT.1.01.AND.S(J1-5,8).LT.1.08.AND.S(J1+1,8).LT.1.08.
              3 AND.S(J1+2,8).LT.1.08.AND.S(J1+2,1).GT.1.01.AND.S(J1+3,8).LT.1.08
              4 .AND.S(J1+3,1).GT.1.01.AND.S(J1+1,1).GT.1.01)  BALL=.TRUE.
               IF (BALL)  J1=J1-4
               IF (BALL)  V2=34
               IF (BALL)  GO TO 300
C
C      *** 35 ***
               BALL=.FALSE.
               IF (S(J1,8).LT.1.08.AND.S(J1-4,8).GT.1.29.AND.S(J1-5,6).GT.1.10.
              1 AND.S(J1-3,8).LT.1.08.AND.S(J1-2,8).LT.1.08.AND.S(J1-1,8).LT.1.08
```

```
     2  .AND.S(J1-2,1).GT.1.01.AND.S(J1-3,1).GT.1.01.AND.S(J1+1,8).LT.
     3 1.08.AND.S(J1+1,1).GT.1.01.AND.S(J1+2,8).GE.2.02.AND.S(J1+3,8).GT
     4 .1.08.AND.S(J1+4,8).GT.1.08)  BALL=.TRUE.
      IF (BALL)  J1=J1-4
      IF (BALL)  V2=35
      IF (BALL)  GO TO 300
C
C     *** 36 ***
      BALL=.FALSE.
      IF (S(J1,8).LT.1.08.AND.S(J1,1).GT.1.01.AND.S(J1+1,8).LT.1.08.AND.
     1 S(J1+2,8).LE.S(J1-4,8).AND.S(J1-4,8).GT.1.47.AND.S(J1-5,8).GT.1.4
     2 7.AND.S(J1-1,1).GT.1.01.AND.S(J1-2,1).GT.0.69.AND.S(J1-3,1).GT.1.
     3 01.AND.S(J1-1,8).LT.2.01.AND.S(J1-2,8).LT.2.01.AND.S(J1-3,8).LT.
     4 2.01)  BALL=.TRUE.
      IF (BALL)  V2=36
      IF (BALL)  J1=J1-4
      IF (BALL)  GO TO 300

C ... J1 = J1+3 .....
C ...................
C
C     *** 37 ***
   40 BALL=.FALSE.
      IF ((J1-3).LE.0)  GO TO 300
      IF (S(J1,1).LE.0.69.AND.S(J1,8).LT.1.08.AND.S(J1-1,8).LT.1.08.AND.
     1 S(J1-2,8).LT.1.08.AND.S(J1-3,8).LT.1.08.AND.S(J1+1,8).LT.1.08.AND
     2 .S(J1+2,8).LT.1.08.AND.S(J1+3,8).GE.1.08.AND.S(J1+3,1).GT.1.01)
     3  BALL=.TRUE.
      IF (BALL)  J1=J1+3
      IF (BALL)  V2=37
      IF (BALL)  GO TO 300
C
C     *** 38 ***
      BALL=.FALSE.
      IF (S(J1,8).LT.1.08.AND.S(J1+1,8).LT.1.08.AND.S(J1-1,8).LT.1.08.AN
     1D.S(J1-2,8).LT.1.08.AND.S(J1-3,8).LT.1.08.AND.S(J1+3,8).GT.S(J1+2,
     2 8).AND.S(J1+2,1).LT.S(J1+3,1))  BALL=.TRUE.
      IF (BALL)  J1=J1+3
      IF (BALL)  V2=38
      IF (BALL)  GO TO 300
C
C     *** 39 ***
      IF ((J1-4).LE.0)  GO TO 50
      BALL=.FALSE.
      IF (S(J1,8).LT.1.47.AND.S(J1+1,8).LT.1.08.AND.S(J1+1,1).LE.0.69.AN
     1 D.S(J1+2,6).GT.1.10.AND.S(J1+3,8).GT.1.47.AND.S(J1-1,8).LT.1.08.A
     2 ND.S(J1-3,8).LT.1.08.AND.S(J1-2,8).LT.1.08.AND.S(J1-4,8).LT.1.08)
     3  BALL=.TRUE.
      IF (BALL)  J1=J1+3
      IF (BALL)  V2=39
```

```
201                 IF (BALL)  GO TO 300
202       C
203       C      *** 40 ***
204                 BALL=.FALSE.
205                 IF (S(J1,8).LT.1.08.AND.S(J1,1).LE.0 .69.AND.S(J1-1,1).LE.0.69.AND
206                1 .M(J1-1).NE.15.AND.S(J1-2,8).LT.1.08.AND.S(J1-3,8).LT.S(J1+3,8).A
207                2 ND.S(J1-4,8).LT.1.08.AND.S(J1+1,8).LT.1.08.AND.S(J1+2,8).LT.1.08
208                3 .AND.S(J1+3,8).GT.1.08)  BALL=.TRUE.
209                 IF (BALL)  J1=J1+3
210                 IF (BALL)  V2=40
211                 IF (BALL)  GO TO 300
212       C
213       C      *** 41 ***
214                 BALL=.FALSE.
215                 IF (S(J1,8).GE.1.08.AND.S(J1+3,8).GT.S(J1,8).AND.S(J1+1,1).LE.0.69
216                1 .AND.M(J1+1).NE.15.AND.S(J1+2,8).LT.1.08.AND.S(J1-1,8).LT.1.08.AN
217                2 D.S(J1-2,8).LT.1.08.AND.S(J1-3,8).LT.1.08.AND.(J1-4).LT.(K3-2))
218                3  BALL=.TRUE.
219                 IF (BALL)  V2=41
220                 IF (BALL)  J1=J1+3
221                 IF (BALL)  GO TO 300
222       C
223       C      *** 42 ***
224                 BALL=.FALSE.
225                 IF (S(J1+3,8).GT.1.47.AND.S(J1,8).LT.1.47.AND.S(J1+1,8).LT.1.08.AN
226                1 D.S(J1+2,8).LT.S(J1+3,8).AND.S(J1+2,1).LT.1.16.AND.S(J1-1,8).LT.
227                2 1.08.AND.S(J1-2,8).LT.1.08.AND.S(J1-3,8).LT.1.08.AND.S(J1-4,8).LT
228                3  .1.08)  BALL=.TRUE.
229                 IF (BALL)  J1=J1+3
230                 IF (BALL)  V2=42
231                 IF (BALL)  GO TO 300
232       C
233       C ... J1 = J1-3 .....
234       C ....................
235       C
236       C      *** 43 ***
237                 BALL=.FALSE.
238                 IF (S(J1,8).GT.1.47.AND.S(J1-3,8).GE.S(J1,8).AND.S(J1-1,8).LT.1.08
239                1 .AND.S(J1-2,1).GT.1.01.AND.S(J1-2,8).LT.S(J1-3,8).AND.S(J1-4,8).
240                2 LT.1.08.AND.S(J1+2,8).LT.1.08.AND.S(J1+1,8).LT.1.08.AND.S(J1+3,8)
241                3 .GT.1.47)  BALL=.TRUE.
242                 IF (BALL)  J1=J1-3
243                 IF (BALL)  V2=43
244                 IF (BALL)  GO TO 300
245       C
246       C      *** 44 ***
247                 BALL=.FALSE.
248                 IF (S(J1,8).LT.1.08.AND.S(J1+1,8).LT.1.08.AND.S(J1+2,8).GT.1.29.
249                1 AND.S(J1-1,8).GE.0.81.AND.S(J1-2,8).GE.1.08.AND.(S(J1-2,1).GT.1.0
250                2 1.OR.S(J1-2,8).GT.2.01) .AND.S(J1-3,8).GT.1.08.AND.S(J1-4,8).LE.
```

```
251               3 S(J1-3,8))  BALL=.TRUE.
252                IF (BALL)  J1=J1-3
253                IF (BALL)  V2=44
254                IF (BALL)  GO TO 300
255      C
256      C      *** 45 ***
257                BALL=.FALSE.
258                IF (S(J1,8).LT.1.08.AND.S(J1-3,8).GE.1.47.AND.S(J1-1,8).LT.1.08.AN
259               1 D.S(J1-2,8).LT.1.08.AND.S(J1-4,8).LT.1.08.AND.S(J1+1,8).LT.1.08
260               2  .AND.S(J1+2,8).LE.1.08.AND.S(J1+3,8).LT.1.08.AND.S(J1,1).GT.1.01
261               3  .AND.S(J1+1,1).GT.1.01.AND.S(J1+3,1).GT.1.01)  BALL=.TRUE.
262                IF (BALL)  J1=J1-3
263                IF (BALL)  V2=45
264                IF (BALL)  GO TO 300
265      C
266      C      *** 46 ***
267                BALL=.FALSE.
268                IF (M(J1).EQ.1.AND.S(J1-3,8).GE.1.08.AND.M(J1-4).EQ.8.AND.S(J1-1,8
269               1 ).LT.1.08.AND.S(J1-1,1).GT.1.01.AND.S(J1-2,8).LT.1.08.AND.S(J1-2,
270               2 1).GT.0.69.AND.S(J1+1,8).LT.1.08.AND.S(J1+2,8).LE.S(J1-3,8).AND.
271               3 S(J1+3,8).LT.1.08)  BALL=.TRUE.
272                IF (BALL)  J1=J1-3
273                IF (BALL)  V2=46
274                IF (BALL)  GO TO 300
275      C
276      C      *** 47 ***
277                BALL=.FALSE.
278                IF (S(J1,8).LT.1.08.AND.M(J1-1).EQ.15.AND.S(J1-2,8).LT.S(J1-3,8).
279               1 AND.M(J1-3).EQ.1.AND.S(J1-4,8).LT.1.08.AND.S(J1+1,8).LT.1.08.AND.
280               2 M(J1+2).EQ.1.AND.M(J1+4).EQ.1)  BALL=.TRUE.
281                IF (BALL)  J1=J1-3
282                IF (BALL)  V2=47
283                IF (BALL)  GO TO 300
284      C
285      C      *** 48 ***
286                BALL=.FALSE.
287                IF (M(J1).EQ.1.AND.S(J1-3,8).GE.1.08.AND.S(J1-4,6).GE.1.22.AND.
288               1 S(J1-1,8).LT.1.08.AND.S(J1-1,1).GT.0.69.AND.S(J1-2,8).LT.1.08.AND
289               2 .S(J1-2,1).GT.1.01.AND.S(J1+1,8).LT.1.08.AND.S(J1+2,8).LT.S(J1-3
290               3  ,8).AND.S(J1+3,8).LT.1.08)  BALL=.TRUE.
291                IF (BALL)  J1=J1-3
292                IF (BALL)  V2=48
293                IF (BALL)  GO TO 300
294      C
295      C      *** 49 ***
296                BALL=.FALSE.
297                IF (S(J1,8).GE.1.08.AND.S(J1+1,8).GE.1.08.AND.S(J1+2,8).LT.1.08.A
298               1 ND.S(J1+3,8).LT.1.08.AND.M(J1-3).EQ.15.AND.S(J1-1,8).LT.1.08.AND.
299               2 S(J1-1,1).GT.1.01.AND.S(J1-2,8).LT.1.08.AND.S(J1-2,1).GT.0.69.AND
300               3 .S(J1-4,8).LT.1.08)  BALL=.TRUE.
```

```
301               IF (BALL)  J1=J1-3
302               IF (BALL)  V2=49
303               IF (BALL)  GO TO 300
304     C
305     C      *** 50 ***
306          50   BALL=.FALSE.
307               IF ((J1-3).LE.0)  GO TO 300
308               IF (M(J1).EQ.1.AND.S(J1-3,8).GT.1.47.AND.S(J1-1,8).LT.1.08.AND.S
309              1 (J1-2,8).LT.1.08.AND.(M(J1-1).EQ.3.OR.M(J1-1).EQ.16.OR.M(J1-1).EQ
310              2 .8).AND.S(J1+1,8).LT.1.08.AND.S(J1+2,8).LT.1.08.AND.S(J1+3,8).GE.
311              3 S(J1-3,8).AND.S(J1+4,8).GT.1.08)  BALL=.TRUE.
312               IF (BALL)  J1=J1-3
313               IF (BALL)  V2=50
314               IF (BALL)  GO TO 300
315     C
316     C      *** 51 ***
317               IF ((J1-5).LE.0)  GO TO 60
318               BALL=.FALSE.
319               IF (S(J1,8).LT.1.08.AND.S(J1,1).GT.1.01.AND.M(J1-3).EQ.15.AND.S(J1
320              1 -1,2).LE.0.75.AND.S(J1-5,2).LE.0.75.AND.S(J1+1,2).LE.0.75.AND.S(J
321              2 1-2,8).LT.2.01.AND.S(J1+1,8).LT.2.01.AND.S(J1+2,8).LT.2.01.AND.S(
322              3 1+3,1).GT.1.01.AND.S(J1+4,1).GT.1.08)  BALL=.TRUE.
323               IF (BALL)  J1=J1-3
324               IF (BALL)  V2=52
325               IF (BALL)  GO TO 300
326     C
327     C ... J1 = J1+2 .....
328     C ....................
329     C
330     C      *** 53 ***
331          60   BALL=.FALSE.
332               IF ((J1-3).LE.0)  GO TO 300
333               IF (S(J1,8).LT.1.08.AND.S(J1,1).LT.1.06.AND.S(J1+2,8).GT.1.47.AND.
334              1 S(J1+1,8).LT.1.08.AND.S(J1-1,1).LE.0.69.AND.M(J1-1).NE.15.AND.S(
335              2 J1-2,8).LT.1.08.AND.S(J1-3,8).LT.1.08.AND.S(J1-2,1).LT.1.06.AND.S
336              3(J1-3,1).LT.1.06)  BALL=.TRUE.
337               IF (BALL)  J1=J1+2
338               IF (BALL)  V2=53
339               IF (BALL)  GO TO 300
340     C
341     C      *** 54 ***
342               BALL=.FALSE.
343               IF (S(J1,1).LE.0.69.AND.M(J).NE.15.AND.S(J1+1,8).LT.1.08.AND.S(J1+
344              1 2,8).GT.1.47.AND.S(J1-1,1).LE.0.69.AND.M(J1-1).NE.15.AND.S(J1-2,
345              2.8).LT.S(J1+2,8).AND.S(J1-3,8).LT.1.08)  BALL=.TRUE.
346               IF (BALL)  J1=J1+2
347               IF (BALL)  V2=54
348               IF (BALL)  GO TO 300
349     C
350     C      *** 55 ***
```

```
              BALL=.FALSE.
              IF (S(J1,8).LT.1.08.AND.S(J1+1,1).LE.0.69.AND.M(J1+1).NE.15.AND.
             1 S(J1+2,8).GT.1.08.AND.S(J1-1,8).LT.1.08.AND.S(J1-2,8).LT.S(J1+2,8
             2 ).AND.S(J1-3,8).LE.S(J1+2,8).AND.S(J1-4,8).LE.S(J1+2,8)) BALL=
             3  .TRUE.
              IF (BALL)  J1=J1+2
              IF (BALL)  V2=55
              IF (BALL)  GO TO 300
C
C     *** 56 ***
              IF ((J1-4).LE.0)  GO TO 70
              BALL=.FALSE.
              IF (S(J1,8).LT.1.08.AND.S(J1,1).LT.1.06.AND.S(J1+1,8).LT.1.08.AND.
             1 S(J1+2,8).GT.1.08.AND.S(J1+2,1).GT.1.01.AND.S(J1-1,8).LT.1.08.AND
             2 .S(J1-1,2).GE.1.47.AND.S(J1-2,1).LE.0.69.AND.M(J1-3).EQ.1.AND.S(J
             3 1-4,2).GE.1.47)  BALL=.TRUE.
              IF (BALL)  J1=J1+2
              IF (BALL)  V2=56
              IF (BALL)  GO TO 300
C
C     *** 57 ***
   70         BALL=.FALSE.
              IF ((J1-3).LE.0)  GO TO 300
              T1=0
              T2=0
              T5=0
              T1=S(J1-3,1)+S(J1-2,1)+S(J1-1,1)+S(J1,1)
              T2=S(J1-3,2)+S(J1-2,2)+S(J1-1,2)+S(J1,2)
              T5=S(J1-3,5)+S(J1-2,5)+S(J1-1,5)+S(J1,5)
              PRINT  75,T1,T2,T5
   75         FORMAT(' ',30X,'T1,T2,T5',3(F7.3),'    STEP 57, J1+2 ,RMJ1')
              IF (T5.GT.T1.AND.T5.GT.T2.AND.S(J1+1,8).LT.1.08.AND.S(J1+1,2).GT.
             1 1.38.AND.S(J1+2,8).GT.0.81.AND.S(J1+3,8).LT.1.08.AND.S(J1+4,8).LT
             2 .1.08.AND.S(J1+3,2).GT.1.10.AND.S(J1+4,2).GT.1.10.AND.S(J1,1).LT.
             3 1.06.AND.S(J1,5).GT.1.43.AND.S(J1-1,5).GT.1.52.AND.S(J1-2,5).GT.
             4 1.52)  BALL=.TRUE.
              IF (BALL)  J1=J1+2
              IF (BALL)  V2=57
              IF (BALL)  GO TO 300
C
C ... J1 = J1-2 .....
C ...................
C
C     *** 58 ***
              BALL=.FALSE.
              IF (S(J1,8).LE.1.08.AND.S(J1-1,1).GT.1.01.AND.S(J1-1,8).LT.S(J1-2,
             1 8).AND.S(J1-2,8).GT.1.47.AND.S(J1-3,8).LT.S(J1-2,8).AND.S(J1+1,8)
             2 .LT.1.08.AND.S(J1+2,8).LT.1.08.AND.S(J1+3,8).LT.1.08) BALL=.TRUE.
              IF (BALL)  J1=J1-2
              IF (BALL)  V2=58
```

```
401                 IF (BALL)  GO TO 300
402       C
403       C     *** 59 ***
404                 BALL=.FALSE.
405                 IF (S(J1,8).LT.1.08.AND.S(J1,1).GT.1.01.AND.S(J1-2,8).GE.S(J1,8)
406                1 .AND.S(J1-1,1).LT.S(J1-2,1).AND.S(J1+1,8).LT.1.47.AND.S(J1+2,8).L
407                2T.1.08.AND.(S(J1-3,1).LE.0.69.OR.S(J1-3,8).LT.1.08).AND.(S(J1-2,1)
408                3 .GT.1.16.OR.M(J1-1).NE.15))  BALL=.TRUE.
409                 IF (BALL)  J1=J1-2
410                 IF (BALL)  V2=59
411                 IF (BALL)  GO TO 300
412       C
413       C     *** 60 ***
414                 BALL=.FALSE.
415                 IF (S(J1,8).LT.1.08.AND.S(J1-1,8).LT.1.08.AND.S(J1-2,8).GT.1.22.AN
416                1 D.S(J1+1,8).LT.1.08.AND.S(J1+2,8).GE.1.08.AND.S(J1+3,8).GT.2.01
417                2 .AND.S(J1-3,8).LT.S(J1-2,8).AND.S(J1-2,1).GT.0.69)  BALL=.TRUE.
418                 IF (BALL)  J1=J1-2
419                 IF (BALL)  V2=60
420                 IF (BALL)  GO TO 300
421       C
422       C     *** 61 ***
423                 BALL=.FALSE.
424                 IF (S(J1,8).LT.1.08.AND.S(J1,1).GT.1.01.AND.S(J1-2,8).GT.2.01.AND.
425                1 S( J1-1,8).LT.S(J1-2,8).AND.S(J1-3,8).LE.S(J1-2,8).AND.S(J1+1,8).
426                2LE.S(J1-2,8).AND.S(J1+2,8).GE.1.08.AND.S(J1+3,8).GE.S(J1-2,8))
427                3  BALL=.TRUE.
428                 IF (BALL)  J1=J1-2
429                 IF (BALL)  V2=61
430                 IF (BALL)  GO TO 300
431       C
432       C     *** 62 ***
433                 BALL=.FALSE.
434                 IF (S(J1,8).GT.1.47.AND.S(J1-1,8).GT.1.47.AND.S(J1-2,8).GT.1.47.AN
435                1 D.S(J1+1,8).GT.1.47.AND.S(J1+2,8).LT.1.08.AND.S(J1+3,8).LT.1.08
436                2.AND.S(J1-3,8).LT.1.08)  BALL=.TRUE.
437                 IF (BALL)  J1=J1-2
438                 IF (BALL)  V2=62
439                 IF (BALL)  GO TO 300
440       C
441       C     *** 63 ***
442                 BALL=.FALSE.
443                 IF (S(J1,8).GT.1.47.AND.M(J1-2).EQ.1.AND.S(J1-1,8).LT.1.08.AND.S(J
444                1 1-3,8).LT.1.08.AND.S(J1+1,8).LT.S(J1-2,8).AND.S(J1+2,8).LT.S(J1-2
445                2 ,8).AND.S(J1+3,8).LE.S(J1-2,8))  BALL=.TRUE.
446                 IF (BALL)  J1=J1-2
447                 IF (BALL)  V2=63
448                 IF (BALL)  GO TO 300
449       C
450       C     *** 64 ***
```

```
451               BALL=.FALSE.
452               IF (M(J1).EQ.7.AND.S(J1-2,8).GT.1.47.AND.S(J1-3,1).GT.0.69.AND.S(
453              1 J1-1,8).LT.1.08.AND.S(J1+1,8).LT.1.08.AND.S(J1+2,1).GT.1.01.AND.S
454              2 (J1-3,8).LT.1.47)  BALL=.TRUE.
455               IF (BALL)  J1=J1-2
456               IF (BALL)  V2=64
457               IF (BALL)  GO TO 300
458     C
459     C ... J1 = J1+1 .....
460     C ....................
461     C
462     C      *** 65 ***
463               BALL=.FALSE.
464               IF (S(J1,8).LE.1.08.AND.S(J1+1,8).GE.1.08.AND.S(J1+1,1).GT.1.01
465              1.AND.(S(J1+2,8).LT.S(J1+1,8).OR.S(J1+2,8).LT.1.47).AND.S(J1+3,8)
466              2.LT.1.08.AND.S(J1-1,8).LT.1.08.AND.S(J1-2,8).LE.1.08.AND.S(J1-3,8)
467              3.LT.1.08.AND.S(J1+4,8).LT.1.08)  BALL=.TRUE.
468               IF (BALL)  J1=J1+1
469               IF (BALL)  V2=65
470               IF (BALL)  GO TO 300
471     C
472     C      *** 66 ***
473               BALL=.FALSE.
474               IF (S(J1,8).LT.1.08.AND.S(J1+1,8).GT.1.47.AND.S(J1+2,8).LT.S(J1+1,
475              1 8).AND.S(J1+3,8).LT.S(J1+1,8).AND.S(J1-1,8).LT.S(J1+1,8).AND.S(J1
476              2 -2,8).LT.S(J1+1,8).AND.S(J1-3,8).LT.S(J1+1,8)) BALL=.TRUE.
477               IF (BALL)  J1=J1+1
478               IF (BALL)  V2=66
479               IF (BALL)  GO TO 300
480     C
481     C
482     C ... J1 = J1-1 .....
483     C ....................
484     C
485     C      *** 67 ***
486               BALL=.FALSE.
487               IF (S(J1,8).LT.1.08.AND.S(J1-1,8).GT.1.08.AND.S(J1-1,1).GT.1.16.AN
488              1 D.S(J1-2,8).LT.S(J1-1,8).AND.S(J1+1,8).LT.1.08.AND.S(J1+2,8).LE.
489              2 S(J1-1,8).AND.S(J1-3,8).LT.1.08)  BALL=.TRUE.
490               IF (BALL)  J1=J1-1
491               IF (BALL)  V2=67
492               IF (BALL)  GO TO 300
493     C
494     C      *** 68 ***
495               BALL=.FALSE.
496               IF (S(J1,8).LT.1.08.AND.M(J1-1).EQ.15.AND.S(J1+1,8).LT.S(J1-1,8).
497              1 AND.S(J1+2,8).LT.S(J1-1,8).AND.S(J1+3,8).LT.S(J1-1,8).AND.S(J1-2,
498              2 8).LT.S(J1-1,8).AND.S(J1-3,8).LT.1.08.AND.S(J1,1).GT.1.01)  BALL
499              4 =.TRUE.
500               IF (BALL)  J1=J1-1
```

```
501               IF (BALL)  V2=68
502               IF (BALL)  GO TO 300
503     C
504     C      *** 69 ***
505               BALL=.FALSE.
506               IF (S(J1,8).GE.1.08.AND.S(J1,1).GT.1.01.AND.S(J1,8).LT.2.44.AND.S(
507              1 J1+1,8).LE.1.08.AND.S(J1+2,8).GE.1.08.AND.S(J1+3,8).LT.S(J1-1,8)
508              2 .AND.M(J1-1).EQ.15.AND.S(J1-2,1).LE.0.69 .AND.(S(J1-3,1).LE.0.69
509              3 .OR.S(J1-3,8).LT.1.08))  BALL=.TRUE.
510               IF (BALL)  J1=J1-1
511               IF (BALL)  V2=69
512               IF (BALL)  GO TO 300
513     C
514     C      *** 70 ***
515               BALL=.FALSE.
516               IF (M(J1).EQ.7.AND.M(J1-1).EQ.15.AND.(S(J1-2,1).LE.0.69.OR.S(J1-2,
517              1 8).LT.1.08).AND.(S(J1-3,1).LE.0.69.OR.S( J1-3,8).LT.1.08).AND.S(J
518              2 1+1,8).GE.1.08.AND.S(J1+2,8).LT.1.08.AND.S(J1+3,8).GE.1.08)  BALL
519              3  =.TRUE.
520               IF (BALL)  J1=J1-1
521               IF (BALL)  V2=70
522               IF (BALL)  GO TO 300
523     C
524     C      *** 71 ***
525               BALL=.FALSE.
526               IF (S(J1,8).LT.1.08.AND.S(J1+1,8).LT.1.08.AND.S(J1+2,8).LT.1.08.AN
527              1 D.S(J1+3,8).LT.1.08.AND.S(J1-1,1).GT.1.16.AND.S(J1-1,8).GE.1.08.
528              2 AND.S(J1-2,8).LT.1.08.AND.S(J1-3,8).LE.S(J1-1,8))  BALL=.TRUE.
529               IF (BALL)  J1=J1-1
530               IF (BALL)  V2=71
531               IF (BALL)  GO TO 300
532     C
533     C      *** 72 ***
534               BALL=.FALSE.
535               IF (S(J1,8).GE.1.08.AND.S(J1-1,8).GT.S(J1,8).AND.S(J1+1,8).LT.1.08
536              1 .AND.S(J1+2,8).LT.1.08.AND.S(J1+3,8).LT.1.08.AND.S (J1-2,1).LE.
537              2 0.69.AND.S(J1-3,8).LE.S(J1-1,8).AND.S(J1,1).GT.1.01)  BALL=.TRUE.
538               IF (BALL)  J1=J1-1
539               IF (BALL)  V2=72
540               IF (BALL)  GO TO 300
541     C
542     C      *** 73 ***
543               BALL=.FALSE.
544               IF (S(J1,8).GT.1.08.AND.S(J1,1).GT.1.01.AND.S(J1+1,8).LE.1.08.AND.
545              1 S(J1+2,8).LT.1.08.AND.S(J1+3,8).LT.S(J1-1,8).AND.M(J1-1).EQ.15.AN
546              2 D.(S(J1-2,1).LE.0.69.OR.S(J1-2,8).LT.S(J1-1,8)).AND.(S(J1-3,1).LE
547              3 .0.69.OR.S(J1-3,8).LT.1.08))  BALL=.TRUE.
548               IF (BALL)  J1=J1-1
549               IF (BALL)  V2=73
550               IF (BALL)  GO TO 300
```

```
    551
    552
    553           END OF N-BOUNDARY ADJUSTMENT. TO CALL SUBROUTINE MOJ2 FOR C-BOUNDA
    554           RY ADJUSTMENT
    555
    556         300   CALL MOJ2
    557               RETURN
    558               END
End of File
```

```
C
C
      SUBROUTINE  MOJ2
C     .................
C
C
C     ..........................................................
C     .                                                        .
C     .            BOUNDARY MOVE OF THE C-TERMINAL             .
C     .                                                        .
C     ..........................................................
C
C
C     ..................................................................
C
C     PURPOSE
C        TO ADJUST THE C-TERMINAL RESIDUE BASED ON THE BOUNDARY CONFOR
C        MATIONAL PARAMETERS AND ON THE POTENTIAL OF TURN OR SHEET OF
C        THE ADJACENT REGIONS
C
C     ..................................................................
C
C
      REAL S,T1,T2,A1,A2 ,T3,T4,T5,TT,P
      INTEGER G,F,H,U,D,V1,V2,W, V3,V4,V5,V6,V7,V8,Q
      LOGICAL HELLO,BYE  ,BALL,MOVE
      DIMENSION S(1000,20),M(1000),H(1000),D(1000,16),P(1000,10)
      COMMON S,T1,T2,T3,T4,T5,TT,A1,A2,P,F,H,U,D,W,M,M1,M2,M3,M4,M5,M6,
     1L,I,K,L1,L2,NZ,NY,JA,JB,JC,JD,J1,J2,KM,N1,N2,NN,J,G,K3,V1,V2,V3,V4
     2,V5,V6,V7,Q,HELLO,BYE,BALL,MOVE
C
C
C     DESCRIPTION OF PARAMETERS
C        V1  - NUMBER OF BREAKERS IN THE PREDICTED HELIX BEFORE THE BOUN
C              DARY ADJUSTMENT
C        V2  - COUNTER USED IN N-BOUNDARY ADJUSTMENT
C        V3  - COUNTER USED IN C-BOUNDARY ADJUSTMENT
C
C
C     V2=80 WHEN THE N-TERMINAL ADJUSTMENT IS DUE TO STRONG B-TURN POTEN
C     TIAL (THROUGH THE PROCEDURE OF REPEATING THE B-TURN CHECK).
C     IF V2=0 NONE OF THE CONDITIONS LISTED IN  THE N-TERMINAL ADJUSTMENT
C     FIT THE CURRENTLY TESTED SEGMENT. IN OTHER WORDS J1 HAS NOT CHANGED
C
      PRINT 1
    1 FORMAT('0', 30X,'BOUNDARY ANALYSIS OF THE C-TERMINAL')
      IF (.NOT. BALL.AND. V2.NE. 80)  V2=0
C
C
C ..... SITUATION WITH J2 CLOSE TO THE C- BOUNDARY .....
```

```
      C
      C     TO TAKE INTO ACCOUNT THE POSITION OF J2 WHEN IT IS CLOSE TO THE C-
      C     TERMINAL OF THE PROTEIN SINCE THERE IS LESS FREEDOM TO MOVE IT TO
      C     WARDS  THIS END
      C
      C     *** 1 ***
            BALL=.FALSE.
            IF (J2.EQ.NN.AND.S(J2,9).GT.1.10.AND.S(J2,1).GT.1.01.AND.(S(J2,9)
           1 .GT.S(J2-1,9).OR.S(J2-1,1).LT.S(J2,1))) BALL=.TRUE.
            IF (BALL)  J2=J2
            IF (BALL)  V3=1
            IF (BALL)  GO TO 300
      C
      C     *** 2 ***
            BALL=.FALSE.
            IF ((J2+3).GT.NN)  GO TO 20
            IF (S(J2,9).GT.0.98.AND.S(J2,1).GT.0.69.AND.S(J2+1,1).LE.0.69.AND.
           1 S(J2+2,1).LE.0.69.AND.S(J2+3,9).LT.1.57.AND.S(J2-1,1).GT.0.69.AND
           2 .S(J2-2,9).GT.1.10.AND.S(J2-3,1).GT.1.16)  BALL=.TRUE.
            IF (BALL)  J2=J2
            IF (BALL)  V3=2
            IF (BALL)  GO TO 300
      C
      C     *** 3 ***
            BALL=.FALSE.
            T1=0
            T2=0
            T1=S(J2,1)+S(J2+1,1)+S(J2+2,1)+S(J2+3,1)
            T2=S(J2,2)+S(J2+1,2)+S(J2+2,2)+S(J2+3,2)
            PRINT  10,T1,T2
        10  FORMAT(' ',30X,'T1,T2    ',2(F7.3),7X,'    STEP 3, J2 CLOSE TO O')
            IF (T2.GT.T1.AND.S(J2,2).GT.1.38.AND.S(J2+2,2).GT.1.38.AND.S(J2+1,
           1 1).LE.0.69.AND.S(J2,9).GT.1.20.AND.S(J2-1,1).GT.1.16.AND.S(J2-2,
           2 2).LE.0.75)  BALL=.TRUE.
            IF (BALL)  J2=J2
            IF (BALL)  V3=3
            IF (BALL)  GO TO 300
      C
      C     *** 4 ***
            BALL=.FALSE.
            IF ((J2+4).GT.NN)  GO TO 20
            IF (S(J2,9).GT.1.08.AND.S(J2-1,1).GT.1.16.AND.S(J2-2,1).GT.1.16.AN
           1 D.S(J2+1,1).GT.0.77.AND.S(J2+2,1).GT.1.01.AND.S(J2+2,9).GT.S(J2+1
           2,9).AND.S(J2+3,7).GT.1.58.AND.S(J2+4,7).GT.1.58)  BALL=.TRUE.
            IF (BALL)  J2=J2+2
            IF (BALL)  V3=4
            IF (BALL)  GO TO 300
      C
      C     *** 5 ***
        20  BALL=.FALSE.
```

```
101               IF ((J2+1).GT.NN)  GO TO 60
102               IF ((J2+1).EQ.NN.AND.M(J2+1).NE.15.AND.V1.LT.((J2+1-J1)/3).AND.S(J
103              1 2,9).GT.1.10.AND.S(J2-1,9).GT.1.10)  BALL=.TRUE.
104               IF (BALL)  J2=J2+1
105               IF (BALL)  V3=5
106               IF (BALL)  GO TO 300
107      C
108      C      *** 6 ***
109               BALL=.FALSE.
110               IF ((J2+2).GT.NN)  GO TO 60
111               T1=0
112               T2=0
113               T5=0
114               TT=0
115               T1=S(J2-1,1)+S(J2,1)+S(J2+1,1)+S(J2+2,1)
116               T2=S(J2-1,2)+S(J2,2)+S(J2+1,2)+S(J2+2,2)
117               T5=S(J2-1,5)+S(J2,5)+S(J2+1,5)+S(J2+2,5)
118               TT=P(J2-1,1)*P(J2,2)*P(J2+1,3)*P(J2+2,4)
119               PRINT  25,T1,T2,T5,TT
120           25  FORMAT(' ',30X,'T1,T2,T5,TT',3(F7.3),F13.9,'  STEP 6, J2 CLOSE O')
121               IF (T5.GT.T1.AND.T5.GT.T2.AND.TT.GT.0.00007500.AND.S(J2-1,9).GT.1.
122              1 57.AND.S(J2-1,9).GT.S(J2-2,9).AND.(S(J2-3,1).GT.1.16.OR.S(J2-3,9)
123              2.GT.1.20))  BALL=.TRUE.
124               IF (BALL)  J2=J2-1
125               IF (BALL)  V3=6
126               IF (BALL)  GO TO 300
127      C
128      C
129      C
130      C      THE DIFFERENT COMMENTS J2=J2,J2=J2+10,...,J2=J2-4 INDICATE THE EVE
131      C      NTUAL POSITION OF J2 IF ITS ENVIRONMENT MEETS ONE OF THE CONDITIONS
132      C      DESCRIBED BELOW
133      C
134      C ..... J2 = J2-10 .....
135      C ......................
136      C
137               IF ((J2-10).LE.0)  GO TO 50
138               IF ((J2+3).GT.NN)  GO TO 60
139               J3=J2
140           30  J4=J3+3
141               BALL=.FALSE.
142               T1=0
143               T2=0
144               T5=0
145               TT=0
146               DO 40 N=J3,J4
147               T1=T1+S(N,1)
148               T2=T2+S(N,2)
149               T5=T5+S(N,5)
150           40  CONTINUE
```

```
151               TT = P(J3,1)*P(J3+1,2)*P(J3+2,3)*P(J3+3,4)
152               PRINT  45,T1,T2,T5,TT
153          45   FORMAT(' ',30X,'T1,T2,T5,TT',3(F7.3),F13.9,'   STEP7, J2-10 , MOJ2')
154               IF (T5.GT.T1.AND.T5.GT.T2.AND.TT.GE.0.00007500.AND.S(J3+1,1).LE.0.
155              1 69.AND.S(J3+2,1).LE.0.69.AND.S(J3+3,1).LT.1.06.AND.S(J3,1).LT.0.
156              2 98.AND.S(J3-1,1).LT.0.98)  J3=J2-7
157               IF (T5.GT.T1.AND.T5.GT.T2.AND.TT.GE.0.00007500.AND.S(J3+1,1).LE.0.
158              1 69.AND.S(J3+2,1).LE.0.69.AND.S(J3+3,1).LT.1.06.AND.S(J3,1).LT.0.
159              2 98.AND.S(J3-1,1).LT.0.98)  GO TO 30
160               IF (T5.GT.T1.AND.T5.GT.T2.AND.TT.GE.0.00007500.AND.J3.EQ.(J2-7).AN
161              1 D.S(J2-8,8).LT.1.10.AND.S(J2-8,1).LT.0.98.AND.S(J2-9,9).LT.1.10.A
162              2ND.S(J2-10,9).GT.1.25.AND.S(J2-10,1).GT.1.16)  BALL=.TRUE.
163               IF (BALL)  J2=J2-10
164               IF (BALL)  V3=7
165               IF (BALL)  GO TO 300
166     C
167     C ..... J2 = J2+10 .....
168     C .......................
169     C
170     C      *** 8 ***
171          50   IF ((J2+11).GT.NN)  GO TO 60
172               BALL=.FALSE.
173               IF (M(J2+1).EQ.16.AND.M(J2).EQ.16.AND.M(J2+3).EQ.16.AND.M(J2+8).EQ
174              1 .16.AND.(P(J2+8,1)*P(J2+9,2)*P(J2+10,3)*P(J2+11,4)).GT.0.000100.A
175              2 ND.S(J2-1,1).GT.1.16.AND.S(J2+2,1).GT.1.01.AND.S(J2+5,1).GT.1.01
176              3 .AND.S(J2+6,1).GT.1.16.AND.S(J2+4,1).GT.0.77.AND.S(J2+7,1).GT.0.
177              4 77.AND.S(J2-3,1).GT.1.13.AND.S(J2-2,1).GT.1.11)  BALL=.TRUE.
178               IF (BALL)  J2=J2+8
179               IF (BALL)  V3=8
180               IF (BALL)  GO TO 300
181     C
182     C
183     C      *** 9 ***
184               BALL=.FALSE.
185               IF ((J2+12).GT.NN)  GO TO 60
186               IF (S(J2,9).GT.1.57.AND.S(J2-2,1).GT.1.16.AND.S(J2+10,1).GT.1.16
187              1 .AND.S(J2+11,7).GT.1.49.AND.S(J2+12,7).GT.1.58.AND.S(J2+2,1).GT.1
188              2 .16.AND.S(J2+3,1).GT.1.16.AND.S(J2+6,1).GT.1.16.AND.S(J2+7,1).GT.
189              3 1.21.AND.S(J2+8,9).GT.1.20.AND.S(J2+8,1).GT.1.01.AND.S(J2+9,7).GT
190              4 .1.57.AND.S(J2+1,9).GT.0.98.AND.S(J2+4,2).EQ.0.75.AND.S(J2+5,1).
191              5 GT.0.77)  BALL=.TRUE.
192               IF (BALL)  J2=J2+10
193               IF (BALL)  V3=9
194               IF (BALL)  GO TO 300
195     C
196     C ..... TO REPEAT THE B-TURN CHECK .....
197     C
198     C      TO CHECK THE PRESENCE OF TURNS IN THE VICINITY OF THE HELIX BOUNDA
199     C      RIES WHICH MAY FORCE THE PREDICTED BOUNDARIES TO BE MOVED TO A NEW
200     C      POSITION.  THIS PROCEDURE STARTS FROM POSITION J2-4 (I=0) TO J2+2
```

```
C      (I=6)
C
   60     I=0
          LE=J2-4
   70     LF=LE+3
          IF ((LE+3).GT.NN)  GO TO 210
          HELLO=.FALSE.
C
C      TO COMPARE PA (T1),PB (T2),AND PT (T5) AND TO CALCULATE THE PROBA
C      BILITY OF B-TURN OCCURRENCE (TT) OF THE TETRAPEPTIDE LE-LF
C
          T1=0
          T2=0
          T5=0
          TT=0
          DO 75  L=LE,LF
          T1=T1+S(L,1)
          T2=T2+S(L,2)
          T5=T5+S(L,5)
   75     CONTINUE
          TT=P(LE,1)*P(LE+1,2)*P(LE+2,3)*P(LE+3,4)
          PRINT  78,LE,T1,T2,T5,TT,I
   78     FORMAT(' ',10X,'LE,T1,T2,T5,TT,I',I5,3(F7.4,2X),F13.9,I4,3X,
         1 'B-TURN SEARCH AT C-TERMINAL')
          IF (T5.GT.T1.AND.T5.GT.T2.AND.TT.GE.0.00007500)  HELLO=.TRUE.
C
C      *** 1 ***
          IF ((J2+1).GT.NN)  GO TO 80
          IF (HELLO.AND.LE.EQ.(J2+1).AND.S(J2,9).GT.1.10.AND.S(J2,1).GT.1.01
         1 .AND.S(J2+1,1).LE.0.69.AND.S(J2-1,9).LE.S(J2,9).AND.S(J2-1,1).GT.
         2 0.67.AND.((S(J2-2,5)+S(J2-1,5)+S(J2,5)+S(J2+1,5)).LT.(S(J2-2,1)+
         3 S(J2-1,1)+S(J2,1)+S(J2+1,1)).OR.(P(J2-2,1)*P(J2-1,2)*P(J2,3)*P(J2
         4 +1,4)).LT.0.00007500))  GO TO 100
C
C      *** 2 ***
          IF (HELLO.AND.LE.EQ.J2.AND.S(J2-1,9).GT.1.10.AND.S(J2-1,1).GT.1.16
         1 .AND.S(J2,9).LT.1.10.AND.S(J2-1,1).GT.S(J2,1).AND.S(J2+1,1).LE.0.
         2 69)  GO TO 101
C
C      *** 3 ***
   80     IF (HELLO.AND.LE.EQ.(J2-1).AND.M(J2-1).EQ.16.AND.S(J2-2,9).GT.1.10
         1 .AND.S(J2-3,1).GT.1.16.AND.S(J2,5).GT.1.19.AND.S(J2+2,5).GT.1.19)
         2  GO TO 101
C
C      *** 4 ***
          IF (HELLO.AND.LE.EQ.(J2-1).AND.S(J2-2,9).GE.1.10.AND.(S(J2-3,9).LT
         1.S(J2-2,9).OR.(S(J2-3,9)-S(J2-2,9)).LT.0.15).AND.S(J2-1,9).LT.S(J2
         2 -2,9).AND.S(J2+1,5).GT.1.19.AND.S(J2+2,5).GT.1.19)  GO TO 102
C
C      *** 5 ***
```

```
251               IF (HELLO.AND.LE.EQ.(J2-2).AND.S(J2-2,9).GT.S(J2-3,9).AND.S(J2-3,2
252              1 ).LE.0.75.AND.S(J2-2,1).GT.0.69.AND.S(J2-5,2).LE.0.75)  GO TO 102
253       C
254       C      *** 6 ***
255               IF (HELLO.AND.LE.EQ.(J2-2).AND. S(J2-2,9).EQ.0.98.AND.S(J2-3,1).GT
256              1 .1.16.AND.S(J2-4,1).GT.1.16.AND.S(J2-3,9).LT.1.57.AND.S(J2-1,9).
257              2 LT.S(J2-2,9))  GO TO 102
258       C
259       C      *** 7 ***
260               IF (HELLO.AND.LE.EQ.J2.AND.(S(J2,9).LT.0.98.OR.S(J2,2).GT.1.38).AN
261              1 D.(S(J2-1,2).GT.1.60.OR.(S(J2-1,9).LT.1.10.AND.S(J2-1,2).GT.1.38)
262              2 ).AND.S(J2-2,9).GT.1.10.AND.S(J2-2,1).GT.1.16)  GO TO 102
263       C
264       C      *** 8 ***
265               IF (HELLO.AND.LE.EQ.(J2-1).AND.S(J2-1,9).LT.S(J2-2,9).AND.S(J2-2,1
266              1 ).GT.1.16.AND.S(J2-3,1).GT.1.16.AND.S(J2-1,1).LT.S(J2-2,1))  GO
267              2 TO 102
268       C
269       C      *** 9 ***
270               IF (HELLO.AND.LE.EQ.(J2-2).AND.S(J2-2,9).GT.1.10.AND.S(J2-3,9).LT.
271              1 1.10.AND.S(J2-4,9).LT.S(J2-2,9))  GO TO 102
272       C
273       C      *** 10 ***
274               IF ((J2+4).GT.NN)  GO TO 90
275               IF (HELLO.AND.LE.EQ.(J2-2).AND.M(J2-2).EQ.16.AND.S(J2-3,9).GT.0.98
276              1 .AND.S(J2-4,9).GT.1.10.AND.S(J2+2,2).GT.1.38.AND.S(J2-1,2).GT.1.0
277              2 5.AND.S(J2+3,2).GT.0.75.AND.S(J2+4,2).GT.1.10)  GO TO 102
278       C
279       C      *** 11 ***
280            90    IF (HELLO.AND.LE.EQ.J2.AND.(P(J2-3,1)*P(J2-2,2)*P(J2-1,3)*P(J2,4))
281              1 .GT.0.00007500.AND.S(J2-4,9).GT.S(J2-3,9).AND.S(J2-4,9).GT.1.10
282              2 .AND.S(J2-4,9).GT.S(J2-5,9))  GO TO 104
283       C
284       C      *** 12 ***
285               IF (HELLO.AND.LE.EQ.(J2-3).AND.S(J2-4,1).LT.1.00.AND.S(J2-3,1).GT.
286              1 0.98.AND.S(J2-4,9).GT.0.98.AND.S(J2-4,9).LT.1.57.AND.S(J2-2,7).GT
287              2 .1.06)  GO TO 103
288       C
289       C      *** 13 ***
290               IF (HELLO.AND.LE.EQ.(J2-1).AND.S(J2-1,1).LE.0.69.AND.S(J2-2,1).GT.
291              1 1.01.AND.S(J2-2,9).GT.0.98.AND.S(J2-3,9).LT.1.57.AND.S(J2-1,5).GT
292              2 .1.19.AND.S(J2,5).GT.0.98.AND.S(J2+1,5).GE.1.56)  GO TO 102
293       C
294       C      *** 14 ***
295               IF (HELLO.AND.LE.EQ.(J2-1).AND.S(J2-2,1).LE.0.69.AND.S(J2-2,7).GT.
296              1 1.49.AND.S(J2-3,9).GT.1.10.AND.S(J2-3,1).GT.1.16)  GO TO 103
297       C
298       C      *** 15 ***
299               IF (HELLO.AND.LE.EQ.(J2-2).AND.S(J2-3,9).GT.0.98.AND.S(J2-3,1).GT.
300              1 1.01.AND.S(J2-4,9).LE.S(J2-3,9).AND.(S(J2-5,9)-S(J2-3,9)).LE.0.16
```

```
     2 .AND.S(J2-1,9).LT.1.77.AND.S(J2-2,9).LT.1.77.AND.S(J2-1,5).GT.1.1
     3 9.AND.S(J2-2,5).GT.1.19)  GO TO 103
C
C      *** 16 ***
       IF (HELLO.AND.LE.EQ.(J2-3).AND.S(J2-3,9).GT.1.57.AND.S(J2-4,1).GT.
     1 1.16.AND.S(J2-2,7).GT.1.24.AND.S(J2-5,1).GT.1.16.AND.S(J2-1,1).
     2 LE.0.69)  GO TO 103
C
C      *** 17 ***
       IF (HELLO.AND.LE.EQ.(J2-2).AND.S(J2-3,9).LT.1.10.AND.S(J2-3,1).LE.
     1 0.69.AND.S(J2-4,9).GT.0.98.AND.S(J2-4,1).GT.1.01.AND.(S(J2-5,1).
     2 LT.S(J2-4,1).OR.(S(J2-5,9)-S(J2-4,9)).LE.0.15))  GO TO 104
C
C      *** 18 ***
       IF (HELLO.AND.LE.EQ.(J2-4).AND.S(J2-4,1).GT.1.16.AND.S(J2-4,9).GT.
     1 0.98.AND.S(J2-5,1).LT.S(J2-4,1).AND.S(J2-6,1).LT.S(J2-4,1).AND.
     2 S(J2-3,9).LT.0.98)  GO TO 104
C
C      *** 19 ***
       IF (HELLO.AND.LE.EQ.(J2-4).AND.S(J2-4,1).GT.0.98.AND.S(J2-5,9).LT.
     1 1.10.AND.S(J2-6,9).LT.1.10)  GO TO 104
C
C      *** 20 ***
       IF (HELLO.AND.LE.EQ.(J2-4).AND.S(J2-5,9).GT.1.25.AND.S(J2-5,1).GT.
     1 1.16.AND.S(J2-4,9).LT.S(J2-5,9).AND.S(J2-4,1).LT.S(J2-5,1).AND.S(
     2 J2-6,9).LT.S(J2-5,9))  GO TO 105
C
C      *** 21 ***
       IF ((J2+2).GT.NN)  GO TO 95
       IF (HELLO.AND.LE.EQ.(J2+1).AND.(S(J2-3,1)+S(J2-1,1)+S(J2-1,1)+S(J2
     1 ,1)+S(J2+1,1)+S(J2+2,1)).LT.(S(J2-3,2)+S(J2-2,2)+S(J2-1,2)+S(J2,2
     2 )+S(J2+1,2)+S(J2+2,2)).AND.S(J2-4,1).GT.1.16.AND.S(J2-4,9).GT.1.
     3 08.AND.S(J2-5,9).GT.1.10)  GO TO 104
C
C
   95  IF (I.EQ.0)  GO TO 200
       IF (I.EQ.1)  GO TO 200
       IF (I.EQ.2)  GO TO 200
       IF (I.EQ.3)  GO TO 200
       IF (I.EQ.4)  GO TO 200
       IF (I.EQ.5)  GO TO 200
       IF (I.EQ.6)  GO TO 210
C
  100  J2=J2
       GO TO 110
  101  J2=J2-1
       GO TO 110
  102  J2=J2-2
       GO TO 110
  103  J2=J2-3
```

```
351               GO TO 110
352          104  J2=J2-4
353               GO TO 110
354          105  J2=J2-5
355               GO TO 110
356       C
357          110  V3=80
358               GO TO 300
359       C
360          200  I=I+1
361               LE=LE+1
362               GO TO 70
363       C
364       C
365       C
366       C      THE CURRENT  POSITION OF J2 MAY BE THE MOST FAVORABLE ONE,HENCE NO
367       C      NEED TO ADJUST IT
368       C
369       C ..... J2 = J2 .....
370       C .....................
371       C
372       C      *** 10 ***
373          210  BALL=.FALSE.
374               IF ((J2+4).GT.NN)  GO TO 220
375               T1=0
376               T2=0
377               T5=0
378               T1=S(J2+1,1)+S(J2+2,1)+S(J2+3,1)+S(J2+4,1)
379               T2=S(J2+1,2)+S(J2+2,2)+S(J2+3,2)+S(J2+4,2)
380               T5=S(J2+1,5)+S(J2+2,5)+S(J2+3,5)+S(J2+4,5)
381               PRINT 215,T1,T2,T5
382          215  FORMAT(' ',30X,'T1,T2,T5',3(F7.3),'    STEP 10, J2=J2 , MOJ2')
383               IF (T5.GT.T1.AND.T5.GT.T2.AND.S(J2,9).GT.0.98.AND.S(J2,1).GT.0.69
384              1 .AND.S(J2-1,9).GE.1.57.AND.(S(J2-2,1).GT.1.16.OR.S(J2-2,9).GT.1.
385              2 20))  BALL=.TRUE.
386       C
387               IF (T5.GT.T1.AND.T5.GT.T2.AND.S(J2+1,1).LE.0.69.AND.S(J2,9).GE.1.
388              1 57.AND.S(J2-1,9).GT.0.98.AND.S(J2-2,1).GT.1.16)  BALL=.TRUE.
389               IF (BALL)  J2=J2
390               IF (BALL)  V3=10
391               IF (BALL)  GO TO 300
392       C      **** 11 ***
393               BALL=.FALSE.
394               IF (T5.GT.T1.AND.T5.GT.T2.AND.S(J2,9).LT.0.73.AND.S(J2-1,9).GE.1.
395              1 57.AND.S(J2-1,1).GT.1.01.AND.S(J2-2,9).GT.1.10)  BALL=.TRUE.
396               IF (BALL)  J2=J2-1
397               IF (BALL)  V3=11
398               IF (BALL)  GO TO 300
399       C
400       C      *** 12 ***
```

```
401               BALL=.FALSE.
402               IF ((J2+6).GT.NN)  GO TO 220
403               IF (S(J2,9).GT.1.10.AND.S(J2,1).GT.1.01.AND.S(J2-1,9).LT.S(J2,9).
404              1AND.S(J2,2).GT.0.93.AND.S(J2+1,2).GT.1.05.AND.S(J2+2,2).GT.1.38.AN
405              2 D.S(J2+4,2).GT.0.75.AND.S(J2+5,2).GT.1.38.AND.S(J2+6,2).GT.1.05)
406              3  BALL=.TRUE.
407               IF (BALL)  J2=J2
408               IF (BALL)  V3=12
409               IF (BALL)  GO TO 300
410     C
411     C      *** 13 ***
412          220  IF ((J2+3).GT.NN)  GO TO 230
413               BALL=.FALSE.
414               IF (S(J2,9).GT.1.10.AND.S(J2,1).GT.1.16.AND.S(J2+1,2).GT.1.38.AND.
415              1 S(J2-1,2).GT.1.38.AND.S(J2-2,2).GT.1.38.AND.S(J2+3,2).GT.1.10.AND
416              2 .S(J2-3,2).GT.1.05)  BALL=.TRUE.
417               IF (BALL)  J2=J2
418               IF (BALL)  V3=13
419               IF (BALL)  GO TO 300
420     C
421     C      *** 14 ***
422          230  BALL=.FALSE.
423               IF ((J2+1).GT.NN)  GO TO 270
424               IF (S(J2,9).GT.1.57.AND.M(J2+1).EQ.15.AND.S(J2-1,9).LT.S(J2,9).AND
425              1 .S(J2-2,9).GT.1.10.AND.(S(J2-3,9).GT.1.10.OR.S(J2-3,1).GT.0.69))
426              2  BALL=.TRUE.
427               IF (BALL)  J2=J2
428               IF (BALL)  V3=14
429               IF (BALL)  GO TO 300
430     C
431     C      *** 15 ***
432               BALL=.FALSE.
433               IF ((J2+5).GT.NN)  GO TO 240
434               IF (M(J2).EQ.16.AND.S(J2+1,1).LE.0.69.AND.S(J2+2,2).GT.1.38.AND.
435              1 S(J2+3,2).GT.1.38.AND.S(J2+4,2).GT.1.38.AND.S(J2+5,2).GT.1.38.AND
436              2 .S(J2-1,9).GT.1.10.AND.S(J2-2,9).GT.1.10)  BALL=.TRUE.
437               IF (BALL)  J2=J2
438               IF (BALL)  V3=15
439               IF (BALL)  GO TO 300
440     C
441     C      *** 16 ***
442          240  BALL=.FALSE.
443               IF ((J2+4).GT.NN)  GO TO 250
444               T1=0
445               T2=0
446               T1=S(J2+1,1)+S(J2+2,1)+S(J2+3,1)+S(J2+4,1)
447               T2=S(J2+1,2)+S(J2+2,2)+S(J2+3,2)+S(J2+4,2)
448               PRINT 245,T1,T2
449          245  FORMAT(' ',30X,'T1,T2    ',2(F7.3),7X,'     STEP 16 , J2=J2')
450                IF (T2.GT.T1.AND.S(J2,1).GT.1.16.AND.S(J2+1,9).LT.1.10.AND.S(J2-1
```

```
451              1 ,1).LT.S(J2,1).AND.S(J2+2,2).GT.0.75.AND.S(J2+3,2).GT.1.38.AND.
452              2 S(J2+4,2).GT.1.38)  BALL=.TRUE.
453               IF (BALL)  J2=J2
454               IF (BALL)  V3=16
455               IF (BALL)  GO TO 300
456      C
457      C      *** 17 ***
458               BALL=.FALSE.
459               IF (S(J2,1).GT.1.16.AND.S(J2,9).GT.1.10.AND.S(J2-1,9).GT.1.10.AND.
460              1 S(J2+1,1).LE.0.69.AND.(S(J2+2,9).LT.1.10.OR.S(J2+2,1).LT.1.06).AN
461              3 D.S(J2+3,9).LT.1.10.AND.(S(J2+4,9).LT.1.10.OR.M(J2+3).EQ.15))  BA
462              4  LL=.TRUE.
463               IF (BALL)  J2=J2
464               IF (BALL)  V3=17
465               IF (BALL)  GO TO 300
466      C
467      C
468      C ..... J2 = J2-1 .....
469      C .....................
470      C
471      C      *** 18 ***
472          250   IF ((J2+2).GT.NN)  GO TO 260
473               BALL=.FALSE.
474               IF ((P(J2-1,1)*P(J2,2)*P(J2+1,3)*P(J2+2,4)).GT.0.00007500.AND.S(J2
475              1 -1,9).GT.1.10.AND.S(J2-1,1).GT.1.16.AND.S(J2-2,1).LT.S(J2-1,1).AN
476              2 D. S(J2-2,1).GT.0.69.AND.S(J2,1).LT.1.06.AND.S(J2,7).GT.0.84.AND.
477              3 S(J2+1,7).GE.1.64)  BALL=.TRUE.
478               IF (BALL)  J2=J2-1
479               IF (BALL)  V3=18
480               IF (BALL)  GO TO 300
481      C
482      C      *** 19 ***
483               BALL=.FALSE.
484               IF ((J2+3).GT.NN)  GO TO 260
485               T1=0
486               T2=0
487               T5=0
488               T1=S(J2,1)+S(J2+1,1)+S(J2+2,1)+S(J2+3,1)
489               T2=S(J2,2)+S(J2+1,2)+S(J2+2,2)+S(J2+3,2)
490               T5=S(J2,5)+S(J2+1,5)+S(J2+2,5)+S(J2+3,5)
491               PRINT 255,T1,T2,T5
492          255  FORMAT(' ',30X,'T1,T2,T5',3(F7.3),'    STEP 19 , J2-1 ,MOJ2')
493               IF (T5.GT.T1.AND.T5.GT.T2.AND.S(J2-1,9).GE.1.57.AND.S(J2-1,1).GT.
494              1 1.08.AND.S(J2,9).LT.1.10.AND.S(J2,2).GT.1.38)  BALL=.TRUE.
495      C
496               IF (T5.GT.T1.AND.T5.GT.T2.AND.S(J2,1).LE.0.69.AND.S(J2-1,1).GT.1.1
497              1 6.AND.S(J2-1,9).GT.1.10)  BALL=.TRUE.
498      C
499               IF (T5.GT.T1.AND.T5.GT.T2.AND.S(J2,9).LT.1.10.AND.S(J2,1).LT.1.01
500              1 .AND.S(J2-1,9).GE.1.10.AND.S(J2-1,1).GT.1.01)  BALL=.TRUE.
```

```
501               IF (BALL)  J2=J2-1
502               IF (BALL)  V3=19
503               IF (BALL)  GO TO 300
504     C
505     C      *** 20 ***
506               BALL=.FALSE.
507               IF (S(J2,9).LT.1.10.AND.S(J2,2).GT.1.38.AND.S(J2+1,2).GT.1.38.AND.
508              1 S(J2+2,2).GT.1.38.AND.S(J2+3,2).GT.0.75.AND.S(J2-1,9).GE.1.57.AND
509              2 .S(J2-2,9).GT.1.10)  BALL=.TRUE.
510               IF (BALL)  J2=J2-1
511               IF (BALL)  V3=20
512               IF (BALL)  GO TO 300
513     C
514     C      *** 21 ***
515         260   BALL=.FALSE.
516               IF ((J2+1).GT.NN)  GO TO 270
517               IF (S(J2,1).LE.0.69.AND.M(J2+1).EQ.15.AND.S(J2-1,1).GT.1.01.AND.S(
518              1 J2-1,9).GT.1.10.AND.S(J2-2,9).GE.1.08)  BALL=.TRUE.
519               IF (BALL)  J2=J2-1
520               IF (BALL)  V3=21
521               IF (BALL)  GO TO 300
522     C
523     C      *** 22 ***
524               BALL=.FALSE.
525               IF ((J2+3).GT.NN)  GO TO 270
526               IF (S(J2,9).LT.1.10.AND.S(J2-1,9).GT.1.10.AND.S(J2-1,1).GT.1.01.AN
527              1 D.S(J2,2).GT.1.38.AND.S(J2+1,2).GT.1.38.AND.S(J2+2,2).GT.0.93.AND
528              2 .S(J2+3,2).GT.1.10.AND.S(J2-2,2).GT.1.38)  BALL=.TRUE.
529               IF (BALL)  J2=J2-1
530               IF (BALL)  V3=22
531               IF (BALL)  GO TO 300
532     C
533     C
534     C ..... J2 = J2+1 .....
535     C ......................
536     C
537     C      *** 23 ***
538               BALL=.FALSE.
539               IF ((J2+4).GT.NN)  GO TO 270
540               IF (S(J2,9).LT.1.10.AND.S(J2+1,9).GT.1.10.AND.S(J2+1,1).GT.1.01.AN
541              1 D.S(J2,1).GT.1.01.AND.S(J2-1,1).GT.1.13.AND.S(J2-1,2).LE.0.75.AND
542              2 .S(J2-2,2).LE.0.75.AND.(P(J2+1,1)*P(J2+2,2)*P(J2+3,3)*P(J2+4,4))
543              3 .GT.0.000100.AND.S(J2-2,1).GT.1.13)  BALL=.TRUE.
544               IF (BALL)  J2=J2+1
545               IF (BALL)  V3=23
546               IF (BALL)  GO TO 300
547     C
548     C      *** 24 ***
549               BALL=.FALSE.
550               IF ((J2+5).GT.NN)  GO TO 270
```

```
      T1=0
      T2=0
      T5=0
      T1=S(J2+2,1)+S(J2+3,1)+S(J2+4,1)+S(J2+5,1)
      T2=S(J2+2,2)+S(J2+3,2)+S(J2+4,2)+S(J2+5,2)
      T5=S(J2+2,5)+S(J2+3,5)+S(J2+4,5)+S(J2+5,5)
      PRINT 265,T1,T2,T5
265   FORMAT(' ',30X,'T1,T2,T5',3(F7.3),'    STEP 24, J2+1 ,MOJ2')
      IF (T5.GT.T1.AND.T5.GT.T2      .AND.S(J2,9).GT.1.57.AND.S(J2+1,9).
     1 GT.1.20.AND.S(J2+1,1).GT.1.01)     BALL=.TRUE.
      IF (BALL)  J2=J2+1
      IF (BALL)  V3=24
      IF (BALL)  GO TO 300
C
C
C
C     TO CALL SUBROUTINE RMJ2 TO KEEP ON CHECKING FOR C-TERMINAL ADJUST
C     MENT,RMJ2 IS A CONTINUATION OF THIS SUBROUTINE
C
270   CALL RMJ2
      RETURN
C
C
C     THE C-TERMINAL HAS BEEN ADJUSTED ACCORDING TO ONE OF THE SITUATIONS
C     MENTIONED ABOVE. TO PRINT OUT THE FINAL VALUES FOR J1,J2 AND TO RE
C     TURN TO SUBROUTINE ONE TO START THE WHOLE PROCEDURE AGAIN
C
300   K3=J2
      PRINT 301,J1,J2,V2,V3
301   FORMAT('0',20X,'EVENTUAL HELIX FROM J1:',I5,5X,'TO J2:',I5,14X,
     1  ' ***  V2,V3:',2(I5),'  ***'//)
      RETURN
      END
End of File
```

```
C
C
      SUBROUTINE  RMJ2
C     .................
C
C
C     ..................................................
C     .                                                .
C     .          RMJ2 = REMAINING OF MOVE OF J2        .
C     .                                                .
C     ..................................................
C
C
C     ..................................................................
C
C     PURPOSE
C        TO KEEP ON CHECKING FOR OTHER POSSIBILITIES OF ADJUSTING THE C-
C        BOUNDARY OF THE PREDICTED HELIX
C
C     ..................................................................
C
C
C      REMARK
C        ALL THE PARAMETERS STILL HAVE THE SAME DEFINITION AS IN THE PRE
C        VIOUS SUBROUTINES
C
C
       REAL S,T1,T2,A1,A2 ,T3,T4,T5,TT,P
       INTEGER G,F,H,U,D,V1,V2,W, V3,V4,V5,V6,V7,V8,Q
       LOGICAL HELLO,BYE  ,BALL,MOVE
       DIMENSION S(1000,20),M(1000),H(1000),D(1000,16),P(1000,10)
       COMMON S,T1,T2,T3,T4,T5,TT,A1,A2,P,F,H,U,D,W,M,M1,M2,M3,M4,M5,M6,
      1L,I,K,L1,L2,NZ,NY,JA,JB,JC,JD,J1,J2,KM,N1,N2,NN,J,G,K3,V1,V2,V3,V4
      2,V5,V6,V7,Q,HELLO,BYE,BALL,MOVE
C
C
C ..... J2 = J2-2 .....
C .....................
C
C     *** 25 ***
       BALL=.FALSE.
       IF ((J2+2).GT.NN)  GO TO 20
       T1=0
       T2=0
       T5=0
       T1=S(J2-1,1)+S(J2,1)+S(J2+1,1)+S(J2+2,1)
       T2=S(J2-1,2)+S(J2,2)+S(J2+1,2)+S(J2+2,2)
       T5=S(J2-1,5)+S(J2,5)+S(J2+1,5)+S(J2+2,5)
       PRINT   5,T1,T2,T5
     5 FORMAT(' ',30X,'T1,T2,T5',3(F7.3),'    STEP 25, J2-2 , RMJ2')
```

```
      IF (T5.GT.T1.AND.T5.GT.T2.AND.S(J2-2,9).GE.1.57.AND.S(J2-2,1).GT.
     1 1.08.AND.((S(J2-1,1).LT.S(J2-2,1).AND.S(J2-1,9).LT.1.10).OR.S(J2-
     2 1,9).LE.S(J2-2,9)).AND.S(J2-1,5).GT.1.19.AND.S(J2+2,5).GT.1.43.AN
     3 D.S(J2+3,1).LT.1.06.AND.S(J2+4,1).LT.1.06)  BALL=.TRUE.
      IF (T5.GT.T1.AND.T5.GT.T2.AND.S(J2-1,9).LT.0.98.AND.S(J2-2,9).GT.
     1 1.10.AND.S(J2-2,1).GT.S(J2-1,1).AND.S(J2-3,9).GT.1.10.AND.S(J2,9)
     2  .LT.0.98)  BALL=.TRUE.
      IF ((J2+3).GT.NN)  GO TO 10
      IF (T1.LT.T2.AND.S(J2,2).GT.1.37.AND.S(J2-1,2).GT.1.37.AND.S(J2+1,
     1 2).GT.1.38.AND.S(J2-2,1).GT.1.16.AND.S(J2+3,1).LE.0.69)  BALL=
     3 .TRUE.
10    IF (BALL)  J2=J2-2
      IF (BALL)  V3=25
      IF (BALL)  GO TO 300
C
C     *** 26 ***
      BALL=.FALSE.
      IF (T5.GT.T1.AND.T5.GT.T2.AND.S(J2-3,9).GT.1.10.AND.S(J2-3,1).GT.
     1 1.16.AND.S(J2-2,9).LT.1.10.AND.S(J2-1,9).LT.S(J2-3,9).AND.S(J2-1,
     2 1).LT.1.01.AND.S(J2,5).GT.1.52.AND.S(J2+1,5).GT.1.46.AND.S(J2-1,5
     3 ).GT.1.14)  BALL=.TRUE.
      IF (BALL)  J2=J2-3
      IF (BALL)  V3=26
      IF (BALL)  GO TO 300
C
C     *** 27 ***
      BALL=.FALSE.
      IF ((J2+3).GT.NN)  GO TO 20
      IF (S(J2-2,9).GT.S(J2,9) .AND.S(J2-2,9).GT.S(J2-3,9).AND.S(J2-1,2)
     1 .GT.1.38.AND.S(J2+1,2).GT.1.38.AND.S(J2+2,2).GT.1.38.AND.S(J2+3,2
     2 ).GT.1.38.AND.S(J2-2,2).LE.0.75) BALL=.TRUE.
      IF (BALL)  J2=J2-2
      IF (BALL)  V3=27
      IF (BALL)  GO TO 300
C
C     *** 28 ***
      BALL=.FALSE.
      T1=0
      T2=0
      T1=S(J2-1,1)+S(J2,1)+S(J2+1,1)+S(J2+2,1)+S(J2+3,1)
      T2=S(J2-1,2)+S(J2,2)+S(J2+1,2)+S(J2+2,2)+S(J2+3,2)
      PRINT  15,T1,T2
15    FORMAT(' ',30X,'T1,T2   ',2(F7.3),7X,'    STEP 28, J2-2, RMJ2')
      IF (T2.GT.T1.AND.S(J2-1,2).GT.1.38.AND.S(J2-1,9).LT.1.10.AND.S(J2-
     1 2,9).GT.0.98.AND.S(J2-2,1).GT.1.01.AND.S(J2-3,9).LT.S(J2-2,9).AND
     2 .S(J2,2).GT.1.38)  BALL=.TRUE.
      IF (BALL)  J2=J2-2
      IF (BALL)  V3=28
      IF (BALL)  GO TO 300
C
```

```
      C      *** 29 ***
           20     BALL=.FALSE.
                  IF ((J2+1).GT.NN)  GO TO 60
                  IF (S(J2,9).LT.0.98.AND.M(J2+1).EQ.15.AND.S(J2-2,9).GT.1.57.AND.S(
                 1 J2-1,2).GT.1.38.AND.S(J2-3,9).LT.S(J2-2,9))  BALL=.TRUE.
                  IF (BALL)  J2=J2-2
                  IF (BALL)  V3=29
                  IF (BALL)  GO TO 300
      C
      C      *** 30 ***
                  BALL=.FALSE.
                  IF (S(J2-2,9).GE.1.10.AND.S(J2-1,2).GT.1.38.AND.S(J2-1,9).LT.1.10
                 1 .AND.M(J2+1).EQ.15.AND.S(J2,2).GT.0.75)  BALL=.TRUE.
                  IF (BALL)  J2=J2-2
                  IF (BALL)  V3=30
                  IF (BALL)  GO TO 300
      C
      C
      C
      C ..... J2 = J2+2 .....
      C .....................
      C
      C      *** 31 ***
                  BALL=.FALSE.
                  IF ((K+1).GT.KM)  GO TO 30
                  IF (S(J2,9).GT.1.25.AND.S(J2+2,9).GT.1.20.AND.S(J2+2,1).GT.1.01.A
                 1 ND.M(J2+1).NE.15.AND.S(J2-1,9).GE.S(J2,9).AND. (J2+3).GE.H(K+1))
                 2  BALL=.TRUE.
                  IF (BALL)  J2=J2+2
                  IF (BALL)  V3=31
                  IF (BALL)  GO TO 300
      C
      C      *** 32 ***
                  BALL=.FALSE.
                  IF (S(J2,9).GT.1.10.AND.S(J2+2,9).GT.1.57.AND.M(J2+1).NE.15.AND.S
                 1 (J2-1,9).LE.S(J2+2,9).AND.(J2+3).GE.H(K+1))  BALL=.TRUE.
                  IF (BALL)  J2=J2+2
                  IF (BALL)  V3=32
                  IF (BALL)  GO TO 300
      C
      C      *** 33 ***
           30     IF ((J2+6).GT.NN)  GO TO 40
                  BALL=.FALSE.
                  IF (S(J2,9).LT.1.25.AND.S(J2+4,9).GT.1.57.AND.S(J2+5,7).GT.1.49.AN
                 1 D.S(J2+1,1).GT.1.01.AND.S(J2+2,1).GT.1.16.AND.S(J2+3,1).GT.1.08
                 2 .AND.S(J2+3,1).LT.1.57.AND.S(J2+2,2).LT.0.87.AND.S(J2+6,2).LT.0.7
                 3 4.AND.S(J2-1,1).GT.1.16)  BALL=.TRUE.
                  IF (BALL)  J2=J2+4
                  IF (BALL)  V3=33
                  IF (BALL)  GO TO 300
```

```
C
C     *** 34 ***
   40     BALL=.FALSE.
          IF ((K+1).GT.KM)  GO TO 50
          IF (S(J2,9).GT.0.98.AND.S(J2,1).GT.1.01.AND.S(J2+2,9).GT.S(J2,9)
         1 .AND.S(J2+2,1).GT.1.16.AND.M(J2+1).NE.15.AND.(J2+3).GE.H(K+1))
         2  BALL=.TRUE.
          IF (BALL)  J2=J2+2
          IF (BALL)  V3=34
          IF (BALL)  GO TO 300
C
C     *** 35 ***
   50     BALL=.FALSE.
          IF ((J2+4).GT.NN)  GO TO 60
          IF (S(J2,9).GE.1.57.AND.S(J2+2,9).GT.1.10.AND.S(J2+2,1).GT.1.16.AN
         1 D.S(J2+1,1).GT.1.01.AND.S(J2+1,9).GT.0.98.AND.S(J2+3,9).LT.1.10.
         2 AND.S(J2+4,1).LE.0.69)  BALL=.TRUE.
          IF (BALL)  J2=J2+2
          IF (BALL)  V3=35
          IF (BALL)  GO TO 300
C
C     *** 36 ***
          BALL=.FALSE.
          IF ((J2+6).GT.NN)  GO TO 60
          IF (S(J2,9).GT.1.10.AND.S(J2+2,9).GT.1.10.AND.S(J2+2,1).GT.1.16.AN
         1 D.M(J2+1).NE.15.AND.(P(J2+3,1)*P(J2+4,2)*P(J2+5,3)*P(J2+6,4)).GT.
         2 0.000100.AND.S(J2-1,9).LT.1.10.AND.S(J2-2,9).GE.1.57) BALL=.TRUE.
          IF (BALL)  J2=J2+2
          IF (BALL)  V3=36
          IF (BALL)  GO TO 300
C.
C     *** 37 ***
          BALL=.FALSE.
          IF (S(J2,9).GE.1.10.AND.S(J2+2,9).GE.S(J2,9).AND.M(J2+1).NE.15.AND
         1 .S(J2-1,1).GT.1.06.AND.(P(J2+3,1)*P(J2+4,2)*P(J2+5,3)*P(J2+6,4))
         2 .GT.0.000100)  BALL=.TRUE.
          IF (BALL)  J2=J2+2
          IF (BALL)  V3=37
          IF (BALL)  GO TO 300
C
C     *** 38 ***
          BALL=.FALSE.
          IF ((P(J2+3,1)*P(J2+4,2)*P(J2+5,3)*P(J2+6,4)).GT.0.00007500.AND.M(
         1 J2+1).NE.15.AND.S(J2,9).GT.0.98.AND.S(J2+2,9).GT.1.24.AND.S(J2+2,
         2 1).GT.1.01.AND.S(J2-1,9).GT.1.57.AND.S(J2+3,9).LT.1.10)  BALL=
         3  .TRUE.
          IF (BALL)  J2=J2+2
          IF (BALL)  V3=38
          IF (BALL)  GO TO 300
C
```

```
C
C ..... J2 = J2-3 .....
C ......................
C
C     *** 39 ***
   60     BALL=.FALSE.
          IF (S(J2,9).LT.1.10.AND.M(J2-2).EQ.15.AND.S(J2-3,9).GT.0.98.AND.S(
         1 J2-4,9).GT.0.98.AND.S(J2+1,9).LT.1.77.AND.S(J2-3,1).GT.1.16)
         2 BALL=.TRUE.
          IF (BALL)  J2=J2-3
          IF (BALL)  V3=39
          IF (BALL)  GO TO 300
C
C     *** 40 ***
          BALL=.FALSE.
          IF ((J2+1).GT.NN)  GO TO 90
          IF (S(J2,2).GT.1.19 .AND.S(J2-1,2).GT.1.19.AND.S(J2-3,2).GT.1.38.A
         1 ND.S(J2+1,2).GT.1.38.AND.S(J2,9).LT.1.24.AND.S(J2-3,9).GT.1.57.AN
         2 D.S(J2-4,9).GT.1.10)  BALL=.TRUE.
          IF (BALL)  J2=J2-3
          IF (BALL)  V3=40
          IF (BALL)  GO TO 300
C
C     *** 41 ***
          BALL=.FALSE.
          T1=0
          T2=0
          T5=0
          T1=S(J2-2,1)+S(J2-1,1)+S(J2,1)+S(J2+1,1)
          T2=S(J2-2,2)+S(J2-1,2)+S(J2,2)+S(J2+1,2)
          T5=S(J2-2,5)+S(J2-1,5)+S(J2,5)+S(J2+1,5)
          PRINT  65,T1,T2,T5
   65     FORMAT(' ',30X,'T1,T2,T5',3(F7.3),'    STEP 41, J2-3 , RMJ2')
          IF (T5.GT.T1.AND.T5.GT.T2.AND.S(J2-3,9).GT.0.98.AND.S(J2-3,1).GT.1
         1 .01.AND.S(J2-4,9).LT.1.77.AND.S(J2-5,9).LT.1.77.AND.S(J2-3,9).GT.
         2 S(J2-2,9).AND.S(J2+2,5).GT.0.96.AND.S(J2,5).GT.0.96.AND.S(J2-1,5)
         3 .GT.1.19)  BALL=.TRUE.
          IF (T5.GT.T1.AND.T5.GT.T2.AND.S(J2-3,9).GE.1.10.AND.S(J2-3,1).GT.
         1 1.13.AND.S(J2- 4,1).GT.0.69.AND.S(J2-5,1).GT.1.16.AND.M(J2+1).EQ.
         2 15.AND.S(J2-2,7).GE.1.64.AND.S(J2-1,7).GT.1.24.AND.S(J2,9).LT.1.
         3 10)  BALL=.TRUE.
          IF (BALL)  J2=J2-3
          IF (BALL)  V3=41
          IF (BALL)  GO TO 300
C
C     *** 42 ***
          BALL=.FALSE.
          IF ((J2+2).GT.NN)  GO TO 90
          IF (S(J2,2).GT.1.30.AND.S(J2-1,2).GT.1.30.AND.M(J2-2).EQ.1.AND.S(J
         1 2+1,2).GT.1.38.AND.M(J2+2).EQ.1.AND.S(J2-3,9).GT.1.10)BALL=.TRUE.
```

```
251               IF (BALL)  J2=J2-3
252               IF (BALL)  V3=42
253               IF (BALL)  GO TO 300
254      C
255      C
256      C ..... J2 = J2+3 .....
257      C ......................
258      C
259      C      *** 43 ***
260                BALL=.FALSE.
261                IF ((J2+4).GT.NN)  GO TO 80
262                IF (S(J2,9).LT.1.10.AND.S(J2-1,9).LT.1.10.AND.S(J2+1,9).LT.1.10.AN
263               1 D.S(J2+2,9).GT.0.98.AND.S(J2+3,9).GT.0.98.AND.S(J2+3,1).GT.1.16.
264               2 AND.S(J2+1,1).GT.0.69)  BALL=.TRUE.
265                IF (BALL)  J2=J2+3
266                IF (BALL)  V3=43
267                IF (BALL)  GO TO 300
268      C
269      C      *** 44 ***
270                BALL=.FALSE.
271                IF (S(J2,9).GT.1.25.AND.S(J2+3,9).GT.0.98.AND.S(J2+3,1).GT.1.16.AN
272               1 D.S(J2+1,9).LT.S(J2+3,9).AND.S(J2+2,9).LT.S(J2+3,9).AND.S(J2+4,7)
273               2 .GT.1.58.AND.S(J2+1,1).GT.0.67.AND.S(J2+2,1).GT.0.67) BALL=.TRUE.
274                IF (BALL)  J2=J2+3
275                IF (BALL)  V3=44
276                IF (BALL)  GO TO 300
277      C
278      C      *** 45 ***
279                BALL=.FALSE.
280                IF (S(J2,9).GT.1.20.AND.S(J2+3,9).GT.1.24.AND.S(J2+3,1).GT.1.01.AN
281               1 D.S(J2+4,7).GT.1.58.AND.S(J2+5,7).GT.1.58.AND.S(J2+1,1).LT.S(J2+3
282               2 ,1).AND.S(J2+2,9).LT.S(J2+3,9))  BALL=.TRUE.
283                IF (BALL)  J2=J2+3
284                IF (BALL)  V3=453
285                IF (BALL)  GO TO 300
286      C
287      C      *** 46 ***
288                BALL=.FALSE.
289                IF (S(J2,9).GT.0.98.AND.S(J2+3,9).GT.1.25.AND.S(J2+3,1).GT.1.16.AN
290               1 D.S(J2+2,9).LT.S(J2+3,9).AND.S(J2+1,9).LT.S(J2+3,9).AND.S(J2+4,7)
291               2 .GT.1.58.AND.S(J2-1,9).LT.S(J2+3,9))  BALL=.TRUE.
292                IF (BALL)  J2=J2+3
293                IF (BALL)  V3=46
294                IF (BALL)  GO TO 300
295      C
296      C      *** 47 ***
297                BALL=.FALSE.
298                IF ((K+1).GT.KM)  GO TO 70
299                IF (S(J2,9).LT.1.10.AND.S(J2+3,9).GT.1.57.AND.S(J2+2,9).GE.S(J2+3,
300               1 9).AND.S(J2+1,9).LT.1.10.AND.M(J2+1).NE.15.AND.S(J2+4,7).GT.0.96
```

```
301              2 .AND.S(J2-1,9).GT.0.98.AND.(J2+3).GE.H(K+1))  BALL=.TRUE.
302               IF (BALL)  J2=J2+3
303               IF (BALL)  V3=47
304               IF (BALL)  GO TO 300
305      C
306      C      *** 48 ***
307                 BALL=.FALSE.
308                 IF (S(J2,9).LT.1.25.AND.S(J2+3,9).GT.S(J2,9).AND.S(J2+3,1).GT.1.16
309                1 .AND.(J2+4).GE.H(K+1).AND.M(J2+1).NE.15.AND.S(J2+2,1).GT.0.69.AND
310                2 .S(J2-1,1).GT.1.01)  BALL=.TRUE.
311                 IF (BALL)  J2=J2+3
312                 IF (BALL)  V3=48
313                 IF (BALL)  GO TO 300
314      C
315      C      *** 49 ***
316           70    BALL=.FALSE.
317                 IF ((J2+5).GT.NN)  GO TO 80
318                 IF (S(J2,9).LT.1.25.AND.S(J2+3,9).GT.S(J2,9).AND.S(J2+3,1).GT.1.08
319                1 .AND.S(J2+2,9).LT.S(J2+3,9).AND.S(J2+1,1).LT.S(J2+3,1).AND.S(J2+4
320                2 ,1).LE.0.69.AND.S(J2+5,7).GT.1.58)  BALL=.TRUE.
321                 IF (BALL)  J2=J2+3
322                 IF (BALL)  V3=49
323                 IF (BALL)  GO TO 300
324      C
325      C      *** 50 ***
326                 BALL=.FALSE.
327                 IF ((K+1).GT.KM)  GO TO 80
328                 IF (S(J2,9).GT.1.25.AND.S(J2+1,1).GT.1.01.AND.S(J2+2,1).GT.1.06.AN
329                1 D.S(J2+3,1).GT.1.16.AND.S(J2+4,1).GT.1.13.AND.S(J2-1,1).GT.1.16.A
330                2 ND.S(J2-2,1).GT.1.01.AND.S(J2-3,1).GT.1.16.AND.(J2+5).GE.H(K+1))
331                3  BALL=.TRUE.
332                 IF (BALL)  J2=J2+3
333                 IF (BALL)  V3=50
334                 IF (BALL)  GO TO 300
335      C
336      C      *** 51 ***
337           80    BALL=.FALSE.
338                 IF ((J2+3).GT.NN)  GO TO 90
339                 IF (S(J2,9).GT.1.57.AND.S(J2+3,9).GE.S(J2,9).AND.S(J2+2,9).GE.S(J2
340                1 ,9).AND.S(J2+1,9).LT.1.57.AND.S(J2+4,9).LT.1.57.AND.M(J2+1).NE.15
341                2 .AND.S(J2+2,1).GT.1.16.AND.S(J2+3,1).GT.1.16)  BALL=.TRUE.
342                 IF (BALL)  J2=J2+3
343                 IF (BALL)  V3=51
344                 IF (BALL)  GO TO 300
345      C
346      C      *** 52 ***
347                 BALL=.FALSE.
348                 IF ((K+1).GT.KM)  GO TO 90
349                 IF (S(J2,9).LT.1.25.AND.S(J2+3,9).GT.S(J2,9).AND.S(J2+3,1).GT.1.08
350                1 .AND.S(J2+2,9).LT.S(J2+3,9).AND.S(J2+1,1).LT.S(J2+3,1).AND.S(J2-1
```

```
351              2 ,9).LT.S(J2+3,9).AND.(J2+3).GE.H(K+1))  BALL=.TRUE.
352               IF (BALL)  J2=J2+3
353               IF (BALL)  V3=52
354               IF (BALL)  GO TO 300
355      C
356      C
357      C ..... J2 = J2-4 .....
358      C .....................
359      C
360      C      *** 53 ***
361           90    BALL=.FALSE.
362                 IF ((J2-6).LE.0)  GO TO 100
363                 IF ((P(J2-4,1)*P(J2-3,2)*P(J2-2,3)*P(J2-1,4)) .GT.0.00007500.AND.
364                1 S(J2-1,9).LT.0.98.AND.S(J2-2,9).LT.0.98.AND.M(J2-5).EQ.12.AND.S(J
365                2 2-4,9).LT.1.57.AND.S(J2-6,9).LT.1.77)  BALL=.TRUE.
366                 IF (BALL)  J2=J2-5
367                 IF (BALL)  V3=53
368                 IF (BALL)  GO TO 300
369      C
370      C      *** 54 ***
371                 BALL=.FALSE.
372                 T1=0
373                 T2=0
374                 T5=0
375                 T1=S(J2-4,1)+S(J2-3,1)+S(J2-2,1)+S(J2-1,1)
376                 T2=S(J2-4,2)+S(J2-3,2)+S(J2-2,2)+S(J2-1,2)
377                 T5=S(J2-4,5)+S(J2-3,5)+S(J2-2,5)+S(J2-1,5)
378                 PRINT  95,T1,T2,T5
379           95    FORMAT(' ',30X,'T1,T2,T5',3(F7.3),'    STEP 54, J2-4 ,RMJ2')
380                 IF (T5.GT.T1.AND.T5.GT.T2.AND.S(J2-5,9).GT.0.96.AND.S(J2-5,1).GT.
381                1 1.01.AND.S(J2-6,9).LT.S(J2-5,9).AND.S(J2-6,1).LT.S(J2-5,1)) BALL
382                2  =.TRUE.
383                 IF (BALL)  J2=J2-5
384                 IF (BALL)  V3=54
385                 IF (BALL)  GO TO 300
386      C
387      C      *** 55 ***
388           100   IF ((J2-5).LE.0)  GO TO 110
389                 BALL=.FALSE.
390                 T1=0
391                 T2=0
392                 T5=0
393                 T1=S(J2-3,1)+S(J2-2,1)+S(J2-1,1)+S(J2,1)
394                 T2=S(J2-3,2)+S(J2-2,2)+S(J2-1,2)+S(J2,2)
395                 T5=S(J2-3,5)+S(J2-2,5)+S(J2-1,5)+S(J2,5)
396                 PRINT 105,T1,T2,T5
397           105   FORMAT(' ',30X,'T1,T2,T5',3(F7.3),'    STEP 55, J2-4 ,RMJ2')
398                 IF (T5.GT.T1.AND.T5.GT.T2      .AND.S(J2-3,1).LE.0.69.AND.S(J2-4,1)
399                1 .GT.1.16.AND.S(J2-4,9).GT.1.10.AND.S(J2-5,1).GT.1.01.AND.S(J2-5,9
400                2 ).LT.S(J2-4,9))  BALL=.TRUE.
```

```
C      *** 12 ***
           IF (T5.GT.T1.AND.T5.GT.T2.AND.S(J2-3,1).LE.0.69.AND.S(J2-4,1).GT.
          11.01.AND.S(J2-4,9).GT.0.96.AND.S(J2-5,9).GT.1.10)  BALL=.TRUE.
           IF (BALL)  J2=J2-4
           IF (BALL)  V3=55
           IF (BALL)  GO TO 300
C
C
C ..... J2 = J2-5 .....
C ......................
C
C      *** 56 ***
           IF ((J2-7).LE.0)  GO TO 110
           BALL=.FALSE.
           IF (S(J2,2).GT.1.47.AND.S(J2-1,2).GT.1.47.AND.S(J2-2,2).GT.1.37.AN
          1 D.S(J2-4,2).GT.1.47.AND.S(J2-3,2).GT.1.47.AND.S(J2-5,1).GT.1.11.A
          2 ND.S(J2-5,9).GT.1.01.AND.S(J2-5,2).LE.0.75.AND.S(J2-6,2).LE.0.75
          3 .AND.S(J2-7,1).GT.1.11.AND.S(J2+1,1).LE.0.69.AND.S(J2+2,1).LE.0.6
          4 9.AND.S(J2+3,1).LE.0.69)  BALL=.TRUE.
           IF (BALL)  J2=J2-5
           IF (BALL)  V3=56
           IF (BALL)  GO TO 300
C
C
C ..... J2 = J2+4 .....
C ......................
C
  110   IF ((J2+5).GT.NN)  GO TO 130
C
C      *** 57 ***
           BALL=.FALSE.
           IF (S(J2,9).LT.1.10.AND.S(J2,1).GT.0.69.AND.S(J2+4,9).GT.S(J2,9).A
          1 ND.S(J2+1,9).LT.S(J2+4,9).AND.S(J2+2,9).LT.S(J2+4,9)
          2  .AND.S(J2+5,9).LT.S(J2+4,9).AND.S(J2+1,1).GT.1.01.AND.S(J2+2,1)
          3 .GT.1.08 .AND.M(J2+3).NE.15.AND.S(J2-1,9).GT.1.10.AND.S(J2+3,9).
          4 LT.S(J2+4,9))  BALL=.TRUE.
           IF (BALL)  J2=J2+4
           IF (BALL)  V3=57
           IF (BALL)  GO TO 300
C
C      *** 58 ***
           BALL=.FALSE.
           IF (S(J2,9).LT.1.25.AND.S(J2+4,9).GT.S(J2,9).AND.S(J2+4,1).GT.1.16
          1 .AND.S(J2+5,7).GT.1.58.AND.S(J2+4,9).GT.S(J2+3,9).AND.S(J2+3,1)
          2 .GT.1.01.AND.S(J2+2,1).GT.0.69.AND.S(J2+1,1).GT.0.67.AND.S(J2-1,9
          3 ).GT.1.10)  BALL=.TRUE.
           IF (BALL)  J2=J2+4
           IF (BALL)  V3=58
           IF (BALL)  GO TO 300
C
```

```
C      *** 59 ***
           BALL=.FALSE.
           IF (S(J2,9).LT.1.10.AND.S(J2,1).GT.0.69.AND.S(J2+4,9).GT.1.08.AND.
          1 S(J2+4,1).GT.1.16.AND.S(J2+3,1).GT.0.69.AND.S(J2+2,1).GT.0.69.AND
          2.M(J2+1).NE.15.AND.S(J2+5,9).LT.S(J2+4,9).AND.S(J2+2,9).GT.0.98.AN
          3 D.S(J2+3,9).GT.0.98.AND.S(J2-1,9).GT.1.10)  BALL=.TRUE.
           IF (BALL)  J2=J2+4
           IF (BALL)  V3=59
           IF (BALL)  GO TO 300
C
C     *** 60 ***
           IF ((K+1).GT.KM)  GO TO 120
           BALL=.FALSE.
           T1=0
           T2=0
           T5=0
           TT=0
           T1=S(J2+1,1)+S(J2+2,1)+S(J2+3,1)+S(J2+4,1)
           T2=S(J2+1,2)+S(J2+2,2)+S(J2+3,2)+S(J2+4,2)
           T5=S(J2+1,5)+S(J2+2,5)+S(J2+3,5)+S(J2+4,5)
           TT=P(J2+1,1)*P(J2+2,2)*P(J2+3,3)*P(J2+4,4)
           PRINT 115,T1,T2,T5,TT
      115  FORMAT(' ',30X,'T1,T2,T5,TT',3(F7.3),F13.9,'  STEP 60, J2+4 ,RMJ2')
C
           IF((T5.LT.T1.OR.T5.LT.T2).AND.TT.LT.0.0000750O.AND.S(J2-1,9).GE.1.
          1 10.AND.S(J2,9).LT.1.10.AND.S(J2,1).GT.0.69.AND.S(J2+4,9).GT.1.08
          2.AND.S(J2+4,1).GT.1.16.AND.S(J2+3,1).GT.0.69.AND.S(J2+2,1).GT.1.16
          3 .AND.M(J2+1).NE.15.AND.(J2+5).GE.H(K+1))  BALL=.TRUE.
           IF (BALL)  J2=J2+4
           IF (BALL)  V3=60
           IF (BALL)  GO TO 300
C
C      *** 61 ***
           BALL=.FALSE.
           IF (S(J2,9).LT.1.25.AND.S(J2+4,9).GT.S(J2,9).AND.S(J2+4,1).GT.1.16
          1 .AND.S(J2+5,7).GT.1.24.AND.S(J2+1,1).GT.0.98.AND.S(J2+2,1).GT.1.0
          2 1.AND.S(J2+3,1).GT.1.01.AND.S(J2+2,7).LT.0.96.AND.S(J2+3,7).LT.0.
          3 96.AND.(J2+4).GE.H(K+1).AND.S(J2-1,9).GT.1.10)  BALL=.TRUE.
           IF (BALL)  J2=J2+4
           IF (BALL)  V3=61
           IF (BALL)  GO TO 300
C
C      *** 62 ***
      120  BALL=.FALSE.
           IF ((J2+6).GT.NN)  GO TO 130
           IF (S(J2,9).LT.1.25.AND.S(J2+4,9).GT.1.57.AND.S(J2+5,7).GT.1.49.AN
          1 D.S(J2+1,1).GT.1.01.AND.S(J2+2,9).GE.S(J2+3,9).AND.S(J2+3,1).LE.0
          2 .69.AND.M(J2+3).NE.15.AND.S(J2+6,7).GT.1.58)  BALL=.TRUE.
           IF (BALL)  J2=J2+4
           IF (BALL)  V3=62
```

```
501                 IF (BALL)  GO TO 300
502       C
503       C      *** 63 ***
504                 BALL=.FALSE.
505                 IF (S(J2,9).GE.1.57.AND.S(J2,1).GT.1.16.AND.M(J2+1).NE.15.AND.S(J2
506                1 +2,9).GE.S(J2,9).AND.S(J2+3,1).GT.1.16.AND.S(J2+4,1).GT.1.16.AND.
507                2 S(J2+4,9).GT.1.10.AND.S(J2+5,7).GT.1.58.AND.S(J2+6,1).LE.0.69)
508                3  BALL=.TRUE.
509                 IF (BALL)  J2=J2+4
510                 IF (BALL)  V3=63
511                 IF (BALL)  GO TO 300
512       C
513       C      *** 64 ***
514                 BALL=.FALSE.
515                 IF (S(J2,1).GT.1.13.AND.S(J2,9).GT.1.10.AND.S(J2-1,1).GT.1.16.AND.
516                1 S(J2-2,1).GT.1.16.AND.S(J2-4,2).LT.0.55.AND.S(J2+4,1).GT.1.16.AND
517                2 .S(J2+5,1).LE.0.69.AND.S(J2+6,1).LE.0.69.AND.S(J2+3,1).GT.1.16.AN
518                3 D.S(J2+2,1).GT.1.13.AND.S(J2+2,2).LT.0.75.AND.M(J2+1).NE.15)
519                4 BALL=.TRUE.
520                 IF (BALL)  J2=J2+4
521                 IF (BALL)  V3=64
522                 IF (BALL)  GO TO 300
523       C
524       C      *** 65 ***
525            130   BALL=.FALSE.
526                 IF ((J2+4).GT.NN)  GO TO 300
527                 IF (S(J2,9).LT.1.10.AND.S(J2+1,1).LE.0.69.AND.M(J2+1).NE.15.AND.S(
528                1 J2+2,1).LE.0.69.AND.M(J2+2).NE.15.AND.S(J2+3,9).GE.1.57.AND.S(J2+
529                2 4,1).GT.1.16.AND.S(J2+4,9).GT.0.98.AND.S(J2+3,1).GT.1.08.AND.S(J2
530                3 -1,1).GT.1.16)  BALL=.TRUE.
531                 IF (BALL)  J2=J2+4
532                 IF (BALL)  V3=65
533                 IF (BALL)  GO TO 300
534       C
535       C ..... J2 = J2+5 .....
536       C ......................
537       C
538                 BALL=.FALSE.
539                 IF ((J2+6).GT.NN)  GO TO 300
540       C
541       C      *** 66 ***
542                 IF (S(J2,9).GT.0.98.AND.S(J2-1,9).LE.S(J2,9).AND.S(J2-2,9).GE.S(J2
543                1 ,9).AND.S(J2,1).GT.1.16.AND.S(J2+5,9).GT.S(J2,9).AND.S(J2+5,1).GT
544                2 .1.01.AND.S(J2+2,1).GT.1.16.AND.S(J2+3,1).GT.1.01.AND.S(J2+6,7).G
545                3 E.1.58.AND.S(J2+4,2).LE.0.75)  BALL=.TRUE.
546                 IF (BALL)  J2=J2+5
547                 IF (BALL)  V3=66
548                 IF (BALL)  GO TO 300
549       C
550       C      *** 67 ***
```

```
551               IF ((J2+7).GT.NN)  GO TO 300
552               BALL=.FALSE.
553               IF (S(J2,9).GT.0.98.AND.S(J2+5,9).GT.1.57.AND.(J2+6).GE.H(K+1).AND
554              1 .M(J2+1).NE.15.AND.M(J2+2).NE.15.AND.S(J2-1,1).GT.1.16.AND.S(J2+3
555              2 ,1).GT.1.01.AND.S(J2+4,1).GT.1.08.AND.S(J2+6,1).GT.1.16.AND.S(J2+
556              3 7,1).GT.1.16)  BALL=.TRUE.
557               IF (BALL)  J2=J2+5
558               IF (BALL)  V3=67
559               IF (BALL)  GO TO 300
560     C
561     C      *** 68 ***
562               BALL=.FALSE.
563               IF (S(J2,9).GT.1.10.AND.S(J2,1).GT.1.16.AND.S(J2-2,1).GT.1.11.AND.
564              1 M(J2+1).EQ.M(J2+2).AND.S(J2+1,9).GT.1.24.AND.S(J2+5,1).GT.1.16.AN
565              2 D.S(J2+5,9).GT.1.10.AND.S(J2+6,7).GT.1.58.AND.S(J2+4,1).GT.1.01
566              3 .AND.S(J2+3,1).GT.0.77.AND.S(J2+7,2).GT.1.38)  BALL=.TRUE.
567               IF (BALL)  J2=J2+5
568               IF (BALL)  V3=68
569               IF (BALL)  GO TO 300
570     C
571     C ..... J2 = J2+6 .....
572     C ......................
573     C
574               BALL=.FALSE.
575               IF ((J2+7).GT.NN)  GO TO 300
576     C
577     C      *** 69 ***
578               IF (S(J2,1).GT.1.16.AND.S(J2,1).LT.1.25.AND.S(J2+6,1).GT.1.00.AND.
579              1 S(J2+2,9).GT.1.57.AND.S(J2+3,1).GT.1.16.AND.S(J2+4,9).GT.1.24.AND
580              2.S(J2+5,1).GT.0.69.AND.S(J2+6,2).LT.0.74.AND.S(J2+1,1).GT.0.69
581              3 .AND.S(J2+7,1).LE.0.69)  BALL=.TRUE.
582               IF (BALL)  J2=J2+6
583               IF (BALL)  V3=69
584               IF (BALL)  GO TO 300
585     C
586     C      *** 70 ***
587               BALL=.FALSE.
588               IF (S(J2,9).GT.0.98.AND.S(J2,1).GT.1.01.AND.S(J2+6,9).GT.S(J2,9)
589              1 .AND.S(J2+6,1).GT.1.08.AND.S(J2+7,1).LE.0.69.AND.S(J2+2,9).GT.1.2
590              2 0.AND.S(J2+3,1).GT.1.16.AND.S(J2+4,1).GT.1.16.AND.S(J2+1,1).GT.0.
591              3 67.AND.S(J2+5,1).GT.0.69.AND.S(J2+5,9).LT.S(J2+6,9).AND.S(J2-1,1)
592              4 .GT.1.01.AND.S(J2-2,1).GT.1.13)  BALL=.TRUE.
593               IF (BALL)  J2=J2+6
594               IF (BALL)  V3=70
595               IF (BALL)  GO TO 300
596     C
597     C
598     C
599     C      V3=80 WHEN THE C-TERMINAL ADJUSTMENT IS DUE TO STRONG B-TURN POTEN
600     C      TIAL (THROUGH THE PROCEDURE OF REPEATING THE B-TURN CHECK).
```

```
C     IF V3=0 NONE OF THE CONDITIONS LISTED IN THE SUBROUTINES MOJ2 AND
C     RMJ2 FIT THE CURRENTLY TESTED SEGMENT. IN OTHER WORDS J2 HAS NOT
C     CHANGED.
C
  300   K3=J2
        IF (.NOT.  BALL.AND. V3.NE. 80)  V3=0
C
C     TO PRINT OUT THE FINAL VALUES FOR J1,J2, TO RETURN TO SUBROUTINE
C     ONE TO START THE WHOLE PROCEDURE AGAIN.
C
        PRINT 301,J1,J2,V2,V3
  301   FORMAT('0',25X,'EVENTUAL HELIX FROM J1:',I5,5X,'TO J2:',I5,14X,
     1  ' ***  V2,V3:',2(I5),'  ***'//)
        RETURN
        END
End of File
```

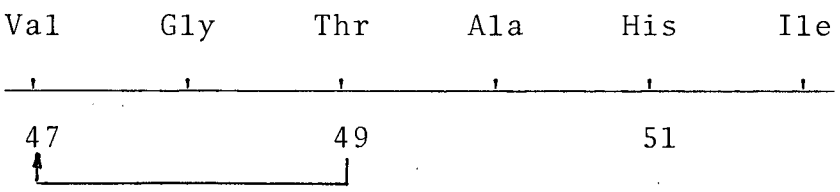
Efficiency of the β-sheet prediction

As indicated in the previous section strict adherence to Chou and Fasman's set of rules led to the missing of a certain number of regions and to some differences between the boundaries of predicted areas from this study and those of Chou and Fasman and X-ray analysis (Table 6). Analysis of the results obtained showed that the observations made for the helical search could also be applied for β-sheet. This involved moving the boundaries J1 and J2 to a more suitable position through consideration of the boundary conformational parameters $P_{\beta N}$, $P_{\beta C}$, $P_{n\beta N}$ and $P_{n\beta C}$ of the neighbouring residues.

Some examples of boundary adjustment for predicted β-sheet regions are listed below:

(1) J1 = J1 - 2:

Concanavalin: 49-57

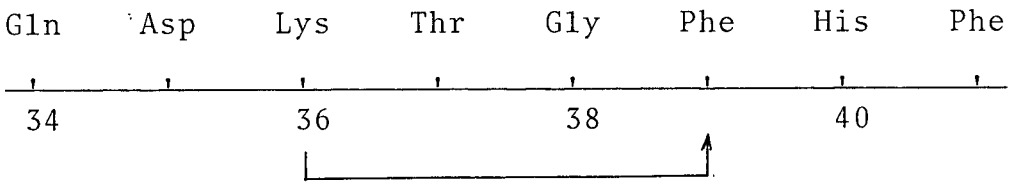


Val (47) is listed second for its $P_{\beta N}$ and it is a strong β-former. Hence, besides the fact that its presence balances the breaker Gly (48), it also ensures a very stable

N-boundary to the predicted -sheet.

(2) J1 = J1 +3

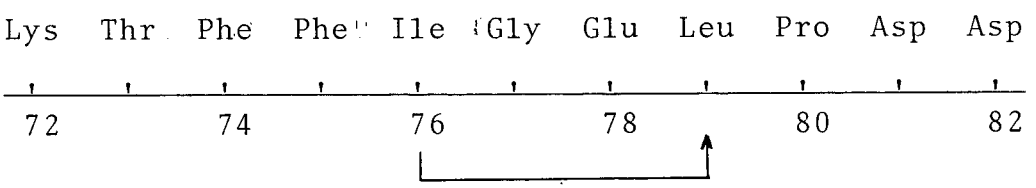
α-Chymotrypsin: 36-42



Besides the good $P_{\beta N}$ of the residue Phe (39), by moving J1 to position J1+3, Phe (39), two β-sheet breakers, Lys (36) and Gly (38) are avoided, as well as, the tetrapeptide 35-38 which exhibits β-turn potential.

(3) J2 = J2 + 3

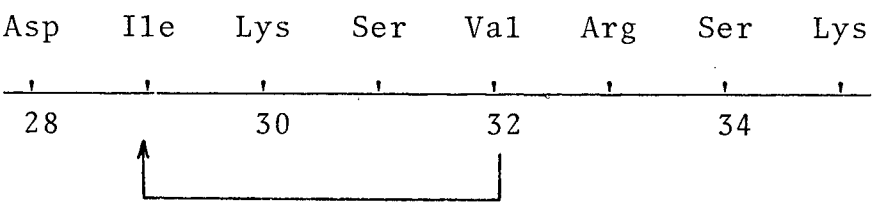
Cytochrome b_5: 73-76



The region 73-76 contains enough β-sheet formers to balance the addition of two breakers, Gly (77) and Glu (78). Leu has been ranked fifth for its $P_{\beta C}$ and the residues Pro (80), Asp (81) and Asp (82) possess good $P_{n\beta C}$.

(4) J2 = J2 - 3:

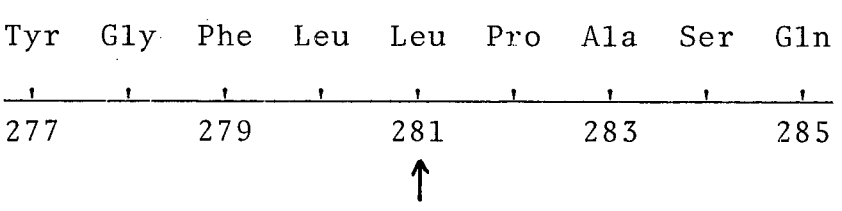
Concanavalin A: 25-32



Although Ile 29 has a lower $P_{\beta C}$ than Val 32, the new region 25-29 still has a stable C-boundary and is itself more stable because of elimination of the two breakers Lys 30, and Ser 31. In fact, region 25-32 has 4 breakers out of 8 residues.

(5) J2 = J2:

Carboxypeptidase A: 277-281



Considering its neighbouring residues, Leu 281 appears to be a good choice for the C-boundary since it is ranked fifth for its $P_{\beta C}$ and is a β-former. The residues Pro 282 and Ser 284 exhibit good $P_{n\beta C}$ which may favor the stabilization of the sheet C-terminal.

These boundary analyses for the β-sheet prediction were elaborated in two extra subroutines added to the end of the propagation procedure (subroutine FOUR deals with the N-boundary adjustment and subroutine FIVE with the C-boundary adjustment). Again, it was recognized that such analyses were quite tedious and did not always ensure completely satisfying results due to the complexity of protein arrangement.

The nucleation procedure was also subjected to some modifications to reduce the number of missing residues. In most cases, once an area with β-sheet potential was located, the nucleation search would start again from its (C-terminal + 1) residue to avoid repetition in the same area (cf. subroutine FIRS). However, for some proteins (e.g. bovine colostrum inhibitor, glucagon, Black Mamba Toxin K and Russell's Viper venom), such a procedure resulted in the omission of some regions (Table 6, p. 205).

The problem was solved by starting the search again every time from the (N-terminal + 1) residue of the previous fragment. The major drawback of such a procedure was the tedious repetition of the search for high molecular weight proteins. The proteins for which the new procedure improved the quality of the prediction were: bovine colostrum inhibitor, glucagon, Black Mamba Toxin K and Russell's

Viper Venom. These proteins have molecular weights of 7,511, 3,483, 6,566 and 6,850, respectively, which are lower than those of other proteins used in this study. Hence it is possible that in low-molecular weight proteins, short range interactions between adjacent residues may lead to the formation of β-sheets under circumstances not encountered in bigger proteins.

The requirement of less than one third β-sheet breakers may sometimes provoke a section or a protein with 3 h's out of 5 residues to be ignored in the nucleation search (e.g. α-chymotrypsin 197-201, ribonuclease 116-124). In addition , the presence of Pro 198 and Pro 117 in α-chymotrypsin and ribonuclease, respectively, is unfavorable to β-sheet nucleation according to rule B.1. Therefore, in the modified program, two distinct decisions were made:

(1) for α-chymotrypsin, once segment 197-201 has been considered as a possible β-sheet, boundary analysis allows shifting of the entire fragment to the right. The final value 199-204, besides having the advantage of being closer to X-ray results (199-203), also has better $P_{\beta N}$ (Leu) and $P_{\beta C}$ (Asn).

(2) for ribonuclease, instead of shifting fragment 116-120, the addition of an extra tetrapeptide (121-124) with an acceptable β-sheet potential makes the presence

of Pro 117 less unfavorable to β-sheet conformation and it eventually leads to the prediction of a β-sheet area with very favorable $P_{\beta N}$(Val) and $P_{\beta C}$(Val).

As Lys did not occur often at the N-terminal of a β-sheet section, the change of its assignment from a β-sheet breaker to a β-sheet former could not readily be made because of the possible result of erroneous predictions. However, in the case of ribonuclease 61-65 it was necessary to have Lys 61 equivalent to a h_β ($P_{\beta N(Lys)}$ = 1.00) so that this section did not violate the requirement of two thirds h's.

The presence of a strong B_β such as Asp could interrupt the preliminary search of areas with β-sheet potential (e.g. papain 4-9). This disruption resulted in an inability to start any nucleation procedure on the two fragments arising from this disruption (papain 3-6, 7-10). These two fragments could not by themselves meet the requirement of two thirds h's. Hence in such conditions the nucleation rule may be slightly modified so that eventually, with the combination of boundary analysis, one could still locate an appropriate β-sheet area. A somewhat similar situation was encountered with subtilisin 44-51. The two adjacent fragments 42-45 and 49-52 could not be the starting point for β-sheet formation. They were separated by a section with quite low β-sheet

potential (Gly-Gly-Ala: $\langle P_\beta \rangle = 0.76$). Nevertheless the entire section 44-51 was detected as β-sheet by Chou and Fasman (1974b) and by X-ray diffraction (Chou and Fasman, 1974b). It also has good end residues, Val (44) and Val (51), and $\langle P_\beta \rangle$ is greater than $\langle P_\alpha \rangle$ (1.045 versus 1.040).

In summary, by taking into account the important contribution of the boundary conformational parameters (subroutines FOUR and FIVE) and the necessity of allowing more flexibility to the nucleation rule (subroutine SECO) under the specific conditions previously mentioned, the following program was adopted for the β-sheet search. Only the different subroutines are presented here since the main program for β-sheet prediction is identical to the one used for α-helix search, except that the β-sheet boundary conformational parameters replace those pertaining to α-helix characterization.

```
C
C
      SUBROUTINE FIRS
C     .................
C
C
C      ..........................................................
C      .                                                        .
C      .        PRELIMINARY SEARCH FOR B-SHEET REGIONS          .
C      .                                                        .
C      ..........................................................
C
C
C     ....................................................................
C
C     PURPOSE
C        PRELIMINARY SEARCH FOR B-SHEET REGIONS BY APPLYING RULE 2: <PB>
C        > 1.05  AND <PA> < <PB>
C
C     ....................................................................
C
C
      REAL S,T1,T2,A1,A2 ,T3,T4,T5,TT,P
      INTEGER G,F,H,U,D,V1,V2,V3,V4,V5,V6,V7,V8,Q
      LOGICAL HELLO,BYE,BALL,MOVE
      DIMENSION S(1000,10),M(1000),H(1000),D(1000,16),P(1000,10)
      COMMON S,T1,T2,T3,T4,T5,TT,A1,A2,P ,V4,V5,V6,V7,V8,Q,G,F,H,U,D,NN,
     1NW,KX,MA,MB,MC,MD,L,I,L1,L2,L3,J1,J2,N,K1,K2,V1,V2,IM,M,K3,K4,V3,
     2BYE,BALL,HELLO,MOVE
C
C
C     DESCRIPTION OF PARAMETERS
C        H    - BOUNDARY RESIDUES OF A PREDICTED REGION
C               H(I)  - N-TERMINAL RESIDUE
C               H(I+1)- C-TERMINAL RESIDUE
C        MB   - FIRST RESIDUE OF A SECTION TO BE CONSIDERED FOR THE PRELI
C               MINARY SEARCH BUT WILL CHANGE DURING N-PROPAGATION (MB-1)
C        MA   - FIRST RESIDUE OF A SECTION TO BE CONSIDERED FOR THE PRELI
C               MINARY SEARCH BUT WILL CHANGE DURING C-PROPAGATION (MA+1)
C        K1   - FIRST RESIDUE OF A SECTION TO BE CONSIDERED FOR THE PRELI
C               MINARY SEARCH
C        K2   - LAST RESIDUE OF A SECTION  TO BE CONSIDERED FOR THE PRELI
C               MINARY SEARCH
C        A1   - AVERAGE <PA>  OF A SECTION
C        A2   - AVERAGE <PB>  OF A SECTION
C        N    - SWITCHING VALUE FOR DECISION MAKING
C               N=1   N-PROPAGATION
C               N=2   C-PROPAGATION
C        I    - COUNTER USED WITH THE ARRAY H TO STORE THE BOUNDARY RESI
C               DUES OF PREDICTED REGIONS
```

```
      C
      C
      C     THE SEARCH WILL STOP WHEN THE LAST SEGMENT AT THE C-TERMINAL HAS
      C     ONLY 2 AMINO ACID RESIDUES. IT IS NOT LONG ENOUGH FOR THE B-SHEET
      C     STATE
      C
   10 I = 2
      H(I)  =0
      H(I-1)=0
      NW = NN-2
      MB = 1
      MA = 1
      LP = 1
   20 N  = 0
   25 K2 = MA+2
      HELLO=.FALSE.
      IF (MB.EQ.0)     HELLO=.TRUE.
      IF (HELLO)  K1=H(I)+1
      IF (.NOT.HELLO)  K1=MB
      C
      C     TO CALCULATE <PA>,<PB> FOR A POLYPEPTIDE CHAIN STARTING AT POSITION
      C     K1 AND ENDING AT POSITION K2
      C
      T1=0
      T2=0
      DO 30  MC=K1,K2
      T1=T1+S(MC,1)
      T2=T2+S(MC,2)
   30 CONTINUE
      A2=T2/(K2-K1+1)
      A1=T1/(K2-K1+1)
      C
      C     IF <PB> IS LESS THAN 1.05 THEN TO START THE SEARCH AGAIN FROM NEXT
      C     POSITION K1+1
      C
      IF (A2.LT.1.05-1.E-6)  GO TO 35
      C
      C     TO START THE SEARCH AGAIN FROM NEXT POSITION MB+1 WHEN <PB> < <PA>
      C     EVEN IF <PB> > 1.05. THE SEARCH IS STOPPED WHEN THE LAST AMINO ACID
      C     RESIDUE HAS BEEN REACHED
      C
      IF (A1.GT.A2.AND.K2.EQ.NN .AND.(K2+1-K1).EQ.3)  GO TO 70
      IF (A1.GT.A2.AND.K2.EQ.NN .AND.(K2+1-K1).GT.3)  GO TO 55
      IF (A1.GT.A2.AND.K2.NE.NN)  GO TO 35
      C
      C     IF <PB> > <PA> AND <PB> > 1.05 TO CONTINUE THE PROPAGATION AT EI
      C     THER N- OR C-TERMINAL SIDE (N=1 INDICATES N-TERMINAL PROPAGATION,
      C     N=2 C-TERMINAL PROPAGATION) UNLESS WE REACHED THE LAST RESIDUE OF
      C     THE SEQUENCE (NN)
      C
```

```
      IF (N.EQ.2 .AND. K2.EQ.NN)  GO TO 45
      IF (N.EQ.2 .AND. K2.NE.NN)  GO TO 40
C
C     TO START N-TERMINAL PROPAGATION WHEN <PB> > 1.05 AND <PB> > <PA>
C
      MB = MB-1
      N  = 1
C
C     AS LONG AS THE N-TERMINAL PROPAGATED PEPTIDE DOES NOT OVERLAP WITH
C     THE PREVIOUS SHEET THE PEPTIDE CAN BE ELONGATED ON THAT SIDE,OTHER
C     WISE TO SWITCH TO C-TERMINAL PROPAGATION
C
      BYE=.FALSE.
      IF (MB.EQ.H(I))  BYE=.TRUE.
      IF (BYE)  N=2
      IF (BYE)  GO TO 40
C
C     N-TERMINAL PROPAGATION IS STOPPED WHEN MB OR K1 = 1,TO SWITCH THEN
C     TO C-TERMINAL PROPAGATION
C
      BALL=.FALSE.
      IF (MB.LE.H(I-1))  BALL=.TRUE.
      IF (BALL)  MB=MB+1
      IF (BALL)  MA=MA+1
      IF (BALL)  N=2
      IF (MA.GT.NW)  GO TO 45
      IF (BALL)  GO TO 25
      IF (MB.GT.H(I-1))  GO TO 25
C
C     TO START C-TERMINAL PROPAGATION WHEN IT IS STOPPED AT THE N-TERMIN
C     AL SIDE. IF BOTH SIDES CANNOT BE ELONGATED ANYMORE THEN THE SEGMENT
C     BEING ANALYZED SO FAR IS RECOGNIZED AS HAVING SHEET POTENTIAL
C
35    MB=MB+1
      IF (N.EQ.2)  GO TO 55
      IF (N.EQ.1)  N=2
40    MA=MA+1
      IF (MA.LE.NW)  GO TO 25
      IF (MA.GT.NW)  GO TO 70
C
C     AFTER PRINTING OUT THE AREA WITH SHEET POTENTIAL THE SEARCH IS STOP
C     PED BECAUSE WE GOT TO THE LAST RESIDUE IN THE SEQUENCE
C
45    I=I+1
      H(I)=K1
      I=I+1
      H(I)=K2
      PRINT 50,H(I-1),H(I)
50    FORMAT('0',30X,I6,10X,I6)
      IM=I
```

```
      GO TO 70
C
C     TO PRINT OUT THE AREA WITH SHEET POTENTIAL (H(I-1),H(I))
C
55    I=I+1
      H(I)=K1
      I=I+1
      H(I)=K2-1
      PRINT 60,H(I-1),H(I)
60    FORMAT('0',30X,I6,10X,I6)
C
C     TO START THE SEARCH AGAIN EITHER FROM (H(I-1) + 1) OR (H(I) + 1)
C
      MB=H(I) + 1
      MA=H(I) + 1
      IM=I
      IF (MA.LE.NW)  GO TO 20
C
C     TO PRINT OUT THE LAST VALUE OF THE COUNTER I (IM) WHICH WILL BE
C     USED IN THE NEXT SUBROUTINE
C
70    PRINT  75,IM
75    FORMAT('0',40X,'IM:',I4)
C
      PRINT 90
90    FORMAT('-',12X,'SEARCH FOR ACTUAL SHEETS FROM THE POTENTIAL REGION
     1S')
      PRINT 95
95    FORMAT(' ',12X,'.....
     1.'//)
C
      I = 2
      Q = 1
      CALL SECO
      RETURN
      END
End of File
```

```
C
C
      SUBROUTINE SECO
C     .............
C
C
C      ..........................................
C      .                                        .
C      .        SEARCH FOR SHEET NUCLEATION     .
C      .                                        .
C      ..........................................
C
C
C     ..................................................................
C     PURPOSE
C        SEARCH FOR NUCLEATING REGIONS WHICH SHOULD CONTAIN THREE BETA-
C        FORMERS OUT OF FIVE RESIDUES
C     ..................................................................
C
      REAL S,T1,T2,A1,A2 ,T3,T4,T5,TT,P
      INTEGER G,F,H,U,D,V1,V2,V3,V4,V5,V6,V7,V8,Q
      LOGICAL HELLO,BYE,BALL,MOVE
      DIMENSION S(1000,10),M(1000),H(1000),D(1000,16),P(1000,10)
      COMMON S,T1,T2,T3,T4,T5,TT,A1,A2,P ,V4,V5,V6,V7,V8,Q,G,F,H,U,D,NN,
     1NW,KX,MA,MB,MC,MD,L,I,L1,L2,L3,J1,J2,N,K1,K2,V1,V2,IM,M,K3,K4,V3,
     2BYE,BALL,HELLO,MOVE
C
C
C     DESCRIPTION OF PARAMETERS
C        G    - FIRST RESIDUE OF THE 5 RESIDUE PEPTIDE SUBJECTED TO THE
C               NUCLEATION SEARCH
C        MA   - FIFTH RESIDUE OF THE 5 RESIDUE PEPTIDE SUBJECTED TO THE
C               NUCLEATION SEARCH
C        Q    - SWITCHING VALUE FOR DECISION MAKING
C               Q=1    THE CURRENT POTENTIAL AREA IS STILL LONG ENOUGH (>
C                      3 RESIDUES) TO BE SUBJECTED TO THE NUCLEATION SEARCH
C               Q=2    THE CURRENT POTENTIAL AREA IS TOO SHORT FOR ANOTHER
C                      SHEET SO TO START WITH THE NEXT POTENTIAL AREA
C
C     REMARKS
C     UNLESS NOTIFIED THE OTHER PARAMETERS STILL HAVE THE SAME DEFINITION
C
C     IF Q=2 THE NUCLEATION SEARCH WILL START ON A NEW POTENTIAL AREA SI
C     NCE THE PREVIOUS ONE HAD BEEN THOROUGHLY SCANNED THROUGH. EACH TI
C     ME THAT I INCREASES BY 1 THE NEXT POTENTIAL AREA WILL BE ANALYZED
C
   10 IF (Q.EQ.2)  GO TO 25
   20 I=I+1
      IF (I.GT.IM)  GO TO 180
      K1=H(I)
```

```
          I=I+1
          K2=H(I)
          G=K1
          KX=K2-3
   25     MA=G+4
          IF (MA.GT.K2)  MA=MA-1
C
C     THE RESIDUE ASN CAN BE CONSIDERED AS A B-FORMER AT THE C-TERMINAL
C     OF THE PEPTIDE CHAIN BECAUSE OF IT GOOD P.BC VALUE
C
          N = G+1+(MA-G)/2
          DO  30  L=N,MA
          IF (M(L).EQ.3)  S(L,2)=1.05
   30     CONTINUE
C
C     TO COUNT THE DIFFERENT TYPES OF ASSIGNMENTS (T3) AND THE NUMBER OF
C     BREAKERS (N) IN THE  SECTION G-MA
C
          T3=0
          N=0
          DO  35  L=G,MA
          S(L,3)=0
          IF (S(L,2).GE.1.05)  S(L,3)=1.0
          T3=T3+S(L,3)
          IF (S(L,2).LE.0.75)  N=N+1
   35     CONTINUE
          PRINT  36,G,MA,T3,N
   36     FORMAT(' ',10X,'G :',I4,5X,'MA:',I4,5X,'T3:',F7.4,5X,'N :',I3,
         1 8X,'SHEET NUCLEATION')
C
C
C     IF THERE IS AT LEAST 3 HB AND LESS THAN 2 BB,THE NUCLEATION RULE
C     IS SATISFIED. WE STILL HAVE TO CHECK FOR THE PRESENCE OF PRO OR
C     GLU IN THE NUCLEATING SEGMENT (THEY ARE STRONG B-BREAKERS)
C
          IF (T3.GE.3.0.AND.N.LT.2)  GO TO 60
C
C
C     SOME MODIFICATIONS OF THE RULE WHICH TAKE INTO ACCOUNT THE PRESEN
C     CE OF NEIGHBORING RESIDUES FAVORABLE TO SHEET NUCLEATION ALTHOUGH
C     THE SEGMENT MAY CONTAIN MORE THAN ONE THIRD OF SHEET-BREAKERS
C
C
          IF (T3.GE.3.0.AND.N.GE.2.AND.S(G,2).GE.1.05.AND.S(MA,9).GE.1.50
         1.AND.S(MA-1,9).GE.1.50)  GO TO 100
C
          IF (T3.GE.3.0.AND.N.GE.2.AND.M(MA).EQ.10.AND.M(G).EQ.10) GO TO 100
C
          IF (T3.GE.3.0.AND.N.GE.2.AND.M(MA).EQ.19.AND.S(G,2).GE.1.05.AND.
         1(M(G+1).EQ.20.OR.M(G+2).EQ.20.OR.M(G+3).EQ.20)) GO TO 100
```

```
C
      IF (T3.GE.3.0.AND.N.GE.2.AND.S(MA,2).LE.0.75)  GO TO 75
C
      IF (S(G,2).LE.0.75.AND.T3.GE.3.0.AND.N.EQ.2.AND.M(G+1).EQ.15.AND.
     1M(MA).EQ.5.AND.M(MA-1).EQ.20.AND.M(MA-2).EQ.11)  GO TO 100
C
      IF (T3.GE.3.0.AND.N.GE.2.AND.S(G,2).LE.0.75)  GO TO 90
C
      IF (T3.GE.2.0.AND.N.LT.2.AND.M(G).EQ.20.AND.M(G+2).EQ.14)  GO TO
     1  120
C
      IF (T3.GE.2.0.AND.N.EQ.1.AND.M(G+2).EQ.20.AND.M(G+4).EQ.5.AND.M(G)
     1.EQ.12)  GO TO 120
C
      IF (M(G).EQ.10.AND.M(MA+1).EQ.20.AND.S(G+1,2).GE.0.93.AND.S(G+2,2)
     1.GE.0.75.AND.S(G+3,2).GE.0.75.AND.M(G-1).EQ.1.AND.M(G-2).EQ.1)
     3  GO TO 130
C
      IF ((G-4).LE.0.AND.(MA+2).GT.NN)  GO TO 45
      IF (T3.GE.2.0.AND.N.EQ.1.AND.S(G-1,2).GE.0.54.AND.M(G-1).NE.15.AND
     1 .S(G-2,2).GE.1.60.AND.S(G-3,2).GE.1.47.AND.S(G-4,2).LE.0.74.AND.S
     2 (MA,2).LE.0.74.AND.S(MA-1,2).GT.0.93.AND.S(MA-3,2).GT.1.30.AND.S(
     3 MA+1,2).LE.0.75.AND.S(MA+2,2).LE.0.83)  GO TO 160
C
45    IF ((G+10).GT.NN.AND.(G-2).LE.0)  GO TO 50
      IF (T3.GE.2.0.AND.N.EQ.1.AND.S(G+1,2).LE.0 .74.AND.S(G+2,8).GE.1.6
     1 9.AND.S(G,8).LT.S(G+2,8).AND.M(G+3).EQ.1.AND.S(G+4,2).GT.0.74.AND
     2 .S(G+5,2).GT.0.74.AND.M(G+6).EQ.1.AND.S(G+7,2).GT.0.74.AND.S(G+8,
     3 2).GT.0.93.AND.S(G+9,2).GE.1.60. AND.S(G+10,2).LE.0.74.AND.S(G-1,
     4 2).LE.0.74.AND.S(G-2,2).LE.0.74)  GO TO 170
C
50    IF ((G-5).LE.0.AND.(MA+2).GT.NN)  GO TO 55
      IF(T2.GE.2.0.AND.N.EQ.2.AND.S(G-3,2).GE.1.60.AND.S(G-4,2).LE.0.75
     1 .AND.S(G-5,2).LE.0.75.AND.S(G-2,2).GE.1.60.AND.M(G-1).EQ.1.AND.S(
     2 G,2).GE.1.60.AND.S(G+1,2).GE.0.75.AND.M(G+2).EQ.1.AND.S(G+3,2).GE
     3 .1.60.AND.S(MA,2).LE.0.55.AND.S(MA+1,2).LE.0.75.AND.S(MA+2,2).LE.
     4 0.75)  GO TO 160
C
C
C     IF THE SEGMENT UNDER CONSIDERATION CANNOT SATISFY ANY OF THE ABOVE
C     CONDITIONS THEN THE SEARCH WILL START AGAIN FROM NEXT POSITION G+1
C
55    G=G+1
      IF (G.LE.KX)  GO TO 25
      GO TO 20
C
C     IF THERE IS NO GLU NOR PRO IN THE NUCLEATING SEGMENT THEN SUBROUTI
C     NE THIR IS CALLED TO CARRY OUT THE PROPAGATION PROCEDURE
C
60    DO  61  L=G,MA
```

```
151               IF (M(L).EQ.7 .OR .M(L).EQ.15)  GO TO 65
152          61   CONTINUE
153               CALL  THIR
154               GO TO 10
155     C
156     C     IN SOME INSTANCES,DESPITE THE PRESENCE OF PRO OR GLU THE NUCLEATING
157     C     AREA REMAINS STABLE BECAUSE OF STRONG B-FORMER RESIDUES
158     C
159          65   IF (T3.GE.3.0.AND.N.EQ.1.AND.S(G,2).GE.1.30.AND.M(G).EQ.M(G+2)
160               1.AND.M(G).EQ.M(G+3).AND.(G-2).EQ.K3.AND.S(G-1,2).GE.0.75)  GO TO
161               2  150
162     C
163               IF ((G+8).GT.NN)  GO TO 70
164               IF (T3.GE.3.0.AND.N.EQ.1.AND.S(G,8).GE.1.65.AND.S(G+1,2).GE.1.19.
165               1AND.S(G-1,2).LE.0.75.AND.S(G+2,2).GE.1.30.AND.S(G+4,9).GE.1.50.AND
166               2 .S(G+5,9).GT.0.79.AND.S(G+6,9).GT.1.79.AND.S(G+7,2).LE.0.75.AND.
167               3 S(G+8,2).LE.0.75)  GO TO 140
168     C
169     C
170     C     NUCLEATION SEARCH STARTS AGAIN FROM NEXT POSITION G+1
171     C
172          70   G=L+1
173               IF (G.LE.KX)  GO TO 25
174               GO TO 20
175     C
176     C
177     C     TO START N-TERMINAL PROPAGATION WHEN THE PRESENCE OF A SHEET-BREA
178     C     KER AT THE C-TERMINAL (MA) IMPEDES THE ELONGATION ON THAT SIDE
179     C
180          75   MV=MA-1
181               DO 76  L=G,MV
182               IF (M(L).EQ.7.OR.M(L).EQ.15)  GO TO 65
183          76   CONTINUE
184          80   BALL=.FALSE.
185               IF ((G-1).LE.K3)  GO TO 85
186               IF (S(G-1,2).GE.1.05)  BALL=.TRUE.
187               IF (BALL)  G=G-1
188               IF (BALL)  GO TO 80
189          85   PRINT 86,G,MA
190          86   FORMAT('0',10X,'PSEUDO-SHEET FROM G TO MA-1',5X,'G:',I5,5X,'MA:',
191               1I5/)
192               J1=G
193               J2=MA-1
194               GO TO 115
195     C
196     C
197     C     TO START C-TERMINAL PROPAGATION WHEN THE PRESENCE OF A SHEET-BREA
198     C     KER AT THE N-TERMINAL (G) IMPEDES THE ELONGATION ON THAT SIDE
199     C
200          90   MU=G+1
```

```
      DO 92  L=MU,MA
      IF (M(L).EQ.7.OR.M(L).EQ.15)  GO TO 65
92    CONTINUE
      NV=MA +1
      NU=MA+4
      DO 94  L=NV,NU
      IF (S(L,2).GE.1.05)  MA=MA+1
      IF (S(L,2).LT.1.05)  GO TO 95
94    CONTINUE
95    PRINT 96,G,MA
96    FORMAT('O',10X,'PSEUDO-SHEET FROM G+1 TO MA',5X,'G:',I5,5X,'MA:',
     1I5/)
      J1=G+1
      J2=MA
      GO TO 115
C
C
C     N- THEN C-TERMINAL PROPAGATION BY ADDING ONE RESIDUE AT A TIME TO
C     THE NUCLEATING SEGMENT. IT IS DIFFERENT FROM THE PROCEDURE IN SUB
C     ROUTINE THIR WHERE TETRAPEPTIDES INSTEAD OF SINGLE RESIDUES ARE CON
C     SIDERED FOR ELONGATING THE SEGMENT
C
100   BALL=.FALSE.
      IF ((G-1).LE.K3)  GO TO 110
      IF (S(G-1,2).GE.1.05)  BALL=.TRUE.
      IF (BALL)  G=G-1
      IF (BALL)  GO TO 100
C
110   HELLO=.FALSE.
      IF (S(MA+1,2).GE.1.05)  HELLO=.TRUE.
      IF (HELLO)  MA=MA+1
      IF (HELLO)  GO TO 110
      J1=G
      J2=MA
      PRINT 112,J1,J2
112   FORMAT('O',10X,'PSEUDO-SHEET FROM J1:',I5,5X,'TO  J2:',I5/)
      GO TO 115
C
C
C     WHEN THE PROPAGATION HAS BEEN STOPPED ON BOTH SIDES THEN SUBROUTINE
C     FOUR IS CALLED FOR ADJUSTING THE BOUNDARIES TO THEIR MOST FAVORABLE
C     POSITIONS. WHEN RETURNING FROM THE BOUNDARY ANALYSIS IF THE CURRENT
C     POTENTIAL AREA IS NOT LONG ENOUGH FOR ANOTHER SHEET FRAGMENT THEN
C     THE NEXT POTENTIAL AREA WILL BE ANALYZED (Q=1)
C
115   CALL FOUR
118   IF (J2.LT.KX)  G=J2+1
      IF (J2.LT.KX)  Q=2
      IF (J2.GE.KX)  Q=1
      GO TO 10
```

```
      C
      C
      C     TO PRINT OUT THE NUCLEATING SEGMENTS WHICH DO NOT FOLLOW THE COM
      C     MON NUCLEATION RULE. SUBROUTINE FOUR IS THEN CALLED TO CARRY OUT
      C     THE BOUNDARY ADJUSTMENT
      C
         120   J1 = G
               J2 = MA
               PRINT 125,J1,J2
         125   FORMAT('O',10X,'PSEUDO-SHEET FROM J1:',I5,5X,'TO  J2:',I5/)
               GO TO 115
      C
         130   J1=G-2
               J2=MA+1
               PRINT 125,J1,J2
               GO TO 115
      C
         140   J1=G
               J2=MA+3
               PRINT 125,J1,J2
               GO TO 115
      C
      C
      C     TO CHECK THE NUMBER OF B-BREAKERS (JC) WHICH SHOULD BE LESS THAN
      C     ONE THIRD OF THE LENGTH OF THE SEGMENT (JCC)
      C
         150   JC = 0
               J2 = MA
               DO 155  L=J1,MA
               IF (S(L,2).LT.0.83)  JC=JC+1
         155   CONTINUE
               JCC=(MA+1-J1)/3
               IF (JC.LE.2.AND.JC.LT.JCC)  PRINT 125,J1,J2
               GO TO 115
      C
         160   J1=G-3
               J2=MA-1
               PRINT 125,J1,J2
               GO TO 118
      C
         170   J1=G+2
               J2=MA+6
               PRINT 125,J1,J2
               GO TO 118
      C
         180   PRINT 185
         185   FORMAT('O','END OF PROGRAM')
               RETURN
               END
End of File
```

```
C
C
      SUBROUTINE THIR
      ...............
C
C
C
C     ....................................................
C     .                                                  .
C     .          PROPAGATION OF THE BETA-SHEET           .
C     .                                                  .
C     ....................................................
C
C
C     ..................................................................
C     PURPOSE
C        TO ADD TO THE NUCLEATING FRAGMENT TETRAPEPTIDES WHICH HAVE <PB>
C        > 1.00 AND WHICH SATISFY THE PROPAGATION SET OF RULES
C     ..................................................................
C
C
       REAL S,T1,T2,A1,A2 ,T3,T4,T5,TT,P
       INTEGER G,F,H,U,D,V1,V2,V3,V4,V5,V6,V7,V8,Q
       LOGICAL HELLO,BYE,BALL,MOVE
       DIMENSION S(1000,10),M(1000),H(1000),D(1000,16),P(1000,10)
       COMMON S,T1,T2,T3,T4,T5,TT,A1,A2,P ,V4,V5,V6,V7,V8,Q,G,F,H,U,D,NN,
      1NW,KX,MA,MB,MC,MD,L,I,L1,L2,L3,J1,J2,N,K1,K2,V1,V2,IM,M,K3,K4,V3,
      2BYE,BALL,HELLO,MOVE
C
C
C     DESCRIPTION OF PARAMETERS
C        MB   - WHETHER IT IS N- OR C-PROPAGATION JB WILL ALWAYS BE THE
C               FIRST LEFT RESIDUE OF THE ADJACENT TETRAPEPTIDE
C        MC   - WHETHER IT IS N- OR C-PROPAGATION JC WILL ALWAYS BE THE
C               FOURTH RESIDUE OF THE ADJACENT TETRAPEPTIDE
C        K1   - N-TERMINAL RESIDUE OF THE CURRENT POTENTIAL AREA
C        N2   - C-TERMINAL RESIDUE OF THE CURRENT POTENTIAL AREA
C        F    - SWITCHING VALUE FOR DECISION MAKING
C               F=1    N-PROPAGATION
C               F=2    C-PROPAGATION
C        V1   - COUNTER
C        V2   - COUNTER
C
C
C     AS LONG AS MB BELONGS TO THE CURRENT POTENTIAL AREA THE N-PROPAGA
C     TION CAN BE CARRIED OUT
C
   10     V1=0
          V2=0
          V3=0
          V4=0
```

```
      V7=0
      F =1
   15 V1=V1+1
      MB=G-(4*V1)
      IF (MB.GT.O .AND.MB.GE.K1)  GO TO  20
C
C     TO SWITCH TO C-TERMINAL PROPAGATION WHEN THE N-TERMINAL SIDE IS STOP
C     PED.THE VARIABLE MB THEN BECOMES THE FIRST RESIDUE IN THE TETRAPEP
C     TIDE ADDED TO THE C-TERMINAL SIDE
C
      IF (V1.EQ.1)  J1=G
      IF (V1.NE.1)  J1=G-4*(V1-1)
      F =2
   20 T2=0
      T1=0
      IF (F.EQ.1)  GO TO 25
      IF (V2.NE.O)  MB=MA+1+(4*V2)
      IF (V2.EQ.O)  MB=MA+1
      V2=V2 + 1
      IF (MB.GT.K2) GO TO 50
   25 MC=MB + 3
      IF (MC.GT.K2 .AND. MB.LE.K2)  GO TO 50
C
C     CALCULATION OF THE PA,PB OF THE TETRAPEPTIDE MB-MC
C
      DO 30  L=MB,MC
      T1=T1+S(L,1)
      T2=T2+S(L,2)
   30 CONTINUE
      PRINT  35,MB,MC,T1,T2
   35 FORMAT(' ',10X,'MB:',I4,5X,'MC:',I4,5X,'T1:',F7.4,5X,'T2:',F7.4,4X,
     1'SHEET PROPAGATION')
C
C     IF PA > PB THEN TO SWITCH TO C-PROPAGATION IF N-PROPAGATION HAS
C     BEEN CARRIED OUT,OTHERWISE TO START ELONGATING BOTH SIDES BY ONE
C     RESIDUE AT A TIME
C
      IF (T1.GT.T2)  GO TO 45
C
C     IF PB > PA AND <PB> <1.00 TO TAKE INTO CONSIDERATION THE TYPES OF
C     SHEET RESIDUES IN THE SEGMENT SINCE IT MAY STILL BE VALID FOR THE
C     PROPAGATION
C
      IF (T2.LT.4.0000)  GO TO 130
C
C     IF PRO OR GLU OCCURS IN THE PROPAGATED TETRAPEPTIDE THEN EITHER TO
C     SWITCH FROM N-PROPAGATION TO C-PROPAGATION (F=2) OR TO START ELON
C     GATING BOTH SIDES BY ONE RESIDUE AT A TIME
C
      DO 40   L=MB,MC
```

```
101               IF (M(L).EQ.15.OR.M(L).EQ.7)  GO TO 45
102          40   CONTINUE
103               GO TO 150
104          45   BALL=.FALSE.
105               IF (F.EQ.1)  BALL=.TRUE.
106               IF (BALL)  J1=MB+4
107               IF (BALL)  F=2
108               IF (BALL)  GO TO  20
109     C
110     C
111     C      ADDITION OF ONE RESIDUE (HB OR IB) AT A TIME TO THE N-TERMINAL SIDE
112     C      WHEN ADDING IB TO EACH END TO CHECK IMMEDIATELY WHETHER THE RULE
113     C      OF AT LEAST HALF OF FORMERS IS STILL SATISFIED OR NOT
114     C
115          50   L1=J1-1
116               IF (L1.LT.(G -4))  GO TO 55
117               IF (M(L1).EQ.12)  S(L1,2)=1.05
118               IF (S(L1,2).GE.1.05)  J1=L1
119               IF (S(L1,2).LE.0.93 .AND.S(L1,2).GE.0.83)  J1=L1
120               IF (S(L1,2).GE.1.05 )  GO TO 50
121     C
122     C      ADDITION OF ONE RESIDUE (HB OR IB) AT A TIME TO THE C-TERMINAL SIDE
123     C
124          55   J2=MB-1
125          60   L2=J2+1
126               IF (L2.GT.(MA+4)) GO TO 65
127               IF (M(L2).EQ.3)  S(L2,2)=1.05
128               IF (S(L2,2).GE.1.05)  J2=L2
129               IF (S(L2,2).LE.0.93 .AND.S(L2,2).GE.0.83)  J2=L2
130               IF (S(L2,2).GE.1.05)  GO TO 60
131     C
132     C      TO COUNT THE NUMBER OF SHEET-FORMERS IN THE ENTIRE SHEET AREA
133     C      TO COMPARE THE ACTUAL NUMBER OF FORMERS (T4) TO ITS THEORITICAL
134     C      ONE (TT : EQUAL TO AT LEAST ONE HALF OF THE SECTION)
135     C
136          65   T4=0
137               DO  70  L=J1,J2
138               S(L,4)=0
139               IF (S(L,2) .GE.1.05)  S(L,4)=1.0
140               T4=T4+S(L,4)
141          70   CONTINUE
142               TT = (J2-J1+1) / 2.0
143               PRINT 75,J1,J2,T4,TT
144          75   FORMAT(' ',10X,'J1:',I4,5X,'J2:',I4,5X,'T4:',F7.4,5X,'TT:',F7.4,
145              1 4X,'ACTUAL AND THEORIT. # FORMERS FROM J1 TO J2')
146     C
147     C      IF THE RULE OF MORE THAN HALF OF SHEET-FORMERS IS SATISFIED THEN
148     C      TO KEEP ON ADDING HB OR IB TO EACH SIDE OF THE SHEET SECTION
149     C
150               IF (T4.GE.TT .AND. S(J1-1,2).LE.0.75)  GO TO 80
```

```
151               IF (T4.GE.TT .AND. S(J1-1,2).GT.0.75.AND.L1.GE.(G -4)) GO TO 50
152          80   IF (T4.GE.TT .AND. S(J2+1,2).GT.0.75.AND.L2.LE.(MA+4)) GO TO 60
153               IF((T4.GE.TT .AND. S(J2+1,2).LE.0.75) .OR.(T4.GE.TT.AND.S(J2+1,2).
154              1GT.0.75.AND.L2.GT.(MA+4)))  GO TO 115
155     C
156     C      IF THE RULE IS NOT SATISFIED THEN TO TAKE AWAY RESIDUES WHICH ARE
157     C      NOT HB SO THAT EVENTUALLY THERE IS ENOUGH HB IN THE SECTION
158     C
159          85   IF (S(J2,2).LT.1.05)  GO TO 90
160               IF (S(J1,2).LT.1.05)  GO TO 95
161          90   J2=J2-1
162               IF (S(J2+1,2).LT.1.05)  GO TO 100
163          95   J1=J1+1
164               IF (S(J1-1,2).LT.1.05)  GO TO 100
165     C
166     C      EVERY TIME A RESIDUE IS TAKEN AWAY THE RULE OF MORE THAN HALF OF
167     C      SHEET-FORMERS IS CHECKED AGAIN ON THE SHORTENED SECTION
168     C
169          100  T4=0
170               DO 105  L=J1,J2
171               S(L,4)=0
172               IF (S(L,2).GE.1.05)  S(L,4)=1.0
173               T4=T4+S(L,4)
174          105  CONTINUE
175               TT=(J2-J1+1)/2.0
176               PRINT 75,J1,J2,T4,TT
177               IF (T4.GE.TT)  GO TO 115
178               IF (T4.LT.TT)  GO TO 85
179          115  PRINT 120,J1,J2
180          120  FORMAT('0',10X,'PSEUDO-SHEET FROM J1:',I5,5X,'TO  J2:',I5/)
181     C
182     C      WHEN THE PROPAGATION IS TERMINATED ON BOTH SIDES TO CALL SUBROUTINE
183     C      FOUR FOR THE BOUNDARY ADJUSTMENT
184     C
185          125  CALL FOUR
186               IF (J2.LT.KX)  G=J2+1
187               IF (J2.LT.KX)  Q=2
188               IF (J2.LT.KX)  RETURN
189               Q=1
190               RETURN
191     C
192     C
193     C      PRESENCE OF B-BREAKER OR OF CHARGED RESIDUE (ARG,LYS) IS NOT FAVO
194     C      RABLE TO PROPAGATED TETRAPEPTIDES WITH <PB> < 1.00. SO EITHER TO
195     C      SWITCH TO C-PROPAGATION OR TO START ADDING HB OR IB TO EACH SIDE
196     C      OF THE SHEET AREA
197     C
198          130  DO 135  L=MB,MC
199               IF (S(L,2).LE.0.75)  GO TO 45
200          135  CONTINUE
```

```
      DO 140  L=MB,MC
      IF (M(L).EQ.2 .OR. M(L).EQ.9)  GO TO 45
  140 CONTINUE
C
C     IF THE TETRAPEPTIDE WITH <PB> <1.00 ONLY HAS IB THEN IT CANNOT BE
C     ADDED TO THE PROPAGATED SHEET
C
      L3=0
      DO 145  L=MB,MC
      IF (M(L).EQ.3 .OR. M(L).EQ.1)  L3=L3+1
  145 CONTINUE
      IF (L3.EQ.4)  GO TO 45
C
C
C     ... TO CHECK THE NUMBER OF BREAKERS IN THE ENTIRE POLYPEPTIDE ...
C
C     TO COUNT THE NUMBER OF BB IN THE ENTIRE SECTION (V8). IT SHOULD NOT
C     BE GREATER THAN ONE THIRD OF THE LENGTH (V6). IF V8 IS LESS THAN V6
C     THEN THE SECTION IS CONSIDERED TO BE VALID AND SUBROUTINE FOUR IS
C     CALLED TO CARRY OUT THE BOUNDARY ADJUSTMENT. IF NOT EITHER HB OR IB IS
C     ADDED TO BOTH SIDES TO SATISFY THE REQUIREMENT OR C-PROPAGATION
C    .WILL REPLACE N-PROPAGATION
C
C     DESCRIPTION OF PARAMETERS
C        V3   - COUNTER
C        V4   - COUNTER
C        MB   - N-TERMINAL OF THE SHEET REGION
C        MD   - C-TERMINAL OF THE SHEET REGION
C
  150 V6=0
      V8=0
      IF (F.EQ.1)  GO TO 155
      MB=MB-5-(4*V4)
  155 MD=MB+8+(4*V3)
      V3=V3+1
      V4=V4+1
      V6=(MD-MB+1)/3
      DO 160  L=MB,MD
      IF (S(L,2).LE.0.75)  V8=V8+1
  160 CONTINUE
C
      PRINT 165,MB,MD,V6,V8
  165 FORMAT(' ',10X,'MB:',I4,5X,'MD:',I4,5X,'V6:',I7,5X,'V8:',I3,8X,
     1'THEORITIC. AND ACTUAL # BREAKERS FROM MB TO MD')
      IF (V8.LT.V6.AND.F.EQ.1)  GO TO 170
      IF (V8.LT.V6.AND.F.EQ.2)  GO TO 180
      V7=V2
      IF (F.EQ.2)  MB=MB+5+(4*V7)
      IF (F.EQ.2)  GO TO 50
      F=2
```

```
      J1=MB+4
      GO TO 20
C
C
C     TO PRINT OUT THE POSSIBLE SHEET AREAS THEN TO CALL SUBROUTINE FOUR
C     FOR THE BOUNDARY ADJUSTMENT
C
  170 J2=MA
      J1=MB
      PRINT 120,J1,J2
      GO TO 125
  180 J1=G
      J2=G+8
      PRINT 120,J1,J2
      GO TO 125
      END
End of File
```

```
C
C
      SUBROUTINE FOUR
C     .............
C
C
C     ..........................................................
C     .                                                        .
C     .           BOUNDARY MOVE OF THE N-TERMINAL              .
C     .                                                        .
C     ..........................................................
C
C
C     ..........................................................................
C     PURPOSE
C        TO FIND OUT THE MOST FAVORABLE N-BOUNDARY RESIDUE FOR THE PREDIC
C        TED SHEET BASED ON THE BOUNDARY CONFORMATIONAL PARAMETERS OF THE
C        ADJACENT RESIDUES
C     ..........................................................................
C
C
      REAL S,T1,T2,A1,A2 ,T3,T4,T5,TT,P
      INTEGER G,F,H,U,D,V1,V2,V3,V4,V5,V6,V7,V8,Q
      LOGICAL HELLO,BYE,BALL,MOVE
      DIMENSION S(1000,10),M(1000),H(1000),D(1000,16),P(1000,10)
      COMMON S,T1,T2,T3,T4,T5,TT,A1,A2,P ,V4,V5,V6,V7,V8,Q,G,F,H,U,D,NN,
     1NW,KX,MA,MB,MC,MD,L,I,L1,L2,L3,J1,J2,N,K1,K2,V1,V2,IM,M,K3,K4,V3,
     2BYE,BALL,HELLO,MOVE
C
C
C     DESCRIPTION OF PARAMETERS
C        V8  - ACTUAL NUMBER OF BREAKERS IN THE PREDICTED SHEET
C        K3  - C-TERMINAL RESIDUE OF THE PREVIOUS PREDICTED SHEET
C        V2  - COUNTER USED IN THE N-BOUNDARY ADJUSTMENT
C        V3  - COUNTER USED IN THE C-BOUNDARY ADJUSTMENT
C
C
      V2 = 0
      V3 = 0
C
C
C ..... J1 = J1 .....
C .....................
C
C     THE POSITION J1 APPEARS TO BE THE MOST FAVORABLE COMPARED TO ITS
C     ADJACENT RESIDUES,NO NEED TO ADJUST IT.
C
C  *** 1 ***
      BALL=.FALSE.
      IF (M(J1).EQ.1.AND.M(J1+1).EQ.1.AND.S(J1-1,8).LT.1.07.AND.S(J1-2,
```

```
     18).LT.1.07.AND.M(J1+2).EQ.10)  BALL=.TRUE.
      IF (BALL)  J1=J1
      IF (BALL)  V2=1
      IF (BALL)  GO TO 200
C
C  *** 2 ***
      BALL = .FALSE.
      IF (M(J1).EQ.1.AND.S(J1-1,2).LT.1.05.AND.M(J1+1).EQ.11)  BALL=
     1 .TRUE.
      IF (BALL)  J1 = J1
      IF (BALL)  V2=2
      IF (BALL)  GO TO 200
C
C  *** 3 ***
      BALL=.FALSE.
      IF (M(J1).EQ.1.AND.S(J1-1,2).LT.1.05.AND.V8.EQ.1.AND.S(J1+2,8).
     1LT.1.69.AND.(J2-J1).GE.8)  BALL=.TRUE.
      IF (BALL)  J1 = J1
      IF (BALL)  V2=3
      IF (BALL)  GO TO 200
C
C  *** 4 ***
      BALL=.FALSE.
      IF (M(J1).EQ.1.AND.S(J1-1,2).LE.0.75.AND.S(J1+1,8).GE.1.30.AND.
     1S(J1+1,8).LT.1.69)  BALL=.TRUE.
      IF (BALL)  J1=J1
      IF (BALL)  V2=4
      IF (BALL)  GO TO 200
C
C
C ..... MOVE OF J1 .....
C ......................

C     THE POSITION OF J1 IS LESS FAVORABLE THAN THAT OF ITS ADJACENT RE
C     SIDUES
C
C  *** 5 ***
      BALL=.FALSE.
      IF (S(J1,2).GT.0.89.AND.V8.EQ.0.AND.M(J1-1).EQ.1.AND.(J2-J1).GE.8)
     1 BALL=.TRUE.
      IF (BALL)  J1=J1-1
      IF (BALL)  V2=5
      IF (BALL)  GO TO 200
C
C  *** 6 ***
      BALL=.FALSE.
      IF (M(J1).EQ.12     .AND.S(J1-1,2).LT.0.74.AND.S(J1+1,8).LT.1.30.
     1AND.S(J1+2,2).LE.0.75.AND.S(J1+3,8).GE.1.50.AND.(S(J1-2,8)-S(J1+3,
     28)).LT.0.20)  BALL=.TRUE.
      IF (BALL)  J1=J1+3
```

```
101                IF (BALL)  V2=6
102                IF (BALL)  GO TO 200
103      C
104      C  *** 7 ***
105                BALL=.FALSE.
106                IF (S(J1,8).LT.1.50.AND.S(J1+1,8).LE.S(J1-2,8).AND.S(J1-2,8).GE.
107               11.50.AND.S(J1-1,2).LT.0.74.AND.(J2-J1+3).GT.8.AND.(J1-2).GE.K3)
108               3BALL=.TRUE.
109                IF (BALL)  J1=J1-2
110                IF (BALL)  V2=7
111                IF (BALL)  GO TO 200
112      C
113      C  *** 8 ***
114                BALL=.FALSE.
115                IF (S(J1,2).LT.1.05.AND.S(J1-1,2).LT.1.05.AND.S(J1-2,2).LT.1.05
116               1.AND.S(J1-3,8).LT.1.65.AND.S(J1+1,8).GE.1.07)  BALL=.TRUE.
117                IF (BALL)  J1=J1+1
118                IF (BALL)  V2=8
119                IF (BALL)  GO TO 200
120      C
121      C  *** 9 ***
122                BALL=.FALSE.
123                IF ((S(J1+1,8)-S(J1,8)).GT.0.20.AND.(M(J1-1).EQ.15.OR.M(J1-1).EQ.
124               17.OR.M(J1-1).EQ.4).AND.(J2-J1+2).LE.8.AND.S(J1-2,8).LT.1.69.AND.
125               2S(J1,8).LT.1.07)  BALL=.TRUE.
126                IF (BALL)  J1=J1+1
127                IF (BALL)  V2=9
128                IF (BALL)  GO TO 200
129      C
130      C  *** 10 ***
131                BALL=.FALSE.
132                IF (S(J1,2).LT.1.05.AND.S(J1-1,2).LT.1.05.AND.S(J1-2,8).LT.1.65
133               1.AND.S(J1+1,8).GE.1.07)  BALL=.TRUE.
134                IF (BALL)  J1=J1+1
135                IF (BALL)  V2=10
136                IF (BALL)  GO TO 200
137      C
138      C  *** 11 ***
139                BALL=.FALSE.
140                IF (S(J1,2).LT.1.05.AND.M(J1+1).EQ.1.AND.(M(J1+2).EQ.11.OR.S(J1+2,
141               12).GE.1.30).AND.S (J1-1,2).LT.0.83)  BALL=.TRUE.
142                IF (BALL)  J1=J1+1
143                IF (BALL)  V2=11
144                IF (BALL)  GO TO 200
145      C
146      C  *** 12 ***
147                BALL=.FALSE.
148                IF (S(J1,2).LT.1.05.AND.S(J1-1,2).LT.1.05.AND.S(J1-2,8).GE.1.69
149               1.AND.S(J1+1,8).LT.1.65.AND.(J1-2).GT.K3)  BALL=.TRUE.
150                IF (BALL)  J1=J1-2
```

```
151                IF (BALL)  V2=12
152                IF (BALL)  GO TO 200
153       C
154       C  *** 13 ***
155                BALL=.FALSE.
156                IF (S(J1,8).GE.1.07.AND.S(J1-2,8).GE.1.65.AND.S(J1-1,2).LT.0.74
157               1.AND.(J2-J1+3).GT.8.AND.(J1-2).GT.K3)  BALL=.TRUE.
158                IF (BALL)  J1=J1-2
159                IF (BALL)  V2=13
160                IF (BALL)  GO TO 200
161       C
162       C  *** 14 ***
163                BALL=.FALSE.
164                IF (S(J1,8).GE.1.07.AND.S(J1-2,8).GT.S(J1,8).AND.S(J1-1,2).GE.
165               10.74 .AND.(J1-2).GT.K3.AND.(S(J1-2,8)-S(J1,8)).GE.0.50)  BALL=
166               2.TRUE.
167                IF (BALL)  J1=J1-2
168                IF (BALL)  V2=14
169                IF (BALL)  GO TO 200
170       C
171       C  *** 15 ***
172                BALL=.FALSE.
173                IF ((S(J1+3,8)-S(J1,8)).GE.0.35 .AND.S(J1+1,8).LT.1.07.AND.S(J1+2,
174               18).LT.1.07.AND.S(J1-1,8).LT.1.07.AND.S(J1-2,8).LT.S(J1+3,8)) BALL=
175               2.TRUE.
176                IF (BALL)  J1=J1+3
177                IF (BALL)  V2=15
178                IF (BALL)  GO TO 200
179       C
180       C  *** 16 ***
181                BALL=.FALSE.
182                IF (S(J1,8).GE.1.07.AND.S(J1-2,8).GT.S(J1,8).AND.S(J1-1,2).GE.
183               10.74 .AND.(J1-2).GT.K3)  BALL=.TRUE.
184                IF (BALL)  J1=J1-2
185                IF (BALL)  V2=16
186                IF (BALL)  GO TO 200
187       C
188       C  *** 17 ***
189                BALL=.FALSE.
190                IF (S(J1,8).GE.1.07.AND.S(J1-2,8).GT.S(J1,8).AND.(S(J1-1,2).GE.0.
191               174.OR.M(J1-1).EQ.4).AND.(J2-J1+3).GE.8.AND.(J1-2).GT.K3)  BALL=
192               2 .TRUE.
193                IF (BALL)  J1=J1-2
194                IF (BALL)  V2=17
195                IF (BALL)  GO TO 200
196       C
197       C  *** 18 ***
198                BALL=.FALSE.
199                IF (S(J1,2).LT.1.05.AND.S(J1-1,2).LT.1.05.AND.S(J1+1,8).GE.1.07
200               1.AND.(J1-2).LE.K3)  BALL=.TRUE.
```

```
201               IF (BALL)  J1=J1+1
202               IF (BALL)  V2=18
203               IF (BALL)  GO TO 200
204     C
205     C  *** 19 ***
206               BALL=.FALSE.
207               IF (S(J1,2).LT.1.05.AND.S(J1+1,2).LT.1.05.AND.(M(J1-1).EQ.15.OR.M(
208              1J1-1).EQ.7.OR.M(J1-1).EQ.4).AND.S(J1+2,8).GE.1.07)  BALL=.TRUE.
209               IF (BALL)  J1=J1+2
210               IF (BALL)  V2=19
211               IF (BALL)  GO TO 200
212     C
213     C  *** 20 ***
214               BALL=.FALSE.
215               IF(S(J1-1,2).LT.1.05.AND.S(J1-2,8).LT.1.07.AND.S(J1+1,2).LT.1.05
216              1.AND.(S(J1+2,8)-S(J1,8)).GE.0.55 .AND.S(J1+2,8).GE.1.07)  BALL=
217              2.TRUE.
218               IF (BALL)  J1=J1+2
219               IF (BALL)  V2=20
220               IF (BALL)  GO TO 200
221     C
222     C  *** 21 ***
223               BALL=.FALSE.
224               IF (S(J1,2).LT.1.05.AND.S(J1+1,2).LT.1.05.AND.(S(J1+2,8)-S(J1-1,8)
225              1).GT.0.25.AND.S(J1+2,8).GE.1.50)  BALL=.TRUE.
226               IF (BALL)  J1=J1+2
227               IF (BALL)  V2=21
228               IF (BALL)  GO TO 200
229     C
230     C  *** 22 ***
231               BALL=.FALSE.
232               IF (M(G).EQ.19.AND.S(G-1,8).GE.1.69.AND.M(G+1).EQ.4.AND.S(G+2,8).L
233              1T.S(G-1,8))  BALL=.TRUE.
234               IF (BALL)  J1=J1-1
235               IF (BALL)  V2=22
236               IF (BALL)  GO TO 200
237     C
238     C  *** 23 ***
239               BALL=.FALSE.
240               IF (S(J1+2,8).GT.S(J1,8).AND.(M(J1+1).EQ.15.OR.M(J1+1).EQ.7.OR.
241              1M(J1+1).EQ.4).AND.S(J1+2,8).GE.1.07.AND.S(J1,8).LT.1.42)  BALL=
242              2.TRUE.
243               IF (BALL)  J1=J1+2
244               IF (BALL)  V2=23
245               IF (BALL)  GO TO 200
246     C
247     C  *** 24 ***
248               BALL=.FALSE.
249               IF (S(J1,2).LT.1.05.AND.S(J1+1,2).LE.0.75.AND.S(J1-1,2).LT.1.05
250              1.AND.S(J1+2,8).GE.1.07)  BALL=.TRUE.
```

```
251               IF (BALL)  J1=J1+2
252               IF (BALL)  V2=24
253               IF (BALL)  GO TO 200
254     C
255     C  *** 25 ***
256               BALL=.FALSE.
257               IF (S(J1,8).LT.1.07.AND.S(J1-1,8).LT.1.07.AND.S(J1-2,8).LT.1.65
258              1.AND.S(J1+1,8).GE.1.07)  BALL=.TRUE.
259               IF (BALL)  J1=J1+1
260               IF (BALL)  V2=25
261               IF (BALL)  GO TO 200
262     C
263     C  *** 26 ***
264               BALL=.FALSE.
265               IF (S(J1,8).GE.1.07.AND.S(J1-2,8).GE.1.07.AND.S(J1-1,2).GE.0.75
266              1.AND.S(J1,8).LE.S(J1-2,8).AND.M(J1+1).EQ.20.AND.M(J1+3).EQ.20)
267              2  BALL=.TRUE.
268               IF (BALL)  J1=J1-2
269               IF (BALL)  V2=26
270               IF (BALL)  GO TO 200
271     C
272     C  *** 27 ***
273               BALL=.FALSE.
274               IF ((S(J1+1,8)-S(J1,8)).GT.0.20.AND.(S(J1-1,8)-S(J1,8)).LT.0.20
275              1.AND.S(J1+1,8).GE.1.07.AND.(J1-2).LE.K3)  BALL=.TRUE.
276               IF (BALL)  J1=J1+1
277               IF (BALL)  V2=27
278               IF (BALL)  GO TO 200
279     C
280     C  *** 28 ***
281               BALL=.FALSE.
282               IF ((S(J1+3,8)-S(J1,8)).GE.0.35 .AND.S(J1+1,8).LT.1.07.AND.S(J1+2,
283              18).LT.1.07.AND.S(J1-1,8).LT.1.07.AND.S(J1-2,8).LT.S(J1+3,8).AND.
284              2(J1-2).LE.K3)  BALL=.TRUE.
285               IF (BALL)  J1=J1+3
286               IF (BALL)  V2=28
287               IF (BALL)  GO TO 200
288     C
289     C  *** 29 ***
290               BALL=.FALSE.
291               IF (S(J1,8).LT.1.07.AND.(J1-1).LE.K3.AND.S(J1+1,8).LT.1.07.AND.
292              1S(J1+2,8).LT.1.07.AND.S(J1+3,8).GE.1.07)  BALL=.TRUE.
293               IF (BALL)  J1=J1+3
294               IF (BALL)  V2=29
295               IF (BALL)  GO TO 200
296     C
297     C  *** 30 ***
298               BALL=.FALSE.
299               IF(((S(J1+1,2).LT.0.74.AND.S(J1+2,8).LT.1.07).OR.(S(J1+1,8).LT.
300              11.07.AND.S(J1+2,2).LT.0.74)).AND.S(J1+3,8).GE.1.07.AND.S(J1,8).LE.
```

```
              21.65)  BALL=.TRUE.
               IF (BALL)  J1=J1+3
               IF (BALL)  V2=30
               IF (BALL)  GO TO 200
C
C  *** 31 ***
               BALL=.FALSE.
               IF (S(J1,8).GE.1.69.AND.S(J1-2,8).GE.1.69.AND.S(J1+1,8).LT.S(J1-2,
              18).AND.S(J1-1,2).LT.0.74.AND.(J2-J1+3).GE.7.AND.(J1-2).GT.K3)
              2 BALL=.TRUE.
               IF (BALL)  J1=J1-2
               IF (BALL)  V2=31
               IF (BALL)  GO TO 200
C
C  *** 32 ***
               BALL=.FALSE.
               IF (S(J1,2).LT.0.83.AND.S(J1+1,8).GE.1.07.AND.(S(J1+1,8)-S(J1-1,8)
              2).GT.0.30.AND.S(J1-1,8).LT.1.30)  BALL=.TRUE.
               IF (BALL)  J1=J1+1
               IF (BALL)  V2=32
               IF (BALL)  GO TO 200
C
C  *** 33 ***
               LK=J1-3
               IF (LK.LE.0)  GO TO 100
               V8=0
               DO  50  L=LK,J2
               IF (S(L,1).LE.0.75). V8=V8+1
    50         CONTINUE
               LY=(J2+1-LK)/3
C
               BALL=.FALSE.
               IF (V8.LE.LY.AND.M(J1).EQ.6 .AND.M(J1-1).EQ.4.AND.S(J1-2,2).GE.
              10.75.AND.S(J1-3,2).GE.1.30.AND.S(J1-3,8).GE.1.30)  BALL=.TRUE.
               IF (BALL)  J1=J1-3
               IF (BALL)  V2=33
               IF (BALL)  GO TO 200
C
C  *** 34 ***
               BALL=.FALSE.
               IF (S(J1,8).GE.1.69.AND.S(J1-3,8).EQ.1.94.AND.S(J1-1,2).GE.0.74
              1.AND.S(J1-2,2).GT.0.75.AND.S(J1+1,8).LT.S(J1-3,8).AND.N.LE.LY)
              2 BALL=.TRUE.
               IF (BALL)  J1=J1-3
               IF (BALL)  V2=34
               IF (BALL)  GO TO 200
C
C
C ..... J1 = J1 .....
C .....................
```

```
C
C  *** 35 ***
 100   BALL=.FALSE.
       IF (S(J1,8).GE.1.50.AND.S(J1-1,8).LT.1.50.AND.S(J1-2,8).LT.1.50
      1.AND.S(J1+1,8).GE.1.07)  BALL=.TRUE.
       IF (BALL)  J1=J1
       IF (BALL)  V2=35
       IF (BALL)  GO TO 200
C
C  *** 36 ***
       BALL=.FALSE.
       IF (S(J1,8).GE.1.07.AND.S(J1-1,2).LE.0.75.AND.S(J1-2,8).LT.S(J1,8)
      1.AND.S(J1+1,8).LT.S(J1,8).AND.S(J1+2,8).LT.S(J1,8).AND.S(J1+1,2)
      2.GE.0.74.AND.S(J1+2,2).GE.0.74)  BALL=.TRUE.
       IF (BALL)  J1=J1
       IF (BALL)  V2=36
       IF (BALL)  GO TO 200
C
C  *** 37 ***
       BALL=.FALSE.
       IF (S(J1,8).GE.1.07.AND.S(J1+1,8).GE.1.07.AND.(S(J1+1,8)-S(J1,8))
      1.LT.0.20.AND.S(J1-1,2) .LT.1.05.AND.S(J1-2,2).LT.1.05) BALL=.TRUE.
       IF (BALL)  J1=J1
       IF (BALL)  V2=37
       IF (BALL)  GO TO 200
C
C  *** 38 ***
       BALL=.FALSE.
       IF ((M(J1-1).EQ.15.OR.M(J1-1).EQ.7.OR.M(J1-1).EQ.4).AND.(J2-J1+2)
      1.LE.8.AND.(S(J1+1,8)-S(J1,8)).LT.0.20.AND.S(J1,8).GE.1.07)  BALL=
      2.TRUE.
       IF (BALL)  J1=J1
       IF (BALL)  V2=38
       IF (BALL)  GO TO 200
C
C  *** 39 ***
       BALL=.FALSE.
       IF ((J1-1).LE.K3.AND.S(J1,8).GE.1.07.AND.(S(J1+1,8)-S(J1,8)).LT.
      10.20)  BALL=.TRUE.
       IF (BALL)  J1=J1
       IF (BALL)  V2=39
       IF (BALL)  GO TO 200
C
C  *** 40 ***
       BALL=.FALSE.
       IF (S(J1,8).GE.1.07.AND.(S(J1+1,8)-S(J1,8)).LT.0.20.AND.S(J1-2,2)
      1.LT.1.05.AND.S(J1-1,2).LT.1.05.AND.S(J1+2,8).LT.S(J1,8).AND.S(J1+1
      2,2).GE.0.74.AND.S(J1+2,2).GE.0.74)  BALL=.TRUE.
       IF (BALL)  J1=J1
       IF (BALL)  V2=40
```

```
401                IF (BALL)  GO TO 200
402      C
403      C
404      C      TO CALL SUBROUTINE FIVE TO CARRY OUT THE ADJUSTMENT OF THE C-BOUN
405      C      DARY
406      C
407           200   CALL FIVE
408                 RETURN
409                 END
End of File
```

```
C
C
      SUBROUTINE FIVE
C     .............
C
C
C     ..........................................................
C     .                                                        .
C     .            BOUNDARY MOVE OF THE C-TERMINAL             .
C     .                                                        .
C     ..........................................................
C
C
C     ..................................................................
C     PURPOSE
C        TO ADJUST THE C-TERMINAL RESIDUE BASED ON THE BOUNDARY CONFOR
C        MATIONAL PARAMETERS OF THE ADJACENT RESIDUES
C     ..................................................................
C
C
      REAL S,T1,T2,A1,A2 ,T3,T4,T5,TT,P
      INTEGER G,F,H,U,D,V1,V2,   V3,V4,V5,V6,V7,V8,Q
      LOGICAL HELLO,BYE  ,BALL,MOVE
      DIMENSION S(1000,10),M(1000),H(1000),D(1000,16),P(1000,10)
      COMMON S,T1,T2,T3,T4,T5,TT,A1,A2,P ,V4,V5,V6,V7,V8,Q,G,F,H,U,D,NN,
     1NW,KX,MA,MB,MC,MD,L,I,L1,L2,L3,J1,J2,N,K1,K2,V1,V2,IM,M,K3,K4,V3,
     2BYE,BALL,HELLO,MOVE
C
C
C    ... TO CHECK THE NUMBER OF BREAKERS IN THE SECTIONS J1-J2+4,J1-J2+3 ...
C
C     DESCRIPTION OF PARAMETERS
C        V8  - ACTUAL NUMBER OF BREAKERS IN THE SEGMENT J1 TO J2+4
C        JJ  - THEORITICAL NUMBER OF BREAKERS IN THE SECTION J1 TO J2+4
C        MM  - ACTUAL NUMBER OF BREAKERS IN THE SEGMENT J1 TO J2+3
C        JM  - THEORITICAL NUMBER OF BREAKERS IN THE SECTION J1 TO J2+3
C        K3  - C-TERMINAL RESIDUE OF THE PREVIOUS PREDICTED SHEET
C
      J3=J2+4
      V8=0
      DO 90  JC=J1,J3
      IF (S(JC,2).LE.0.75)  V8=V8+1
   90 CONTINUE
      JJ=(J3+1-J1)/3
C
      J5=J2+3
      MM=0
      DO 102  JC=J1,J5
      IF (S(JC,2).LE.0.75)  MM=MM+1
  102 CONTINUE
```

```
      JM=(J5+1-J1)/3
C
C
C ..... J2 = J2 .....
C .....................
C
C   *** 1 ***
      BYE=.FALSE.
      IF (M(J2).EQ.1.AND.S(J2+1,2).LT.1.05.AND.S(J2-1,2).GE.1.47)  BYE=
     1.TRUE.
      IF (BYE)  J2=J2
      IF (BYE)  V3=1
      IF (BYE)  GO TO 300
C
C   *** 2 ***
      BYE=.FALSE.
      IF (S(J2,9).GE.1.11.AND.M(J2).EQ.5.AND.M(J2-1).EQ.1.AND.S(J2-2,2)
     1.GE.1.30.AND.S(J2+1,2).LT.1.05)  BYE=.TRUE.
      IF (BYE)  J2=J2
      IF (BYE)  V3=2
      IF (BYE)  GO TO 300
C
C   *** 3 ***
      BYE=.FALSE.
      IF (M(J2).EQ.1.AND.S(J2-1,9).LT.1.79.AND.T3.EQ.3.O.AND.N.EQ.O.AND.
     1(J2+1-J1).EQ.5)  BYE=.TRUE.
      IF (BYE)  J2=J2
      IF (BYE)  V3=3
      IF (BYE)  GO TO 300
C
C   *** 4 ***
      BYE=.FALSE.
      IF (M(J2-1).EQ.19.AND.M(J2-2).EQ.20.AND.M(J2-3).EQ.13.AND.S(J2,2)
     1.GE.O.75)  BYE=.TRUE.
      IF (BYE)  J2=J2
      IF (BYE)  V3=4
      IF (BYE)  GO TO 300
C
C.
C ..... MOVE OF J2 .....
C ......................
C
C
C   *** 6 ***
      BYE=.FALSE.
      IF (S(J2,2).GE.1.05.AND.S(J2+1,9).LT.1.11.AND.S(J2+2,9).LT.1.11
     1.AND.S(J2+3,9).GE.1.96.AND.((M(J2+1).EQ.4.AND.S(J2+2,2).GE.O.74)
     2.OR.(M(J2+2).EQ.4.AND.S(J2+1,2).GE.O.74)).AND.(J2-J1+5).GE.7.AND.
     3MM.LE.JM)  BYE=.TRUE.
      IF (BYE)  J2=J2+3
```

```
101                IF (BYE)  V3=6
102                IF (BYE)  GO TO 300
103      C
104      C    *** 7 ***
105                BYE=.FALSE.
106                IF (S(J2,9).LT.1.11.AND.S(J2-1,9).GT.S(J2,9).AND.S(J2+1,9).LT.1.11
107               1.AND.S(J2+2,9).LT.1.11.AND.S(J2-2,9).LE.S(J2-1,9))  BYE=.TRUE.
108                IF (BYE)  J2=J2-1
109                IF (BYE)  V3=7
110                IF (BYE)  GO TO 300
111      C
112      C    *** 8 ***
113                BYE=.FALSE.
114                IF (S(J2+3,9).GE.1.96.AND.S(J2+2,2).GE.1.05.AND.S(J2,9).LT.1.11
115               1.AND.S(J2-1,9).LT.1.11.AND.S(J2-2,9).LE.S(J2+3,9).AND.S(J2+1,9).
116               2LT.1.11.AND.M(J2+1).NE.15.AND.M(J2+1).NE.7.AND.M(J2+1).NE.4.AND.
117               3MM.LE.JM)  BYE=.TRUE.
118                IF (BYE)  J2=J2+3
119                IF (BYE)  V3=8
120                IF (BYE)  GO TO 300
121      C
122      C    *** 9 ***
123                BYE=.FALSE.
124                IF (S(J2,9).GE.1.11.AND.S(J2+3,9).GE.S(J2,9).AND.S(J2-1,9).LE.
125               1S(J2+3,9).AND.M(J2+1).NE.15.AND.M(J2+1).NE.4.AND.M(J2+1).NE.7.AND.
126               2M(J2+2).NE.15.AND.M(J2+2).NE.4.AND.M(J2+2).NE.7.AND.(S(J2-1,2).GE.
127               30.74 .AND.S(J2-2,2).GE.0.74).AND.MM.LT.JM)  BYE=.TRUE.
128                IF (BYE)  J2=J2+3
129                IF (BYE)  V3=9
130                IF (BYE)  GO TO 300
131      C
132      C   *** 10 ***
133                BYE=.FALSE.
134                IF (S(J2,9).GE.1.11.AND.S(J2-1,9).LT.1.11.AND.S(J2+4,2).GE.1.05.AN
135               1D.S(J2+1,2).GE.0.74.AND.S(J2+2,2).GE.0.74.AND.S(J2+3,2).GE.0.74
136               2.AND.S(J2-1,2).GE.0.74.AND.(S(J2-2,9)-S(J2,9)).LT.0.60.AND.V8.LE.
137               3JJ.AND.S(J2+2,9).LT.1.50)  BYE=.TRUE.
138                IF (BYE)  J2=J2+4
139                IF (BYE)  V3=10
140                IF (BYE)  GO TO 300
141      C
142      C   *** 11 ***
143                BYE=.FALSE.
144                IF ((J2+5).GT.NN)  GO TO 100
145                IF (S(J2,9).GE.1.11.AND.S(J2-1,9).GE.1.50.AND.S(J2+4,9).GT.1.11
146               1.AND.S(J2+5,9).GT.S(J2+4,9).AND.S(J2+1,2).GT.0.74.AND.S(J2+2,2)
147               2.GT.0.74.AND.S(J2+3,2).GT.0.74.AND.(S(J2-1,9)-S(J2+5,9)).LT.0.30
148               3.AND.V8.LE.((J2+6-J1)/3))  BYE=.TRUE.
149                IF (BYE)  J2=J2+5
150                IF (BYE)  V3=11
```

```
151                IF (BYE)  GO TO 300
152       C
153       C  *** 12 ***
154            100   BYE=.FALSE.
155                  IF (S(J2,2).LE.0.75.AND.S(J2-1,2).LE.0.75.AND.M(J2-2).EQ.19.AND.
156                 1(S(J2-2,9)-S(J2+1,9)).GE.0.84)  BYE=.TRUE.
157                  IF (BYE)  J2=J2-2
158                  IF (BYE)  V3=12
159                  IF (BYE)  GO TO 300
160       C
161       C  *** 13 ***
162                  BYE=.FALSE.
163                  IF (S(J2-1,9).LT.1.11.AND.S(J2+1,9).LT.1.11.AND.S(J2-2,9).GT.S(J2,
164                 19).AND.S(J2-2,9).GE.1.79.AND.(J2-2).GT.J1)  BYE=.TRUE.
165                  IF (BYE)  J2=J2-2
166                  IF (BYE)  V3=13
167                  IF (BYE)  GO TO 300
168       C
169       C  *** 14 ***
170                  BYE=.FALSE.
171                  IF (S(J2,9).LT.1.11.AND.S(J2+1,9).GE.1.11.AND.(S(J2-1,9)-S(J2+1,9)
172                 1).LE.0.60)  BYE=.TRUE.
173                  IF (BYE)  J2=J2+1
174                  IF (BYE)  V3=14
175                  IF (BYE)  GO TO 300
176       C
177       C  *** 15 ***
178                  BYE=.FALSE.
179                  IF (S(J2,9).LT.1.11.AND.S(J2+1,9).GE.1.11.AND.(S(J2-1,9)-S(J2+1,9)
180                 1).GT.0.60)  BYE=.TRUE.
181                  IF (BYE)  J2=J2-1
182                  IF (BYE)  V3=15
183                  IF (BYE)  GO TO 300
184       C
185       C  *** 16 ***
186                  BYE=.FALSE.
187                  IF (S(J2,9).LE.1.11.AND.S(J2-1,9).LT.1.11.AND.S(J2-2,9).LT.1.11
188                 1.AND.S(J2+1,9).LT.1.11.AND.S(J2+2,9).LT.1.11.AND.(S(J2-3,9).GE.1.
189                211.OR.M(J2-3).EQ.19).AND.(J2-3).GT.J1.AND.S(J2+3,9).LT.1.96)  BYE=
190                3.TRUE.
191                  IF (BYE)  J2=J2-3
192                  IF (BYE)  V3=16
193                  IF (BYE)  GO TO 300
194       C
195       C  *** 17 ***
196                  BYE=.FALSE.
197                  IF (S(J2-1,9).GE.1.79.AND.(S(J2-1,9)-S(J2,9)).GE.0.70.AND.S(J2+1,
198                 19).LT.1.11.AND.S(J2-2,9).LT.S(J2-1,9))  BYE=.TRUE.
199                  IF (BYE)  J2=J2-1
200                  IF (BYE)  V3=17
```

```
              IF (BYE)  GO TO 300
C
C  *** 18 ***
              BYE=.FALSE.
              IF (S(J2-1,9).LT.1.11.AND.S(J2+1,9).GT.S(J2,9).AND.S(J2,9).GE.1.11
             1)  BYE=.TRUE.
              IF (BYE)  J2=J2+1
              IF (BYE)  GO TO 300
C
C  *** 19 ***
              BYE=.FALSE.
              IF ((S(J2-1,9).LT.1.11.OR.S(J2-1,9).LE.S(J2+1,9)).AND.S(J2+1,9)
             1.GT.S(J2,9).AND.S(J2,9).GE.1.11)  BYE=.TRUE.
              IF (BYE)  J2=J2+1
              IF (BYE)  V3=19
              IF (BYE)  GO TO 300
C
C  *** 20 ***
              BYE=.FALSE.
              IF (S(J2+2,9).GE.S(J2,9).AND.(S(J2,9).GE.1.11.OR.M(J2+2).EQ.5)
             1 .AND.(M(J2+1).NE.15.AND.M(J2+1).NE.7).AND.S(J2+2,9).GT.S(J2+1,9))
             2 BYE=.TRUE.
              IF (BYE)  J2=J2+2
              IF (BYE)  V3=20
              IF (BYE)  GO TO 300
C
C  *** 21 ***
              BYE=.FALSE.
              IF (S(J2,2).LT.1.05.AND.S(J2+1,2).LT.1.05.AND.S(J2-1,9).GT.S(J2+2,
             19).AND.S(J2-1,9).GE.1.11)  BYE=.TRUE.
              IF (BYE)  J2=J2-1
              IF (BYE)  V3=21
              IF (BYE)  GO TO 300
C
C  *** 22 ***
              BYE=.FALSE.
              IF (S(J2,2).LT.1.05.AND.S(J2-2,9).GT.S(J2-1,9).AND.S(J2-2,9).GE.
             11.07.AND.S(J2+1,9).LT.1.11.AND.S(J2+2,9).LT.1.11.AND.(J2-2).GT.J1)
             2BYE=.TRUE.
              IF (BYE)  J2=J2-2
              IF (BYE)  V3=22
              IF (BYE)  GO TO 300
C
C  *** 23 ***
              BYE=.FALSE.
              IF (S(J2,9).GE.1.11.AND.(S(J2+1,2).GE.0.74.OR.M(J2+1).EQ.4).AND.
             1S(J2+2,9).GE.1.11.AND.S(J2-1,9).GE.1.11.AND.S(J2,9).LT.1.96)  BYE
             2=.TRUE.
              IF (BYE)  J2=J2+2
              IF (BYE)  V3=23
```

```
251                IF (BYE)  GO TO 300
252      C
253      C  *** 24 ***
254                BYE=.FALSE.
255                IF (S(J2,2).GE.1.05.AND.S(J2+3,9).GE.1.79.AND.S(J2+2,2).GE.1.05
256               1.AND.S(J2-1,2).GE.1.05.AND.(J2-J1+5).GE.7.AND.(M(J2+1).EQ.7 .OR.
257               2M(J2+1).EQ.4).AND.MM.LE.JM)  BYE=.TRUE.
258                IF (BYE)  J2=J2+3
259                IF (BYE)  V3=24
260                IF (BYE)  GO TO 300
261      C
262      C  *** 25 ***
263                BYE=.FALSE.
264                IF ((J2-J1+1).LE.5.AND.T3.GE.3.0.AND.N.LE.1.AND.M(J2+1).EQ.4.AND.
265               1S(J2+2,9).GE.1.11.AND.S(J2+3,9).GE.1.11.AND.S(J2+1,2).GE.1.05.AND.
266               2S(J2+3,2).GE.1.05.AND.MM.LE.JM)  BYE=.TRUE.
267                IF (BYE)  J2=J2+3
268                IF (BYE)  V3=25
269                IF (BYE)  GO TO 300
270      C
271      C  *** 26 ***
272                BYE=.FALSE.
273                IF (S(J2,9).LT.1.11.AND.S(J2+1,9).LT.1.11.AND.(J2+2).GT.NN.AND.
274               1S(J2-1,9).GE.1.11)  BYE=.TRUE.
275                IF (BYE)  J2=J2-1
276                IF (BYE)  V3=26
277                IF (BYE)  GO TO 300
278      C
279      C  *** 27 ***
280                BYE=.FALSE.
281                IF (S(J2,9).LT.1.11.AND.S(J2+1,9).LT.1.11.AND.S(J2+2,9).GE.1.79.AN
282               1D.S(J2-2,9).LT.S(J2+2,9))  BYE=.TRUE.
283                IF (BYE)  J2=J2+2
284                IF (BYE)  V3=27
285                IF (BYE)  GO TO 300
286      C
287      C  *** 28 ***
288                BYE=.FALSE.
289                IF (S(J2,9).GE.1.79.AND.S(J2-1,9).GE.1.79.AND.S(J2-2,9).GE.1.79
290               1.AND.S(J2-3,9).GE.1.79.AND.M(J2+1).EQ.1.AND.M(J2+2).EQ.1)  BYE=
291               2.TRUE.
292                IF (BYE)  J2=J2+2
293                IF (BYE)  V3=28
294                IF (BYE)  GO TO 300
295      C
296      C  *** 29 ***
297                BYE=.FALSE.
298                IF (M(J2).EQ.5.AND.S(J2-1,9).GE.1.79.AND.MM.LE.JM.AND.S(J2-2,9).GE
299               1.1.27.AND.S(J2+3,9).GE.1.21.AND.S(J2+1,9).GE.0.74.AND.S(J2+2,2).GE
300               2.0.74)  BYE=.TRUE.
```

```
301                IF (BYE)  J2=J2+3
302                IF (BYE)  V3=29
303                IF (BYE)  GO TO 300
304       C
305       C  *** 30 ***
306                BYE=.FALSE.
307                IF (S(J2,2).GE.1.60.AND.S(J2-1,2).GE.1.60.AND.S(J2-2,2).GE.1.38.AN
308               1D.MM.LE.JM.AND.S(J2+1,2).GE.0.75.AND.(M(J2+2).EQ.7.OR.M(J2+2).EQ.
309               24).AND.S(J2+3,9).GE.1.27) BYE=.TRUE.
310                IF (BYE)  J2=J2+3
311                IF (BYE)  V3=30
312                IF (BYE)  GO TO 300
313       C
314       C  *** 31 ***.
315                BYE=.FALSE.
316                IF (M(J2).EQ.20.AND.M(J2+4).EQ.20.AND.M(J2-2).EQ.20.AND.S(J2+1,2)
317               1.GE.0.75.AND.S(J2+2,2).GE.0.93.AND.S(J2+3,2).GE.0.75)  BYE=.TRUE.
318                IF (BYE)  J2=J2+4
319                IF (BYE)  V3=31
320                IF (BYE)  GO TO 300
321       C
322       C
323       C ..... J2 = J2 .....
324       C .....................
325       C
326       C
327       C  *** 32 ***
328                BYE=.FALSE.
329                IF (S(J2,9).GE.1.11.AND.S(J2-1,9).LT.1.11.AND.S(J2+1,9).LT.1.11
330               1.AND.S(J2-2,9).LE.S(J2,9))  BYE=.TRUE.
331                IF (BYE)  J2=J2
332                IF (BYE)  V3=32
333                IF (BYE)  GO TO 300
334       C
335       C  *** 33 ***
336                BYE=.FALSE.
337                IF (S(J2,9).GE.1.11.AND.S(J2+1,9).LT.1.11.AND.S(J2+2,9).LT.S(J2,9)
338               1.AND.S(J2-1,9).LE.S(J2,9).AND.S(J2-2,9).LT .1.11)  BYE=.TRUE.
339                IF (BYE)  J2=J2
340                IF (BYE)  V3=33
341                IF (BYE)  GO TO 300
342       C
343       C  *** 34 ***
344                BYE=.FALSE.
345                IF (S(J2,9).GE.1.27.AND.S(J2+1,9).LT.1.11.AND.(S(J2+2,9).LT.1.11
346               1.OR.S(J2+2,9).LT.S(J2,9)).AND.S(J2-1,9).LE.S(J2,9).AND.S(J2-2,9)
347               2.LE.S(J2,9))  BYE=.TRUE.
348                IF (BYE)  J2=J2
349                IF (BYE)  V3=34
350                IF (BYE)  GO TO 300
```

```
      C
      C  *** 35 ***
                BYE=.FALSE.
                IF (S(J2,9).GE.1.11.AND.(J2-2).LE.J1.AND.S(J2-1,9).LT.1.11.AND.
               1S(J2+1,9).LT.1.11.AND.S(J2+2,9).LT.1.11)  BYE=.TRUE.
                IF (BYE)  J2=J2
                IF (BYE)  V3=35
                IF (BYE)  GO TO 300
      C
      C  *** 36 ***
                BYE=.FALSE.
                IF (S(J2,9).GE.1.27.AND.S(J2-1,9).GE.1.21.AND.S(J2-2,9).GE.1.21
               1.AND.S(J2+1,9).LT.1.11.AND.S(J2+2,9).LT.1.11)  BYE=.TRUE.
                IF (BYE)  J2=J2
                IF (BYE)  V3=36
                IF (BYE)  GO TO 300
      C
      C  *** 37 ***
                BYE=.FALSE.
                IF (S(J2,9).GE.1.11.AND.S(J2+1,9).LT.1.11.AND.S(J2+2,9).LT.1.11
               1.AND.(J2-1).LE.J1)  BYE=.TRUE.
                IF (BYE)  J2=J2
                IF (BYE)  V3=37
                IF (BYE)  GO TO 300
      C
      C
      C     TO PRINT OUT THE FINAL VALUES J1,J2 OF THE PREDICTED SHEET. THEN TO
      C     RETURN TO SUBROUTINE FIRS TO START THE SEARCH AGAIN
      C
          300   K3=J2
                PRINT 301,J1,J2,V2,V3
          301   FORMAT('0',25X,'EVENTUAL SHEET FROM J1:',I5,5X,'TO   J2:',I5,14X,
               1'  ***  V2,V3 :',2I5,'  ***'//)
                RETURN
                END
End of File
```

Efficiency of the β-turn prediction

Although some small differences existed between the results in this study and those reported by Chou and Fasman (1977, 1979) (e.g. carbonic anhydrase 71-74, 109-112; α-chymotrypsin 148-151; α-hemoglobin 81-84; thermolysin 19-22, 43-46), in general the results in this study agreed very well with those of Chou and Fasman (1977, 1979). Therefore,no modification was needed for β-turn prediction. The program used for β-turn prediction consisted of the main program and one subroutine. The subroutine was the only part presented in this study because the main program was similar to the one used for α-helix and β-sheet search.

```
C
C
      SUBROUTINE  TURN
C     ................
C
C
C     ..................................................................
C     PURPOSE
C        TO LOCATE B-TURNS BY APPLYING THE RULE: <PA> < <PT>  > <PB>
C        AND THE PROBABILITY OF TURN OCCURRENCE SHOULD BE GREATER THAN
C        0.000075. AND FOR 2 ADJACENT TURNS THE ONE WITH THE HIGHEST PRO
C        BABILITY OF OCCURRENCE WILL BE CHOSEN
C     ..................................................................
C
C
      INTEGER G,F,H,U,D
      REAL S,T1,T2,A1,A2 ,T3,T4,T5,TT,PRB,P,PRBO,A3
      LOGICAL HELLO,BYE,BALL
      DIMENSION S(1000,8),M(1000),H(100) ,D(100,16),P(1000,8)
      COMMON S,T1,T2,A1,A2,T3,T4,T5,TT,PRB,P,G,F,H,U,D,M,IM,I,K1,K2,MB,
     1NN,NW,N,K,J,MC,HELLO,BYE,BALL
C
C
C     DESCRIPTION OF PARAMETERS
C        I      - COUNTER
C        H      - ARRAY TO STORE THE BOUNDARY VALUES OF TURNS
C                 H(I)     - N-BOUNDARY VALUE
C                 H(I+1)   - C-BOUNDARY VALUE
C        MB     - FIRST RESIDUE OF A TETRAPEPTIDE (=K1)
C        K2     - FOURTH RESIDUE OF A TETRAPEPTIDE (=K1+3)
C        A1     - AVERAGE PA OF A TETRAPEPTIDE
C        A2     - AVERAGE PB OF A TETRAPEPTIDE
C        A3     - AVERAGE PT OF A TETRAPEPTIDE
C        PRB    - PROBABILITY OF B-TURN OCCURRENCE
C        PRBO   - PROBABILITY OF B-TURN OCCURRENCE OF THE ADJACENT TETRA
C                 PEPTIDE STARTING AT K1-1
C
C
   10 I=1
      H(I)=0
      NW=NN-3
      MB=1
   20 K2=MB+3
      K1=MB
      PRB=0
      T1=0
      T2=0
      T3=0
      A1=0
      A2=0
```

```
      A3=O
C
C     TO CALCULATE THE AVERAGE PA,PB,PT OF A TETRAPEPTIDE
C
      DO 25   MC=K1,K2
      T1=T1+S(MC,1)
      T2=T2+S(MC,2)
      T3=T3+S(MC,6)
25    CONTINUE
      A1=T1/4.O
      A2=T2/4.O
      A3=T3/4.O
      PRB=P(K1,1)*P(K1+1,2)*P(K1+2,3)*P(K2,4)
      PRINT  30,A1,A2,A3,PRB,MB
30    FORMAT(' ',1OX,'A1:',F6.3,5X,'A2:',F6.3,5X,'A3:',F6.3,5X,'PRB:',F1
     1 3.1O,I8)
      IF ((A3.GT.A2.AND.A3.GT.A1).AND.(PRB.GT.O.OOOO75).AND.A3.GT.1.OOOO
     1 O)  GO TO 5O
40    MB=MB+1
      IF (MB.LE.NW)  GO TO 2O
      IF (MB.GT.NW)  GO TO 7O
50    I=I+1
      H(I)=K1
      I=I+1
      H(I)=K2
      PRINT  55,H(I-1),H(I)
55    FORMAT('O',1OX,'POTENTIAL  BETA-TURN',5X,I4,5X,I4)
C
C
C     TO CHECK FOR THE POSSIBLE PRESENCE OF AN ADJACENT TURN
C
      IF (I.LE.3)  GO TO 6O
      IF (K1.EQ.(H(I-3)+1))  GO TO 8O
C
60    MB=K1+1
      IM=I
      IF (MB.LE.NW)  GO TO 2O
70    PRINT  75,IM
75    FORMAT('O',1OX,'END OF PROGRAM',5X,I6)
      GO TO 9O
C
C
C     TO CALCULATE THE PROBABILITY OF OCCURRENCE OF THE ADJACENT TURN
C
80    KO=H(I-3)
      PRBO=O
      PRBO=P(KO,1)*P(KO+1,2)*P(KO+2,3)*P(KO+3,4)
      IF (PRBO.GT.PRB)  PRINT  85,PRBO,PRB,K1,KO
85    FORMAT('O',2OX,'PRBO:',F11.8,4X,'PRB:',F11.8,6X,'B-TURN NOT AT',I5
     1 ,'     BUT AT',I5,/)
```

```
      IF (PRBO.LT.PRB)  PRINT  88,PRBO,PRB,KO,K1
   88 FORMAT('O',20X,'PRBO:',F11.8,4X,'PRB:',F11.8,6X,'B-TURN NOT AT',I5
     1 ,'     BUT AT',I5,/)
      GO TO 60
C
   90 RETURN
      END
End of File .
```

Efficiency of the resolution of overlapping α- and β- areas

In general, the procedure outlined by Chou and Fasman (1978a, 1978b) was effective to solve the dilemma. In the present program, if more than half of the conditions tested (P_α, P_β; character analysis; boundary analysis; ratio of helix length to β-sheet length) favored one of the conformations, then this conformation would be adopted. However, it happened that some cases could not be easily solved because both conformations (α-helix and β-sheet) were equally favored. Although the calculations showed that $\langle P_\alpha \rangle < \langle P_\beta \rangle$, the overlapping section may contain more H_α than H_β, or less B_α than B_β. This may be explained by the higher values of β-sheet conformational parameters compared to helix; thus, they compensate the lower number of occurrence of H_β in the overlapping regions.

For ambiguous situations (e.g. papain 26-33, ribonuclease 49-59, myoglobin 100-119, lysozyme 107-114, subtilisin 269-275, thermolysin 138-150, thermolysin 160-175, thermolysin 175-180, thermolysin 261-274, thermolysin 234-246), more weight was given to factors such as: a) presence of antiparallel β-sheets. According to rule 3 for solving overlapping areas, antiparallel β-sheets are preferentially predicted due to interactions which enhance conformational stability. Thus, in case antiparallel β-sheets are absent,

preference for long α-helix over shorter β-sheet is one of the major factors to be considered; especially when the helical conformation is supported by only half or less than half of the conditions tested. b) ratio of helix length to -sheet length ($R_1 \geq 2.0$) and c) character analysis (to take into account the different types of residues, former, indifferent to, or breaker of α- and β-conformation).

Staphylococcal nuclease 13-18, 30-39; concanavalin 125-133; ribonuclease 94-110; α-chymotrypsin 85-91; papain 161-166 are examples of antiparallel β-sheets being predicted instead of longer helices, although these regions also exhibit good potential for helical conformation.

The reference to known proteins in the prediction of unknown ones is very useful, especially when some homology exists between the known and unknown proteins (Argos et al., 1976). This was observed in the present study for the prediction of proteinase inhibitors.

The following program was written to assess the different important factors contributing to the resolution of overlapping α- and β-regions. An extra part to read pairs of overlapping helices and β-sheets was added to the main program common to the search of α-helix,β-sheet and β-turn.

```
C
C
C      ..... EXTRA PART FOR OVERLAPPING AREAS .....
C      .............................................
C            TO READ IN PAIRS OF OVERLAPPING HELIX AND SHEETS
C
C
C     DESCRIPTION OF PARAMETERS
C        NR  - NUMBER OF LINES OF DATA (16 DATA PER LINE)
C        NT  - TOTAL NUMBER OF DATA
C        AH  - ARRAY TO STORE THE HELICAL VALUES
C              AH(I)     - N-TERMINAL VALUE
C              AH(I+1)   - C-TERMINAL VALUE
C        SH  - ARRAY TO STORE THE SHEET VALUES
C              SH(I)     - N-TERMINAL VALUE
C              SH(I+1)   - C-TERMINAL VALUE
C
C
             PRINT 35
        35   FORMAT('-','PAIRS OF OVERLAPPING HELICES AND SHEETS')
             PRINT 36
        36   FORMAT(' ','.....'/)
             READ(5,40)  NT,NR
        40   FORMAT(6X,I4,6X,I4)
        41   FORMAT(16I5)
             WRITE (6,42)  ((R(J,K),K=1,16),J=1,NR)
        42   FORMAT(' ',16I5)
             IM=NT/2
C
             I=1
             DO 52  J=1,NR
             DO 51  K=1,16,2
             AH(I)=R(J,K)
             IF (AH(I).EQ.0)  GO TO 54
             I=I+1
        51   CONTINUE
        52   CONTINUE
C
        54   I=1
             DO 56  J=1,NR
             DO 55  K=2,16,2
             SH(I)=R(J,K)
             IF (SH(I).EQ.0)  GO TO 60
             I=I+1
        55   CONTINUE
        56   CONTINUE
C
C      TO CALL SUBROUTINE OLA1 TO CARRY OUT THE COMPARISON OF PA,PB OF
C      EACH REGION AND THAT OF THEIR OVERLAPPING AREA
C
```

```
    52          60    CALL OLA1
    53     C
    54                STOP
    55                END
End of File
```

```
      C
      C
      C
                SUBROUTINE  OLA1
      C         .................
      C        OLA1  - PROCEDURE OF OVERLAPPING # 1
      C
      C
      C     ..................................................................
      C     PURPOSE
      C     TO COMPARE THE AVERAGE PA,PB OF THE PREDICTED HELIX (H1-H2),OF THE
      C     SHEET (S1-S2),AND OF THEIR OVERLAPPING AREA AND TO CALCULATE THE
      C     RATIO HELIX LENGTH/SHEET LENGTH
      C     ..................................................................
      C
      C
                REAL A1,A2,S,T1,T2,TTH,TTS,P,HN,HC,NHN,NHC,SN,SC,NSN,NSC,HHF,HF,
                1 IIH,IH,BH,BBH,SSF,SF,IS,BS,BBS
                INTEGER  H1,H2,S1,S2,AH,SH,IT1,IT2,D,R
                DIMENSION  S(1000,20),AH(1000),SH(1000),M(1000),R(1000,16),D(1000,
                1 16),P(1000,10)
                COMMON  A1,A2,S,T1,T2,TTH,TTS,P,HN,HC,NHN,NHC,SN,SC,NSN,NSC,HHF,HF
                1 ,IIH,IH,BH,BBH,SSF,SF,IS,BS,BBS,H1,H2,S1,S2,AH,SH,IT1,IT2,D,R,NR,
                2 NT,NN,N,M,IM,I,K,J
      C
      C
      C     DESCRIPTION OF PARAMETERS
      C        I    - COUNTER
      C        H1   - N-TERMINAL OF THE PREDICTED HELIX
      C        H2   - C-TERMINAL OF THE PREDICTED HELIX
      C        S1   - N-TERMINAL OF THE PREDICTED SHEET
      C        S2   - C-TERMINAL OF THE PREDICTED SHEET
      C        LH   - HELIX LENGTH
      C        LS   - SHEET LENGTH
      C        A1   - AVERAGE PA OF A SECTION
      C        A2   - AVERAGE PB OF A SECTION
      C
      C
      C     EVERY TIME I INCREASES BY 1 A NEW SET OF OVERLAPPING HELIX AND SH
      C     EET IS SUBJECTED TO THE ANALYSIS
      C
                I=1
            1   I=I+1
                IF (I.GT.IM)  GO TO 300
      C
                H1=AH(I-1)
                H2=AH(I)
                S1=SH(I-1)
                S2=SH(I)
                LH=H2-H1+1
```

```
      LS=S2-S1+1
      A1=LH/LS
      PRINT 5,LH,LS,A1
    5 FORMAT('-',20X,'*** COMPARISON OF THEIR LENGTH ***',5X,'L-HELIX:',
     1 I4,3X,'L-SHEET:',I4,5X,'RATIO=LH/LS:',F4.1)
      PRINT 8
    8 FORMAT('0',30X,'***** COMPARISON OF P-HELIX AND P-SHEET *****')
C
      K=1
      GO TO 110
   10 IF (A1.GT.A2)  PRINT 11,H1,H2,A1,A2
   11 FORMAT('0',15X,'H1 :',I4,3X,'H2 :',I4,5X,'A1:',F6.3,3X,'A2:',F6.3,
     1 10X,'A1 > A2   FROM H1 TO H2')
      IF (A1.LT:A2)  PRINT 12,H1,H2,A1,A2
   12 FORMAT('0',15X,'H1 :',I4,3X,'H2 :',I4,5X,'A1:',F6.3,3X,'A2:',F6.3,
     1 10X,'A1 < A2   FROM H1 TO H2')
C
      K=2
      GO TO 120
   20 IF (A1.GT.A2)  PRINT 21,S1,S2,A1,A2
   21 FORMAT('0',15X,'S1 :',I4,3X,'S2 :',I4,5X,'A1:',F6.3,3X,'A2:',F6.3,
     1 10X,'A1 > A2   FROM S1 TO S2')
      IF (A1.LT.A2)  PRINT 22,S1,S2,A1,A2
   22 FORMAT('0',15X,'S1 :',I4,3X,'S2 :',I4,5X,'A1:',F6.3,3X,'A2:',F6.3,
     1 10X,'A1 < A2   FROM S1 TO S2')
C
      IF (SH(I-1).LT.AH(I-1).AND.SH(I).GT.AH(I-1).AND.SH(I).LT.AH(I))
     1  K=3
      IF (SH(I-1).LT.AH(I-1).AND.SH(I).GT.AH(I-1).AND.SH(I).LT.AH(I))
     1  GO TO 130
C
      IF (AH(I-1).LT.SH(I-1).AND.AH(I).GT.SH(I-1).AND.AH(I).LT.SH(I))
     1  K=4
      IF (AH(I-1).LT.SH(I-1).AND.AH(I).GT.SH(I-1).AND.AH(I).LT.SH(I))
     1  GO TO 140
C
C     TO CALL SUBROUTINE OLA2 TO ANALYZE THE TYPES OF RESIDUES WITHIN
C     EACH SECTION
C
   50 CALL OLA2
      GO TO 1
C
C
  110 L1=H1
      L2=H2
      GO TO 200
  120 L1=S1
      L2=S2
      GO TO 200
  130 L1=AH(I-1)
```

```
101                 L2=SH(I)
102                 GO TO 200
103           140   L1=SH(I-1)
104                 L2=AH(I)
105                 GO TO 200
106       C
107       C      TO CALCULATE PA,PB OF THE  REGION L1-L2
108       C
109           200   A1=0
110                 A2=0
111                 T1=0
112                 T2=0
113                 DO 210 L=L1,L2
114                 T1=T1+S(L,1)
115                 T2=T2+S(L,2)
116           210   CONTINUE
117                 A1=T1/(L2+1-L1)
118                 A2=T2/(L2+1-L1)
119       C
120                 IF (K.EQ.1)  GO TO 10
121                 IF (K.EQ.2)  GO TO 20
122                 IF (K.EQ.3)  GO TO 230
123                 IF (K.EQ.4)  GO TO 240
124       C
125           230   PRINT 232
126           232   FORMAT('0',25X,'*** P-HELIX AND P-SHEET OF INTERS. AREA : H1 TO S2
127                1 ***')
128                 IF (A1.GT.A2) PRINT 233, L1, L2,A1,A2
129           233   FORMAT('0',15X,'OL1:',I4,3X,'OL2:',I4,5X,'A1:',F6.3,3X,'A2:',F6.3,
130                1 10X,'A1 > A2  FROM H1 TO S2',/)
131                 IF (A1.LT.A2) PRINT 234, L1, L2,A1,A2
132           234   FORMAT('0',15X,'OL1:',I4,3X,'OL2:',I4,5X,'A1:',F6.3,3X,'A2:',F6.3,
133                1 10X,'A1 < A2  FROM H1 TO S2',/)
134                 GO TO 50
135       C
136           240   PRINT 242
137           242   FORMAT('0',25X,'*** P-HELIX AND P-SHEET OF INTERS. AREA : S1 TO H2
138                1 ***')
139                 IF (A1.GT.A2) PRINT 243, L1, L2,A1,A2
140           243   FORMAT('0',15X,'OL1:',I4,3X,'OL2:',I4,5X,'A1:',F6.3,3X,'A2:',F6.3,
141                1 10X,'A1 > A2  FROM S1 TO H2',/)
142                 IF (A1.LT.A2) PRINT 244, L1, L2,A1,A2
143           244   FORMAT('0',15X,'OL1:',I4,3X,'OL2:',I4,5X,'A1:',F6.3,3X,'A2:',F6.3,
144                1 10X,'A1 < A2  FROM S1 TO H2',/)
145                 GO TO 50
146       C
147           300   PRINT 305
148           305   FORMAT('0',10X,'END OF PROGRAM')
149                 RETURN
150                 END
End of file
```

```
      C
      C
            SUBROUTINE  OLA2
      C     ................
      C     OLA2  - PROCEDURE OF OVERLAPPING # 2
      C
      C
      C     ..................................................................
      C     PURPOSE
      C       TO COMPARE THE TYPES OF RESIDUES (BREAKER,FORMER,INDIFFERENT)
      C       CONTAINED IN THE PREDICTED HELIX (H1-H2),THE SHEET (S1-S2),AND
      C       IN THEIR OVERLAPPING AREA
      C     ..................................................................
      C
      C
            REAL A1,A2,S,T1,T2,TTH,TTS,P,HN,HC,NHN,NHC,SN,SC,NSN,NSC,HHF,HF,
           1 IIH,IH,BH,BBH,SSF,SF,IS,BS,BBS
            INTEGER  H1,H2,S1,S2,AH,SH,IT1,IT2,D,R
            DIMENSION  S(1000,20),AH(1000),SH(1000),M(1000),R(1000,16),D(1000,
           1 16),P(1000,10)
            COMMON  A1,A2,S,T1,T2,TTH,TTS,P,HN,HC,NHN,NHC,SN,SC,NSN,NSC,HHF,HF
           1 ,IIH,IH,BH,BBH,SSF,SF,IS,BS,BBS,H1,H2,S1,S2,AH,SH,IT1,IT2,D,R,NR,
           2 NT,NN,N,M,IM,I,K,J
      C
      C
      C     DESCRIPTION OF PARAMETERS
      C        HHF   - COUNTER FOR STRONG HELIX-FORMER
      C        HF    - COUNTER FOR HELIX-FORMER
      C        IIH   - COUNTER FOR WEAK HELIX-FORMER
      C        IH    - COUNTER FOR HELIX-INDIFFERENT
      C        BH    - COUNTER FOR HELIX-BREAKER
      C        BBH   - COUNTER FOR STRONG HELIX-BREAKER
      C        SSF   - COUNTER FOR STRONG SHEET-FORMER
      C        SF    - COUNTER FOR SHEET-FORMER
      C        IS    - COUNTER FOR SHEET-INDIFFERENT
      C        BS    - COUNTER FOR SHEET-BREAKER
      C        BBS   - COUNTER FOR STRONG SHEET-BREAKER
      C        TTH   - TOTAL OF THE DIFFERENT HELIX COUNTERS =HHF+HF+IIH+IH+BH+
      C                BBH
      C        TTS   - TOTAL OF THE DIFFERENT SHEET COUNTERS =SSF+SF+IS+BS+BBS
      C        K     - COUNTER
      C
      C
            PRINT 5
        5   FORMAT('O',32X,'*** COMPARISON OF ASSIGNMENTS TYPES ***'/)
      C
            K=1
            GO TO 110
       10   IF (TTH.GT.TTS)  PRINT 13,H1,H2,TTH,TTS
       13   FORMAT('O',15X,'H1 :',I4,3X,'H2 :',I4,5X,'TTH:',F6.3,3X,'TTS:',F6.
```

```
     1 3,10X,'TTH > TTS  FROM H1 TO H2'/)
       IF (TTH.LT.TTS)  PRINT 14,H1,H2,TTH,TTS
14     FORMAT('O',15X,'H1 :',I4,3X,'H2 :',I4,5X,'TTH:',F6.3,3X,'TTS:',F6.
     1 3,10X,'TTH < TTS  FROM H1 TO H2'/)
C
       K=2
       GO TO 120
20     IF (TTH.GT.TTS)  PRINT 23,S1,S2,TTH,TTS
23     FORMAT('O',15X,'S1 :',I4,3X,'S2 :',I4,5X,'TTH:',F6.3,3X,'TTS:',F6.
     1 3,10X,'TTH > TTS  FROM S1 TO S2'/)
       IF (TTH.LT.TTS)  PRINT 24,S1,S2,TTH,TTS
24     FORMAT('O',15X,'S1 :',I4,3X,'S2 :',I4,5X,'TTH:',F6.3,3X,'TTS:',F6.
     1 3,10X,'TTH < TTS  FROM S1 TO S2'/)
C
C      TO CHECK THE BOUNDARIES OF THE OVERLAPPING AREA. IN CASE B-SHEET
C      IS CONTAINED WITHIN A-HELIX NO NEED TO CARRY OUT THE ANALYSIS FOR
C      THE OVERLAPPING AREA AGAIN.
C
       IF (SH(I-1).LT.AH(I-1).AND.SH(I).GT.AH(I-1).AND.SH(I).LT.AH(I))
     1  K=3
       IF (SH(I-1).LT.AH(I-1).AND.SH(I).GT.AH(I-1).AND.SH(I).LT.AH(I))
     1  GO TO 130
C
       IF (AH(I-1).LT.SH(I-1).AND.AH(I).GT.SH(I-1).AND.AH(I).LT.SH(I))
     1  K=4
       IF (AH(I-1).LT.SH(I-1).AND.AH(I).GT.SH(I-1).AND.AH(I).LT.SH(I))
     1  GO TO 140
C
C      TO CALL SUBROUTINE OLA3 TO CARRY OUT THE BOUNDARY ANALYSIS OF EACH
C      REGION
C
50     CALL OLA3
       RETURN
C
110    L1=H1
       L2=H2
       GO TO 200
120    L1=S1
       L2=S2
       GO TO 200
130    L1=AH(I-1)
       L2=SH(I)
       GO TO 200
140    L1=SH(I-1)
       L2=AH(I)
       GO TO 200
C
C      TO CALCULATE THE DIFFERENT TYPES OF RESIDUES IN THE REGION L1-L2
C
200    HHF=0
```

```
101                HF=0
102                IIH=0
103                IH=0
104                BH=0
105                BBH=0
106                SSF=0
107                SF=0
108                IS=0
109                IS=0
110                BS=0
111                BBS=0
112                TTH=0
113                TTS=0
114       C
115                DO 210  L=L1,L2
116                IF (S(L,1).GT.1.16)  HHF=HHF+2.00
117                IF (S(L,1).GT.1.01.AND.S(L,1).LE.1.16)  HF=HF+1.00
118                IF (S(L,1).GT.0.98.AND.S(L,1).LE.1.01)  IIH=IIH+0.50
119                IF (S(L,1).GT.0.69.AND.S(L,1).LE.0.98)  IH=IH+0.25
120                IF (S(L,1).GT.0.57.AND.S(L,1).LE.0.69)  BH=BH-0.50
121                IF (S(L,1).LE.0.57)  BBH=BBH-1.00
122                IF (S(L,2).GT.1.38)  SSF=SSF+2.00
123                IF (S(L,2).GT.0.93.AND.S(L,2).LE.1.38)  SF=SF+1.00
124                IF (S(L,2).GT.0.75.AND.S(L,2).LE.0.93)  IS=IS+0.25
125                IF (S(L,2).GT.0.55.AND.S(L,2).LE.0.75)  BS=BS-0.50
126                IF (S(L,2).LE.0.55)  BBS=BBS-1.00
127           210  CONTINUE
128                TTH=HHF+HF+IIH+IH+BH+BBH
129                TTS=SSF+SF+0.0+IS+BS+BBS
130                PRINT 211
131           211  FORMAT(' ',11X,'HHF',6X,'HF',5X,'IIH',6X,'IH',6X,'BH',5X,'BBH',5X,
132               1 'SSF',6X,'SF',6X,'IS',6X,'BS',5X,'BBS')
133                PRINT 212,HHF,HF,IIH,IH,BH,BBH,SSF,SF,IS,BS,BBS
134           212  FORMAT(' ',10X,11(F5.2,3X))
135       C
136                IF (K.EQ.1)  GO TO 10
137                IF (K.EQ.2)  GO TO 20
138                IF (K.EQ.3)  GO TO 230
139                IF (K.EQ.4)  GO TO 240
140       C
141           230  PRINT 231
142           231  FORMAT('0',28X,'*** ASSIGNM. TYPES IN OVERL. AREAS : H1 TO S2  ***
143               1')
144                IF (TTH.GT.TTS) PRINT 235, L1, L2,TTH,TTS
145           235  FORMAT('0',15X,'OL1:',I4,3X,'OL2:',I4,5X,'TTH:',F6.3,3X,'TTS:',F6.
146               1 3,10X,'TTH > TTS  FROM H1 TO S2'/)
147                IF (TTH.LT.TTS) PRINT 236, L1, L2,TTH,TTS
148           236  FORMAT('0',15X,'OL1:',I4,3X,'OL2:',I4,5X,'TTH:',F6.3,3X,'TTS:',F6.
149               1 3,10X,'TTH < TTS  FROM H1 TO S2'/)
150                GO TO 50
```

```
    C
  240   PRINT 241
  241   FORMAT('O',28X,'*** ASSIGNM. TYPES IN OVERL. AREAS : S1 TO H2 ***
       1')
        IF (TTH.GT.TTS) PRINT 245, L1, L2,TTH,TTS
  245   FORMAT('O',15X,'OL1:',I4,3X,'OL2:',I4,5X,'TTH:',F6.3,3X,'TTS:',F6.
       1 3,10X,'TTH > TTS  FROM S1 TO H2'/)
        IF (TTH.LT.TTS) PRINT 246, L1, L2,TTH,TTS
  246   FORMAT('O',15X,'OL1:',I4,3X,'OL2:',I4,5X,'TTH:',F6.3,3X,'TTS:',F6.
       1 3,10X,'TTH < TTS  FROM S1 TO H2'/)
        GO TO 50
        END
End of File
```

```
      1
      C
      C
            SUBROUTINE  OLA3
      C     .................
      C     OLA3  - PROCEDURE OF OVERLAPPING # 3
      C
      C
      C     ....................................................................
      C     PURPOSE
      C        TO COMPARE THE SUM OF THE BOUNDARY CONFORMATIONAL PARAMETERS OF
      C        THE PREDICTED HELIX AND SHEET. ONLY THE 3 RESIDUES BELONGING TO
      C        THE BOUNDARIES OF EACH SECTION AND THOSE 3 ADJACENT TO THE BOUN
      C        DARIES ARE CONSIDERED
      C     ....................................................................
      C
      C
            REAL A1,A2,S,T1,T2,TTH,TTS,P,HN,HC,NHN,NHC,SN,SC,NSN,NSC,HHF,HF,
           1 IIH,IH,BH,BBH,SSF,SF,IS,BS,BBS
            INTEGER  H1,H2,S1,S2,AH,SH,IT1,IT2,D,R
            DIMENSION  S(1000,20),AH(1000),SH(1000),M(1000),R(1000,16),D(1000,
           1 16),P(1000,10)
            COMMON  A1,A2,S,T1,T2,TTH,TTS,P,HN,HC,NHN,NHC,SN,SC,NSN,NSC,HHF,HF
           1 ,IIH,IH,BH,BBH,SSF,SF,IS,BS,BBS,H1,H2,S1,S2,AH,SH,IT1,IT2,D,R,NR,
           2 NT,NN,N,M,IM,I,K,J
      C
      C
      C     DESCRIPTION OF PARAMETERS
      C        HN  - SUM OF THE BOUNDARY CONFORMATIONAL PARAMETERS OF THE 3 RE
      C              SIDUES BELONGING TO THE HELIX N-TERMINAL
      C        HC  - SUM OF THE BOUNDARY CONFORMATIONAL PARAMETERS OF THE 3 RE
      C              SIDUES BELONGING TO THE HELIX C-TERMINAL
      C        NHN - SUM OF THE BOUNDARY CONFORMATIONAL PARAMETERS OF THE 3 RE
      C              SIDUES ADJACENT TO THE HELIX N-TERMINAL
      C        NHC - SUM OF THE BOUNDARY CONFORMATIONAL PARAMETERS OF THE 3 RE
      C              SIDUES ADJACENT TO THE HELIX C-TERMINAL
      C
      C     REMARKS
      C        THE DEFINITIONS OF SN,SC,NSN,SNC ARE SIMILAR TO HN,HC,NHN,HHC
      C        EXCEPT THAT SHEET IS CONSIDERED INSTEAD OF HELIX
      C
            HN=0
            HC=0
            NHN=0
            NHC=0
            HN=S(H1,8)+S(H1+1,8)+S(H1+2,8)
            HC=S(H2,9)+S(H2-1,9)+S(H2-2,9)
            IF ((H1-3).LE.0)  NHN=0
            IF ((H1-3).LE.0)  GO TO 1
            NHN=S(H1-1,6)+S(H1-2,6)+S(H1-3,6)
```

```
      1      IF ((H2+3).GT.NN)  NHC=0
      2      IF ((H2+3).GT.NN)  GO TO 2
             NHC=S(H2+1,7)+S(H2+2,7)+S(H2+3,7)
C
      2      SN=0
             SC=0
             NSN=0
             NSC=0
             SN=S(S1,10)+S(S1+1,10)+S(S1+2,10)
             SC=S(S2,11)+S(S2-1,11)+S(S2-2,11)
             IF ((S1-3).LE.0)  NSN=0
             IF ((S1-3).LE.0)  GO TO 3
             NSN=S(S1-1,12)+S(S1-2,12)+S(S1-3,12)
      3      IF ((S2+3).GT.NN)  NSC=0
             IF ((S2+3).GT.NN)  GO TO 4
             NSC=S(S2+1,13)+S(S2+2,13)+S(S2+3,13)
C
      4      PRINT 10,H1,H2,S1,S2
     10      FORMAT('0',12X,' BOUNDARY ANALYS. FOR HELIX FROM:',I5,'  TO:',
     1 I5,3X,'AND FOR SHEET FROM:',I5,'  TO:',I5/)
             PRINT 11
     11      FORMAT(' ',12X,'HN',7X,'SN',7X,'HC',7X,'SC',6X,'NHN',6X,'NSN',6X,
     1 'NHC',6X,'NSC')
             PRINT 12,HN,SN,HC,SC,NHN,NSN,NHC,NSC
     12      FORMAT(' ',10X,8(F5.2,4X)//)
             I=I+1
             RETURN
             END
End of File
```

Comparison of the predictive accuracy

In Tables 5 and 6 (p. 189 and 198) the results of Chou and Fasman (1974b), those of X-ray analysis (Chou and Fasman, 1974b) and those obtained before and after refinement of the program of this study are pooled together. The prediction of the different conformations of lysozyme (egg white) was chosen as an example of the output yielded by the present program (Appendix). The quality of prediction was assessed by the parameters Q_α, Q_β and the coefficients C_α, C_β calculated for most of the proteins used in this study (X-ray data were not available for some proteins). For the β-turn search, as results of the present study were almost the same as those of Chou and Fasman (1979), it is reasonable to assume that the accuracy obtained in this study is comparable to that of Chou and Fasman (1979). Although Chou and Fasman reported their results and compared them to X-ray data (Chou and Fasman, 1979), this entire procedure will not be repeated again.

Tables 7 and 8 (p. 207 and 209) list the values of Q_α, C_α and Q_β, C_β, respectively as obtained by Chou and Fasman (1974b) and by the present program. As an extra reference, values reported by Argos et al. (1976), who used joint prediction histograms resulting from the combination of five computerized methods (including the method of Chou and Fasman) were also used

The good agreement between X-ray data and the prediction from this study ($C \geq 0.40$), except for concanavalin A and α-chymotrypsin ($C_\alpha = 0.39$), was expected since this was the aim in refining the program. In general, the predictive accuracy reported by Chou and Fasman (1974b) was clearly superior to that of Argos et al. (1976), except for cytochrome c and myogen (Q_α, C_α). The paired-sample t-test revealed that the values of C_α ($P < 0.01$) and C_β ($P \leq 0.05$) calculated for the present prediction were significantly improved from the values of Chou and Fasman (1974b). One may argue about the validity of the parameters Q and C in this study, and the reliability of the present program when applied to unknown proteins since the present program was adjusted to fit X-ray data on the basis of a limited number of samples (24 proteins). Chou and Fasman (1978b) studied the influence of neighbouring residues (n-1) and (n+1) in dipeptides and tripeptides on the conformation of amino acid n. They noted that the interactions of some residues with high α-helix or β-sheet potential may result in dipeptides or tripeptides with much lower conformational parameters (e.g., the combination of Lys and Glu). Hence efforts to improve the quality of the predictive methods are still necessary and one may expect that the eventual program will become more and more complicated since so many different factors must be considered in order to obtain good agreement with X-ray data. Argos et al.

(1976) suggested that a perfect predictive algorithm should include a consideration of energy minimization, thermalization, and long-range interactions. In their study, the use of joint prediction histograms, which were shown to be superior to any individual prediction, did not always yield good agreement with X-ray data. Hence, in the present study, the modifications made to the present program are not completely useless because if the models used are adjusted to fit experimental data, they can still provide some useful guidelines for unknown systems. In fact, in refining the present program, more consideration to the influence of the neighbouring residues, especially at the boundaries, and to the conformational potential of the adjacent segments, was emphasized. As a result, the number of overlapping areas between α-helix and β-sheet or between α-helix and β-turn was decreased (Table 5). In general, the predicted regions also had boundary residues with favorable conformational parameters. Hence, at least one may be confident that areas with strong potential for a specific conformation will not be missed when using the present program. Special situations may not permit the attainment of satisfactory results.

Argos et al. (1976) and Matthews (1975a) agreed that no favorable prediction can be expected for unknown proteins unless they possess some common organization with the known ones through sequence homology. Further modifications

of the present program will be made when additional data or extra rules for the predictive algorithm are reported by Chou and Fasman or other researchers. As it has been emphasized by Fasman (1980), the lack of high accuracy of the present predictive methods should not stop researchers from using them to obtain a suggestive model for proteins. This will partially help to get an insight on protein behavior while X-ray data are not yet available.

In summary, in the present study, the major framework for the secondary structure search based on the method of Chou and Fasman (1978a, 1978b) has been computerized. Extra modifications which will be necessitated by the advent of improvements in the predictive methods are not perceived as being of any great obstacle to the use of the basic programs developed in this study.

Table 5. Comparison of Experimental[a] (X-ray) and Predicted Helical Regions Obtained by Chou and Fasman[b] and by the Present Program Before and After its Refinement.

	Present Program		Chou & Fasman	X-Ray
	Before	After		
Adenylate Kinase	1-14	1-9	1-8	1-9
(194 aa)	23-31	23-28	-[c]	23-31
	39-49	41-48	40-48	41-48
	51-67	52-67	55-68	52-64
	69-88	69-86	69-86	70-86
	97-109	98-108	97-109	99-108
	123-132	123-132	123-132	124-133
	138-152	143-156	142-151	142-157
	157-167	157-165	157-164	159-162
	178-194	180-194	186-194	178-194
Carboxypeptidase A	19-25	14-29	13-29	14-28
(307 aa)	79-85	72-88	72-88	72-88
	97-110	98-102	98-102	94-103
	-	116-122	116-122	112-122
	170-182	173-186	173-184	173-187
	215-233	215-233	215-233	215-231
	-	-	-	254-262
	286-292	288-305	289-305	288-306
	297-302			

(cont'd)

Table 5. (cont'd)

Comparison of Experimental[a] (X-Ray) and Predicted Helical Regions Obtained by Chou and Fasman[b] and by the Present Program Before and After its Refinement

	Present Program		Chou & Fasman	X-Ray
	Before	After		
Concanavalin A-	32-40	38-42	38-43	-
Jack Bean	42-47[d]	-	-	-
(237 aa)	-	80-85	81-86	81-85
	-	155-160	155-160	-
	180-188	178-190	180-189	-
α-Chymotrypsin	53-58	55-60	55-60	-
(245 aa)	76-90	78-84	78-84	-
	111-116	111-116	111-116	-
	-	-	-	164-173
	238-244	233-245	233-245	234-245
Cytochrome b_5	1-6	-	-	-
(93 aa)	7-15	8-15	9-15	8-15
	31-39	33-39	34-39	33-38
	42-51	42-49	43-50	42-49
	53-76	54-74	54-61	55-62
			65-74	64-74

(cont'd)

Table 5. (cont'd)

Comparison of Experimental[a] (X-Ray) and Predicted Helical Regions Obtained by Chou and Fasman[b] and by the Present Program Before and After its Refinement

	Present Program		Chou & Fasman	X-Ray
	Before	After		
Cytochrome c (104 aa)	2-22	2-20	2-13	9-13
			14-21	14-18
	-	-	-	49-54
	55-69	59-69	59-69	62-70
	-	-	-	71-75
	77-102	89-101	88-101	91-101
α-Hemoglobin (141 aa)	1-8	4-17	4-17	3-18
	8-17			
	25-34	21-36	20-36	20-35
	-	-	-	36-42
	45-64	53-73	53-73	52-71
	68-76			
	79-94	79-94	79-84	80-89
			86-93	
	98-103	94-113	96-113	94-112
	120-129	120-138	120-138	118-138

(cont'd)

Table 5. (cont'd)

Comparison of Experimental[a] (X-Ray) and Predicted Helical Regions Obtained by Chou and Fasman[b] and by the Present Program Before and After its Refinement.

	Present Program		Chou & Fasman	X-Ray
	Before	After		
β-Hemoglobin	1-23	6-23	6-23	4-18
(146 aa)	26-35	26-34	26-34	19-34
	37-45	-	-	35-41
	-	51-56	51-55	50-56
	59-71	58-78	59-71	57-76
	73-82		73-78	
	82-99	85-97	85-98	85-94
	101-106	98-117	101-118	99-117
	106-118			
	122-129	123-143	122-135	123-143
	129-135		137-145	
	137-144			
Lysozyme	3-15	7-15	7-15	5-15
(129 aa)	27-36	27-35	27-35	25-35
	-	79-84	79-84	79-84
	90-98	89-99	88-99	88-99
	105-112	107-114	107-114	108-115
	-	119-124	119-125	119-124

(cont'd)

Table 5. (cont'd)

Comparison of Experimental[a] (X-Ray) and Predicted Helical Regions Obtained by Chou and Fasman[b] and by the Present Program Before and After its Refinement.

	Present Program		Chou & Fasman	X-Ray
	Before	After		
Myogen	1-6	1-9	1-6	-
(108 aa)	5-24	9-19	8-21	7-15
	24-55	26-33	26-33	26-33
		40-50	40-52	40-51
	59-79	57-77	57-77	67-71
	81-92	81-88	81-88	78-89
	96-108	100-108	99-108	102-107
Myoglobin	1-11	4-22	4-22	3-18
(153 aa)	13-22			
	24-36	22-36	24-36	20-35
	38-64	37-43	38-43	36-42
		48-77	48-57	51-57
	66-87		58-77	58-77
		81-96	81-85	-
	89-99		86-97	86-95
	101-119	100-119	101-119	100-118
	123-145	123-149	123-128	124-149
			130-149	

(cont'd)

Table 5. (cont'd)

Comparison of Experimental[a] (X-Ray) and Predicted Helical Regions Obtained by Chou and Fasman[b] and by the Present Program Before and After its Refinement.

	Present Program		Chou & Fasman	X-Ray
	Before	After		
Papain	5-10	-	-	-
(212 aa)	24-30	26-35	26-35	24-41
	47-60	50-58	50-57	50-57
	69-74	68-77	68-77	67-78
	-	118-126	120-126	117-126
	133-143	136-143	136-143	138-143
Ribonuclease S	1-23	2-13	2-13	3-13
(124 aa)	26-33	28-35	28-35	24-35
	45-61	49-59	49-59	50-59
Staphylococcal	3-10	5-10	5-10	-
Nuclease	-	56-76	56-67	54-67
(149 aa)	57-78		69-76	-
	94-106	98-106	98-110	99-106
	120-137	122-137	121-142	122-134

(cont'd)

Table 5. (cont'd)

Comparison of Experimental[a] (X-Ray) and Predicted Helical Regions Obtained by Chou and Fasman[b] and by the Present Program Before and After its Refinement.

	Present Program		Chou & Fasman	X-Ray
	Before	After		
Subtilisin BPN'	8-13	-	-	5-10
(275 aa)	15-20	15-19	13-19	14-20
	69-75	66-75	64-75	64-73
	110-120	111-116	111-116	103-117
	130-145	132-145	132-145	132-145
	195-200	195-200	195-200	-
	226-238	223-238	222-238	223-238
	-	-	-	242-252
	267-275	269-275	267-275	269-275
Thermolysin	53-59	55-60	53-58	-
(316 aa)	67-74	67-77	67-74	65-88
	136-144	137-150	137-150	137-152
	163-172	160-180	158-180	159-180
	175-180			
	236-241	234-246	238-246	235-246
	261-267	261-273	261-271	259-274
	280-295	281-295	281-295	280-296
	299-313	302-313	301-313	302-313

(cont'd)

Table 5. (cont'd)

Comparison of Experimental[a] (X-Ray) and Predicted Helical Regions Obtained by Chou and Fasman[b] and by the Present Program Before and After its Refinement.

	Present Program		Chou & Fasman	X-Ray
	Before	After		
Pancreatic Trypsin Inhibitor (58 aa)	-	2-7	2-7	3-6
	44-55	45-55	45-54	45-56
Myohemerithrin (118 aa)	19-29	22-37	19-37	19-38
	33-39			
	44-51	46-63	-	40-62
	53-66		58-65	
	68-85	68-84	70-84	69-87
	91-104	92-110	86-96	93-110
	106-115		100-108	
Thioredoxin (108 aa)	12-19	10-19	12-19	11-18
	38-48	38-48	38-48	34-49
	-	-	-	59-63
	84-90	85-91	85-91	-
	98-108	98-108	98-108	95-107

(cont'd)

Table 5. (cont'd)

Comparison of Experiment[a] (X-Ray) and Predicted Helical Regions Obtained by Chou and Fasman and by the Present Program Before and After its Refinement.

	Present Program		Chou & Fasman	X-Ray
	Before	After		
Glucagon[e] (29 aa)	14-27	15-27	19-27	
Bovine Colostrum Inhibitor[e] (67 aa)	9-14	5-10	5-10	
	17-23	-	-	
	48-59	49-59	48-56	
Russell's Viper Toxin[e] (60 aa)	27-36	-	-	
	-	47-55	47-55	
Black Mamba Toxin K[e] (57 aa)	44-53	44-53	45-51	

[a]References to the X-ray data are given by Chou and Fasman (1974b).

[b]Predicted values reported by Chou and Fasman (1974b)

[c]Region omitted in prediction

[d]Overpredicted region

[e]The results of Chou and Fasman (1978b) serve as reference values.

Table 6. Comparison of Experimental[a] (X-Ray) and Predicted β-Sheet Regions Obtained by Chou and Fasman[b] and by the Present Program Before and After its Refinement.

	Present Program		Chou & Fasman	X-Ray
	Before	After		
Adenylate Kinase (194 aa)	9-14	10-14	10-15	10-15
	27-39	29-39	26-35	34-39
	-	-	80-85	-
	89-92	90-95	88-93	89-95
	113-118	113-118	110-118	114-118
	-	-	151-157	-
	169-175	169-174	169-175	169-175
	182-188	-	182-187	-
Carboxypeptidase A (307 aa)	33-42	32-38	32-38	32-36
	47-52	47-52	47-52	49-53
	62-66	61-66	61-68	60-67
	105-107	103-110	103-111	104-109
	125-133	125-132	-	-
	137-141	137-141	137-141	-
	189-195	189-195	191-195	190-196
	200-204	200-204	200-204	200-204
	206-211	206-211	206-211	-
	233-234	-	234-238	-
	-	-	-[c]	239-241
	243-248	243-249	243-249	-
	263-269	263-269	261-269	265-271
	277-281	277-281[d]	277-281	-

(cont'd)

Table 6. (cont'd)

Comparison of Experimental[a] (X-Ray) and Predicted β-Sheet Regions Obtained by Chou and Fasman[b] and by the Present Program Before and After its Refinement.

	Present Program		Chou & Fasman	X-Ray
	Before	After		
Concanavalin A	1-7	3-7	3-12	4-9
Jack Bean	9-12	9-12		
(237 aa)	25-32	25-29	25-29	25-29
	49-57	47-55	47-55	48-55
	60-65	61-67	60-67	60-67
	79-82	73-79	73-80	73-78
	88-93	88-97	88-96	92-97
	106-109	105-115	106-113	106-116
	125-132	125-133	124-134	124-132
	137-143	140-143	140-144	140-144
	172-177	173-177	173-177	173-177
	193-199	191-199	190-200	190-199
	209-217	210-215	209-215	209-215
	226-230	228-232	229-234	-
α-Chymotrypsin	29-33	29-34	29-34	29-35
(245 aa)	34-42	39-47	39-47	39-47
	51-54	51-54	50-54	50-54
	61-67	61-67	61-68	65-68
	88-91	85-91	85-89	86-91

(cont'd)

Table 6. (cont'd)

Comparison of Experimental[a] (X-Ray) and Predicted β-Sheet Regions Obtained by Chou and Fasman[b] and by the Present Program Before and After its Refinement.

	Present Program		Chou & Fasman	X-Ray
	Before	After		
α-Chymotrypsin (245 aa) (cont'd)	103-108	103-108	103-108	103-108
	117-122	117-122	117-123	119-122
	134-143	134-146	134-146	134-140
	140-146			
	154-158	155-163	155-163	155-163
	180-182	179-183	179-184	179-184
	-	199-204	197-201	199-203
	207-213	206-213	206-214	206-214
	227-232	227-232	226-232	226-230
Cytochrome b_5 (93 aa)	2-9	4-7	4-8	4-6
	20-28	21-29	21-25	21-25
	30-33	29-33	29-33	28-32
	-	-	-	50-54
	72-76	73-79	75-79	75-79
Cytochrome C (104 aa)	31-36	32-36	-	-
	45-49	-	46-50	-
	78-83	78-82	80-85	-

(cont'd)

Table 6. (cont'd)

Comparison of Experimental[a] (X-Ray) and Predicted β-Sheet Regions Obtained by Chou and Fasman[b] and by the Present Program Before and After its Refinement.

	Present Program		Chou & Fasman	X-Ray
	Before	After		
α-Hemoglobin (141 aa)	36-39	38-43	38-43	-
	40-43	-	-	-
β-Hemoglobin (146 aa)	37-45	37-42	35-42	-
Lysozyme	1-6	2-6	2-6	1-3
	38-46	38-46	38-43	38-46
	53-59	51-59	50-58	50-54
	56-65	-	-	57-60
Myogen (108 aa)	-	-	-	-
Myoglobin (153 aa)	-	-	-	-

(cont'd)

Table 6. (cont'd)

Comparison of Experimental[a] (X-Ray) and Predicted β-Sheet Regions Obtained by Chou and Fasman[b] and by the Present Program Before and After its Refinement.

	Present Program		Chou & Fasman	X-Ray
	Before	After		
Papain	-	4-9	4-9	5-7
(212 aa)	37-45	37-40	37-42	-
	78-82	78-82	-	-
	91-94	91-95	91-95	-
	110-113	110-113	110-114	111-112
	130-136	130-134	130-135	-
	161-166	161-166	161-167	162-167
	170-173	170-175	170-174	169-175
	186-188	184-189	185-189	185-191
	197-201	199-208	199-208	206-208
	202-205			
Ribonuclease S	43-48	43-47	43-48	41-48
(124 aa)	-	61-65	60-65	60-65
	69-82	69-76	69-76	69-76
	-	79-84	79-85	79-87
	94-110	95-110	95-102	96-110
			105-110	
	-	116-124	115-124	116-124

(cont'd)

Table 6. (cont'd)

Comparison of Experimental[a] (X-Ray) and Predicted β-Sheet Regions Obtained by Chou and Fasman[b] and by the Present Program Before and After its Refinement.

	Present Program		Chou & Fasman	X-Ray
	Before	After		
Staphylococcal Nuclease (149 aa)	12-15	13-18	12-18	12-19
	22-27	22-27	22-27	21-27
	32-41	30-39	32-41	30-36
	87-94	89-94	88-94	-
	108-115	111-115	111-115	-
Subtilisin BPN′ (275 aa)	8-11	4-11	4-11	-
	26-31	26-32	28-32	28-32
	-	44-51	44-51	45-50
	81-84	81-84	79-84	-
	90-96	89-96	89-96	89-94
	103-111	103-108	103-108	-
	116-124	119-124	119-124	120-124
	147-150	147-152	147-152	148-152
	-	174-180	174-180	-
	203-207	203-209	205-209	-
	241-246	241-246	241-246	-
	250-257	250-255	250-255	-

(cont'd)

Table 6. (cont'd)

Comparison of Experimental[a] (X-Ray) and Predicted β-Sheet Regions Obtained by Chou and Fasman[b] and by the Present Program Before and After its Refinement.

	Present Program		Chou & Fasman	X-Ray
	Before	After		
Thermolysin (316 aa)	1-4	4-13	4-17	4-13
	7-9	14-20	-	-
	17-33	21-33	20-32	15-32
	39-42	39-42	37-50	37-46
	41-50	41-50		
	-	-	-	52-58
	61-66	61-66	61-66	60-63
	71-84	78-84	75-84	-
	-	98-106	98-110	97-106
		108-110		
	110-116	-	-	112-116
	120-122	120-123	120-124	119-123
	128-131	128-131	127-131	-
	148-157	151-157	151-157	-
	192-193	-	192-197	-
	-	-	221-225	-
	249-258	249-258	251-260	-
	266-274	-	272-276	-
Pancreatic Trypsin Inhibitor (58 aa)	18-24	16-24	16-23	16-24
	29-35	27-35	27-38	27-36

(cont'd)

Table 6. (cont'd)

Comparison of Experimental[a] (X-Ray) and Predicted β-Sheet Regions Obtained by Chou and Fasman[b] and by the Present Program Before and After its Refinement.

	Present Program		Chou & Fasman	X-Ray
	Before	After		
Myohemerithrin	13-21	14-21	14-18	-
(118 aa)	47-51	-	44-52	-
Thioredoxin	4-7	4-8	4-8	2-8
(108 aa)	22-25	22-29	22-29	22-29
	52-55	53-60	53-60	53-58
	54-60			
	77-81	77-81	77-81	77-81
	-	-	-	88-91
Glucagon[e]	3-7	6-10	5-10	-
(29 aa)	20-29	20-26	19-27	-
Bovine Colostrum	21-29	21-26	21-26	-
Inhibitor (67 aa)	-	36-38	36-38	-
Russell's Viper	-	5-10	5-9	-
Toxin[e]	20-27	20-27	23-27	--
(60 aa)	34-37	31-37	32-37	-

(cont'd)

Table 6. (cont'd)

Comparison of Experimental[a] (X-Ray) and Predicted β-Sheet Regions Obtained by Chou and Fasman[b] and by the Present Program Before and After its Refinement.

	Present Program		Chou & Fasman	X-Ray
	Before	After		
Black Mamba	-	4-7	4-9	-
Toxin K[e]	18-23	21-25	21-25	-
(57 aa)	23-31	22-35	29-35	-

[a]References to the X-ray data are given by Chou and Fasman (1974b).

[b]Predicted values reported by Chou and Fasman (1974b).

[c]Region omitted in prediction.

[d]Overpredicted region.

[e]The results of Chou and Fasman (1978b) serve as reference values.

Table 7. Agreement Factors Q_α, C_α obtained by Chou and Fasman[a], Argos et al.[b], and the Present Program[c]

	Q_α			C_α		
Carboxypeptidase A (bovine)	89[a]	82[b]	90[c]	.81[a]	.70[b]	.83[c]
Concanavalin A (Jack bean)	95	95	95	.40	.37	.39
α-Chymotrypsin (bovine)	73	64	73	.39	.21	.39
Cytochrome b_5 (bovine)	84	82	89	.69	.67	.79
Cytochrome c (horse)	73	89	74	.45	.78	.48
α-Hemoglobin (horse)	81	72	79	.59	.38	.58
β-Hemoglobin (horse)	83	64	84	.52	.25	.65
Lysozyme (hen egg white)	94	79	94	.89	.59	.91
Myogen (carp)	66	85	69	.35	.72	.42
Myoglobin (sperm whale)	81	72	79	.67	.43	.71

(cont'd)

Table 7. (cont'd)

Agreement Factors Q_α, C_α obtained by Chou and Fasman[a], Argos et al.[b], and the Present Program[c]

	Q_α			C_α		
Papain[d] (papaya)	88	-	89	.81	-	.82
Ribonuclease S[d] (bovine)	93	-	92	.87	-	.87
Staphylococcal nuclease[d]	85	-	87	.60	-	.66
Subtilisin BPN' (B. amylolique-faciens)	80	76	80	.64	.55	.67
Thermolysin (B. thermopro-teolyticus)	85	81	89	.74	.64	.80
Pancreatic trypsin inhibitor (bovine)	90	71	94	.82	.51	.87
Myohemerithrin (T. pyroides)	73	61	87	.42	.20	.70
Thioredoxin[d] (E. coli)	77	-	77	.54	-	.54

[a]Results obtained by Chou and Fasman (1974b)

[b]Results obtained by Argos et al. (1976)

[c]Results obtained by our program.

[d]Proteins not tested by Argos et al. (1976)

Table 8. Agreement Factors Q_β, C_β obtained by Chou and Fasman[a], Argos et al.[b], and the Present Program[c]

	Q_β			C_β		
Carboxypeptidase A (bovine)	83[a]	70[b]	84[c]	.54[a]	.33[b]	.70[c]
Concanavalin A (Jack bean)	90	72	90	.77	.45	.78
α-Chymotrypsin (bovine)	92	75	92	.80	.49	.82
Cytochrome b_5 (bovine)	85	82	86	.73	.67	.74
Cytochrome c[d] (horse)	89	-	90	-	-	-
α-Hemoglobin[d] (horse)	96	-	96	-	-	-
β-Hemoglobin[d] (horse)	95	-	96		--	-
Lysozyme (hen egg white)	83	61	90	.68	.20	.78
Myogen[d] (carp)	100	-	100	-	-	-
Myoglobin[d] (sperm whale)	100	-	100	-	-	-

(cont'd)

Table 8. (cont'd)

Agreement Factors Q_β, C_β obtained by Chou and Fasman[a], Argos et al.[b], and the Present Program[c]

	Q_β			C_β		
Papain[e] (papaya)	88	-	89	.81	-	.82
Myohemerithrin[d] (T. pyroides)	88	-	93	-	-	-
Ribonuclease S[e] (bovine)	93	-	93	.87	-	.87
Staphylococcal nuclease[e]	85	-	88	.57	-	.64
Subtilisin BPN' (B. amyloliquefaciens)	91	63	89	.54	.17	.52
Thermolysin (B. thermoproteolyticus)	75	75	80	.44	.47	.54
Thioredoxin[e]	89	-	85	.81	-	.74
Pancreatic trypsin inhibitor (bovine)	95	79	97	.89	.61	.96

[a]Results obtained by Chou and Fasman (1974b).

[b]Results obtained by Argos et al. (1976).

[c]Results obtained by our program.

[d]Proteins with little or no sheet conformations were not tested by Argos et al. (1976).

[e]Proteins not tested by Argos et al. (1976).

Conformations of some food related proteins

The second objective of this study was to obtain some information on the conformation of food related proteins such as bovine serum albumin (BSA), α_{s1}-casein, β-casein, κ-casein, chymosin, α-lactalbumin, β-lactoglobulin, ovalbumin, pepsin, and trypsinogen. Table 9 lists the percentage of α-helix, β-sheet, and β-turn found for each protein using the modified program. Table 10 shows the possible locations of the different conformations. The schematic diagram corresponding to each of the tested proteins can be found in Figures I to X. Some references were found to corroborate the reliability of the prediction from the present study. Loucheux-Lefebvre et al. (1978), using the method of Chou and Fasman (1974b), obtained 23% α-helix, 31% β-sheet, and 21% β-turn for κ-casein (bovine). These results are quite comparable to those of the present study (20, 33, and 29%). The locations of the different conformations were almost the same, except for helix 90-97 which was predicted as β-sheet by the present program, and helix 62-68 which was not predicted by Loucheux-Lefebvre et al. (1978). α-Lactalbumin was predicted by the method of Lim (1974b) to contain 43% helix and 12% β-sheet compared to 38% helix and 15% β-sheet obtained in the present study. Ovalbumin was reported to be composed of 40% helix by Yang and Doty (1957) using ORD, while a value of 25-30% helix was found by Gorbunoff (1969). Extra ref-

erences would be useful to evaluate the precision of the results of Yang and Doty (1957), of Gorbunoff (1969), and of the present study (44% helix).

The bovine gastric proteases, chymosin and pepsin, are very homologous in their amino acid sequence and their zymogens may even be activated by a similar mechanism (Foltmann et al., 1973). This is partially reflected in the prediction from the present study which yielded high percentages of β-sheet and very low percentages of α-helix for both (40.2 versus 3.7% for chymosin, and 33.4 versus 1.8% for pepsin). The difference in the values between the two enzymes may be explained by difference in their source, chymosin from bovine source and pepsin from porcine source. The pancreatic proteases, α-chymotrypsin and trypsin, also exhibit homology in their primary structure (Huang and Tang, 1970). Hence, it was not surprising to observe a very similar conformational pattern between the two enzymes: 33.5% β-sheet versus 9.0% α-helix for α-chymotrypsin with X-ray diffraction (Chou and Fasman, 1974b), and 31.0% β-sheet versus 13.5% α-helix for trypsin with the present program.

Although no reference was found for BSA, it may be reasonable to compare it to ovalbumin as they both belong to the albumin group. The high percentage of α-helix predicted for BSA (52.1%) may be comparable to that of ovalbumin

(44.1%). However, the percentage of β-sheet was much lower for BSA (2.2%) compared to 20.5% for ovalbumin. α_{s1}-Casein and β-casein were predicted to contain very similar percentages of the three types of conformation (14.6, 26.1, and 30.1% for α_{s1}-casein versus 13.9, 23.0, and 33.0% for β-casein). Unfortunately, there is no reference to check the present results. No reference was found to assess the precision of the prediction for β-lactoglobulin (35.8% α-helix and 30.9% β-sheet)

All the results concerning food related proteins should be considered as suggestive and should be confirmed by other techniques (CD, ORD, X-ray). Nevertheless, one advantage of the method of Chou and Fasman (1987a, 1978b) is that it allows the detection of areas exhibiting potential for both α-helix and β-sheet conformations. Hence, conformational changes observed with CD may be explained by the transitions that those sensitive areas have undergone. All the conformational transition phenomena may help to understand protein functionalities such as gelation, foaming,and emulsifying activity. It has been observed that denaturation of proteins must occur to some extent before those properties are actually exhibited. For instance, for glucagon (29 amino acid residues), it has been hypothesized that the transition from α- to β-conformation of the region 19-27 is necessary for the receptor binding because of the more compact structure

resulting from such a transition (Chou and Fasman, 1978b). It was also observed that glucagon in the gel state has a higher percentage of β-sheet (52%) than glucagon in solution (21%) (Gratzer et al., 1967; Epand, 1971). The predictive method can help to locate the sensitive area 19-27 (Chou and Fasman, 1978b). In sickle cell hemoglobin, the replacement of some α-formers or β-breakers by strong β-formers (Val) results in the transition from α- to β-conformation of the section 1-6. This leads to the aggregation of hemoglobin cells due to interchain interactions replacing intrachain ones (Chou and Fasman, 1978b).

In summary, even though a complete picture of protein behavior cannot be expected without the consideration of the three-dimensional organization which has a great impact on the whole problem, the knowledge of the secondary structure remains one of the useful means to explore the complex nature of proteins.

Table 9. Percentages of Helix, β-Sheet and β-Turn of Some Food Related Proteins Obtained from the Present Program

	Helix (%)	Sheet (%)	Turn (%)
Bovine serum albumin (582 aa)	52.1	2.2	29.6
α_{s1}-Casein (bovine) (199 aa)	14.6	26.1	30.1
β-Casein (bovine) (209 aa)	13.9	23.0	33.0
κ-Casein (bovine) (169 aa)	20.1	33.1	29.0
Chymosin (bovine) (323 aa)	3.7	40.2	36.2
α-Lactalbumin (bovine) (123 aa)	38.2	14.6	37.4
β-Lactoglobulin (bovine) (162 aa)	35.8	30.9	17.3
Ovalbumin (385 aa)	44.1	20.5	19.5
Pepsin (porcine) (326 aa)	1.8	33.4	46.9
Trypsinogen (bovine) (229 aa)	13.5	31.0	35.4

Table 10. Helix, Sheet, and Turn Regions of Some Food Related Proteins as Predicted by the Present Program

	Helix	Sheet
Bovine serum albumin (582 aa)	6-33	403-415
	38-58	
	63-70	
	72-81	
	100-106	
	122-134	
	140-145	
	164-170	
	179-187	
	192-201	
	206-221	
	223-242	
	289-295	
	305-313	
	318-329	
	341-361	
	373-381	
	418-423	
	450-463	
	497-512	
	517-533	
	535-552	
	573-581	

(cont'd)

Table 10. (cont'd)

Helix, Sheet, and Turn Regions of Some Food Related Proteins as Predicted by the Present Program

	Helix	Sheet
Bovine serum albumin (cont'd) (582 aa)		
Turns: 1-4, 34-37, 59-62, 82-85, 88-91, 95-98, 105-108, 107-110, 109-112, 116-119, 118-121, 135-138, 145-148, 155-158, 157-160, 171-174, 188-191, 202-205, 243-246, 245-248, 263-266, 265-268, 270-273, 276-279, 278-281, 284-287, 296-299, 301-304, 314-317, 332-335, 336-339, 363-366, 382-385, 424-427, 431-434, 435-438, 437-440, 443-446, 446-449, 464-467, 471-474, 474-477, 480-483, 482-485, 489-492, 513-516, 553-556, 559-562, 569-572.		
α_{s1}-Casein (bovine) (209 aa)	13-18 34-42 52-65	20-26 30-32 91-95 97-101 135-140 142-146 149-158 163-173

(cont'd)

Table 10. (cont'd)

Helix, Sheet, and Turn Regions of Some Food Related Proteins as Predicted by the Present Program

	Helix	Sheet
α_{s1}-casein (bovine) (cont'd)		
Turns: 1-4, 8-11, 27-29, 43-46, 45-48, 48-51, 66-69, 72-75, 87-89, 87-90, 112-115, 159-162, 174-176, 176-179, 182-185, 184-187, 188-191, 190-193.		
β-casein (bovine) (209 aa)	1-6	23-27
	11-16	39-41
	29-37	52-60
	43-50	92-95
		123-130
		138-143
		160-165
		187-193
Turns: 8-11, 17-20, 61-63, 62-65, 66-69, 71-74, 75-78, 85-88, 104-107, 109-112, 111-114, 136-317, 146-149, 152-155, 158-160, 166-169, 178-181, 180-183, 201-204, 203-206.		

(cont'd)

Table 10. (cont'd)

Helix, Sheet, and Turn Regions of Some Food Related Proteins as Predicted by the Present Program

	Helix	Sheet
κ-casein (bovine)	1-7	22-26
(169 aa)	9-16	28-32
	62-68	38-43
	102-108	48-56
	137-147	72-79
		93-98
		121-126
		159-169
Turns:	18-21, 33-36, 57-60, 69-72, 80-82, 85-88, 99-101, 109-112, 113-116, 127-129, 129-132, 133-136, 149-152, 156-158.	
Chymosin (bovine)	2-6	8-12
(323 aa)	318-323	20-22
		29-33
		40-42
		45-47
		65-69
		82-86
		91-97
		94-103
		105-108
		113-116
		122-126

(cont'd)

Table 10. (cont'd)

Helix, Sheet, and Turn Regions of Some Food Related Proteins as Predicted by the Present Program

	Helix	Sheet
Chymosin (bovine) (cont'd)		136-143
		148-156
		165-171
		180-183
		185-194
		198-204
		212-215
		229-240
		253-255
		275-277
		296-298
		301-303
		306-310

Turns: 13-16, 24-27, 34-37, 36-39, 47-50, 50-53, 52-55, 59-62, 61-64, 76-79, 78-81, 87-90, 109-112, 127-130, 132-135, 144-147, 158-161, 161-164, 172-175, 176, 179, 207-210, 208-211, 216-219, 218-221, 224-227, 226-228, 241-244, 247-250, 250-252, 272-274, 278-280, 279-282, 283-286, 291-294, 293-295, 312-315.

(cont'd)

Table 10. (cont'd)

Helix, Sheet, and Turn Regions of Some Food Related Proteins as Predicted by the Present Program

	Helix	Sheet
α-Lactalbumin (bovine)	1-16	26-31
(123 aa)	89-99	52-59
	104-123	72-75
Turns: 17-20, 32-35, 33-36, 34-37, 43-46, 45-48, 47-50, 48-51, 61-64, 64-67, 66-69, 68-71, 76-79, 82-85, 85-88, 100-103.		
β-Lactoglobulin (bovine)	22-37	1-5
(162 aa)	67-78	12-20
	80-87	39-43
	129-143	56-61
	156-162	92-95
		102-107
		115-123
		145-151
Turns: 6-9, 49-52, 63-66, 88-91, 96-99, 125-128, 152-155.		

(cont'd)

Table 10. (cont'd)

Helix, Sheet, and Turn Regions of Some Food Related Proteins as Predicted by the Present Program

	Helix	Sheet
Ovalbumin (385 aa)	5-23	27-29
	31-41	51-56
	102-109	77-79
	133-143	86-91
	169-189	117-121
	198-206	145-149
	221-232	156-161
	239-245	194-196
	248-259	208-219
	259-268	276-282
	284-290	291-305
	319-334	364-371
	340-362	
	373-379	

Turns: 24-27, 45-48, 47-50, 62-65, 65-68, 71-74, 73-76, 80-83, 92-95, 95-98, 97-100, 125-128, 152-155, 162-165, 165-168, 190-193, 235-238, 245-48, 269-272, 307-310, 311-314.

(cont'd)

Table 10. (cont'd)

Helix, Sheet, and Turn Regions of Some Food Related Proteins as Predicted by the Present Program

	Helix	Sheet
Pepsin (porcine) (326 aa)	65-70	15-21
		26-31
		38-40
		71-75
		83-91
		99-103
		111-115
		140-146
		151-155
		164-167
		179-182
		191-194
		203-205
		211-214
		228-231
		245-249
		259-267
		274-277
		298-313

Turns: 11-14, 22-25, 32-35, 34-37, 35-38, 45-48, 50-53, 52-55, 54-57, 57-60, 59-62, 76-79, 79-82, 94-97, 96-99, 107-110, 116-119, 125-128, 129-132, 137-140, 147-150, 156-159, 158-161, 160-163, 171-174, 175-178, 187-190,

(cont'd)

Table 10. (cont'd)

Helix, Sheet, and Turn Regions of Some Food Related Proteins as Predicted by the Present Program

	Helix	Sheet
Pepsin (porcine) (cont'd) (326 aa)		
Turns:	198-201, 200-203, 206-209, 207-210, 215-218, 217-220, 221-224, 223-226, 232-235, 238-241, 240-243, 250-253, 251-254, 255-258, 268-270, 270-273, 278-281, 279-282, 282-285, 288-291, 292-295, 293-296, 315-318.	
Trypsinogen (bovine) (229 aa)	92-102	12-18
	106-111	21-25
	141-146	28-30
	223-228	52-58
		61-64
		68-71
		82-87
		120-125
		161-172
		193-199
		211-221

(cont'd)

Table 10. (cont'd)

Helix, Sheet, and Turn Regions of Some Food Related Proteins as Predicted by the Present Program

	Helix	Sheet
Trypsinogen (bovine) (cont'd) (229 aa)		
Turns: 3-6, 7-10, 26-27, 32-35, 46-49, 48-51, 65-67, 78-81, 88-91, 103-105, 112-115, 117-119, 126-129, 129-132, 132-135, 134-137, 149-152, 151-154, 154-157, 158-161, 173-175, 175-178, 177-180, 179-182, 181-184, 182-185, 200-203, 205-208, 208-210.		

CONCLUSIONS

A computer program has been written in Fortran language to predict the secondary structure of proteins based on the method of Chou and Fasman (1978a, 1978b), which mainly relies on the frequency of occurrence of each amino acid residue in a certain conformation. This led to the classification of the 20 amino acids as either former, indifferent to, or breaker of the conformations.

Four programs have been designed to locate each type of conformation involved in the secondary structure (α-helix, β-sheet and β-turn) and to solve the possible overlapping α- and β-areas. Each program consists of the main program and several subroutines which correspond to the various steps to be followed in the method (nucleation, propagation and termination), or to the various conditions to be checked ($\langle P_\alpha \rangle$, $\langle P_\beta \rangle$, character assignment, conformational parameters of the boundary residues, and possible presence of antiparallel β-sheets). For the β-turn search, because of the constant number of residues involved (four) and the less complicated predictive rule, the program corresponding to it is much simpler than the other ones.

On testing the present program on 24 different proteins, some missing areas and differences in the boundary residues between the results of the present study and

those of Chou and Fasman were observed. After a thorough analysis of the problem, some modifications were added to the program of this study, including the following. The condition that at least two thirds of formers for helix nucleation may not be satisfied in some cases, although the eventually predicted area met the general requirement of being comprised of one half or more helix formers. Similarly the requirement of less than one third of breakers for β-sheet nucleation may lead to the omission of a potential β-sheet area, although it contains enough β-formers. Hence, the type of residues in the nucleation area , as well as, the surrounding residues may stabilize the area conformation such that the presence of some breakers cannot provoke its disruption. For the boundary residues of the predicted α- and β-areas, the use of the boundary conformational parameters ($P_{\alpha C}$, $P_{\alpha N}$, $P_{\beta C}$, $P_{\beta N}$) results in predictive values closer to those of Chou and Fasman (1974b, 1978b) and of X-ray data (Chou and Fasman, 1974b, 1978b). The use of those parameters also helps to avoid predicting too many overlapping α- and β-areas, or overlapping α-helix and β-turn.

The method outlined by Chou and Fasman (1978a, 1978b) to solve the problem of overlapping α- and β-regions proved to be useful in most cases. However, ambiguous situations may occur where the area under consideration exhib-

its strong potential for both conformations. In such cases more emphasis should be given to the presence of antiparallel β-sheets and to the type of residues present in the area although it may happen that the average $<P_\alpha>$ or $<P_\beta>$ does not support the same conformation as the residue assignment. The ration of length of the predicted α-helix and β-sheet is another useful factor to evaluate the importance of each one. It is not unexpected that for the prediction of unknown proteins which exhibit some homology with known ones, this procedure gives less problems than for completely unknown proteins.

Comparing the predictive accuracy parameters $Q_{\alpha(\beta)}$, and $C_{\alpha(\beta)}$ obtained by Chou and Fasman (1974b), Argos et al. (1976) and the present program, it appears that predictions from the present study and those of Chou and Fasman (1974b) are in general better than those of Argos et al. (1976). The paired-sample t-test revealed that the values of C_α ($P < 0.01$) and C_β ($P \leq 0.05$) calculated for the present prediction were significantly improved from the values of Chou and Fasman (1974b). For most of the proteins used in this study, except for concanavalin A and α-chymotrypsin ($C_\alpha = 0.39$), good agreement with X-ray data ($C \geq 0.40$) was observed as an expected consequence of the modifications given to the present program.

Stimulated by those positive results, the program developed in this study was applied to food related proteins so as to provide a possible means of explaining and predicting food protein behavior under various conditions. Although references could not be found for all of the proteins tested (bovine serum albumin, α_{s1}-casein, β-casein, κ-casein, chymosin, α-lactalbumin, β-lactoglobulin, ovalbumin, pepsin, and trypsinogen), the predicted regions for κ-casein and α-lactalbumin were very similar to those reported by other researchers. They either used the method of Chou and Fasman (Loucheux-Lefebvre et al., 1978), or their own method (Lim, 1974b).

In summary, the main objective of this study to computerize the method of Chou and Fasman (1978a, 1978b) was attained. Extra modifications of the program will be made when additional data or new set of rules (incorporating long-range interaction and energy minimization factors) are published by Chou and Fasman or other researchers. So far most of the predictive methods do not always ensure high predictive accuracy and caution should be given to the prediction of unknown proteins. Nevertheless considering the cost and the lengthy and complex operations involved in the X-ray technique, the predictive algorithms still remain a valuable tool for access to the complicated organization

of proteins while awaiting for confirmation by X-ray analysis. Furthermore, the accuracy of the predictive methods may be improved by combining them with CD or ORD techniques which constitute an additional means to solve ambiguous cases of overlapping α- and β-areas. The percentage of each conformation in proteins can be obtained using these techniques.

LITERATURE CITED

Anglemier, A. F. and Montgomery, M. N., 1976. In "Principles of Food Science", p. 205-284, Ed. Fennema, O. R., Part I, Marcel Dekker Inc.

Argos, P., Schwarz, J. and Schwarz, J., 1976. An assessment of protein secondary structure prediction methods based on amino acid sequence. Biochim. Biophys. Acta 439: 261-273

Anfinsen, C. B., Haber, E., Sela, M. and White, F. H., Jr., 1961. The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. Proc. Natl. Acad. Sci. U.S. 47: 1309-1314

Birktoft, J. J. and Blow, D. M., 1972. Structure of crystalline α-chymotrypsin. V. The atomic structure of tosyl-α-chymotrypsin at 2Å resolution. J. Mol. Biol. 68: 187-240

Blout, E. R., de Loze, C., Bloom, S. M. and Fasman, G. D., 1960. The dependence of the conformation of synthetic polypeptides on amino acid composition. J. Amer. Chem. Soc. 82: 3787-3789

Blout, E. R., 1962. In "Polyamino Acids, Polypeptides and Proteins", p. 275-279, Ed. Stahmann, M. A., University of Wisconsin Press, Madison.

Blundell, T., Dodson, G., Hodgkin, D. and Mercola, D., 1972. Insulin: the structure in the crystal and its reflection in chemistry and biology. Adv. Protein Chem. 26: 279-402

Bradbury, A. F., Smyth, D. G. and Snell, C. R., 1976. Lipotropin: precursor to two biologically active peptides. Biochem. Biophys. Res. Commun. 69: 950-956

Chou, P. Y., Wells, M. and Fasman, G. D., 1972. Conformational studies on copolymers of hydroxypropyl-L-glutamine and L-leucine. Circular dichroism studies. Biochemistry 11: 3028-3043

Chou, P. Y. and Fasman, G. D., 1973. Structural and functional role of leucine residues in proteins. J. Mol. Biol. 74: 263-281

Chou, P. Y. and Fasman. G. D., 1974a. Conformational parameters for amino acids in helical, β-sheet, and random coil regions calculated from proteins. Biochemistry 13: 211-221

Chou, P. Y. and Fasman, G. D., 1974b. Prediction of protein conformation. Biochemistry 13: 222-245

Chou, P. Y. and Fasman, G. D., 1977. β-turns in proteins. J. Mol. Biol. 115: 135-175

Chou, P. Y. and Fasman, G. D., 1978a. Empirical predictions of protein conformation. Ann. Rev. Biochem. 47: 251-276

Chou, P. Y. and Fasman, G. D., 1978b. Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. 47: 45-148

Chou, P. Y. and Fasman, G. D., 1979. Prediction of β-turns. Biophys. J. 26: 367-384

Davies, D. R., 1964. A correlation between amino acid composition and protein structure. J. Mol. Biol. 9: 605-609

Dayhoff, M. O., 1972. Atlas of Protein Swquence and Structure, National Biomedical Research Foundation, Georgetown University Medical Centre, Washington, D.C.

Dayhoff, M. O., 1973. Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Georgetown University Medical Centre, Washington, D.C.

Dayhoff, M. O., 1976. Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Georgetown University Medical Centre, Washington, D.C.

Dayhoff, M. O., 1978. Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Georgetown University Medical Centre, Washington, D.C.

Deber, C. M., Madison, Y. and Blout, E. R., 1976. Why cyclic peptides? Complementary approaches to conformations. Acc. Chem. Res. 9: 106-113

Dickerson, R. F., Takano, T., Eisenberg, D., Kallai., O. B., Samson, L., Cooper, A. and Margoliash, E., 1971. Ferricytochrome c. I. General features of the horse and bonito proteins at 2.8A resolution. J. Biol. Chem. 246: 1511-1533

Dunn, B. M. and Chaiken, I. M., 1975. Relation between α-helical propensity and formation of the ribonuclease S complex. J. Mol. Biol. 95: 497-511

Dzionara, M., Robinson, S. M.L. and Wittman-Liebold, B., 1977. Secondary structure of proteins from the 30S subunit of the Escherichia coli ribosome. Hoppe - Seyler's Z.Physiol. Chem. 358: 1003-1019

Edelman, G. M., Cunningham, B. A., Reeke, G. N., Jr., Beeker, J. W., Waxdal, M. J. and Wang, J. L., 1972. The covalent and three-dimensional structure of concanavalin A. Proc. Natl. Acad. Sci. U. S. 69: 2580-2584

Epand, R. M., 1971. Studies of the conformation of glucagon. Can. J. Biochem. 49: 166-169

Fasman, G. D., 1980. Prediction of protein conformation from the primary structure. Ann. N.Y. Acad. Sci. 348: 147-159

Fink, M. L. and Bodanszky, M. J., 1976. Secretin. VI. Simultaneous "in situ" syntheses of three analogues of the C-terminal tricosapeptide and a study of their conformation. J. Amer. Chem. Soc. 98: 974-977

Foltmann, B., Kauffman, D., Parl, M. and Maack Andersen, P., 1973. Comparison between the primary structure of chymosin (rennin), pepsin and of their zymogens. Neth. Milk Dairy J. 27: 288-297

Garel, A., Kovacs, A. M., Champagne, M. and Daune, M., 1975. Comparison between histones F_{2a} and F_{2a2} of chicken erythrocyte. I. Structure, stability and conformation of the free proteins. Biochim. Biophys. Acta 395: 5-15

Garnier, J., Pernollet, J. C., Tertrin-Clary, C., Salerse, R., Casteing, M., Barnavon, M., de la Llosa and Jutisz, M., 1975. Conformational studies of ovine lutropin (luteinizing hormone) and its native and chemically modified subunits by circular dichroism and ultraviolet absorption spectroscopy. Eur. J. Biochem. 53: 243-254

Goldsack, D. E., 1969. Relation of amino acid composition and the Moffit parameters to the secondary structure of proteins. Biopolymers 7: 299-313

Gorbunoff, M. J., 1969. Exposure of tyrosine residues in proteins. III. The reaction of cyanic fluoride and N-acetyl imidazole with ovalbumin, chymotrypsinogen, and trypsinogen. Biochemistry 8: 2591-2598

Gratzer, W. B., Bailey, E. and Beaven, G. H., 1967. Conformational states of glucagon. Biochem. Biophys. Res. Commun. 28: 914-919

Green, N. M., 1975. Avidin. Adv. Protein Chem. 29: 85-133

Guzzo, A. V., 1965. The influence of amino acid sequence on protein structure. Biophys. J. 5: 809-822

Havsteen, B. H., 1966. A study of the correlation between the amino acid composition and the helical content of proteins. J. Theor. Biol. 10: 1-10

Holladay, I. A. and Puett, D., 1976. Somatostatin conformation: evidence for a stable intramolecular structure from circular dichroism, diffusion, and sedimentation equilibrium. Proc. Natl. Acad. Sci. U.S. 73: 1199-1202

Huang, W. Y. and Tang, J., 1970. Carboxyl-terminal sequence of human gastricsin and pepsin. J. Biol. Chem. 245: 2189-2193

IUPAC - IUB, 1970. Commission on biochemical nomenclature - Abbreviations and symbols for the description of the conformation of polypeptide chains. Tentative rules (1969). Biochemistry 9: 3471-3479

Kabat, E. A. and Wu, T. T., 1973a. The influence of nearest neighbouring amino acid residues on aspects of secondary structure of proteins. Attempt to locate α-helices and β-sheets. Biopolymers 12: 751-774

Kabat, E. A. and Wu, T. T., 1073b. The influence of nearest neighbouring amino acids on the conformation of the middle amino acid in predicted and experimental determination of β-sheets in concanavalin A. Proc. Natl. Acad. Sci. U.S. 70: 1473-1477

Kato, A. and Nakai, S., 1980. Hydrophobicity determined by a fluorescence probe method and its correlation with surface properties of proteins. Biochim. Biophys. Acta 624: 13-20

Kawauchi, H. and Li, C. H., 1974. Reaction of human chorionic somatomammotropin and human pituitary growth hormone with tetranitromethane at 0° C. Arch. Biochem. Biophys. 165: 255-262

Kendrew, J. C., Dickerson, R. E.,Strandberg, B. E. and Davies, D. R., 1960. Structure of myoglobin: a three-dimensional Fourier synthesis at 2Å resolution. Nature 185: 422-427

Kinsella, J. E., 1976. Functional properties of proteins in foods: a survey. Crit. Rev. Food Sci. Nutrit. 7: 219-280

Kopple, K. D., Go, A. and Pilipauskas,D. R., 1975. Studies of peptide conformation. Evidence for β-structures in solutions of linear tetrapeptides containing proline. J. Amer. Chem. Soc. 97: 6830-6838

Kotelchuck, D. and Scheraga, H. A., 1969. The influence of short-range interactions on protein conformation. II. A model for predicting the α-helical regions of proteins. Proc. Natl. Acad. Sci. U.S. 62: 14-21

Leberman, R. J., 1971. Secondary structure of tobacco mosaic virus protein. J. Mol. Biol. 55: 23-30

Lewis, P. N., Go, N., Go, M., Kotelchuck, D. and Scheraga, H. A., 1970. Helix probability profiles of denatured proteins and their correlation with native structures. Proc. Natl. Acad. Sci. U.S. 65: 810-815

Lewis, P. N. and Scheraga, H. A., 1971. Predictions of structural homologies in cytochromes c proteins. Arch. Biochem. Biophys. 144: 576-583

Lewis, P. N., Momany, F. A. and Scheraga, H. A., 1971. Folding of polypeptide chains in proteins: a proposed mechanism for folding. Proc. Natl. Acad. Sci. U.S. 68: 2293-2297

Lewis, P. N., Momany, F. A. and Scheraga, H. A., 1973. Chain reversals in proteins. Biochim. Bio;hys. Acta 303: 211-229

Liljas, A. and Rossman, M. G., 1974. X-ray studies of protein interactions. Ann. Rev. Biochem. 43:475-507

Lim, V. I., 1974a. Structural principles of the globular organization of protein chains. A stereochemical theory of globular protein secondary structure. J. Mol. Biol. 88: 857-872

Lim, V. I., 1974b. Algorithms for prediction of α-helical and β-structural regions in globular proteins. J. Mol. Biol. 88: 873-894

Loucheux-Lefebvre, M. H., Aubert, J. P. and Jolles, P., 1978. Prediction of the conformation of the cow and sheep κ-caseins. Biophys. J. 23: 323-334

Low, B. W., Lovell, F. M. and Rudko, A. D., 1968. Prediction of α-helical regions in proteins of known sequence. Proc. Natl. Acad. Sci. U. S. 60: 1519-1526

Matthews, B. W., 1975a. Comparison of the predicted and observed secondary structure of the T4 phage lysozyme. Biochim. Biophys. Acta 405: 442-451

Matthews, B. W., 1975b. In "The Proteins", 3rd ed., p. 403-590, Eds. Neurath, H. and Hill, R. L., Academic Press, New York

Mathews, F. S., Levine, M. and Argos, P., 1972. Three-dimensional Fourier sybthesis of calf liver cytochrome b_5 at 2.8Å resolution. J. Mol. Biol. 64: 449-464

Muñoz, P. A., Warren, J. R. and Noelken, M. E., 1976. β-Structure of aqueous staphyloccocal enterotoxin B by spectropolarimetry and sequence-based conformational predictions. Biochemistry 15: 4666-4671

Nagano, K., 1973. Logical analysis of the mechanism of protein folding. I. Prediction of helices, loops, and β-structures from primary structure. J. Mol. Biol. 75: 401-420

Nishikawa, K. and Ooi, T., 1972. Tertiary structure of proteins. II. Freedom of dihedral angles and energy calculations. J. Phys. Soc. Jap. 32: 1338-1347

Nishikawa, K., and Ooi, T., 1973. In "Conformation of molecules and polymers", p. 173-188, Eds. Bergmann, E. D. and Pullman, B., Academic Press, New York.

Peña, C., Stewart , J. M., Paladini, A. C., Dellacha, J. M. and Santome, J. A., 1973. In "Peptides: chemistry, structure and biology", p. 523-528, Eds. Walter, R. and Meienhofer, J., Ann Arbor Science Publishers, Ann Arbor.

Periti, P F., Quagliarotti, G. and Liquori, A. M., 1967. Recognition of α-helical segments in proteins of Known primary structure. J. Mol. Biol. 24: 313-322

Prothero, J. N., 1966. Correlation between the distribution of amino acids and alpha-helices. Biophys. J. 6: 367-370

Ptitsyn, O. B. and Finkelshtein, A. V., 1970. Connexion between the secondary and primary structures of globular proteins. Biofizika 15: 757-767

Scanu, A. M., Edelstein, C. and Kein, P., 1975. In "The Plasma Proteins", 2nd ed., Vol. I, p. 317-391, Academic Press, New York.

Schiffer, M. and Edmunson, A. B., 1967. Use of helical wheels to represent the structure of proteins and to identify segments with helical potential. Biophys. J. 7: 121-135

Scheraga, H. A., 1960. Structural studies of ribonuclease. III. A model for the secondary and tertiary structure. J. Amer. Chem. Soc. 82:3847-3852

Szent-Györgyi, A. G. and Cohen, C., 1957. Role of proline in polypeptide chain configuration of proteins. Science 126: 697-698

Takano, T., Kallai, O. B., Swanson, R. and Dickerson, R. E., 1973. The structure of ferrocytochrome c at 2.45Å resolution. J. Biol. Chem. 248: 5234-5255

Venkatachalam, C. M., 1968. Stereochemical criteria for polypeptides and proteins. V. Conformation of a system of three linked peptide units. Biopolymers 6: 1425-2436

Wallace, D. G., 1976. Prediction of the secondary and tertiary structure of plastocyanin. Biophys. Chem. 4: 123-130

Yang, J. T. and Doty, P., 1957. The optical rotatory dispersion of polypeptides and proteins in relation to configuration. J. Amer. Chem. Soc. 79: 761-775

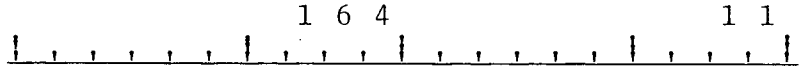
Zimm, B. H. and Bragg, J. K., 1959. Theory of the phase transition between helix and random coil in polypeptide chains. J. Chem. Phys. 31: 526-535

APPENDIX

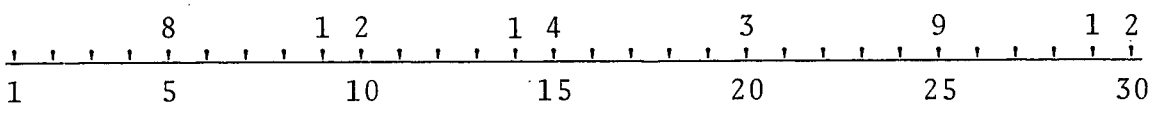
How to Use the Programs

After converting the entire protein sequence into a series of corresponding numbers, the following set of cards must be prepared as input data for the program of helix, β-sheet and β-turn prediction. The first card of the set gives the total number of amino acid residues of the protein in question (NN) and the number of data cards (N). Each of those data cards is composed of 16 numbers or amino acid residues, except the last data card which may or may not be filled with 16 numbers. The following format has been used for NN and N: (6X, 14, 6X, 14).

An example of how the first card looks like for a protein of 164 amino acid residues (NN = 164 and N = 11):



The protein sequence is reported on the subsequent cards (16 data per card) whose format has arbitrarily been chosen as 16 I5. In other words, each of the 16 numbers will occupy five columns on a current IBM card of 80-column width. To keep all the numbers right justified, the one-digit data should be located at columns 5xn (n = 1, 2, 3, 4, ..., 16), and the two-digit ones should start at (5n-1) columns. A typical data card may look like:



An echo print of the input data in the prediction output enables the detection of any typographical error.

In summary, in order to use the programs for helix, sheet, and turn prediction, one has to enter the protein sequence in the form of an "introductory" card (which provides the total number of amino acids and the total number of data cards) followed by the actual data cards (16 data per card).

In addition to the protein sequence, extra information concerning the positions of the overlapping helices and sheets are necessary for the utilization of the

overlapping program. For this reason, the last data card of the protein sequence will immediately be followed by a second set of cards which consists of:

- an "introductory" card of the same format similar to the first one (6X, I4, 6X, I4). The two numbers in question are the total number of values giving the positions of the overlapping helices and β-sheets (it will always be a multiple of four because pairs of helices and β-sheets are involved in the procedure), and the total number of data cards (16 data per card).

- for convenience, keep the format of 16 I5 (5 columns for each datum) for the data cards carrying the information on the positions of the different pairs of overlapping helices and β-sheets. On each card, the boundary values of helices and β-sheets were arranged according to the following ways:

```
          H1      S1      H2      S2      H3      S3
          |-------|-------|-------|-------|-------|-------|---
Column    1       5       10      15      20      25      30
```

H1 : N-boundary of the helix starting from H1 to H2

S1 : N-boundary of the β-sheet starting from S1 to S2

H2 : C-boundary of the helix H1-H2

S2 : C-boundary of the β-sheet-S1-S2

H3 : N-boundary of the helix starting from H3 to H4

Hence the relative positions of helices and β-sheets alternate with each other. An 80-column card can contain up to eight pairs of values.

In summary, in order to use the overlapping program, two sets of cards must be prepared. The first set provides the computer with the information on the protein sequence and the second set, which immediately follows the first , contains data on the relative positions of the overlapping helices and β-sheets. For convenience, the same type of format is used in each set of cards.

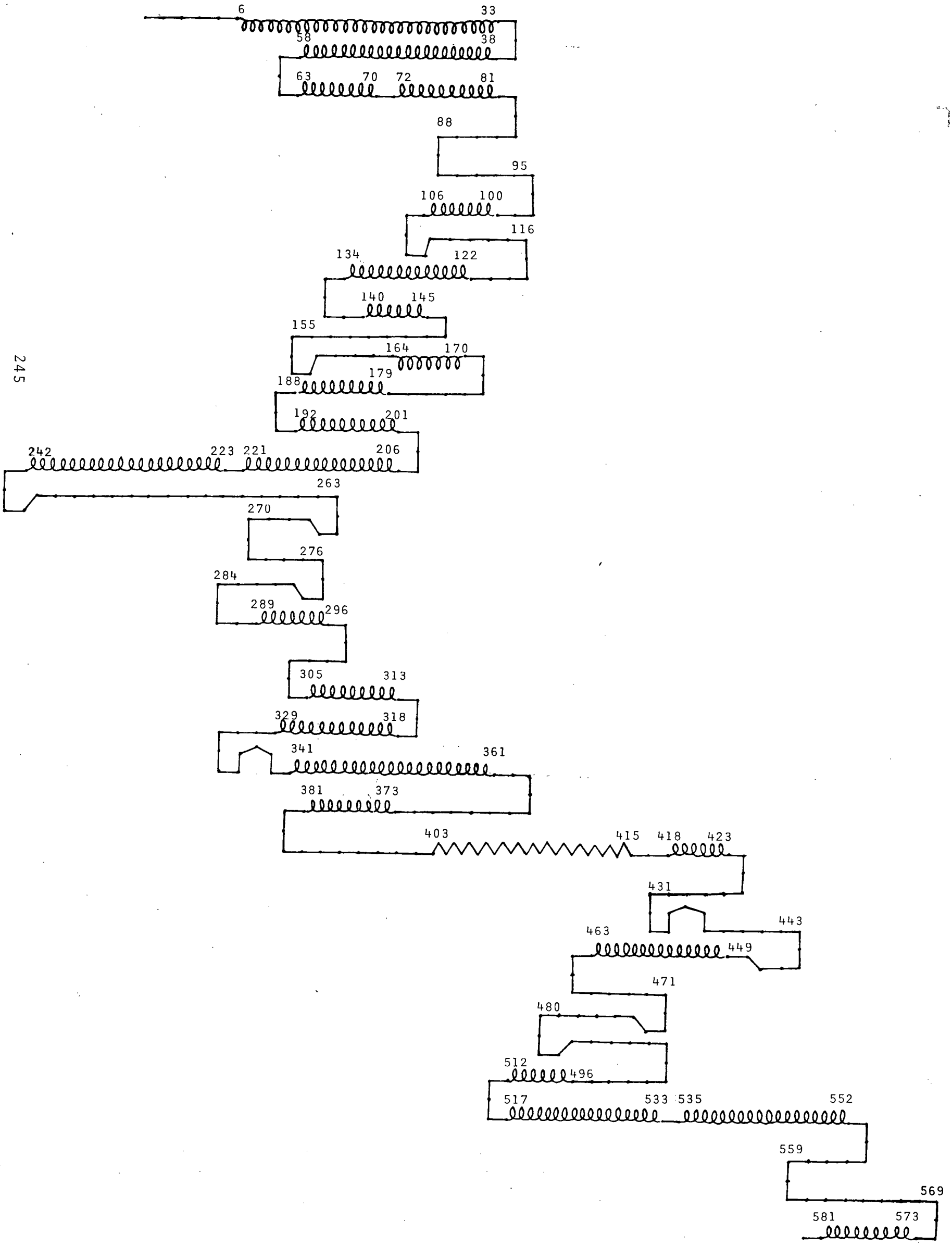


Fig. I - Schematic diagram of the predicted secondary structure of bovine serum albumin

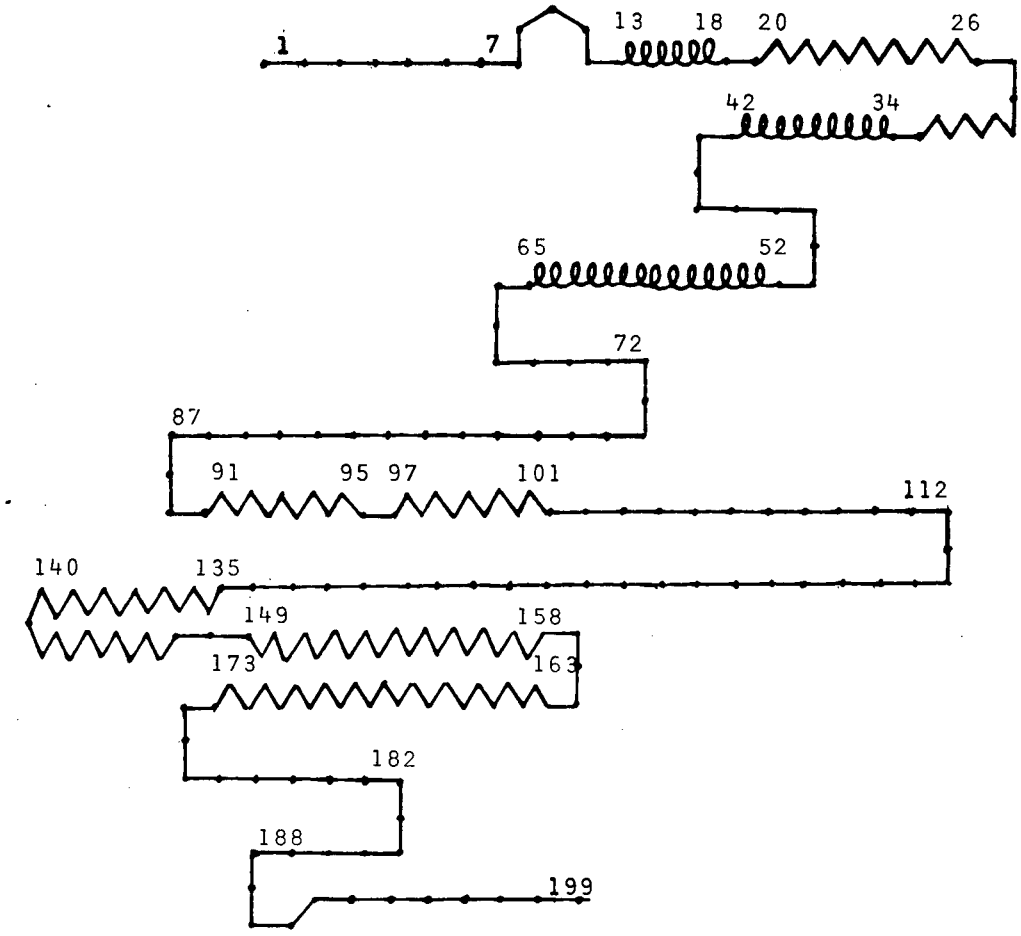


Fig. II - Schematic diagram of the predicted secondary structure of α_{S1}-casein (bovine)

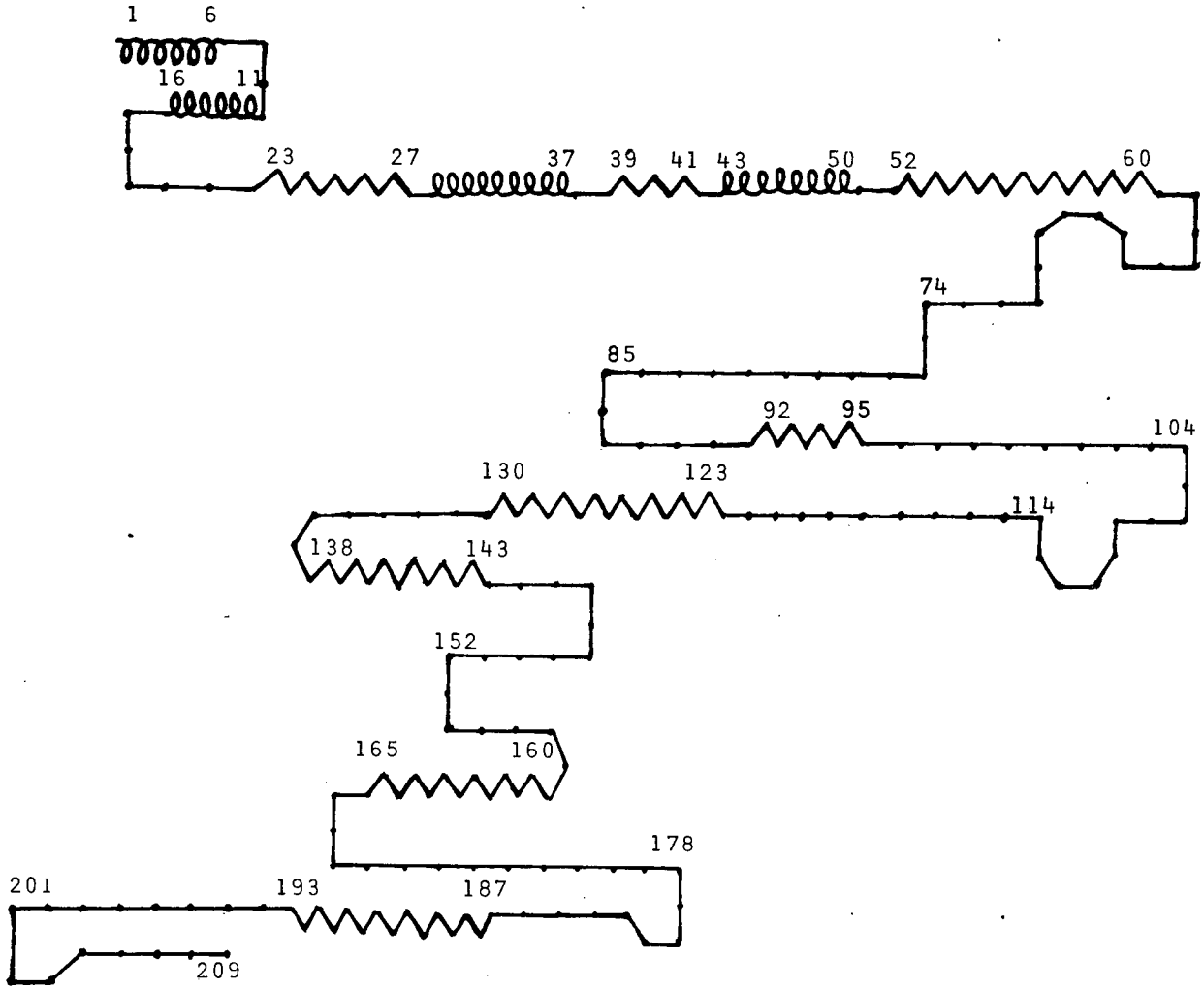


Fig. III - Schematic diagram of the predicted secondary structure of β-casein (bovine)

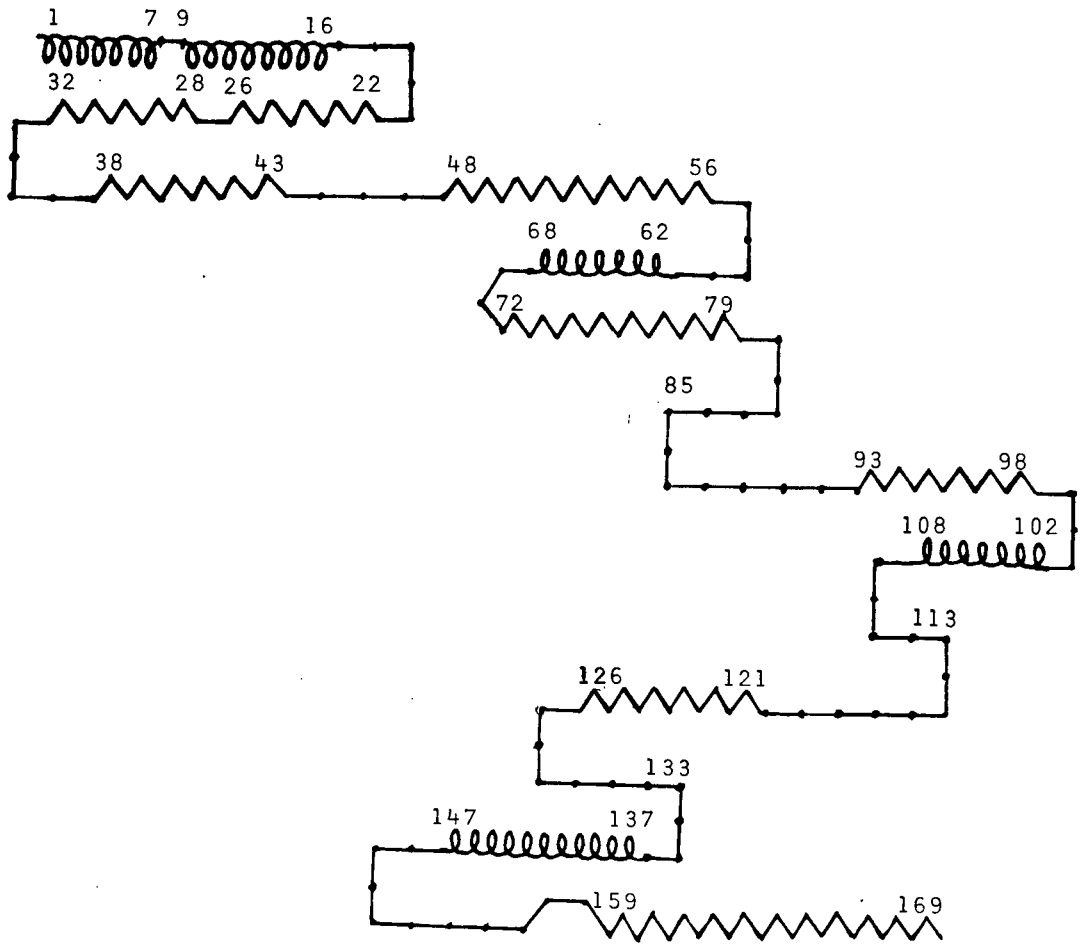


Fig. IV - Schematic diagram of the predicted secondary structure of κ-casein (bovine)

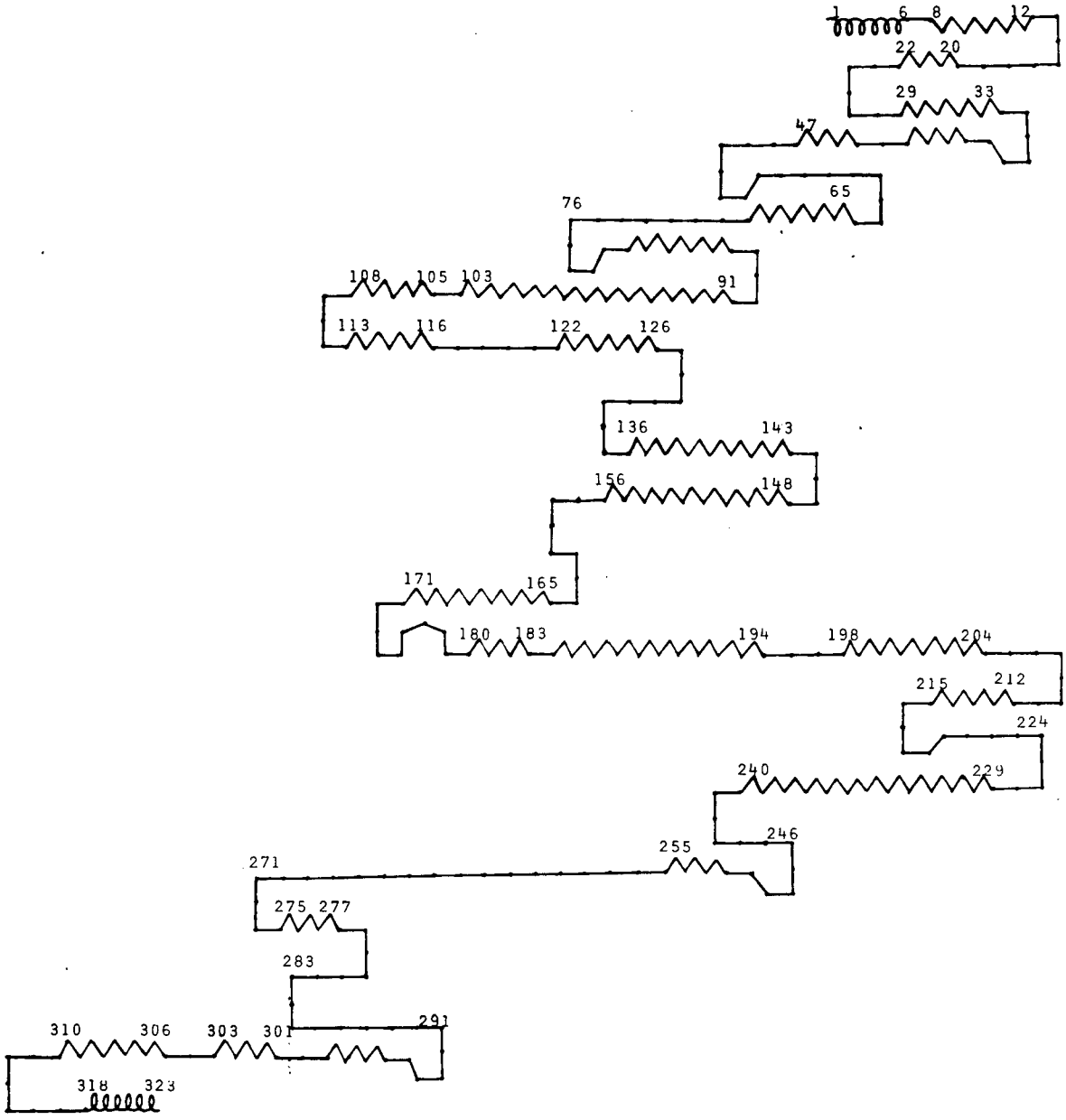


Fig. V - Schematic diagram of the predicted secondary structure of chymosin (bovine)

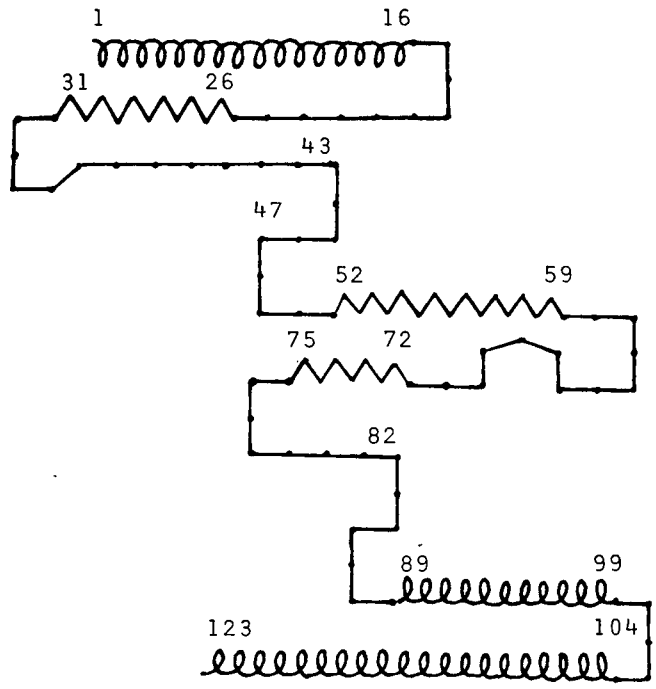


Fig. VI - Schématic diagram of the predicted secondary structure of α-Lactalbumin (bovine)

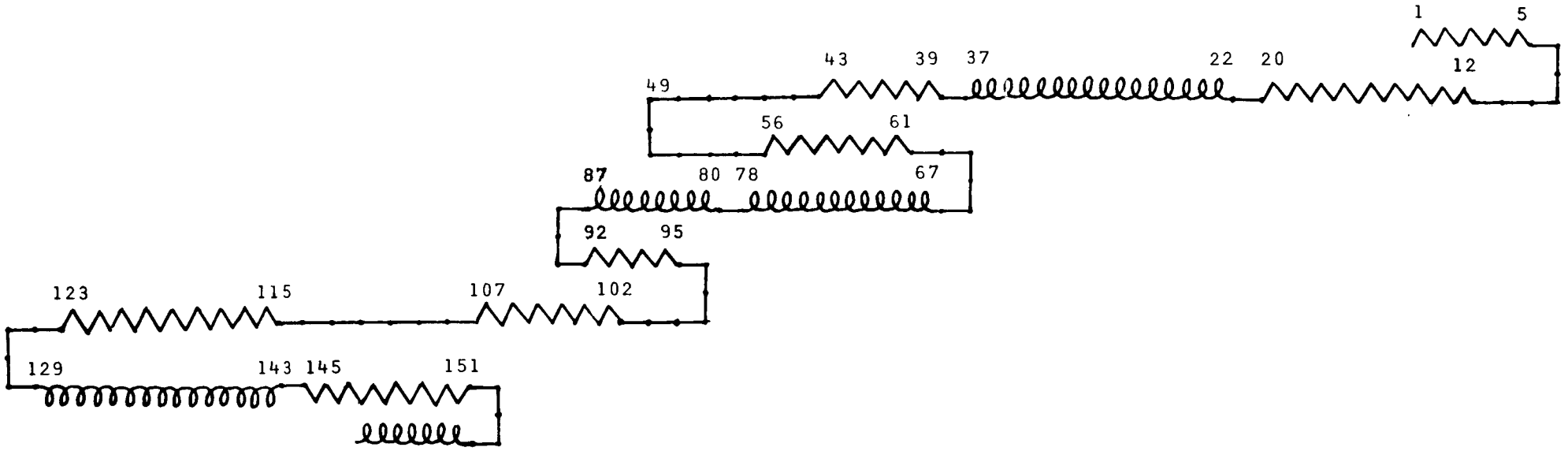


Fig. VII - Schematic diagram of the predicted secondary structure of β-Lactoglobulin (bovine)

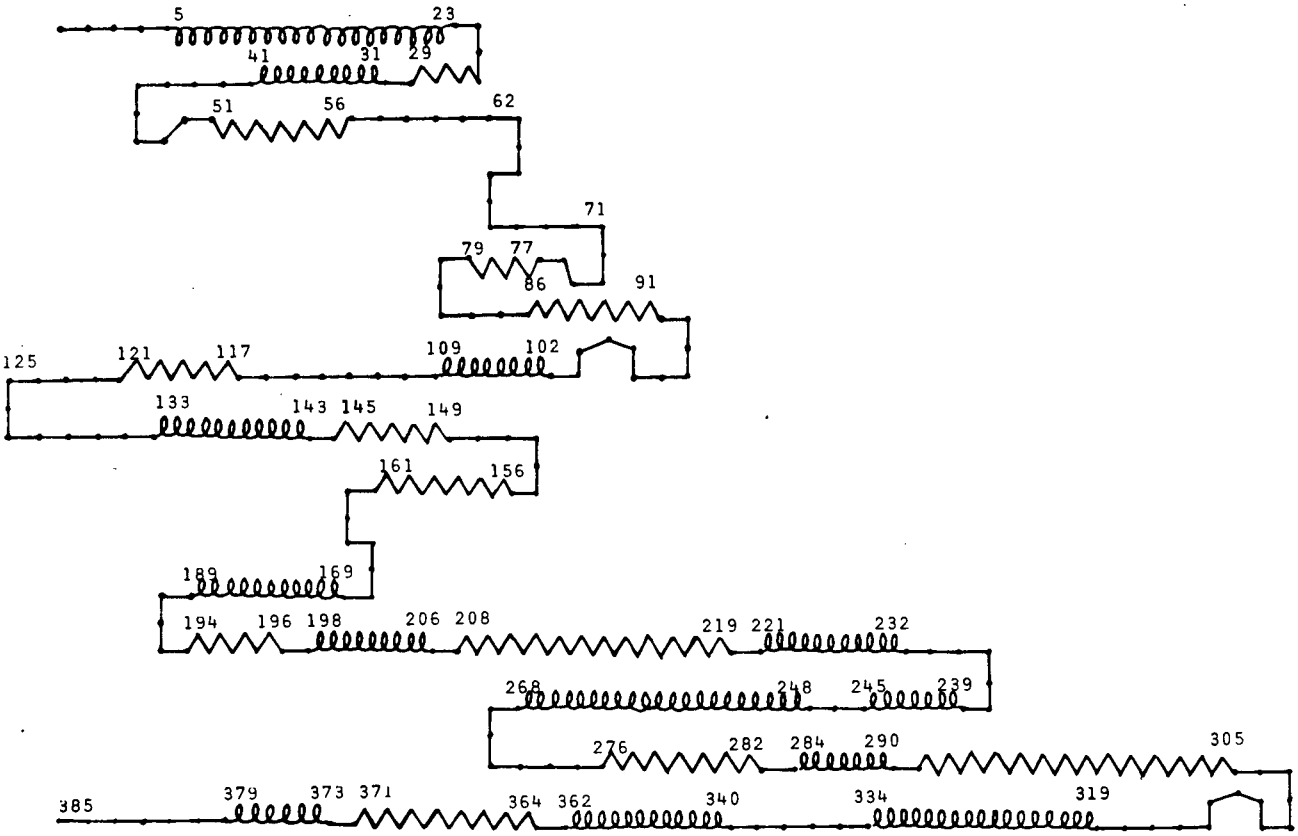


Fig. VIII - Schematic diagram of the predicted secondary structure of ovalbumin

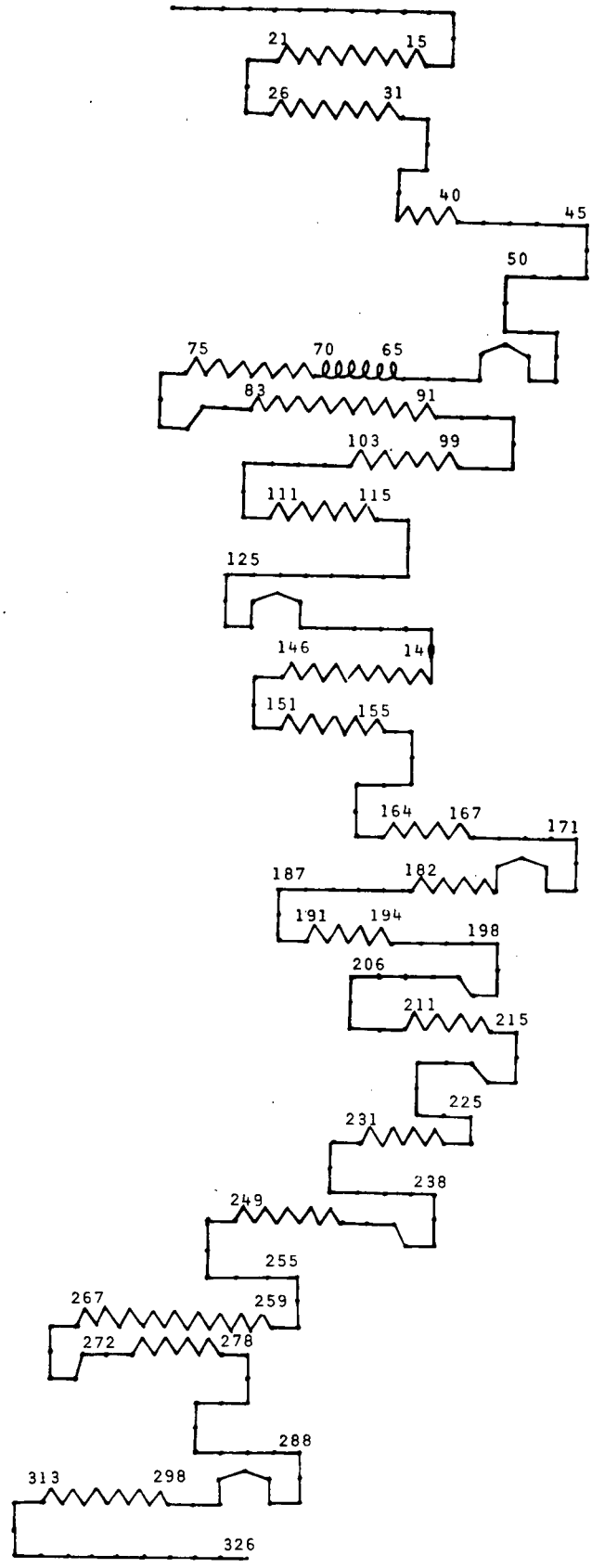


Fig. IX - Schematic diagram of the predicted secondary structure of pepsin (porcine)

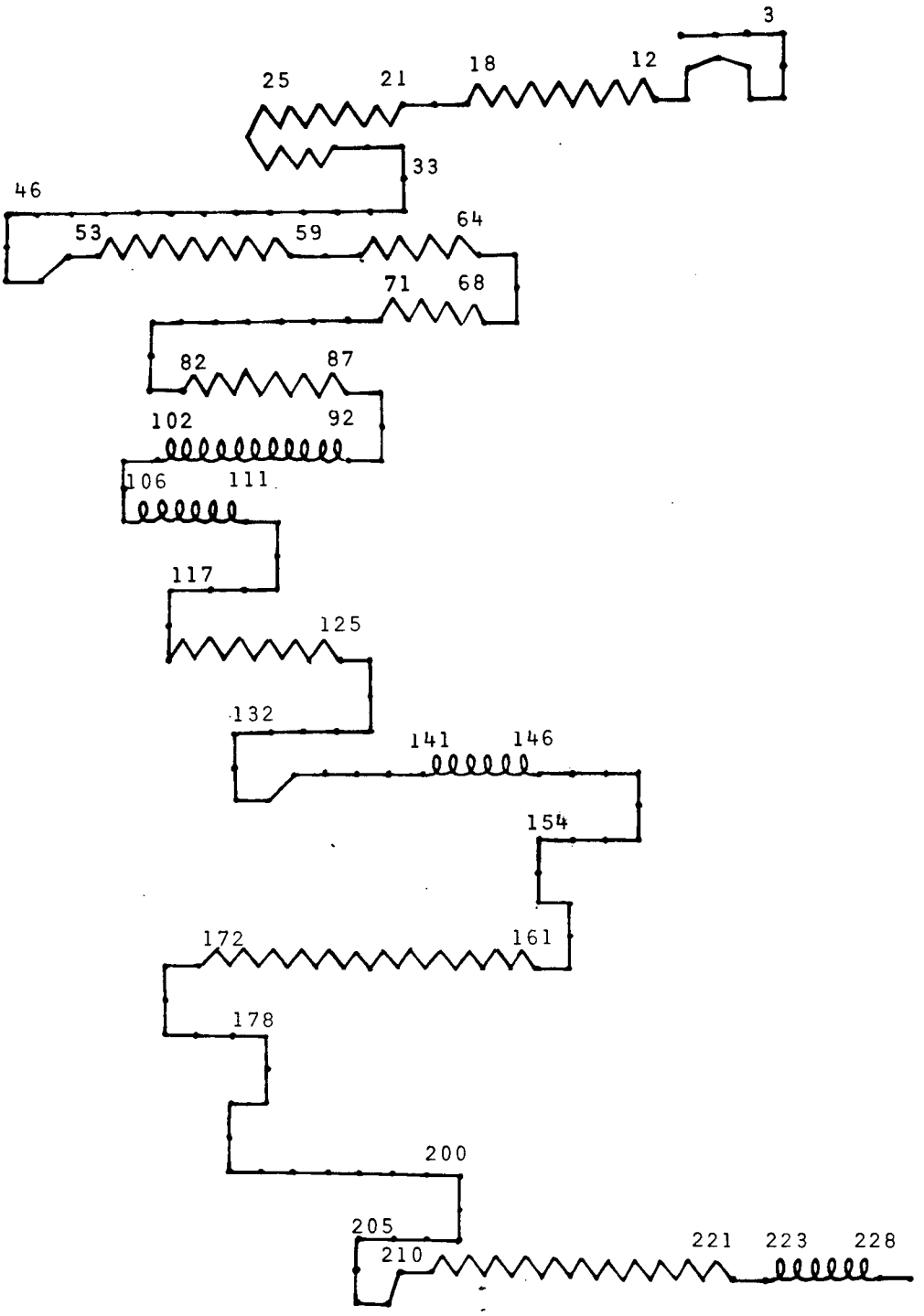


Fig. X - Schematic diagram of the predicted secondary structure of trypsinogen (bovine)

```
**********************************
*                                *
*    ALPHA-HELIX PREDICTION      *
*                                *
**********************************
```

TOTAL NUMBER OF AA: 129
NUMBER OF DATA LINES: 9

PROTEIN SEQUENCE

12	20	14	8	2	5	7	11	1	1	1	13	12	2	9	8
11	4	3	19	2	8	19	16	11	8	3	18	20	5	1	1
12	14	7	16	3	14	3	17	6	1	17	3	2	3	17	4
8	16	17	4	19	8	10	11	6	11	3	16	2	18	18	5
3	4	8	2	17	15	8	16	2	3	11	5	3	10	15	5
16	1	11	11	16	16	4	10	17	1	16	20	3	5	1	12
12	10	20	16	4	8	4	8	13	3	1	18	20	1	18	2
3	2	5	12	8	17	4	20	6	1	18	10	2	8	5	2
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

PRELIMINARY SEARCH FOR REGIONS WITH HELIX POTENTIAL - RULE 1

3	23
26	33
24	38
80	85
90	98
103	116
119	124

KM: 16

SEARCH FOR ACTUAL HELICES FROM THE POTENTIAL REGIONS

J : 3 JA: 8 T3: 4.0000 L: 1 HELIX NUCLEATION

```
JB:    9     JC:   12     T3: 5.7100                HELIX PROPAGATION
JB:    7     JC:   12     T4: 6.0000                HELIX FORMERS IN 6 OVERLAPPING RESIDUES
JB:    3     JD:   12     M5:      3     L:  1      THEORIT. AND ACTUAL # BREAKERS FROM JB TO JD
JB:   13     JC:   16     T3: 3.7300                HELIX PROPAGATION
J1:    3     J2:   14     T4: 9.5000     TT: 6.0000     ACTUAL AND THEORIT. # FORMERS FROM J1 TO J2
J1:    3     J2:   15     T4:10.0000     TT: 6.5000     ACTUAL AND THEORIT. # FORMERS FROM J1 TO J2

PSEUDO-HELIX FROM J1:    3   TO  J2:   15


                    BOUNDARY ANALYSIS OF THE N-TERMINAL
                    T1,T2,T5  4.080  3.240  4.440     STEP 5,MOJ1 CLOSE TO O

                    BOUNDARY ANALYSIS OF THE C-TERMINAL
                    T1,T2      3.790  3.460            STEP 3, J2 CLOSE TO O
                    T1,T2,T5,TT  3.780  3.850  4.050  0.000043757  STEP 6, J2 CLOSE O
                    T1,T2,T5,TT  3.790  3.460  4.560  0.000034700   STEP7, J2-10 , MOJ2
LE,T1,T2,T5,TT,I   11 5.0300   3.5500   3.2200    0.000030110   0   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I   12 4.6100   3.5900   3.5100    0.000041806   1   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I   13 3.7300   3.2900   4.4700    0.000082413   2   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I   14 3.7800   3.8500   4.0500    0.000043757   3   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I   15 3.7900   3.4600   4.5600    0.000034700   4   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I   16 3.4600   3.4800   5.1700    0.000041537   5   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I   17 3.5800   4.2000   4.7500    0.000160201   6   B-TURN SEARCH AT C-TERMINAL
                    T1,T2,T5  3.460  3.480  5.170     STEP 10, J2=J2 , MOJ2
                    T1,T2      3.460  3.480            STEP 16 , J2=J2
                    T1,T2,T5  3.790  3.460  4.560     STEP 19 , J2-1 ,MOJ2
                    T1,T2,T5  3.580  4.200  4.750     STEP 24, J2+1 ,MOJ2
                    T1,T2,T5  3.780  3.850  4.050     STEP 25, J2-2 , RMJ2
                    T1,T2      4.790  4.390            STEP 28, J2-2, RMJ2
                    T1,T2,T5  3.730  3.290  4.470     STEP 41, J2-3 , RMJ2
                    T1,T2,T5  5.030  3.550  3.220     STEP 54, J2-4 ,RMJ2
                    T1,T2,T5  4.610  3.590  3.510     STEP 55, J2-4 ,RMJ2
                    T1,T2,T5,TT  3.460  3.480  5.170  0.000041537 . STEP 60, J2+4 ,RMJ2

              EVENTUAL HELIX FROM J1:    7     TO J2:   15              ***  V2,V3:   11    0  ***


J :   16     JA:   21     T3: 2.0000     L:  3     HELIX NUCLEATION
J :   17     JA:   22     T3: 2.0000     L:  3     HELIX NUCLEATION
J :   18     JA:   23     T3: 1.0000     L:  4     HELIX NUCLEATION
J :   19     JA:   24     T3: 0.5000     L:  4     HELIX NUCLEATION
J :   20     JA:   25     T3: 1.5000     L:  3     HELIX NUCLEATION
J :   21     JA:   26     T3: 1.5000     L:  3     HELIX NUCLEATION
J :   22     JA:   27     T3: 1.0000     L:  4     HELIX NUCLEATION
J :   23     JA:   28     T3: 2.0000     L:  3     HELIX NUCLEATION
J :   24     JA:   29     T3: 3.0000     L:  2     HELIX NUCLEATION
J :   25     JA:   30     T3: 3.5000     L:  2     HELIX NUCLEATION
J :   26     JA:   31     T3: 3.5000     L:  2     HELIX NUCLEATION
J :   27     JA:   32     T3: 4.5000     L:  1     HELIX NUCLEATION
J1:   27     J2:   33     T4: 5.5000     TT: 3.5000     ACTUAL AND THEORIT. # FORMERS FROM J1 TO J2
```

```
PSEUDO-HELIX FROM J1:   27   TO  J2:   33


                    BOUNDARY ANALYSIS OF THE N-TERMINAL
                    T1,T2,T5  4.560  5.090  3.310    STEP 5,MOJ1 CLOSE TO O
LE,T1,T2,T5,TT,I   24 3.2200   3.6900   5.1400    0.000051870   0   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I   25 3.5300   4.3100   4.6700    0.000165386   1   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I   26 3.3800   4.7100   4.5800    0.000028717   2   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I   27 3.8100   5.1500   4.2100    0.000007501   3   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I   28 4.5600   5.0900   3.3100    0.000025081   4   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I   29 4.9000   4.5500   3.0100    0.000006671   5   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I   30 5.0000   3.5900   3.5200    0.000037652   6   B-TURN SEARCH AT N-TERMINAL
                    T1,T2,T5  4.560  5.090  3.310  LC:  1   STEP 12 ,MOJ1,   B-TURN PROBLEM
                    T1,T2,T5  3.530  4.310  4.670    STEP 14,MOJ1 B-T PROBL.
                    T1,T2,T5  3.240  4.270  4.720    STEP 22,MOJ1 B-T PROBL.
                    T1,T2,T5  3.380  4.710  4.580    STEP 26,MOJ1 B-T PROBL.
                    T1,T2,T5  3.530  4.310  4.670    STEP 29, J1+5 ,RMJ1
                    T1,T2,T5  3.220  3.690  5.140    STEP 57, J1+2 ,RMJ1

                    BOUNDARY ANALYSIS OF THE C-TERMINAL
                    T1,T2     4.570  3.240          STEP 3, J2 CLOSE TO O
                    T1,T2,T5,TT  5.220  3.320  3.010  0.000028704  STEP 6, J2 CLOSE O
                    T1,T2,T5,TT  4.570  3.240  3.780  0.000018405   STEP7, J2-10 , MOJ2
LE,T1,T2,T5,TT,I   29 4.9000   4.5500   3.0100    0.000006671   0   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I   30 5.0000   3.5900   3.5200    0.000037652   1   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I   31 5.1300   3.7800   2.9300    0.000021341   2   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I   32 5.2200   3.3200   3.0100    0.000028704   3   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I   33 4.5700   3.2400   3.7800    0.000018405   4   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I   34 4.0800   3.3900   4.3300    0.000040267   5   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I   35 4.0800   3.3900   4.3300    0.000096638   6   B-TURN SEARCH AT C-TERMINAL
                    T1,T2,T5  4.080  3.390  4.330    STEP 10, J2=J2 , MOJ2
                    T1,T2     4.080  3.390          STEP 16 , J2=J2
                    T1,T2,T5  4.570  3.240  3.780    STEP 19 , J2-1 ,MOJ2
                    T1,T2,T5  4.080  3.390  4.330    STEP 24, J2+1 ,MOJ2
                    T1,T2,T5  5.220  3.320  3.010    STEP 25, J2-2 , RMJ2
                    T1,T2     5.990  4.070          STEP 28, J2-2, RMJ2

          EVENTUAL HELIX FROM J1:   27     TO J2:   35              ***  V2,V3:    0   35  ***


J :  24     JA:  29     T3: 3.0000     L:  2     HELIX NUCLEATION
J :  25     JA:  30     T3: 3.5000     L:  2     HELIX NUCLEATION
J :  26     JA:  31     T3: 3.5000     L:  2     HELIX NUCLEATION
J :  27     JA:  32     T3: 4.5000     L:  1     HELIX NUCLEATION
JB:  33     JC:  36     T3: 4.5700               HELIX PROPAGATION
JB:  31     JC:  36     T4: 5.0000               HELIX FORMERS IN 6 OVERLAPPING RESIDUES
JB:  27     JD:  36     M5:     3     L:  1     THEORIT. AND ACTUAL # BREAKERS FROM JB TO JD
J1:  27     J2:  36     T4: 7.5000     TT: 5.0000    ACTUAL AND THEORIT. # FORMERS FROM J1 TO J2

PSEUDO-HELIX FROM J1:   27   TO  J2:   36
```

```
                    BOUNDARY ANALYSIS OF THE N-TERMINAL
                    T1,T2,T5  4.560  5.090  3.310    STEP 5,MOJ1 CLOSE TO O
LE,T1,T2,T5,TT,I   24 3.2200   3.6900   5.1400    0.000051870   0   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I   25 3.5300   4.3100   4.6700    0.000165386   1   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I   26 3.3800   4.7100   4.5800    0.000028717   2   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I   27 3.8100   5.1500   4.2100    0.000007501   3   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I   28 4.5600   5.0900   3.3100    0.000025081   4   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I   29 4.9000   4.5500   3.0100    0.000006671   5   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I   30 5.0000   3.5900   3.5200    0.000037652   6   B-TURN SEARCH AT N-TERMINAL
                    T1,T2,T5  4.560  5.090  3.310  LC:  1   STEP 12 ,MOJ1,   B-TURN PROBLEM
                    T1,T2,T5  3.530  4.310  4.670    STEP 14,MOJ1 B-T PROBL.
                    T1,T2,T5  3.240  4.270  4.720    STEP 22,MOJ1 B-T PROBL.
                    T1,T2,T5  3.380  4.710  4.580    STEP 26,MOJ1 B-T PROBL.
                    T1,T2,T5  3.530  4.310  4.670    STEP 29, J1+5 ,RMJ1
                    T1,T2,T5  3.220  3.690  5.140    STEP 57, J1+2 ,RMJ1

                    BOUNDARY ANALYSIS OF THE C-TERMINAL
                    T1,T2     3.240  3.910           STEP 3, J2 CLOSE TO O
                    T1,T2,T5,TT  4.080  3.390  4.330  0.000096638  STEP 6, J2 CLOSE O
                    T1,T2,T5,TT  3.240  3.910  5.150  0.000058913   STEP7, J2-10 , MOJ2
LE,T1,T2,T5,TT,I   32 5.2200   3.3200   3.0100    0.000028704   0   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I   33 4.5700   3.2400   3.7800    0.000018405   1   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I   34 4.0800   3.3900   4.3300    0.000040267   2   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I   35 4.0800   3.3900   4.3300    0.000096638   3   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I   36 3.2400   3.9100   5.1500    0.000058913   4   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I   37 3.3000   4.3500   4.6800    0.000099602   5   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I   38 3.7400   4.5600   4.1000    0.000031194   6   B-TURN SEARCH AT C-TERMINAL
                    T1,T2,T5  3.300  4.350  4.680    STEP 10, J2=J2 , MOJ2
                    T1,T2     3.300  4.350           STEP 16 , J2=J2

          EVENTUAL HELIX FROM J1:   27     TO J2:   35              ***  V2,V3:    0   18  ***


J :  80     JA:  85     T3: 3.5000     L:  0     HELIX NUCLEATION

PSEUDO-HELIX FROM     J1:   79   TO J2:   84          SPECIAL CASE

J :  90     JA:  95     T3: 3.5000     L:  1     HELIX NUCLEATION
J :  91     JA:  96     T3: 3.5000     L:  1     HELIX NUCLEATION
J :  92     JA:  97     T3: 4.5000     L:  1     HELIX NUCLEATION
J1:  91     J2:  98     T4: 5.5000     TT: 4.0000    ACTUAL AND THEORIT. # FORMERS FROM J1 TO J2
J1:  90     J2:  98     T4: 6.5000     TT: 4.5000    ACTUAL AND THEORIT. # FORMERS FROM J1 TO J2

PSEUDO-HELIX FROM J1:   90   TO  J2:   98


                    BOUNDARY ANALYSIS OF THE N-TERMINAL
                    T1,T2,T5  3.500  4.530  4.680    STEP 5,MOJ1 CLOSE TO O
LE,T1,T2,T5,TT,I   87 4.3400   4.1600   3.5500    0.000018842   0   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I   88 4.1000   4.3700   3.5200    0.000017229   1   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I   89 4.0800   4.4700   3.5500    0.000043301   2   B-TURN SEARCH AT N-TERMINAL
```

```
LE,T1,T2,T5,TT,I   90 3.9200   4.1700   4.1500    0.000021250   3   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I   91 3.5000   4.5300   4.6800    0.000140820   4   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I   92 4.1500   4.6100   3.9100    0.000034921   5   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I   93 4.2500   3.6500   4.4200    0.000028372   6   B-TURN SEARCH AT N-TERMINAL
                    T1,T2,T5  3.500  4.530  4.680  LC:  2   STEP 12 ,MOJ1,   B-TURN PROBLEM
                    T1,T2,T5  4.100  4.370  3.520    STEP 14,MOJ1 B-T PROBL.
                    T1,T2,T5  3.690  4.080  4.320    STEP 22,MOJ1 B-T PROBL.
                    T1,T2,T5  4.080  4.470  3.550    STEP 26,MOJ1 B-T PROBL.
                    T1,T2,T5  4.100  4.370  3.520    STEP 29, J1+5 ,RMJ1
                    T1,T2,T5  4.340  4.160  3.550    STEP 57, J1+2 ,RMJ1

                    BOUNDARY ANALYSIS OF THE C-TERMINAL
                    T1,T2     3.920  4.590           STEP 3, J2 CLOSE TO O
                    T1,T2,T5,TT  4.070  4.790  3.410  0.000005550  STEP 6, J2 CLOSE O
                    T1,T2,T5,TT  3.920  4.590  3.860  0.000020898   STEP7, J2-1O , MOJ2
LE,T1,T2,T5,TT,I   94 4.7400   3.5000   3.8700    0.000077456   0   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I   95 4.8200   3.9100   3.1500    0.000027821   1   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I   96 4.4600   4.7800   2.9900    0.000004358   2   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I   97 4.0700   4.7900   3.4100    0.000005550   3   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I   98 3.9200   4.5900   3.8600    0.000020898   4   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I   99 3.4100   3.7400   4.9500    0.000234478   5   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I  100 3.3600   2.5800   5.9100    0.000203148   6   B-TURN SEARCH AT C-TERMINAL
                    T1,T2,T5  3.410  3.740  4.950    STEP 10, J2=J2 , MOJ2
                    T1,T2     3.410  3.740           STEP 16 , J2=J2
                    T1,T2,T5  3.920  4.590  3.860    STEP 19 , J2-1 ,MOJ2

          EVENTUAL HELIX FROM J1:   90     TO J2:   99              ***  V2,V3:    0   23  ***


J : 103     JA: 108     T3: 3.5000     L:  2     HELIX NUCLEATION
J : 104     JA: 109     T3: 4.0000     L:  2     HELIX NUCLEATION
J : 105     JA: 110     T3: 5.0000     L:  1     HELIX NUCLEATION
JB: 111     JC: 114     T3: 3.7500               HELIX PROPAGATION
J1: 105     J2: 112     T4: 6.5000     TT: 4.0000    ACTUAL AND THEORIT. # FORMERS FROM J1 TO J2

PSEUDO-HELIX FROM J1:  105   TO  J2:  112


                    BOUNDARY ANALYSIS OF THE N-TERMINAL
                    T1,T2,T5  4.230  4.790  3.680    STEP 5,MOJ1 CLOSE TO O
LE,T1,T2,T5,TT,I  102 3.6000   3.0900   5.1800    0.000117249   0   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I  103 3.7000   3.2300   5.1800    0.000015919   1   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I  104 4.1100   3.5200   4.3800    0.000092656   2   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I  105 4.6200   4.1400   3.7800    0.000032989   3   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I  106 4.2300   4.7900   3.6800    0.000041504   4   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I  107 4.9800   4.7300   2.7800    0.000001267   5   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I  108 4.6400   5.2700   3.0800    0.000021603   6   B-TURN SEARCH AT N-TERMINAL
                    T1,T2,T5  4.230  4.790  3.680  LC:  0   STEP 12 ,MOJ1,   B-TURN PROBLEM
                    T1,T2,T5  3.700  3.230  5.180    STEP 14,MOJ1 B-T PROBL.
                    T1,T2,T5  3.160  2.580  6.040    STEP 22,MOJ1 B-T PROBL.
```

```
                    BOUNDARY ANALYSIS OF THE C-TERMINAL
                    T1,T2      3.670  3.940            STEP 3, J2 CLOSE TO 0
                    T1,T2,T5,TT  3.750  4.120  4.420  0.000132510  STEP 6, J2 CLOSE 0
                    T1,T2,T5,TT  3.670  3.940  4.650  0.000073624   STEP7, J2-10 , MOJ2
LE,T1,T2,T5,TT,I  108 4.6400   5.2700   3.0800    0.000021603   0   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I  109 4.5600   4.8300   3.0700    0.000025633   1   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I  110 4.1700   4.0200   4.1300    0.000007027   2   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I  111 3.7500   4.1200   4.4200    0.000132510   3   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I  112 3.6700   3.9400   4.6500    0.000073624   4   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I  113 3.8300   3.7500   4.7100    0.000189688   5   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I  114 3.7300   3.6100   4.7100    0.000040602   6   B-TURN SEARCH AT C-TERMINAL
                    T1,T2,T5   3.830  3.750  4.710     STEP 10, J2=J2 , MOJ2
                    T1,T2      3.830  3.750            STEP 16 , J2=J2
                    T1,T2,T5   3.670  3.940  4.650     STEP 19 , J2-1 ,MOJ2
                    T1,T2,T5   3.730  3.610  4.710     STEP 24, J2+1 ,MOJ2
                    T1,T2,T5   3.750  4.120  4.420     STEP 25, J2-2 , RMJ2
                    T1,T2      4.750  5.310            STEP 28, J2-2, RMJ2

              EVENTUAL HELIX FROM J1:  107     TO J2:  114              ***  V2,V3:   24   37  ***


J : 119     JA: 124     T3: 5.5000     L:  0     HELIX NUCLEATION
J1: 119     J2: 124     T4: 5.5000     TT: 3.0000    ACTUAL AND THEORIT. # FORMERS FROM J1 TO J2

PSEUDO-HELIX FROM J1:  119   TO  J2:  124


                    BOUNDARY ANALYSIS OF THE N-TERMINAL
                    T1,T2,T5  4.670  5.000  3.100     STEP 5,MOJ1 CLOSE TO 0
LE,T1,T2,T5,TT,I  116 3.5700   3.2200   4.9900    0.000024614   0   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I  117 3.4700   4.1800   4.4800    0.000104509   1   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I  118 4.0100   4.5300   3.9000    0.000025958   2   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I  119 4.6000   4.1700   3.6000    0.000015142   3   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I  120 4.6700   5.0000   3.1000    0.000035514   4   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I  121 4.6900   4.9000   3.0700    0.000020156   5   B-TURN SEARCH AT N-TERMINAL
LE,T1,T2,T5,TT,I  122 4.5600   4.7300   3.0400    0.000000862   6   B-TURN SEARCH AT N-TERMINAL
                    T1,T2,T5  4.670  5.000  3.100  LC:  1   STEP 12 ,MOJ1,   B-TURN PROBLEM
                    T1,T2,T5  3.470  4.180  4.480     STEP 14,MOJ1 B-T PROBL.
                    T1,T2,T5  3.560  3.870  4.720     STEP 22,MOJ1 B-T PROBL.
                    T1,T2,T5  4.010  4.530  3.900     STEP 26,MOJ1 B-T PROBL.
                    T1,T2,T5  3.470  4.180  4.480     STEP 29, J1+5 ,RMJ1
                    T1,T2,T5  3.570  3.220  4.990     STEP 57, J1+2 ,RMJ1

                    BOUNDARY ANALYSIS OF THE C-TERMINAL
                    T1,T2      3.330  4.470            STEP 3, J2 CLOSE TO 0
                    T1,T2,T5,TT  3.710  4.650  3.940  0.000039396  STEP 6, J2 CLOSE 0
                    T1,T2,T5,TT  3.330  4.470  4.170  0.000110850   STEP7, J2-10 , MOJ2
LE,T1,T2,T5,TT,I  120 4.6700   5.0000   3.1000    0.000035514   0   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I  121 4.6900   4.9000   3.0700    0.000020156   1   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I  122 4.5600   4.7300   3.0400    0.000000862   2   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I  123 3.7100   4.6500   3.9400    0.000039396   3   B-TURN SEARCH AT C-TERMINAL
```

```
LE,T1,T2,T5,TT,I  124 3.3300   4.4700   4.1700    0.000110850   4   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I  125 3.2500   3.8000   4.6500    0.000059173   5   B-TURN SEARCH AT C-TERMINAL
LE,T1,T2,T5,TT,I  126 3.4800   4.1700   4.2900    0.000037464   6   B-TURN SEARCH AT C-TERMINAL
                    T1,T2,T5  3.250  3.800  4.650    STEP 10, J2=J2 , MOJ2
                    T1,T2     3.250  3.800           STEP 16 , J2=J2
                    T1,T2,T5  3.330  4.470  4.170    STEP 19 , J2-1 ,MOJ2
                    T1,T2,T5  3.480  4.170  4.290    STEP 24, J2+1 ,MOJ2
                    T1,T2,T5  3.710  4.650  3.940    STEP 25, J2-2 , RMJ2
                    T1,T2     4.410  5.840           STEP 28, J2-2, RMJ2
                    T1,T2,T5  4.560  4.730  3.040    STEP 41, J2-3 , RMJ2
                    T1,T2,T5  4.670  5.000  3.100    STEP 54, J2-4 ,RMJ2
                    T1,T2,T5  4.690  4.900  3.070    STEP 55, J2-4 ,RMJ2

              EVENTUAL HELIX FROM J1:  119     TO J2:  124              ***  V2,V3:    0    0  ***


END OF PROGRAM
```

```
**********************************
*                                *
*    BETA-SHEET  PREDICTION      *
*                                *
**********************************
```

TOTAL NUMBER OF AA: 129
NUMBER OF DATA LINES: 9

PROTEIN SEQUENCE

12	20	14	8	2	5	7	11	1	1	1	13	12	2	9	8
11	4	3	19	2	8	19	16	11	8	3	18	20	5	1	1
12	14	7	16	3	14	3	17	6	1	17	3	2	3	17	4
8	16	17	4	19	8	10	11	6	11	3	16	2	18	18	5
3	4	8	2	17	15	8	16	2	3	11	5	3	10	15	5
16	1	11	11	16	16	4	10	17	1	16	20	3	5	1	12
12	10	20	16	4	8	4	8	13	3	1	18	20	1	18	2
3	2	5	12	8	17	4	20	6	1	18	10	2	8	5	2
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

PRELIMINARY SEARCH FOR REGIONS WITH SHEET POTENTIAL - RULE 2

1	6
19	21
21	25
25	34
37	41
51	53
53	69
73	76
76	80
83	85
87	89

```
                    89              95
                    95              99
                   104             109
                   111             113
                   118             121
                   121             129

                        IM = 36

  SEARCH FOR ACTUAL SHEETS FROM THE POTENTIAL REGIONS
  ...................................................

G :   1     MA:   5     T3: 2.0000     N:  2         SHEET NUCLEATION
G :   2     MA:   6     T3: 3.0000     N:  1         SHEET NUCLEATION
J1:   1     J2:   6     T4: 4.0000     TT: 3.0000    ACTUAL AND THEORIT. # FORMERS FROM J1 TO J2

PSEUDO-SHEET FROM J1:    1     TO  J2:    6

                EVENTUAL SHEET FROM J1:    2     TO   J2:    6              ***  V2,V3 :   25   32  ***

G :  19     MA:  22     T3: 1.0000     N:  1         SHEET NUCLEATION
G :  21     MA:  25     T3: 2.0000     N:  2         SHEET NUCLEATION
G :  22     MA:  25     T3: 2.0000     N:  2         SHEET NUCLEATION
G :  25     MA:  29     T3: 3.0000     N:  1         SHEET NUCLEATION
MB:  30     MC:  33     T1: 4.7000     T2: 3.5900    SHEET PROPAGATION
J1:  25     J2:  31     T4: 4.0000     TT: 3.5000    ACTUAL AND THEORIT. # FORMERS FROM J1 TO J2
J1:  25     J2:  32     T4: 4.0000     TT: 4.0000    ACTUAL AND THEORIT. # FORMERS FROM J1 TO J2

PSEUDO-SHEET FROM J1:   25     TO  J2:   32

                EVENTUAL SHEET FROM J1:   25     TO   J2:   32              ***  V2,V3 :   36    0  ***

G :  37     MA:  41     T3: 3.0000     N:  0         SHEET NUCLEATION
J1:  37     J2:  42     T4: 3.0000     TT: 3.0000    ACTUAL AND THEORIT. # FORMERS FROM J1 TO J2
J1:  37     J2:  45     T4: 5.0000     TT: 4.5000    ACTUAL AND THEORIT. # FORMERS FROM J1 TO J2
J1:  37     J2:  45     T4: 5.0000     TT: 4.5000    ACTUAL AND THEORIT. # FORMERS FROM J1 TO J2

PSEUDO-SHEET FROM J1:   37     TO  J2:   45
```

```
                EVENTUAL SHEET FROM J1:   38     TO   J2:   46              ***  V2,V3 :    8   14  ***

G :  51     MA:  54     T3: 2.0000     N:  2        SHEET NUCLEATION
G :  53     MA:  57     T3: 4.0000     N:  1        SHEET NUCLEATION
MB:  58     MC:  61     T1: 3.6300     T2: 3.8700   SHEET PROPAGATION
J1:  53     J2:  59     T4: 6.0000     TT: 3.5000   ACTUAL AND THEORIT. # FORMERS FROM J1 TO J2

PSEUDO-SHEET FROM J1:   53     TO  J2:   59

                EVENTUAL SHEET FROM J1:   51     TO   J2:   59              ***  V2,V3 :   17    0  ***

G :  60     MA:  64     T3: 3.0000     N:  1        SHEET NUCLEATION
MB:  56     MC:  59     T1: 4.2000     T2: 4.7500   SHEET PROPAGATION
MB:  56     MD:  64     V6:     3      V8: 1       THEORITIC. AND ACTUAL # BREAKERS FROM MB TO MD

PSEUDO-SHEET FROM J1:   56     TO  J2:   64

                EVENTUAL SHEET FROM J1:   56     TO   J2:   65              ***  V2,V3 :    0   19  ***

G :  66     MA:  69     T3: 1.0000     N:  2        SHEET NUCLEATION
G :  73     MA:  76     T3: 2.0000     N:  0        SHEET NUCLEATION
G :  76     MA:  80     T3: 3.0000     N:  1        SHEET NUCLEATION
G :  83     MA:  86     T3: 2.0000     N:  2        SHEET NUCLEATION
G :  87     MA:  90     T3: 2.0000     N:  1        SHEET NUCLEATION
G :  89     MA:  93     T3: 3.0000     N:  1        SHEET NUCLEATION
J1:  88     J2:  95     T4: 5.0000     TT: 4.0000   ACTUAL AND THEORIT. # FORMERS FROM J1 TO J2

PSEUDO-SHEET FROM J1:   88     TO  J2:   95

                EVENTUAL SHEET FROM J1:   88     TO   J2:   94              ***  V2,V3 :   35   21  ***

G :  95     MA:  99     T3: 2.0000     N:  2        SHEET NUCLEATION
G :  96     MA:  99     T3: 2.0000     N:  2        SHEET NUCLEATION
G : 104     MA: 108     T3: 2.0000     N:  1        SHEET NUCLEATION
G : 105     MA: 109     T3: 3.0000     N:  0        SHEET NUCLEATION
J1: 105     J2: 110     T4: 3.0000     TT: 3.0000   ACTUAL AND THEORIT. # FORMERS FROM J1 TO J2
J1: 105     J2: 112     T4: 4.0000     TT: 4.0000   ACTUAL AND THEORIT. # FORMERS FROM J1 TO J2
J1: 105     J2: 113     T4: 5.0000     TT: 4.5000   ACTUAL AND THEORIT. # FORMERS FROM J1 TO J2

PSEUDO-SHEET FROM J1:  105     TO  J2:  113

                EVENTUAL SHEET FROM J1:  105     TO   J2:  113              ***  V2,V3 :   36   32  ***
```

```
G : 111     MA: 114     T3: 2.0000     N:  0          SHEET NUCLEATION
G : 118     MA: 121     T3: 3.0000     N:  1          SHEET NUCLEATION
J1: 118     J2: 122     T4: 3.0000     TT: 2.5000     ACTUAL AND THEORIT. # FORMERS FROM J1 TO J2
J1: 118     J2: 125     T4: 5.0000     TT: 4.0000     ACTUAL AND THEORIT. # FORMERS FROM J1 TO J2

PSEUDO-SHEET FROM J1:  118     TO  J2:  125

                EVENTUAL SHEET FROM J1:  120     TO   J2:  127              ***  V2,V3 :   23   20  ***

G : 121     MA: 125     T3: 3.0000     N:  0          SHEET NUCLEATION
MB: 126     MC: 129     T1: 3.4600     T2: 4.1700     SHEET PROPAGATION
MB: 121     MD: 129     V6:      3     V8: 1          THEORITIC. AND ACTUAL # BREAKERS FROM MB TO MD

PSEUDO-SHEET FROM J1:  121     TO  J2:  129

                EVENTUAL SHEET FROM J1:  121     TO   J2:  129              ***  V2,V3 :   39   32  ***


END OF PROGRAM
```

```
**********************************
*                                *
*    OVERLAPPING RESOLUTION      *
*                                *
**********************************
```

TOTAL NUMBER OF AA: 129
NUMBER OF DATA LINES: 9

PROTEIN SEQUENCE

12	20	14	8	2	5	7	11	1	1	1	13	12	2	9	8
11	4	3	19	2	8	19	16	11	8	3	18	20	5	1	1
12	14	7	16	3	14	3	17	6	1	17	3	2	3	17	4
8	16	17	4	19	8	10	11	6	11	3	16	2	18	18	5
3	4	8	2	17	15	8	16	2	3	11	5	3	10	15	5
16	1	11	11	16	16	4	10	17	1	16	20	3	5	1	12
12	10	20	16	4	8	4	8	13	3	1	18	20	1	18	2
3	2	5	12	8	17	4	20	6	1	18	10	2	8	5	2
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

PAIRS OF OVERLAPPING HELICES AND SHEETS

27	27	35	32	89	88	99	94	107	107	114	113	119	120	124	124

*** COMPARISON OF THEIR LENGTH *** L-HELIX: 9 L-SHEET: 6 RATIO=LH/LS: 1.0

***** COMPARISON OF P-HELIX AND P-SHEET *****

H1 : 27 H2 : 35 A1: 1.128 A2: 1.033 A1 > A2 FROM H1 TO H2

S1 : 27 S2 : 32 A1: 1.058 A2: 1.135 A1 < A2 FROM S1 TO S2

*** COMPARISON OF ASSIGNMENTS TYPES ***

HHF	HF	IIH	IH	BH	BBH	SSF	SF	IS	BS	BBS
6.00	4.00	0.0	0.25	-0.50	0.0	2.00	3.00	0.75	-0.50	-1.00

H1 : 27 H2 : 35 TTH: 9.750 TTS: 4.250 TTH > TTS FROM H1 TO H2

HHF	HF	IIH	IH	BH	BBH	SSF	SF	IS	BS	BBS
4.00	2.00	0.0	0.25	-0.50	0.0	2.00	2.00	0.75	0.0	0.0

S1 : 27 S2 : 32 TTH: 5.750 TTS: 4.750 TTH > TTS FROM S1 TO S2

BOUNDARY ANALYS. FOR HELIX FROM: 27 TO: 35 AND FOR SHEET FROM: 27 TO: 32

HN	SN	HC	SC	NHN	NSN	NHC	NSC
2.89	3.84	4.17	2.61	3.81	3.33	3.61	2.11

*** COMPARISON OF THEIR LENGTH *** L-HELIX: 11 L-SHEET: 7 RATIO=LH/LS: 1.0

***** COMPARISON OF P-HELIX AND P-SHEET *****

H1 : 89 H2 : 99 A1: 1.030 A2: 1.105 A1 < A2 FROM H1 TO H2

S1 : 88 S2 : 94 A1: 0.933 A2: 1.164 A1 < A2 FROM S1 TO S2

*** P-HELIX AND P-SHEET OF INTERS. AREA : H1 TO S2 ***

OL1: 89 OL2: 94 A1: 0.908 A2: 1.092 A1 < A2 FROM H1 TO S2

*** COMPARISON OF ASSIGNMENTS TYPES ***

HHF	HF	IIH	IH	BH	BBH	SSF	SF	IS	BS	BBS
4.00	5.00	0.0	0.75	-0.50	0.0	6.00	2.00	0.75	-1.50	0.0

H1 : 89 H2 : 99 TTH: 9.250 TTS: 7.250 TTH > TTS FROM H1 TO H2

HHF	HF	IIH	IH	BH	BBH	SSF	SF	IS	BS	BBS
2.00	2.00	0.0	0.75	-0.50	0.0	4.00	2.00	0.50	-0.50	0.0

S1 : 88 S2 : 94 TTH: 4.250 TTS: 6.000 TTH < TTS FROM S1 TO S2

HHF	HF	IIH	IH	BH	BBH	SSF	SF	IS	BS	BBS
2.00	1.00	0.0	0.75	-0.50	0.0	2.00	2.00	0.50	-0.50	0.0

*** ASSIGNM. TYPES IN OVERL. AREAS : H1 TO S2 ***

OL1: 89 OL2: 94 TTH: 3.250 TTS: 4.000 TTH < TTS FROM H1 TO S2

BOUNDARY ANALYS. FOR HELIX FROM: 89 TO: 99 AND FOR SHEET FROM: 88 TO: 94

HN	SN	HC	SC	NHN	NSN	NHC	NSC
3.11	3.97	4.06	4.11	3.31	4.21	3.63	2.38

*** COMPARISON OF THEIR LENGTH *** L-HELIX: 8 L-SHEET: 7 RATIO=LH/LS: 1.0

***** COMPARISON OF P-HELIX AND P-SHEET *****

H1 : 107 H2 : 114 A1: 1.086 A2: 1.106 A1 < A2 FROM H1 TO H2

S1 : 107 S2 : 113 A1: 1.101 A2: 1.131 A1 < A2 FROM S1 TO S2

*** COMPARISON OF ASSIGNMENTS TYPES ***

HHF	HF	IIH	IH	BH	BBH	SSF	SF	IS	BS	BBS
4.00	3.00	0.0	0.50	-0.50	0.0	2.00	2.00	1.25	0.0	0.0

H1 : 107 H2 : 114 TTH: 7.000 TTS: 5.250 TTH > TTS FROM H1 TO H2

HHF	HF	IIH	IH	BH	BBH	SSF	SF	IS	BS	BBS
4.00	3.00	0.0	0.25	-0.50	0.0	2.00	2.00	1.00	0.0	0.0

S1 : 107 S2 : 113 TTH: 6.750 TTS: 5.000 TTH > TTS FROM S1 TO S2

BOUNDARY ANALYS. FOR HELIX FROM: 107 TO: 114 AND FOR SHEET FROM: 107 TO: 113

HN	SN	HC	SC	NHN	NSN	NHC	NSC
3.37	4.04	3.09	3.30	3.66	4.84	4.07	2.40

*** COMPARISON OF THEIR LENGTH *** L-HELIX: 6 L-SHEET: 5 RATIO=LH/LS: 1.0

***** COMPARISON OF P-HELIX AND P-SHEET *****

H1 : 119 H2 : 124 A1: 1.127 A2: 1.190 A1 < A2 FROM H1 TO H2

S1 : 120 S2 : 124 A1: 1.150 A2: 1.320 A1 < A2 FROM S1 TO S2

*** COMPARISON OF ASSIGNMENTS TYPES ***

HHF	HF	IIH	IH	BH	BBH	SSF	SF	IS	BS	BBS
2.00	4.00	0.50	0.0	0.0	0.0	4.00	2.00	0.25	0.0	-1.00

H1 : 119 H2 : 124 TTH: 6.500 TTS: 5.250 TTH > TTS FROM H1 TO H2

HHF	HF	IIH	IH	BH	BBH	SSF	SF	IS	BS	BBS
2.00	4.00	0.0	0.0	0.0	0.0	4.00	2.00	0.25	0.0	0.0

S1 : 120 S2 : 124 TTH: 6.000 TTS: 6.250 TTH < TTS FROM S1 TO S2

BOUNDARY ANALYS. FOR HELIX FROM: 119 TO: 124 AND FOR SHEET FROM: 120 TO: 124

HN	SN	HC	SC	NHN	NSN	NHC	NSC
3.85	4.20	2.58	3.29	3.51	3.94	3.82	3.26

END OF PROGRAM

```
**********************************
*                                *
*    BETA-TURN   PREDICTION      *
*                                *
**********************************
```

TOTAL NUMBER OF AA: 129
NUMBER OF DATA LINES: 9

PROTEIN SEQUENCE

12	20	14	8	2	5	7	11	1	1	1	13	12	2	9	8
11	4	3	19	2	8	19	16	11	8	3	18	20	5	1	1
12	14	7	16	3	14	3	17	6	1	17	3	2	3	17	4
8	16	17	4	19	8	10	11	6	11	3	16	2	18	18	5
3	4	8	2	17	15	8	16	2	3	11	5	3	10	15	5
16	1	11	11	16	16	4	10	17	1	16	20	3	5	1	12
12	10	20	16	4	8	4	8	13	3	1	18	20	1	18	2
3	2	5	12	8	17	4	20	6	1	18	10	2	8	5	2
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

DEFINITION OF PARAMETERS
PRB : PROBABILITY OF OCCURRENCE OF THE B-TURN STARTING FROM I
PRBO : PROBABILITY OF OCCURRENCE OF THE B-TURN STARTING FROM (I-1)

A1: 0.980	A2: 1.142	A3: 0.917	PRB: 0.0000260832	1
A1: 0.935	A2: 1.190	A3: 0.902	PRB: 0.0000410533	2
A1: 0.845	A2: 1.062	A3: 1.075	PRB: 0.0000635500	3
A1: 0.940	A2: 0.810	A3: 1.110	PRB: 0.0000809601	4
POTENTIAL BETA-TURN		4 7		
A1: 1.100	A2: 0.947	A3: 0.867	PRB: 0.0000199969	5
A1: 1.210	A2: 0.922	A3: 0.795	PRB: 0.0000186667	6
A1: 1.390	A2: 0.832	A3: 0.662	PRB: 0.0000028420	7
A1: 1.367	A2: 0.947	A3: 0.642	PRB: 0.0000094111	8
A1: 1.427	A2: 0.885	A3: 0.645	PRB: 0.0000087780	9
A1: 1.362	A2: 0.862	A3: 0.732	PRB: 0.0000060648	10
A1: 1.252	A2: 0.887	A3: 0.805	PRB: 0.0000301103	11
A1: 1.147	A2: 0.897	A3: 0.877	PRB: 0.0000418056	12
A1: 0.927	A2: 0.822	A3: 1.117	PRB: 0.0000824127	13
POTENTIAL BETA-TURN		13 16		
A1: 0.940	A2: 0.962	A3: 1.012	PRB: 0.0000437570	14
A1: 0.947	A2: 0.865	A3: 1.140	PRB: 0.0000347003	15
A1: 0.865	A2: 0.870	A3: 1.292	PRB: 0.0000415369	16

```
A1: 0.895     A2: 1.050     A3: 1.187     PRB: 0.0001602011       17

POTENTIAL  BETA-TURN        17        20
A1: 0.837     A2: 0.957     A3: 1.277     PRB: 0.0001182275       18

POTENTIAL  BETA-TURN        18        21

           PRBO: 0.00016020    PRB: 0.00011823      B-TURN NOT AT   18     BUT AT   17

A1: 0.727     A2: 1.010     A3: 1.302     PRB: 0.0001574771       19

POTENTIAL  BETA-TURN        19        22

           PRBO: 0.00011823    PRB: 0.00015748      B-TURN NOT AT   18     BUT AT   19

A1: 0.732     A2: 1.155     A3: 1.197     PRB: 0.0002064347       20

POTENTIAL  BETA-TURN        20        23

           PRBO: 0.00015748    PRB: 0.00020643      B-TURN NOT AT   19     BUT AT   20

A1: 0.752     A2: 0.975     A3: 1.270     PRB: 0.0000718997       21
A1: 0.810     A2: 1.067     A3: 1.180     PRB: 0.0000580124       22
A1: 0.810     A2: 1.067     A3: 1.180     PRB: 0.0000623697       23
A1: 0.805     A2: 0.922     A3: 1.285     PRB: 0.0000518699       24
A1: 0.882     A2: 1.077     A3: 1.167     PRB: 0.0001653858       25

POTENTIAL  BETA-TURN        25        28
A1: 0.845     A2: 1.177     A3: 1.145     PRB: 0.0000287166       26
A1: 0.877     A2: 1.287     A3: 1.052     PRB: 0.0000075013       27
A1: 1.065     A2: 1.272     A3: 0.827     PRB: 0.0000250810       28
A1: 1.150     A2: 1.137     A3: 0.752     PRB: 0.0000066706       29
A1: 1.175     A2: 0.897     A3: 0.880     PRB: 0.0000376522       30
A1: 1.282     A2: 0.945     A3: 0.732     PRB: 0.0000213408       31
A1: 1.305     A2: 0.830     A3: 0.752     PRB: 0.0000287039       32
A1: 1.142     A2: 0.810     A3: 0.945     PRB: 0.0000184053       33
A1: 1.020     A2: 0.847     A3: 1.082     PRB: 0.0000402674       34
A1: 1.020     A2: 0.847     A3: 1.082     PRB: 0.0000966383       35

POTENTIAL  BETA-TURN        35        38
A1: 0.810     A2: 0.977     A3: 1.287     PRB: 0.0000589133       36
A1: 0.825     A2: 1.087     A3: 1.170     PRB: 0.0000996024       37

POTENTIAL  BETA-TURN        37        40
A1: 0.935     A2: 1.140     A3: 1.025     PRB: 0.0000311938       38
A1: 1.007     A2: 1.002     A3: 1.040     PRB: 0.0000373146       39
A1: 1.047     A2: 1.077     A3: 0.890     PRB: 0.0000233034       40
A1: 1.007     A2: 1.002     A3: 1.040     PRB: 0.0000332659       41
A1: 0.975     A2: 0.960     A3: 1.032     PRB: 0.0001052027       42

POTENTIAL  BETA-TURN        42        45
```

```
A1: 0.787     A2: 0.975     A3: 1.257     PRB: 0.0000643061      43
A1: 0.787     A2: 0.975     A3: 1.257     PRB: 0.0002575084      44

POTENTIAL  BETA-TURN       44       47
A1: 0.872     A2: 0.887     A3: 1.232     PRB: 0.0000305896      45
A1: 0.770     A2: 0.842     A3: 1.385     PRB: 0.0004730921      46
POTENTIAL  BETA-TURN       46       49

A1: 0.795     A2: 0.807     A3: 1.352     PRB: 0.0001905241      47

POTENTIAL  BETA-TURN       47       50

           PRBO: 0.00047309    PRB: 0.00019052      B-TURN NOT AT   47     BUT AT   46

A1: 0.795     A2: 0.807     A3: 1.352     PRB: 0.0001233840      48

POTENTIAL  BETA-TURN       48       51

           PRBO: 0.00019052    PRB: 0.00012339      B-TURN NOT AT   48     BUT AT   47

A1: 0.795     A2: 0.807     A3: 1.352     PRB: 0.0000746471      49
A1: 0.825     A2: 0.987     A3: 1.247     PRB: 0.0002899796      50

POTENTIAL  BETA-TURN       50       53
A1: 0.775     A2: 0.987     A3: 1.280     PRB: 0.0001639226      51

POTENTIAL  BETA-TURN       51       54

           PRBO: 0.00028998    PRB: 0.00016392      B-TURN NOT AT   51     BUT AT   50

A1: 0.837     A2: 1.090     A3: 1.157     PRB: 0.0001016651      52

POTENTIAL  BETA-TURN       52       55

           PRBO: 0.00016392    PRB: 0.00010167      B-TURN NOT AT   52     BUT AT   51

A1: 0.887     A2: 1.280     A3: 0.940     PRB: 0.0000063427      53
A1: 0.992     A2: 1.187     A3: 0.900     PRB: 0.0000122351      54
A1: 1.152     A2: 1.325     A3: 0.657     PRB: 0.0000027842      55
A1: 1.050     A2: 1.147     A3: 0.930     PRB: 0.0000195839      56
A1: 0.940     A2: 1.010     A3: 1.140     PRB: 0.0000374550      57
A1: 0.907     A2: 0.967     A3: 1.132     PRB: 0.0000537943      58
A1: 0.875     A2: 0.985     A3: 1.225     PRB: 0.0003699916      59

POTENTIAL  BETA-TURN       59       62
A1: 0.977     A2: 1.105     A3: 1.075     PRB: 0.0001359511      60
A1: 0.960     A2: 1.215     A3: 1.015     PRB: 0.0000074547      61
A1: 0.882     A2: 1.205     A3: 1.167     PRB: 0.0000106576      62
A1: 0.865     A2: 0.997     A3: 1.292     PRB: 0.0000631370      63
A1: 0.737     A2: 0.842     A3: 1.442     PRB: 0.0003364808      64
```

```
POTENTIAL  BETA-TURN       64       67
A1: 0.807     A2: 0.777     A3: 1.382     PRB: 0.0002860161      65

POTENTIAL  BETA-TURN       65       68

           PRBO: 0.00033648    PRB: 0.00028602      B-TURN NOT AT   65     BUT AT   64

A1: 0.847     A2: 0.852     A3: 1.232     PRB: 0.0000977233      66

POTENTIAL  BETA-TURN       66       69

           PRBO: 0.00028602    PRB: 0.00009772      B-TURN NOT AT   66     BUT AT   65

A1: 0.737     A2: 0.855     A3: 1.247     PRB: 0.0000477890      67
A1: 0.737     A2: 0.855     A3: 1.247     PRB: 0.0000390700      68
A1: 0.685     A2: 0.810     A3: 1.367     PRB: 0.0005213434      69

POTENTIAL  BETA-TURN       69       72
A1: 0.722     A2: 0.745     A3: 1.365     PRB: 0.0000921187      70

POTENTIAL  BETA-TURN       70       73

           PRBO: 0.00052134    PRB: 0.00009212      B-TURN NOT AT   70     BUT AT   69

A1: 0.747     A2: 0.830     A3: 1.375     PRB: 0.0001277295      71

POTENTIAL  BETA-TURN       71       74

           PRBO: 0.00009212    PRB: 0.00012773      B-TURN NOT AT   70     BUT AT   71

A1: 0.907     A2: 0.967     A3: 1.132     PRB: 0.0001700662      72

POTENTIAL  BETA-TURN       72       75

           PRBO: 0.00012773    PRB: 0.00017007      B-TURN NOT AT   71     BUT AT   72

A1: 0.890     A2: 1.077     A3: 1.072     PRB: 0.0000267724      73
A1: 0.812     A2: 1.067     A3: 1.225     PRB: 0.0000428541      74
A1: 0.915     A2: 1.245     A3: 0.952     PRB: 0.0000345801      75
A1: 0.755     A2: 1.057     A3: 1.185     PRB: 0.0000109324      76
A1: 0.755     A2: 1.057     A3: 1.185     PRB: 0.0000238228      77
A1: 0.780     A2: 1.022     A3: 1.152     PRB: 0.0001605189      78

POTENTIAL  BETA-TURN       78       81
A1: 0.865     A2: 0.830     A3: 1.200     PRB: 0.0000391935      79
A1: 1.025     A2: 1.017     A3: 0.967     PRB: 0.0000507419      80
A1: 1.152     A2: 1.045     A3: 0.817     PRB: 0.0000229824      81
A1: 1.152     A2: 1.045     A3: 0.817     PRB: 0.0000057240      82
A1: 0.990     A2: 1.025     A3: 1.010     PRB: 0.0000202062      83
A1: 0.940     A2: 0.835     A3: 1.227     PRB: 0.0000858498      84
```

```
POTENTIAL  BETA-TURN       84       87
A1: 0.907     A2: 0.910     A3: 1.197     PRB: 0.0001672002      85

POTENTIAL  BETA-TURN       85       88

          PRBO: 0.00008585    PRB: 0.00016720     B-TURN NOT AT   84     BUT AT   85

A1: 0.922     A2: 1.020     A3: 1.080     PRB: 0.0000135564      86
A1: 1.085     A2: 1.040     A3: 0.887     PRB: 0.0000188424      87
A1: 1.025     A2: 1.092     A3: 0.880     PRB: 0.0000172292      88
A1: 1.020     A2: 1.117     A3: 0.887     PRB: 0.0000433009      89
A1: 0.980     A2: 1.042     A3: 1.037     PRB: 0.0000212503      90
A1: 0.800     A2: 1.132     A3: 1.170     PRB: 0.0001408203      91

POTENTIAL  BETA-TURN       91       94
A1: 0.962     A2: 1.152     A3: 0.977     PRB: 0.0000349207      92
A1: 0.987     A2: 0.912     A3: 1.105     PRB: 0.0000283722      93
A1: 1.110     A2: 0.875     A3: 0.967     PRB: 0.0000774560      94
A1: 1.205     A2: 0.977     A3: 0.787     PRB: 0.0000278207      95
A1: 1.115     A2: 1.195     A3: 0.747     PRB: 0.0000043579      96
A1: 1.017     A2: 1.197     A3: 0.852     PRB: 0.0000055502      97
A1: 0.980     A2: 1.147     A3: 0.965     PRB: 0.0000208980      98
A1: 0.852     A2: 0.935     A3: 1.237     PRB: 0.0002344783      99

POTENTIAL  BETA-TURN       99      102

A1: 0.840     A2: 0.645     A3: 1.477     PRB: 0.0002031477     100

POTENTIAL  BETA-TURN      100      103

          PRBO: 0.00023448    PRB: 0.00020315     B-TURN NOT AT  100     BUT AT   99

A1: 0.790     A2: 0.645     A3: 1.510     PRB: 0.0003399635     101

POTENTIAL  BETA-TURN      101      104

          PRBO: 0.00020315    PRB: 0.00033996     B-TURN NOT AT  100     BUT AT  101

A1: 0.900     A2: 0.772     A3: 1.295     PRB: 0.0001172489     102

POTENTIAL  BETA-TURN      102      105

          PRBO: 0.00033996    PRB: 0.00011725     B-TURN NOT AT  102     BUT AT  101

A1: 0.925     A2: 0.807     A3: 1.295     PRB: 0.0000159186     103
A1: 1.027     A2: 0.880     A3: 1.095     PRB: 0.0000926563     104

POTENTIAL  BETA-TURN      104      107
A1: 1.155     A2: 1.035     A3: 0.945     PRB: 0.0000329891     105
A1: 1.057     A2: 1.197     A3: 0.920     PRB: 0.0000415044     106
A1: 1.245     A2: 1.182     A3: 0.695     PRB: 0.0000012667     107
```

```
A1: 1.160     A2: 1.317     A3: 0.770     PRB: 0.0000216031     108
A1: 1.135     A2: 1.207     A3: 0.767     PRB: 0.0000256332     109
A1: 1.037     A2: 1.005     A3: 1.032     PRB: 0.0000070270     110
A1: 0.927     A2: 1.030     A3: 1.105     PRB: 0.0001325099     111

POTENTIAL  BETA-TURN      111      114
A1: 0.832     A2: 0.985     A3: 1.162     PRB: 0.0000736242     112
A1: 0.877     A2: 0.937     A3: 1.177     PRB: 0.0001896883     113

POTENTIAL  BETA-TURN      113      116
A1: 0.852     A2: 0.902     A3: 1.177     PRB: 0.0000406022     114
A1: 0.815     A2: 0.967     A3: 1.180     PRB: 0.0002571961     115

POTENTIAL  BETA-TURN      115      118
A1: 0.892     A2: 0.805     A3: 1.247     PRB: 0.0000246138     116
A1: 0.867     A2: 1.045     A3: 1.120     PRB: 0.0001045087     117

POTENTIAL  BETA-TURN      117      120
A1: 1.002     A2: 1.132     A3: 0.975     PRB: 0.0000259582     118
A1: 1.150     A2: 1.042     A3: 0.900     PRB: 0.0000151422     119
A1: 1.167     A2: 1.250     A3: 0.775     PRB: 0.0000355142     120
A1: 1.172     A2: 1.225     A3: 0.767     PRB: 0.0000201564     121
A1: 1.140     A2: 1.182     A3: 0.760     PRB: 0.0000008619     122
A1: 0.927     A2: 1.162     A3: 0.985     PRB: 0.0000393956     123
A1: 0.832     A2: 1.117     A3: 1.042     PRB: 0.0001108504     124
A1: 0.807     A2: 0.950     A3: 1.162     PRB: 0.0000591727     125
A1: 0.865     A2: 1.042     A3: 1.072     PRB: 0.0000374636     126

END OF PROGRAM          77
```