

# ATLAS of PROTEIN SEQUENCE and STRUCTURE 1965

FACILITY FORM 602

N66-23819  
(ACCESSION NUMBER)

~~111~~  
(PAGES)

CR-71805  
(NASA CR OR TMX OR AD NUMBER)

(THRU)

1  
(CODE)

04  
(CATEGORY)

LYS LYS GLU  
ALA THR ASN GLU  
SER CYS GLY ALA CYS  
SER GLN GLY ASP SER ILU  
ASP CYS GLY ASN CYS ALA  
AL GLN GLU DAY HOF FNG RAN  
TYR LYS GLU LEU GLY TYR GLY  
ALA MET GLY LYS ALA LEU GLU LEU  
ARG ALA THR LYS HIS ASP ASP GLU  
ILU PRO ILU LYS TYR GLU HIS LEU  
GLU LYS LYS LYS ILU PHE ILU  
THR VAL LYS GLY GLY  
LYS

GPO PRICE \$ \_\_\_\_\_

CFSTI PRICE(S) \$ \_\_\_\_\_

Hard copy (HC) 4.00

Microfiche (MF) 175

**Margaret O. Dayhoff**  
**Richard V. Eck**  
**Marie A. Chang**  
**R. Sochard**

# 653 July 65

**NATIONAL BIOMEDICAL RESEARCH FOUNDATION**  
8600 16TH STREET  
Silver Spring, Maryland

**ATLAS of  
PROTEIN SEQUENCE  
and STRUCTURE  
1965**

**Margaret O. Dayhoff  
Richard V. Eck  
Marie A. Chang  
Minnie R. Sochard**



**NATIONAL BIOMEDICAL RESEARCH FOUNDATION**

8600 16TH STREET  
*Silver Spring, Maryland*

ATLAS OF PROTEIN SEQUENCE AND STRUCTURE

1965

Library of Congress Card Catalogue Number 65-29342

© Copyright 1965

by

The National Biomedical Research Foundation

Extra copies may be purchased from:

The National Biomedical Research Foundation

8600 16th Street, Silver Spring, Maryland

## DEDICATION

To all the investigators who have developed the techniques necessary for the grand accomplishments represented by this tabulation, and to all those who have spent so much tedious effort in their application.

We would gratefully appreciate receiving suggestions, corrections, new data (even if fragmentary or provisional), and references to any data omitted from this volume.

M. O. D.  
R. V. E.  
M. A. C.  
M. R. S.

## PREFACE

This Atlas voluminously illustrates the triumph of experimental technique over the secretiveness of nature. Perhaps nowhere has the power of the scientific method been more brilliantly demonstrated than in the development of procedures for the study of the chemistry of life. As recently as twenty years ago, it was customary for biologists to have a hopeless attitude about biochemistry. Some details might be elicited, perhaps, but living things were thought to be so very complex and intricate that there surely was no hope of fully "understanding" them in all their chemical detail. Who, if he really comprehended the difficulty of the problem, would dare to think of man's ever knowing the detailed structure of a protein, for example, much less be able to synthesize it? Who would ever understand the mechanism of an enzyme as clearly as a chemist understands the details of an inorganic reaction? How could we ever hope to know the atomic details of a protein crystal?

Today some of these ambitions have already been attained, and the others no longer seem out of reach. We now rationally hope to be able to discover and understand the finest chemical details of living processes. These accomplishments and hopes have been made possible by the combined effect of several new approaches.

Techniques which permit the separation of chemically similar compounds have been developed for microgram samples. Among these are ion-exchange columns, paper chromatography, electrophoresis, and counter-current distribution. Radioactive tracer techniques and other micro-quantitative analytical procedures, often dependent on electronics and automation, have aided the analyses. X-ray crystallography, starting with the art of protein crystal production and ending with the processing of great numbers of experimental observations in the high-speed computer, has permitted a glimpse of three-dimensional structure.

Confidence in our understanding of experimental procedures and relationships among proteins has grown so great that sequences of amino acids are inferred from those found in homologous proteins. This technique requires only a small proportion of the analytical work needed to sequence a protein with no known relatives. The effectiveness of laboratory effort is thus magnified.

Some of the insights which have been developed cannot be attributed to any particular worker or school. Perhaps the greatest of these insights is that nature always uses "building blocks." A living cell is extremely complex and almost unimaginably intricate in detail. But it consists of a limited, understandable number of types of processes, reduplicated with variations. To understand the cell, we must have a few examples of each type of process, from which we can see the overall principles. For understanding, we need not work out the details of all the variations on these principles, although we may eventually choose to do so for medical or other practical reasons. Similarly, the analysis of such large, complex chemical molecules as proteins has been made possible by the recognition of their essential modularity, their building-block nature. Proteins are precise chemical structures built from regular subunits,

not statistical mixtures or hopelessly intricate molecular conglomerates as was once thought. It is by means of the discovery and utilization of such building block principles, combined with the large-scale application of new and improved techniques, that we now dare hope to make all of living nature accessible to our understanding.

Hidden in the amino acid sequence of a protein is the chemical information that produces its three-dimensional structure. In the case of an enzyme, this structure forms locks into which the proper keys—cell chemicals—fit. By these locks, the enzymes bring the proper reactants together quickly, efficiently, and selectively. Uncatalyzed reactions cannot compete with such specifically catalyzed reactions; therefore, the presence of enzymes determines which reactions can take place in living chemistry. In many cases, if not all, this three-dimensional structure is fully determined by the information in the one-dimensional sequence. The folding is the thermodynamically most stable result of all the possible intermolecular forces, such as hydrogen bonds and hydrophobic bonds, which can form between the various links of the chain. In principle, if we knew these forces in detail, and if we had appropriate computer routines, we should then be able to determine the three-dimensional structure of a protein, given only its amino acid sequence.

Also hidden in the sequences is information about the genes which directed their synthesis. For each amino acid there are a small number of possible corresponding nucleotide triplets in the gene. That is, each protein sequence corresponds to a limited number of possible nucleotide sequences. When nucleotide mutations occur, the substitution of alternative amino acids is not random. Analysis of amino acid sequence data, considered as a mathematical puzzle, can help elucidate both the mathematical details of the genetic code and the structural aspects of the genetic mechanism.

Hidden in each family of homologous sequences is the story of its evolution. Simple organisms, caught in their primitive ecological niches, still preserve even today enzymes performing primeval functions, held relatively fixed by natural selection. Even the older proteins of man are preserved as living "fossils" in his metabolism.

Enmeshed also in homologous sequences are the records of the many thousands of mutational steps by which we can quantify a phylogenetic tree. Each amino acid link is a trait by which we can trace species evolution. By comparison, the traditional taxonomic criteria are extremely vague and uncertain. In the case of distant relationships, they often break down completely. A truly quantitative and inclusive system of phylogenetic classification would be of great help to comparative physiologists and other students of evolution.

Conspicuous in comparative human protein sequences is information of great medical-diagnostic value. A long series of abnormalities has been found to be attributable to single amino acid replacements. One such tragically serious disease is sickle-cell anemia.

To facilitate the theoretical study of the protein sequences which have already been so ingeniously and laboriously determined, we have undertaken this compilation.

The information is kept in a compact, uniform format on punched cards. New information and corrections are easily inserted, while the text is kept accurate.

It is our intention to include the currently accepted amino acid sequence of every protein for which complete or substantial data is available. Usually, only the definitive report giving the complete sequence from each laboratory will be referenced. If a substantial amount of work has been done on the same protein in other laboratories, their reports will also be referenced. We have also included some smaller peptides that have come to our attention. Unusual polypeptides which are presumably not produced by the genetic code have been omitted.

The format in which the Atlas is kept on punched cards is suitable for direct use in our computer programs. We use a three-letter code, which is a slight modification of the conventional notation, and also a mnemonic one-letter code which is clearer and much more suitable for certain comparative studies. We use a system of punctuation to describe the degree of confidence in each bond. Brief remarks are also included about the nature and function of the protein, and additional structural information such as the attachment of prosthetic groups, the location of S-S bonds, amino acids involved in active sites, and three-dimensional structures. In later editions we intend to include a section in which the alignment of all sequences of each family is given. Possibly we will also have sections on alleles and on mathematical methods and computer programs to treat the information.

This first edition is incomplete and imperfect and is intended mainly for distribution to investigators who have published protein sequence analyses, to acquaint them with the existence of this Atlas. We would gratefully appreciate their cooperation in making corrections, additions and suggestions for future editions. Since sequences are being reported in great numbers, we plan to bring out supplementary material in six months and a second edition in a year.

We thank all those who have assisted with this compilation, particularly Mr. Javier Albarran for his help with the computer aspects and Miss Lorrie Goldstein for her design of the cover.

The tabulations and computations were made at the University of Maryland Computer Science Center, College Park.

This work was supported by Grants GM-08710 and GM-12168 from the National Institutes of Health to the National Biomedical Research Foundation.

TABLE OF CONTENTS

	PAGE
PREFACE . . . . .	0.001
CONTENTS . . . . .	0.002
EXPLANATION OF NOTATION . . . . .	0.010
 I. CYTOCHROME C	 1.000
BAKER'S YEAST . . . . . CY BY	1.001
CHICKEN . . . . . CY CH	1.002
HORSE . . . . . CY HO	1.003
HUMAN . . . . . CY HU	1.004
PIG AND BOVINE . . . . . CY PG	1.005
PSEUDOMONAS . . . . . CY PS	1.006
TUNA FISH . . . . . CY TF	1.007
BOMBYX MORI (SILKWORM) . . . . . CY SW	1.008
RATTLESNAKE . . . . . CY RS	1.009
RHODOSPIRILLUM RUBRUM . . . . . CY RR	1.010
SALMON . . . . . CY SM	1.011
 II. GLOBINS	 2.000
HEMOGLOBINS	2.000
HUMAN ALPHA . . . . . GL HUHA	2.001
HUMAN BETA . . . . . GL HUHB	2.002
HUMAN GAMMA . . . . . GL HUHG	2.003
GORILLA BETA . . . . . GL GOHB	2.004
HORSE BETA . . . . . GL HOHB	2.005
HORSE ALPHA . . . . . GL HOHA	2.006
LEMUR FULVUS BETA . . . . . GL LEHB	2.007
ABNORMAL HUMAN HEMOGLOBINS . . . . . GL HUH	2.020
MYOGLOBINS	2.100
WHALE . . . . . GL WHMY	2.101
 III. OTHER RESPIRATORY PROTEINS	 3.000
DIHEME PEPTIDE - CHROMATIUM . . . . . DH CH	3.001
FERREDOXIN - CLOSTRIDIUM PASTEURIANUM . . . . . FE CP	3.002
AZURIN - PSEUDOMONAS . . . . . AZ PS	3.003
 IV. RIBONUCLEASE	 4.000
BOVINE . . . . . RN BO	4.001
 V. INHIBITORS	 5.000
TRYPSIN INHIBITOR - BOVINE . . . . . TI BOPA	5.001



PAGE

VI. TOBACCO MOSAIC VIRUS

6.000

TOBACCO MOSAIC VIRUS . . . . .	TM TM	6.001
TOBACCO MOSAIC VIRUS DAHLMENSE . . . . .	TM TMD	6.002

VII. DIGESTIVE ENZYMES

7.000

CHYMOTRYPSINOGEN-A - BOVINE . . . . .	TR BOCH	7.001
TRYPSINOGEN - BOVINE . . . . .	TR BOTR	7.002
PAPAIN . . . . .	PA PA	7.101
LYSOZYME - CHICKEN . . . . .	LS CH	7.201

VIII. HORMONES

8.000

GLUCAGON

BOVINE . . . . .	GN BO	8.001
------------------	-------	-------

PRESSINS

VASOPRESSIN ARGININE . . . . .	PR BOAR	8.101
VASOPRESSIN LYSINE . . . . .	PR PGLS	8.102
OXYTOCIN . . . . .	PR BOOX	8.103
HYPERTENSIN . . . . .	PR BOHY	8.104

CORTICOIDS

ALPHA MELANOCYTE-STIMULATING HORMONE		
BOVINE, PIG, AND HORSE . . . . .	TN BPAM	8.201
BETA MELANOCYTE-STIMULATING HORMONE		
BOVINE . . . . .	TN BOBM	8.202
PIG . . . . .	TN PGBM	8.203
HORSE . . . . .	TN HOBM	8.204
HUMAN . . . . .	TN HUBM	8.205
BETA CORTICOTROPIN - PIG . . . . .	TN PGAC	8.206
ALPHA CORTICOTROPIN - SHEEP AND BOVINE . . . . .	TN SBAC	8.207

INSULINS

INSULIN A		
BOVINE . . . . .	IS BOA	8.301
BONITO . . . . .	IS BNA	8.302
HORSE . . . . .	IS HOA	8.303
SHEEP . . . . .	IS SHA	8.304
SPERM WHALE, FIN-WHALE, PIG, AND HUMAN . . . . .	IS WPA	8.305
SEI-WHALE . . . . .	IH WHA	8.306

INSULIN B		
BOVINE, SHEEP, HORSE, HUMAN, PIG, AND SPERM WHALE . . . . .	IS BOB	8.321
BONITO . . . . .	IS BNB	8.322

PAGE

IX. PLASMA PROTEINS

9.000

FIBRINOPEPTIDE A

BOVINE. . . . .	FB BOA	9.001
SHEEP . . . . .	FB SHA	9.002
GOAT. . . . .	FB GTA	9.003
REINDEER. . . . .	FB RDA	9.004
PIG . . . . .	FB PGA	9.005
HUMAN . . . . .	FB HUA	9.006
RABBIT. . . . .	FB RTA	9.007

FIBRINOPEPTIDE B

BOVINE. . . . .	FB BOB	9.101
SHEEP . . . . .	FB SHB	9.102
GOAT. . . . .	FB GTB	9.103
REINDEER. . . . .	FB RDB	9.104
PIG . . . . .	FB PGB	9.105
HUMAN . . . . .	FB HUB	9.106
RABBIT. . . . .	FB RTB	9.107

X. GLOBULINS

10.000

BENCE-JONES . . . . .	BJ HU	10.001
-----------------------	-------	--------

AUTHOR INDEX

100.000

## THE MEANING OF THE PUNCTUATION IS AS FOLLOWS.

BLANK	SEQUENCE OF AMINO ACIDS HAS BEEN DETERMINED.
( )	ENCLOSE PORTION OF SEQUENCE NOT SPECIFICALLY DETERMINED. TO PRESERVE PROPER SPACING, IS USED INSTEAD OF )(
=	
BUT	, • SEPARATES AMINO ACIDS WITHIN PARENTHESES, SEPARATES AMINO ACIDS, WHERE THE SEQUENCE IS PRESUMED BY HOMOLOGY WITH A KNOWN SEQUENCE.
/ OR ///	FRAGMENT, CONNECTION UNDETERMINED
• OR ***	CARBOXYL END OF PROTEIN

• ASTERISK BEFORE REFERENCE INDICATES THAT THE SEQUENCE WAS COPIED FROM, AND PROOFREAD AGAINST, THE ORIGINAL ARTICLE.

= BEFORE REFERENCE INDICATES THAT WE HAVE NOT SEEN THE ORIGINAL ARTICLE.

NO MARK BEFORE REFERENCE INDICATES OTHER GROUPS WHICH HAVE ALSO REPORTED WORK ON THE SAME PROTEIN.

BOTH SINGLE- AND THREE-LETTER NOTATIONS ARE USED, AS FOLLOWS.

A = ALA = ALANINE	M = MET = METHIONINE
C = CYS = CYSTEINE	N = ASN = ASPARAGINE
D = ASP = ASPARTIC ACID	O = TYR = TYROSINE
E = GLU = GLUTAMIC ACID	P = PRO = PROLINE
F = PHE = PHENYLALANINE	Q = GLN = GLUTAMINE
G = GLY = GLYCINE	R = ARG = ARGININE
H = HIS = HISTIDINE	S = SER = SERINE
I = ILE = ISOLEUCINE	T = THR = THREONINE
K = LYS = LYSINE	W = TRP = TRYPTOPHAN
L = LEU = LEUCINE	V = VAL = VALINE

B = ASX = ASPARTIC ACID OR ASPARAGINE  
 Z = GLX = GLUTAMIC ACID OR GLUTAMINE  
 X = XXX = UNDETERMINED OR OTHERWISE UNUSUAL

#### MNEMONICS OF THE ONE-LETTER CODE

IF POSSIBLE, THE INITIAL LETTER OF THE AMINO ACID IS USED.  
 IF MORE THAN ONE AMINO ACID BEGINS WITH THE SAME LETTER,  
 THE MOST COMMONLY-OCCURRING ONE IS ASSIGNED THE INITIAL.

A = ALANINE	I = ISOLEUCINE	S = SERINE
C = CYSTEINE	L = LEUCINE	T = THREONINE
G = GLYCINE	M = METHIONINE	V = VALINE
H = HISTIDINE	P = PROLINE	

SOME OF THE OTHERS ARE PHONETICALLY SUGGESTIVE.

F = PHENYLALANINE  
 R = ARGININE  
 O = TYROSINE

DOUBLE RING IN THE SIDE CHAIN.

W = TRYPTOPHAN

THE TWO ACIDS ARE ADJACENT, IN ALPHABETICAL ORDER.

D = ASPARTIC ACID  
 E = GLUTAMIC ACID

THE TWO AMINES HAVE LETTERS FROM THE MIDDLE OF THE ALPHABET.

N = ASPARAGINE (CONTAINS N)  
 Q = GLUTAMINE ('Q-TAMINE')

NON-INITIAL LETTER AS CLOSE AS POSSIBLE TO ITS INITIAL.

K = LYSINE

## CYTOCHROME C - BAKER'S YEAST

HEME BONDED TO CYSTEINES AT POSITIONS 19 AND 22.

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 T E F K A G S A K K G A T L F K T R C E L C H T V E K G G P  
 31 H K V G P N L H G I F G R H S G Q A Q G D S O T D A N I K K  
 61 N V L W D E N N M S E O L T N P K K O I P G T K M A F G G L  
 91 K K E K D R N D L I T O L K K A C E \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 THR GLU PHE LYS ALA GLY SER ALA LYS LYS GLY ALA THR LEU PHE  
 LYS THR ARG CYS GLU LEU CYS HIS THR VAL GLU LYS GLY GLY PRO  
 31 HIS LYS VAL GLY PRO ASN LEU HIS GLY ILU PHE GLY ARG HIS SER  
 GLY GLN ALA GLN GLY TYR SER TYR THR ASP ALA ASN ILU LYS LYS  
 61 ASN VAL LEU TRP ASP GLU ASN ASN MET SER GLU TYR LEU THR ASN  
 PRO LYS LYS TYR ILU PRO GLY THR LYS MET ALA PHE GLY GLY LEU  
 91 LYS LYS GLU LYS ASP ARG ASN ASP LEU ILU THR TYR LEU LYS LYS  
 ALA CYS GLU \*\*\*

## COMPOSITION

7 ALA A	2 GLN Q	8 LEU L	4 SER S
3 ARG R	7 GLU E	16 LYS K	8 THR T
7 ASN N	12 GLY G	2 MET M	1 TRP W
4 ASP D	4 HIS H	4 PHE F	5 TYR O
3 CYS C	4 ILU I	4 PRO P	3 VAL V

TOTAL NO. OF ACIDS = 108

## CYTOCHROME C - CHICKEN

ACETYL AT AMINO END.

HEME BONDED TO CYSTEINES AT POSITIONS 14 AND 17.

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 G D I E K G K K I F V Q K C S Q C H T V E K G G K H K T G P  
 31 N L H G L F G R K T G Q A E G F S O T D A N K N K G I T W G  
 61 E D T L M E O L E N P K K O I P G T K M I F A G I K K K S E  
 91 R V D L I A O L K K A T N S \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 GLY ASP ILU GLU LYS GLY LYS LYS ILU PHE VAL GLN LYS CYS SER  
 GLN CYS HIS THR VAL GLU LYS GLY GLY LYS HIS LYS THR GLY PRO  
 31 ASN LEU HIS GLY LEU PHE GLY ARG LYS THR GLY GLN ALA GLU GLY  
 PHE SER TYR THR ASP ALA ASN LYS ASN LYS GLY ILU THR TRP GLY  
 61 GLU ASP THR LEU MET GLU TYR LEU GLU ASN PRO LYS LYS TYR ILU  
 PRO GLY THR LYS MET ILU PHE ALA GLY ILU LYS LYS LYS SER GLU  
 91 ARG VAL ASP LEU ILU ALA TYR LEU LYS LYS ALA THR ASN SER \*\*\*

## COMPOSITION

5 ALA A	3 GLN Q	6 LEU L	4 SER S
2 ARG R	7 GLU E	18 LYS K	8 THR T
5 ASN N	13 GLY G	2 MET M	1 TRP W
4 ASP D	3 HIS H	4 PHE F	4 TYR O
2 CYS C	7 ILU I	3 PRO P	3 VAL V

TOTAL NO. OF ACIDS = 104

## CYTOCHROME C - HORSE

ACETYL AT AMINO END.

HEME BONDED TO CYSTEINES AT POSITIONS 14 AND 17.

OXIDATION-REDUCTION POTENTIAL EQUALS .250 V.

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 G D V E K G K K I F V Q K C A Q C H T V E K G G K H K T G P  
 31 N L H G L F G R K T G Q A P G F T O T D A N K N K G I T W K  
 61 E E T L M E O L E N P K K D I P G T K M I F A G I K K K T E  
 91 R E D L I A O L K K A T N E \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 GLY ASP VAL GLU LYS GLY LYS LYS ILU PHE VAL GLN LYS CYS ALA  
 GLN CYS HIS THR VAL GLU LYS GLY GLY LYS HIS LYS THR GLY PRO  
 31 ASN LEU HIS GLY LEU PHE GLY ARG LYS THR GLY GLN ALA PRO GLY  
 PHE THR TYR THR ASP ALA ASN LYS ASN LYS GLY ILU THR TRP LYS  
 61 GLU GLU THR LEU MET GLU TYR LEU GLU ASN PRO LYS LYS TYR ILU  
 PRO GLY THR LYS MET ILU PHE ALA GLY ILU LYS LYS LYS THR GLU  
 91 ARG GLU ASP LEU ILU ALA TYR LEU LYS LYS ALA THR ASN GLU \*\*\*

## COMPOSITION

6 ALA	A	3 GLN	Q	6 LEU	L	0 SER	S
2 ARG	R	9 GLU	E	19 LYS	K	10 THR	T
5 ASN	N	12 GLY	G	2 MET	M	1 TRP	W
3 ASP	D	3 HIS	H	4 PHE	F	4 TYR	O
2 CYS	C	6 ILU	I	4 PRO	P	3 VAL	V

TOTAL NO. OF ACIDS = 104

\* MARGOLIASH, E., SMITH, E. L., KREIL, G., AND TUPPY, H., NATURE,  
 VOL. 192, NO. 4808, PP. 1121-1127, DEC. 23, 1961

## CYTOCHROME C - HUMAN

ACETYL AT AMINO END.

HEME BONDED TO CYSTEINES AT POSITIONS 14 AND 17.

LEU (L) REPLACES MET (M) AT POSITION 65 IN 10 PERCENT  
YIELD IN POOLED PROTEIN.

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 G D V E K G K K I F I M K C S Q C H T V E K G G K H K T G P  
 31 N L H G L F G R K T G Q A P G O S O T A A N K N K G I I W G  
 61 E D T L M E O L E N P K K O I P G T K M I F V G I K K K E E  
 91 R A D L I A O L K K A T N E \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 GLY ASP VAL GLU LYS GLY LYS LYS ILU PHE ILU MET LYS CYS SER  
 GLN CYS HIS THR VAL GLU LYS GLY GLY LYS HIS LYS THR GLY PRO  
 31 ASN LEU HIS GLY LEU PHE GLY ARG LYS THR GLY GLN ALA PRO GLY  
 TYR SER TYR THR ALA ALA ASN LYS ASN LYS GLY ILU ILU TRP GLY  
 61 GLU ASP THR LEU MET GLU TYR LEU GLU ASN PRO LYS LYS TYR ILU  
 PRO GLY THR LYS MET ILU PHE VAL GLY ILU LYS LYS LYS GLU GLU  
 91 ARG ALA ASP LEU ILU ALA TYR LEU LYS LYS ALA THR ASN GLU \*\*\*

## COMPOSITION

6 ALA	A	2 GLN	Q	6 LEU	L	2 SER	S
2 ARG	R	8 GLU	E	18 LYS	K	7 THR	T
5 ASN	N	13 GLY	G	3 MET	M	1 TRP	W
3 ASP	D	3 HIS	H	3 PHE	F	5 TYR	O
2 CYS	C	8 ILU	I	4 PRO	P	3 VAL	V

TOTAL NO. OF ACIDS = 104

• MATSUBARA, H., AND SMITH, E.L., J. BIOL. CHEM., VOL. 237, NO.11,  
 PC3575-PC3576, NOV., 1962



## CYTOCHROME - C PIG AND BOVINE

ACETYL AT AMINO END.

HEME BONDED TO CYSTEINES AT POSITIONS 14 AND 17.

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 G D V E K G K K I F V Q K C A Q C H T V E K G G K H K T G P  
 31 N L H G L F G R K T G Q A P G F S O T D A N K N K G I T W G  
 61 E E T L M E O L E N P K K O I P G T K M I F A G I K K K G E  
 91 R E D L I A O L K K A T N E \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 GLY ASP VAL GLU LYS GLY LYS LYS ILU PHE VAL GLN LYS CYS ALA  
 GLN CYS HIS THR VAL GLU LYS GLY GLY LYS HIS LYS THR GLY PRO  
 31 ASN LEU HIS GLY LEU PHE GLY ARG LYS THR GLY GLN ALA PRO GLY  
 PHE SER TYR THR ASP ALA ASN LYS ASN LYS GLY ILU THR TRP GLY  
 61 GLU GLU THR LEU MET GLU TYR LEU GLU ASN PRO LYS LYS TYR ILU  
 PRO GLY THR LYS MET ILU PHE ALA GLY ILU LYS LYS LYS GLY GLU  
 91 ARG GLU ASP LEU ILU ALA TYR LEU LYS LYS ALA THR ASN GLU \*\*\*

## COMPOSITION

6 ALA	A	3 GLN	Q	6 LEU	L	1 SER	S
2 ARG	R	9 GLU	E	18 LYS	K	8 THR	T
5 ASN	N	14 GLY	G	2 MET	M	1 TRP	W
3 ASP	D	3 HIS	H	4 PHE	F	4 TYR	O
2 CYS	C	6 ILU	I	4 PRO	P	3 VAL	V

TOTAL NO. OF ACIDS = 104

\* MARGOLIASH, E., NEEDLEMAN, S. B. AND STEWART, J. W., ACTA CHEM. SCAND.,  
 VOL. 17, SUPPL. 1, PP. 250-256, 1963 (PIG)

TUPPY H., AND PALEUS, S., ACTA CHEM. SCAND., VOL. 13, NO.4,  
PP. 641-646, 1959, (HEME ATTACHEMENT REGION ONLY - BOVINE)

YASUNOBU, K. T., NAKASHIMA, T., HIGA, H., MATSUBARA, H., AND  
BENSON, A. , BIOCHIM. BIOPHYS. ACTA VOL. 78, PN1324  
PP. 791-794, 1963 (BOVINE)

## CYTOCHROME C - PSEUDOMONAS

HEME BONDED TO CYSTEINES AT POSITIONS 12 AND 15.  
THE AMINO END IS NOT ACETYLATED.

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 E D P E V L F K N K G C V A C H A I D T K M V G P A O K D V  
 31 A A K F A G Q A G A E A E L A Q R I K N G S Q G V W G P I P  
 61 M P P N A V S D D E A Q T L A K W V L S Q K \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 GLU ASP PRO GLU VAL LEU PHE LYS ASN LYS GLY CYS VAL ALA CYS  
 HIS ALA ILU ASP THR LYS MET VAL GLY PRO ALA TYR LYS ASP VAL  
 31 ALA ALA LYS PHE ALA GLY GLN ALA GLY ALA GLU ALA GLU LEU ALA  
 GLN ARG ILU LYS ASN GLY SER GLN GLY VAL TRP GLY PRO ILU PRO  
 61 MET PRO PRO ASN ALA VAL SER ASP ASP GLU ALA GLN THR LEU ALA  
 LYS TRP VAL LEU SER GLN LYS \*\*\*

## COMPOSITION

13 ALA	A	5 GLN	Q	4 LEU	L	3 SER	S
1 ARG	R	5 GLU	E	8 LYS	K	2 THR	T
3 ASN	N	7 GLY	G	2 MET	M	2 TRP	W
5 ASP	D	1 HIS	H	2 PHE	F	1 TYR	O
2 CYS	C	3 ILU	I	6 PRO	P	7 VAL	V

TOTAL NO. OF ACIDS = 82

• AMBLER, R. P., BIOCHEM. J., VOL.89, P.349-378. 1963

## CYTOCHROME C - TUNA FISH

ACETYL AT AMINO END.

HEME BONDED TO CYSTEINES AT POSITIONS 14 AND 17.

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 G D V A K.G K K.T F V Q K.C A Q(C.H)T V E N G G K.H K(V.G.P.  
 31 N)L W.G L F.G R.K T(G.Q)A E G D.S O T(D.A.N)K.S K.G I V W(N,  
 61 N,D)T L M E O.L E N P K K.O(I.P.G)T K(M.I)F.A G I K K.K G E  
 91 R.Q D L(V.A)O.L K S T A S •

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 GLY ASP VAL ALA LYS.GLY LYS LYS.THR PHE VAL GLN LYS.CYS ALA  
 GLN(CYS.HIS)THR VAL GLU ASN GLY GLY LYS.HIS LYS(VAL.GLY.PRO.  
 31 ASN)LEU TRP.GLY LEU PHE.GLY ARG.LYS THR(GLY.GLN)ALA GLU GLY  
 TYR.SER TYR THR(ASP.ALA.ASN)LYS.SER LYS.GLY ILU VAL TRP(ASN,  
 61 ASN,ASP)THR LEU MET GLU TYR.LEU GLU ASN PRO LYS LYS.TYR(ILU.  
 PRO.GLY)THR LYS.MET.ILU)PHE.ALA GLY ILU LYS LYS.LYS GLY GLU  
 91 ARG.GLN ASP LEU(VAL.ALA)TYR.LEU LYS SER THR ALA SER \*\*\*

## COMPOSITION

7 ALA	A	4 GLN	Q	6 LEU	L	4 SER	S
2 ARG	R	5 GLU	E	16 LYS	K	7 THR	T
6 ASN	N	13 GLY	G	2 MET	M	2 TRP	W
4 ASP	D	2 HIS	H	3 PHE	F	5 TYR	O
2 CYS	C	4 ILU	I	3 PRO	P	6 VAL	V

TOTAL NO. OF ACIDS = 103

\* KREIL, G., Z. PHYSIOL. CHEM., BD. 334, PP.154-166, 1963

## CYTOCHROME C - BOMBYX MORI (SILKWORM)

HEME BONDED TO CYSTEINES AT POSITIONS 4 AND 7 OF FRAGMENT.

1 2 3 4 5 6 7 8 9 0 1  
 / V Q R C A Q C H T(V,E)/

1 2 3 4 5 6 7 8 9 10 11  
 /// VAL GLN ARG CYS ALA GLN CYS HIS THR(VAL,GLU)///

## COMPOSITION OF FRAGMENT

1 ALA	A	2 GLN	Q	0 LEU	L	0 SER	S
1 ARG	R	1 GLU	E	0 LYS	K	1 THR	T
0 ASN	N	0 GLY	G	0 MET	M	0 TRP	W
0 ASP	D	1 HIS	H	0 PHE	F	0 TYR	O
2 CYS	C	0 ILE	I	0 PRO	P	2 VAL	V

TOTAL NO. OF ACIDS IN FRAGMENT = 11

\* TUPPY H., Z. NATURFORSCH., VOL.12, PP.784-788, 1957

## CYTOCHROME C - RATTLESNAKE

ACETYL AT AMINO END.

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 G D V E K G K K I F (I.T.K.C.S.Q.C.H.T.V.E.K.G.G.K.H) K T G P  
 31 N L H.G L F.G R K T G Q A V G D.S O.T A A N K N.K G I I W.G  
 61 D D T L M E O.L E N P K K O.I P G T K M.V F.T G L.S K K K E  
 91 R T N L.I A D.L K E K T A A •

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 GLY ASP VAL GLU LYS GLY LYS LYS ILU PHE (ILU.THR.LYS.CYS.SER.  
 GLN.CYS.HIS.THR.VAL.GLU.LYS.GLY.GLY.LYS.HIS) LYS THR GLY PRO  
 31 ASN LEU HIS.GLY LEU PHE.GLY ARG LYS THR GLY GLN ALA VAL GLY  
 TYR.SER TYR.THR ALA ALA ASN LYS ASN.LYS GLY ILU ILU TRP.GLY  
 61 ASP ASP THR LEU MET GLU TYR.LEU GLU ASN PRO LYS LYS TYR.ILU  
 PRO GLY THR LYS MET.VAL PHE.THR GLY LEU.SER LYS LYS LYS GLU  
 91 ARG THR ASN LEU.ILU ALA TYR.LEU LYS GLU LYS THR ALA ALA \*\*\*

## COMPOSITION

6 ALA A	2 GLN Q	7 LEU L	3 SER S
2 ARG R	6 GLU E	18 LYS K	10 THR T
5 ASN N	13 GLY G	2 MET M	1 TRP W
3 ASP D	3 HIS H	3 PHE F	5 TYR O
2 CYS C	6 ILU I	3 PRO P	4 VAL V

TOTAL NO. OF ACIDS = 104

\* BAHL, D. P. AND SMITH, E. L., J. BIOL. CHEM., VOL.240, NO.9,  
 PP.3585-3593, SEPT., 1965

## CYTOCHROME - C RHODOSPIRILLUM RUBRUM

HEME BONDED TO CYSTEINES AT POSITIONS 1 AND 4 OF FRAGMENT.

1 2 3 4 5 6 7 8 9 0 1 2 3  
 / C L A C H T F B Z G A N K /

1 2 3 4 5 6 7 8 9 10 11 12 13  
 /// CYS LEU ALA CYS HIS THR PHE ASX GLX GLY ALA ASN LYS ///

## COMPOSITION OF FRAGMENT

2 ALA A	0 GLN Q	1 LEU L	0 SER S
0 ARG R	0 GLU E	1 LYS K	1 THR T
1 ASN N	1 GLY G	0 MET M	0 TRP W
0 ASP D	1 HIS H	1 PHE F	0 TYR O
2 CYS C	0 ILU I	0 PRO P	0 VAL V
1 ASX B	1 GLX Z		

TOTAL NO. OF ACIDS IN FRAGMENT = 13

- TUPPY H., AND PALEUS, S., ACTA CHEM. SCAND., VOL. 13, NO.4, PP. 641-646, 1959

## CYTOCHROME C - SALMON

HEME BONDED TO CYSTEINES AT POSITIONS 4 AND 7 OF FRAGMENT.

1 2 3 4 5 6 7 8 9 0 1  
 / V Q K C A Q H C T(V,E)/

1 2 3 4 5 6 7 8 9 10 11  
 /// VAL GLN LYS CYS ALA GLN CYS HIS THR(VAL,GLU)///

## COMPOSITION OF FRAGMENT

1 ALA A	2 GLN Q	0 LEU L	0 SER S
0 ARG R	1 GLU E	1 LYS K	1 THR T
0 ASN N	0 GLY G	0 MET M	0 TRP W
0 ASP D	1 HIS H	0 PHE F	0 TYR O
2 CYS C	0 ILU I	0 PRO P	2 VAL V

TOTAL NO. OF ACIDS IN FRAGMENT = 11

\* TUPPY H., AND PALEUS, S., ACTA CHEM. SCAND., VOL. 9,  
 P.353-364, 1955



## HEMOGLOBIN ALPHA - HUMAN

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 V L S P A D K T N V K A A W G K V G A H A G E O G A E A L E  
 31 R M F L S F P T T K T O F P H F D L S H G S A Q V K G H G K  
 61 K V A D A L T N A V A H V D D M P N A L S A L S D L H A H K  
 91 L R V D P V N F K L L S H C L L V T L A A H L P A E F T P A  
 121 V H A S L D K F L A S V S T V L T S K O R \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 VAL LEU SER PRO ALA ASP LYS THR ASN VAL LYS ALA ALA TRP GLY  
 LYS VAL GLY ALA HIS ALA GLY GLU TYR GLY ALA GLU ALA LEU GLU  
 31 ARG MET PHE LEU SER PHE PRO THR THR LYS THR TYR PHE PRO HIS  
 PHE ASP LEU SER HIS GLY SER ALA GLN VAL LYS GLY HIS GLY LYS  
 61 LYS VAL ALA ASP ALA LEU THR ASN ALA VAL ALA HIS VAL ASP ASP  
 MET PRO ASN ALA LEU SER ALA LEU SER ASP LEU HIS ALA HIS LYS  
 91 LEU ARG VAL ASP PRO VAL ASN PHE LYS LEU LEU SER HIS CYS LEU  
 LEU VAL THR LEU ALA ALA HIS LEU PRO ALA GLU PHE THR PRO ALA  
 121 VAL HIS ALA SER LEU ASP LYS PHE LEU ALA SER VAL SER THR VAL  
 LEU THR SER LYS TYR ARG \*\*\*

## COMPOSITION

21 ALA	A	1 GLN	Q	18 LEU	L	11 SER	S
3 ARG	R	4 GLU	E	11 LYS	K	9 THR	T
4 ASN	N	7 GLY	G	2 MET	M	1 TRP	W
8 ASP	D	10 HIS	H	7 PHE	F	3 TYR	D
1 CYS	C	0 ILE	I	7 PRO	P	13 VAL	V

TOTAL NO. OF ACIDS = 141

\* HILL, R.J., AND KONIGSBERG, W., J. BIOL. CHEM., VOL. 237, NO.10,  
PP. 3151-3156, OCT., 1962

BRAUNITZER, G., GEHRING-MULLER, R., HILSCHMANN, N., HILSE, K.,  
HOBOM, G., RUDLOFF, V., AND WITTMANN-LIEBOLD, B.,  
Z. PHYSIOL. CHEM., VOL. BD 325, PP.283-286, 1961

THE SAME SEQUENCE, WITHOUT DISTINGUISHING AMINES, ALSO  
REPORTED IN THE ARTICLE.

SCHROEDER, W.A., J.R.SHELTON, J.B.SHELTON, AND J.CORMICK  
BIOCHEMISTRY, VOL. 2, NO.6, PP.1353-1357, NOV.-DEC., 1963

FETAL ALPHA CHAIN IS VERY PROBABLY IDENTICAL WITH ADULT ALPHA  
CHAIN. TRYPTIC AND CHYMOTRYPTIC PEPTIDES, MOST OF WHICH  
WERE COMPLETELY SEQUENCED, WERE SHOWN TO FIT EXACTLY INTO  
THE ADULT ALPHA CHAIN SEQUENCE.

## HEMOGLOBIN BETA - HUMAN

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 V H L T P E E K S A V T A L W G K V D V D E V G G E A L G R  
 31 L L V V O P W T E R F F E S F G D L S T P D A V M G D P K V  
 61 K A H G K K V L G A F S D G L A H L D D L K G T F A T L S E  
 91 L H C D K L H V D P E D F R L L G D V L V C V L A H H F G K  
 121 E F T P P V E A A D E K V V A G V A D A L A H K O H \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 VAL HIS LEU THR PRO GLU GLU LYS SER ALA VAL THR ALA LEU TRP  
 GLY LYS VAL ASP VAL ASP GLU VAL GLY GLY GLU ALA LEU GLY ARG  
 31 LEU LEU VAL VAL TYR PRO TRP THR GLU ARG PHE PHE GLU SER PHE  
 GLY ASP LEU SER THR PRO ASP ALA VAL MET GLY ASP PRO LYS VAL  
 61 LYS ALA HIS GLY LYS LYS VAL LEU GLY ALA PHE SER ASP GLY LEU  
 ALA HIS LEU ASP ASP LEU LYS GLY THR PHE ALA THR LEU SER GLU  
 91 LEU HIS CYS ASP LYS LEU HIS VAL ASP PRO GLU ASP PHE ARG LEU  
 LEU GLY ASP VAL LEU VAL CYS VAL LEU ALA HIS HIS PHE GLY LYS  
 121 GLU PHE THR PRO PRO VAL GLU ALA ALA TYR GLU LYS VAL VAL ALA  
 GLY VAL ALA ASP ALA LEU ALA HIS LYS TYR HIS \*\*\*

## COMPOSITION

15 ALA	A	0 GLN	Q	18 LEU	L	5 SER	S
3 ARG	R	11 GLU	E	11 LYS	K	7 THR	T
0 ASN	N	13 GLY	G	1 MET	M	2 TRP	W
13 ASP	D	9 HIS	H	8 PHE	F	3 TYR	O
2 CYS	C	0 ILE	I	7 PRO	P	18 VAL	V

TOTAL NO. OF ACIDS = 146

- BRAUNITZER, G., GEHRING-MULLER, R., HILSCHMANN, N., HILSE, K.,  
 HOBOM, G., RUDLOFF, V., AND WITTMANN-LIEBOLD, B.,  
 Z. PHYSIOL. CHEM., VOL. BD 325, PP.283-286, 1961

## HEMOGLOBIN GAMMA - HUMAN

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 G H F T E E D K A T I T S L W G K V N V E D A G G E T L G R  
 31 L L V V O P W T Q R F F D S F G N L S S A S A I M G N P K V  
 61 K A H G K K V L T S L G D A I K H L D D L K G T F A Q L S E  
 91 L H C D K L H V D P E N F K L L G N V L V T V L A I H F G K  
 121 E F T P E V Q A S W Q K M V T G V A S A L S S R D H \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 GLY HIS PHE THR GLU GLU ASP LYS ALA THR ILU THR SER LEU TRP  
 GLY LYS VAL ASN VAL GLU ASP ALA GLY GLY GLU THR LEU GLY ARG  
 31 LEU LEU VAL VAL TYR PRO TRP THR GLN ARG PHE PHE ASP SER PHE  
 GLY ASN LEU SER SER ALA SER ALA ILU MET GLY ASN PRO LYS VAL  
 61 LYS ALA HIS GLY LYS LYS VAL LEU THR SER LEU GLY ASP ALA ILU  
 LYS HIS LEU ASP ASP LEU LYS GLY THR PHE ALA GLN LEU SER GLU  
 91 LEU HIS CYS ASP LYS LEU HIS VAL ASP PRO GLU ASN PHE LYS LEU  
 LEU GLY ASN VAL LEU VAL THR VAL LEU ALA ILU HIS PHE GLY LYS  
 121 GLU PHE THR PRO GLU VAL GLN ALA SER TRP GLN LYS MET VAL THR  
 GLY VAL ALA SER ALA LEU SER SER ARG TYR HIS \*\*\*

## COMPOSITION

11 ALA	A	4 GLN	Q	17 LEU	L	11 SER	S
3 ARG	R	8 GLU	E	12 LYS	K	10 THR	T
5 ASN	N	13 GLY	G	2 MET	M	3 TRP	W
8 ASP	D	7 HIS	H	8 PHE	F	2 TYR	O
1 CYS	C	4 ILU	I	4 PRO	P	13 VAL	V

TOTAL NO. OF ACIDS = 146

\* SCHROEDER, W.A., SHELTON, J.R., SHELTON, J.B., CORMICK, J., AND JONES, R.T., BIOCHEMISTRY, VOL. 2, NO. 5, PP. 992-1008, SEPT.-OCT., 1963

## HEMOGLOBIN BETA - GORILLA

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 V H L T P E E K S A V T A L W G K V D V D E V G G E A L G R  
 31 L L V V O P W T E R F F E S F G D L S T P D A V M G D P K V  
 61 K A H G K K V L G A F S D G L A H L D D L K G T F A T L S E  
 91 L H C D K L H V D P E D F L L L G D V L V C V L A H H F G K  
 121 E F T P P V E A A O E K V V A G V A D A L A H K O H \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 VAL HIS LEU THR PRO GLU GLU LYS SER ALA VAL THR ALA LEU TRP  
 GLY LYS VAL ASP VAL ASP GLU VAL GLY GLY GLU ALA LEU GLY ARG  
 31 LEU LEU VAL VAL TYR PRO TRP THR GLU ARG PHE PHE GLU SER PHE  
 GLY ASP LEU SER THR PRO ASP ALA VAL MET GLY ASP PRO LYS VAL  
 61 LYS ALA HIS GLY LYS LYS VAL LEU GLY ALA PHE SER ASP GLY LEU  
 ALA HIS LEU ASP ASP LEU LYS GLY THR PHE ALA THR LEU SER GLU  
 91 LEU HIS CYS ASP LYS LEU HIS VAL ASP PRO GLU ASP PHE LEU LEU  
 LEU GLY ASP VAL LEU VAL CYS VAL LEU ALA HIS HIS PHE GLY LYS  
 121 GLU PHE THR PRO PRO VAL GLU ALA ALA TYR GLU LYS VAL VAL ALA  
 GLY VAL ALA ASP ALA LEU ALA HIS LYS TYR HIS \*\*\*

## COMPOSITION

15	ALA	A	0	GLN	Q	19	LEU	L	5	SER	S
2	ARG	R	11	GLU	E	11	LYS	K	7	THR	T
0	ASN	N	13	GLY	G	1	MET	M	2	TRP	W
13	ASP	D	9	HIS	H	8	PHE	F	3	TYR	O
2	CYS	C	0	ILU	I	7	PRO	P	18	VAL	V

TOTAL NO. OF ACIDS = 146

= ZUCKERKANDL, E., SCIENTIFIC AMERICAN, VOL. 212, NO. 5,  
 PP. 110-118, MAY 1965

## HEMOGLOBIN BETA - HORSE

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 V E L S G E E K A A L(V,A,L,W,D)K V D E E E V G(G.E.A)L G R  
 31 L L V V O P W T E R F(F.E.S.F.G.D.L.S.G.P.D.A.V)M(G.D.P)K V  
 61 K A H G K K V L H S F G E G V H H(L.D.D.L)K G T F A(A.L.S.E.  
 91 L.H.C.D.K.L.H.V.D.P.E.D.F)R L L G D V L A L V V A R H F G K  
 121 D F T P E L E A S D E K V V A G V A D A L A H K D H \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 VAL GLU LEU SER GLY GLU GLU LYS ALA ALA LEU(VAL,ALA,LEU,TRP,  
 ASP)LYS VAL ASP GLU GLU GLU VAL GLY(GLY.GLU.ALA)LEU GLY ARG  
 31 LEU LEU VAL VAL TYR PRO TRP THR GLU ARG PHE(PHE.GLU.SER.PHE.  
 GLY.ASP.LEU.SER.GLY.PRO.ASP.ALA.VAL)MET(GLY.ASP.PRO)LYS VAL  
 61 LYS ALA HIS GLY LYS LYS VAL LEU HIS SER PHE GLY GLU GLY VAL  
 HIS HIS(LEU.ASP.ASP.LEU)LYS GLY THR PHE ALA(ALA.LEU.SER.GLU.  
 91 LEU.HIS.CYS.ASP.LYS.LEU.HIS.VAL.ASP.PRO.GLU.ASP.PHE)ARG LEU  
 LEU GLY ASP VAL LEU ALA LEU VAL VAL ALA ARG HIS PHE GLY LYS  
 121 ASP PHE THR PRO GLU LEU GLU ALA SER TYR GLU LYS VAL VAL ALA  
 GLY VAL ALA ASP ALA LEU ALA HIS LYS TYR HIS \*\*\*

## COMPOSITION

15	ALA	A	0	GLN	Q	19	LEU	L	6	SER	S
4	ARG	R	15	GLU	E	11	LYS	K	3	THR	T
0	ASN	N	14	GLY	G	1	MET	M	2	TRP	W
13	ASP	D	9	HIS	H	8	PHE	F	3	TYR	O
1	CYS	C	0	ILU	I	5	PRO	P	17	VAL	V

TOTAL NO. OF ACIDS = 146

\* SMITH, D. B., CAN. J. BIOCHEM., VOL.42, NO.5, PP.755-762, 1964

## HEMOGLOBIN ALPHA - HORSE

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 V L S A A D K T N V K A A W S K V G G H A G E O G A E A L E  
 31 R M F L G F P T T K T D F P H F D L S H G S A Q V K A H G K  
 61 K V A D G L T L A V G H L D D L P G A L S N L S D L H A H K  
 91 L R V D P V N F K L L S H C L L S T L A V H L P N D F T P A  
 121 V H A S L D K F L S S V S T V L T S K O R \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 VAL LEU SER ALA ALA ASP LYS THR ASN VAL LYS ALA ALA TRP SER  
 LYS VAL GLY GLY HIS ALA GLY GLU TYR GLY ALA GLU ALA LEU GLU  
 31 ARG MET PHE LEU GLY PHE PRO THR THR LYS THR TYR PHE PRO HIS  
 PHE ASP LEU SER HIS GLY SER ALA GLN VAL LYS ALA HIS GLY LYS  
 61 LYS VAL ALA ASP GLY LEU THR LEU ALA VAL GLY HIS LEU ASP ASP  
 LEU PRO GLY ALA LEU SER ASN LEU SER ASP LEU HIS ALA HIS LYS  
 91 LEU ARG VAL ASP PRO VAL ASN PHE LYS LEU LEU SER HIS CYS LEU  
 LEU SER THR LEU ALA VAL HIS LEU PRO ASN ASP PHE THR PRO ALA  
 121 VAL HIS ALA SER LEU ASP LYS PHE LEU SER SER VAL SER THR VAL  
 LEU THR SER LYS TYR ARG \*\*\*

## COMPOSITION

16 ALA	A	1 GLN	Q	21 LEU	L	13 SER	S
3 ARG	R	3 GLU	E	11 LYS	K	9 THR	T
4 ASN	N	10 GLY	G	1 MET	M	1 TRP	W
9 ASP	D	10 HIS	H	7 PHE	F	3 TYR	O
1 CYS	C	0 ILE	I	6 PRO	P	12 VAL	V

TOTAL NO. OF ACIDS = 141

- BRAUNITZER, G. AND MATSUDA, G., J. BIOCHEM. (TOKYO), VOL.53, NO.3,  
 PP.262-263, 1963  
 THIS SEQUENCE WAS DETERMINED PARTIALLY BY HOMOLOGY WITH HUMAN ALPHA.

## HEMOGLOBIN BETA - LEMUR FULVUS

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 T L L S A E E D A H V T S L W G K V N V E K V G G E A L G R  
 31 L L V V(O,P,W,T,E,R,F,F,E,S,F,G,D=L,S,S,P,S,A,V,H,G,D,P,K,V,  
 61 K,A,H,G,K,K,V,L,S,A,F,S,E,G=L,H,H,L,D,D,L,K,G,T,F,A,A,L,S,E,  
 91 L,H,C,V,A,L,H,V,D,P,E,D,F,K,L,L,G,D,S,L,S,D,V,L,A,D,H,F,G,K)  
 121 X X X X X X X X X X X X V V A G V(A,D,A,L,A,H,K,D,H)\*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 THR LEU LEU SER ALA GLU GLU ASP ALA HIS VAL THR SER LEU TRP  
 GLY LYS VAL ASN VAL GLU LYS VAL GLY GLY GLU ALA LEU GLY ARG  
 31 LEU LEU VAL VAL(TYR,PRO,TRP,THR,GLU,ARG,PHE,PHE,GLU,SER,PHE,  
 GLY,ASP=LEU,SER,SER,PRO,SER,ALA,VAL,MET,GLY,ASP,PRO,LYS,VAL,  
 61 LYS,ALA,HIS,GLY,LYS,LYS,VAL,LEU,SER,ALA,PHE,SER,GLU,GLY=LEU,  
 HIS,HIS,LEU,ASP,ASP,LEU,LYS,GLY,THR,PHE,ALA,ALA,LEU,SER,GLU,  
 91 LEU,HIS,CYS,VAL,ALA,LEU,HIS,VAL,ASP,PRO,GLU,ASP,PHE,LYS,LEU,  
 LEU,GLY,ASP,SER,LEU,SER,ASP,VAL,LEU,ALA,ASP,HIS,PHE,GLY,LYS)  
 121 XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX VAL VAL ALA  
 GLY VAL(ALA,ASP,ALA,LEU,ALA,HIS,LYS,TYR,HIS)\*\*\*

## COMPOSITION

14 ALA	A	0 GLN	Q	19 LEU	L	11 SER	S
2 ARG	R	9 GLU	E	10 LYS	K	4 THR	T
1 ASN	N	12 GLY	G	1 MET	M	2 TRP	W
11 ASP	D	9 HIS	H	7 PHE	F	2 TYR	O
1 CYS	C	0 ILU	I	4 PRO	P	15 VAL	V
12 XXX	X						

TOTAL NO. OF ACIDS = 146

• BUETTNER-JANUSCH, J. AND HILL, R. L., SCIENCE, VOL. 147,  
 PP. 836-842, FEB. 19, 1965



## ABNORMAL HUMAN HEMOGLOBIN

Normal adult human hemoglobin (hemoglobin A) contains two pairs of polypeptide chains, termed alpha and beta. Each pair is identical. Some modified beta chains have been given other Greek letters, for example, normal fetal hemoglobin is composed of two alpha chains and two "gamma" chains. Usually, however, altered hemoglobins are different in only a single amino acid. A number of hemoglobins bearing these altered amino acid sequences in their polypeptide chains have been described. For example, one of the early reports by Ingram (1957) shows the chemical difference between normal human hemoglobin and sickle cell hemoglobin. By comparison of amino acid sequences of tryptic peptide digests of the two hemoglobins, it was established that hemoglobin A (normal) contains a GLU residue in the locus where hemoglobin S (sickle cell) contains VAL. This replacement of two charged GLU residues for two uncharged VAL residues in the hemoglobin tetramer is sufficient to account for the "sickling" phenomenon in the abnormal hemoglobin. Listed below are a number of known amino acid replacements in abnormal human hemoglobins.

HEMOGLOBIN NAME	CHANGES CHAIN POS. FROM TO	REFERENCE
A NORMAL		
F NORMAL FETAL	BETA (CALLED GAMMA)	1
I	ALPHA 16 LYS-ASP	2
NORFOLK	ALPHA 57 GLY-ASP	3
M BOSTON	ALPHA 58 HIS-TYR	4
M SASKATOON	BETA 63 HIS-TYR	4
M MILWAUKEE	BETA 67 VAL-GLU	4
D PUNJAB	BETA 121 GLU-GLN	5
G SAN JOSE	BETA 7 GLU-GLY	6
ZURICH	BETA 63 HIS-ARG	7
C	BETA 6 GLU-LYS	8
D ARABIA	BETA 121 GLU-LYS	9
O INDONESIA	ALPHA 116 GLU-LYS	9
X	ALPHA 68 ASN-LYS	
	and BETA 6 GLU-LYS	10
S	BETA 6 GLU-VAL	11
D IBADAN	BETA 87 THR-LYS	12
F TEXAS	GAMMA 5 or 6 GLU-LYS	13
KENWOOD	BETA 143 HIS-ASP	14
G	BETA 7 GLU-GLY	15

1. Rhinesmith, H. W., Schroeder, W. A., and Pauling, L., J. Am. Chem. Soc., Vol. 79, p. 4682, 1957  
Rhinesmith, H. W., Schroeder, W. A., and Martin, N., J. Am. Chem. Soc., Vol. 80, p. 3358, 1958
2. Murayama, M., Fed. Proc., Vol. 19, p. 78, 1960
3. Baglioni, C., J. Biol. Chem., Vol. 237, pp. 69-74, 1962
4. Gerald, P. S. and Efron, M. L., Proc. Natl. Acad. Sci. U.S., Vol. 47, pp. 1758-1767, 1958
5. Baglioni, C., Biochim. Biophys. Acta, Vol. 59, pp. 437-440, 1962
6. Hill, R. L. and Schwartz, H. C., Nature, Vol. 184, pp. 641-642, 1959
7. Muller, C. J. and Kingma, S., Biochim. Biophys. Acta, Vol. 50, p. 595, 1961
8. Hunt, J. A. and Ingram, V. M., Nature, Vol. 184, p. 640, 1959  
Ingram, V. M., Nature, Vol. 180, pp. 326-328, 1957
9. Baglioni, C. and Lehmann, H., Nature, Vol. 196, pp. 229-232, 1962
10. Baglioni, C. and Ingram, V. M., Nature, Vol. 189, pp. 465-467, 1961
11. Ingram, V. M., Nature, Vol. 180, pp. 326-328, 1957
12. Watson-Williams, E. J., Nature, Vol. 205, pp. 1273-1276, 1965
13. Schneider, R. G., Science, Vol. 148, pp. 240-242, 1965
14. Beale, D. and Lehmann, H., Nature, Vol. 207, pp. 249-261, 1965
15. Hill, R. L., Swenson, R. T., and Schwartz, H. C., J. Biol. Chem., Vol. 235, pp. 3182-3187, 1960

## MYOGLOBIN - WHALE

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 V L S E G E W Q L V L H V W A K V E A D V A G H G Q D I L I  
 31 R L F K S H P E T L E K F D R F K H L K T E A E M K A S E D  
 61 L K K H G V T V L T A L G A I L K K K G H H E A E L K P L A  
 91 Q S H A T K H K I P I K O L E F I S E A I I H V L H S R H P  
 121 G N F G A D A Q G A M N K A L E L F R K D I A A K O K E L G  
 151 O Q G \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 VAL LEU SER GLU GLY GLU TRP GLN LEU VAL LEU HIS VAL TRP ALA  
 LYS VAL GLU ALA ASP VAL ALA GLY HIS GLY GLN ASP ILU LEU ILU  
 31 ARG LEU PHE LYS SER HIS PRO GLU THR LEU GLU LYS PHE ASP ARG  
 PHE LYS HIS LEU LYS THR GLU ALA GLU MET LYS ALA SER GLU ASP  
 61 LEU LYS LYS HIS GLY VAL THR VAL LEU THR ALA LEU GLY ALA ILU  
 LEU LYS LYS LYS GLY HIS HIS GLU ALA GLU LEU LYS PRO LEU ALA  
 91 GLN SER HIS ALA THR LYS HIS LYS ILU PRO ILU LYS TYR LEU GLU  
 PHE ILU SER GLU ALA ILU ILU HIS VAL LEU HIS SER ARG HIS PRO  
 121 GLY ASN PHE GLY ALA ASP ALA GLN GLY ALA MET ASN LYS ALA LEU  
 GLU LEU PHE ARG LYS ASP ILU ALA ALA LYS TYR LYS GLU LEU GLY  
 151 TYR GLN GLY \*\*\*

## COMPOSITION

17 ALA	A	5 GLN	Q	18 LEU	L	6 SER	S
4 ARG	R	14 GLU	E	19 LYS	K	5 THR	T
2 ASN	N	11 GLY	G	2 MET	M	2 TRP	W
6 ASP	D	12 HIS	H	6 PHE	F	3 TYR	O
0 CYS	C	9 ILU	I	4 PRO	P	8 VAL	V

TOTAL NO. OF ACIDS = 153

## DIHEME PEPTIDE - CHROMATIUM

THE PEPTIDE CONTAINS TWO HEME GROUPS. THE FIRST IS COVALENTLY BONDED TO CYSTEINES 5 AND 8. THERE IS ONLY ONE OTHER CYSTEINE AVAILABLE FOR THE OBSERVED COVALENT BONDING OF THE SECOND HEME.

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7  
/ F A G K C S Q C H T L V A D E G S A K C H T F D E G S /

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
/// PHE ALA GLY LYS CYS SER GLN CYS HIS THR LEU VAL ALA ASP GLU  
GLY SER ALA LYS CYS HIS THR PHE ASP GLU GLY SER ///

## COMPOSITION

3 ALA A	1 GLN Q	1 LEU L	3 SER S
0 ARG R	2 GLU E	2 LYS K	2 THR T
0 ASN N	3 GLY G	0 MET M	0 TRP W
2 ASP D	2 HIS H	2 PHE F	0 TYR O
3 CYS C	0 ILU I	0 PRO P	1 VAL V

TOTAL NO. OF ACIDS IN FRAGMENT = 27

\* DUS, K., BARTSCH, R.G., AND KAMEN, M.D., J. BIOL. CHEM., VOL. 237, NO. 10, PP. 3083-3093, OCT., 1962

## FERREDOXIN - CLOSTRIDIUM PASTEURIANUM

THE PROTEIN CONTAINS 7 SULPHIDE AND 7 IRON ATOMS PER MOLECULE.  
IT DOES NOT CONTAIN HEME.

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
1 A O K I A D S C V S C/G A C/A S E C P V N A I S Q G D S I F/  
31 V I D A D T C I D C G N C A N V C P V G A P V Q E \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
1 ALA TYR LYS ILU ALA ASP SER CYS VAL SER CYS/GLY ALA CYS/ALA  
SER GLU CYS PRO VAL ASN ALA ILU SER GLN GLY ASP SER ILU PHE/  
31 VAL ILU ASP ALA ASP THR CYS ILU ASP CYS GLY ASN CYS ALA ASN  
VAL CYS PRO VAL GLY ALA PRO VAL GLN GLU \*\*\*

## COMPOSITION

8 ALA	A	2 GLN	Q	0 LEU	L	5 SER	S
0 ARG	R	2 GLU	E	1 LYS	K	1 THR	T
3 ASN	N	4 GLY	G	0 MET	M	0 TRP	W
5 ASP	D	0 HIS	H	1 PHE	F	1 TYR	O
8 CYS	C	5 ILU	I	3 PRO	P	6 VAL	V

TOTAL NO. OF ACIDS = 55

- TANAKA, M., NAKASHIMA, T., BENSON, A., MOWER, H.F., AND YASUNOBU, K.T.,  
BIOCHEM. BIOPHYS. RES. COMMUN., VOL. 16, NO. 5,  
PP. 422-427, 1964

## AZURIN - PSEUDOMONAS FLUORESCENS

THE BLUE PROTEIN CONTAINS ONE COPPER ATOM PER MOLECULE.

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 A E C S V D I Q G N D Q M Q F N T N A I T V D K S C K Q F T  
 31 V N L S H P G N L P K N V M G H N W V L S T A A D M Q G V V  
 61 T D G M A S G L D K D D L K P D D S R V I A H T K L I G S G  
 91 E K D S V T F D V S K L K E G E Q O M F F C T F P G H S A L  
 121 M K G T L T L K \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 ALA GLU CYS SER VAL ASP ILU GLN GLY ASN ASP GLN MET GLN PHE  
 ASN THR ASN ALA ILU THR VAL ASP LYS SER CYS LYS GLN PHE THR  
 31 VAL ASN LEU SER HIS PRO GLY ASN LEU PRO LYS ASN VAL MET GLY  
 HIS ASN TRP VAL LEU SER THR ALA ALA ASP MET GLN GLY VAL VAL  
 61 THR ASP GLY MET ALA SER GLY LEU ASP LYS ASP TYR LEU LYS PRO  
 ASP ASP SER ARG VAL ILU ALA HIS THR LYS LEU ILU GLY SER GLY  
 91 GLU LYS ASP SER VAL THR PHE ASP VAL SER LYS LEU LYS GLU GLY  
 GLU GLN TYR MET PHE PHE CYS THR PHE PRO GLY HIS SER ALA LEU  
 121 MET LYS GLY THR LEU THR LEU LYS \*\*\*

## COMPOSITION

7 ALA	A	6 GLN	Q	10 LEU	L	10 SER	S
1 ARG	R	4 GLU	E	11 LYS	K	10 THR	T
7 ASN	N	11 GLY	G	6 MET	M	1 TRP	W
11 ASP	D	4 HIS	H	6 PHE	F	2 TYR	O
3 CYS	C	4 ILU	I	4 PRO	P	10 VAL	V

TOTAL NO. OF ACIDS = 128

## RIBONUCLEASE - BOVINE

DISULPHIDE BONDS ARE FORMED BETWEEN CYSTEINES AT POSITIONS  
26 AND 84, 40 AND 95, 58 AND 110, AND 65 AND 72.

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
1 K E T A A A K F E R Q H M D S S T S A A S S S N D C N Q M M  
31 K S R N L T K D R C K P V N T F V H E S L A D V Q A V C S Q  
61 K N V A C K N G Q T N C D Q S D S T M S I T D C R E T G S S  
91 K O P N C A O K T T Q A N K H I I V A C E G N P O V P V H F  
121 D A S V \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
1 LYS GLU THR ALA ALA ALA LYS PHE GLU ARG GLN HIS MET ASP SER  
SER THR SER ALA ALA SER SER SER ASN TYR CYS ASN GLN MET MET  
31 LYS SER ARG ASN LEU THR LYS ASP ARG CYS LYS PRO VAL ASN THR  
PHE VAL HIS GLU SER LEU ALA ASP VAL GLN ALA VAL CYS SER GLN  
61 LYS ASN VAL ALA CYS LYS ASN GLY GLN THR ASN CYS TYR GLN SER  
TYR SER THR MET SER ILU THR ASP CYS ARG GLU THR GLY SER SER  
91 LYS TYR PRO ASN CYS ALA TYR LYS THR THR GLN ALA ASN LYS HIS  
ILU ILU VAL ALA CYS GLU GLY ASN PRO TYR VAL PRO VAL HIS PHE  
121 ASP ALA SER VAL \*\*\*

## COMPOSITION

12 ALA	A	7 GLN	Q	2 LEU	L	15 SER	S
4 ARG	R	5 GLU	E	10 LYS	K	10 THR	T
10 ASN	N	3 GLY	G	4 MET	M	0 TRP	W
5 ASP	D	4 HIS	H	3 PHE	F	6 TYR	O
8 CYS	C	3 ILU	I	4 PRO	P	9 VAL	V

TOTAL NO. OF ACIDS = 124

\* SMYTH, D.G., STEIN, W.H. AND MOORE, S., J. BIOL. CHEM.,  
VOL. 238, NO. 1, PP. 227-234, JAN., 1963

## TRYPSIN INHIBITOR - BOVINE

DISULPHIDE BONDS ARE FORMED BETWEEN CYSTEINES AT POSITIONS 5-55,  
14-38, AND 30-51.

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 R P D F C L E P P O T G P C K A R I I R D F O N A K A G L C  
 31 Q T F V O G G C R A K R N N F K S A E D C M R T C G G A \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 ARG PRO ASP PHE CYS LEU GLU PRO PRO TYR THR GLY PRO CYS LYS  
 ALA ARG ILU ILU ARG TYR PHE TYR ASN ALA LYS ALA GLY LEU CYS  
 31 GLN THR PHE VAL TYR GLY GLY CYS ARG ALA LYS ARG ASN ASN PHE  
 LYS SER ALA GLU ASP CYS MET ARG THR CYS GLY GLY ALA \*\*\*

## COMPOSITION

6 ALA	A	1 GLN	Q	2 LEU	L	1 SER	S
6 ARG	R	2 GLU	E	4 LYS	K	3 THR	T
3 ASN	N	6 GLY	G	1 MET	M	0 TRP	W
2 ASP	D	0 HIS	H	4 PHE	F	4 TYR	O
6 CYS	C	2 ILU	I	4 PRO	P	1 VAL	V

TOTAL NO. OF ACIDS = 58

- KASSELL, B., RADICEVIC, M., ANSFIELD, M. J., AND LASKOWSKI, M.,  
BIOCHEM. BIOPHYS. RES. COMMUN., VOL. 18, NO. 2, PP. 255-258,  
1965

DLOUHA, V., POSPISILOVA, D., MELOUN, B. AND SORM, F., COLLECTION  
CZECH. CHEM. COMMUN., VOL. 30, PP. 1311-1325, 1965

THE SEQUENCE REPORTED HERE DIFFERS FROM THE ABOVE IN HAVING  
THE ILU (I) DELETED AT POSITION 19.



CHAUVET, J., NOUVEL, G., AND ACHER, R., BIOCHIM. BIOPHYS.  
ACTA , VOL. 92, PP. 200-201, 1964

THE SEQUENCE REPORTED HERE DIFFERS FROM THE ABOVE IN THE  
FOLLOWING RESPECTS.

THE ARG (R) FROM POSITION 42 HAS BEEN REMOVED AND INSERTED  
BETWEEN POSITIONS 20 AND 21. THE GLN (Q) AT POSITION 31 HAS  
BEEN DELETED AND A GLU (E) ADDED BETWEEN POSITIONS 32 AND 33.

KASSELL, B., AND LASKOWSKI, M., BIOCHEM. BIOPHYS. RES.  
COMMUN. VOL 20, NO.4, PP.463-468, 1965

TOBACCO MOSAIC VIRUS

ACETYL - AT AMINO END

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 S O S I T T P S Q F V F L S S A W A D P I E L I N L C T N A  
 31 L G N Q F Q T Q Q A R T V Q V R Q F S Q V W K P S P Q V T V  
 61 R F P D S D F K V O R O N A V L D P L V T A L L G A F D T R  
 91 N R I I Q V Q D Q A N P T T A Q T L D A T R R V D D A T V A  
 121 I R S A D I N L I V E L I R G T G S Q N R S S F E S S S G L  
 151 V W T S G P A T •

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 SER TYR SER ILU THR THR PRO SER GLN PHE VAL PHE LEU SER SER  
 ALA TRP ALA ASP PRO ILU GLU LEU ILU ASN LEU CYS THR ASN ALA  
 31 LEU GLY ASN GLN PHE GLN THR GLN GLN ALA ARG THR VAL GLN VAL  
 ARG GLN PHE SER GLN VAL TRP LYS PRO SER PRO GLN VAL THR VAL  
 61 ARG PHE PRO ASP SER ASP PHE LYS VAL TYR ARG TYR ASN ALA VAL  
 LEU ASP PRO LEU VAL THR ALA LEU LEU GLY ALA PHE ASP THR ARG  
 91 ASN ARG ILU ILU GLN VAL GLN ASP GLN ALA ASN PRO THR THR ALA  
 GLN THR LEU ASP ALA THR ARG ARG VAL ASP ASP ALA THR VAL ALA  
 121 ILU ARG SER ALA ASP ILU ASN LEU ILU VAL GLU LEU ILU ARG GLY  
 THR GLY SER TYR ASN ARG SER SER PHE GLU SER SER SER GLY LEU  
 151 VAL TRP THR SER GLY PRO ALA THR \*\*\*

COMPOSITION

14 ALA A	13 GLN Q	12 LEU L	16 SER S
11 ARG R	3 GLU E	2 LYS K	16 THR T
8 ASN N	6 GLY G	0 MET M	3 TRP W
10 ASP D	0 HIS H	8 PHE F	4 TYR O
1 CYS C	9 ILU I	8 PRO P	14 VAL V

TOTAL NO. OF ACIDS = 158

- ANDERER, F.A., Z. NATURFORSCH., VOL. 17, PP.526-543, 1962

STRUCTURE REVISIONS AND CONFIRMATIONS.

ANDERER, F.A., UHLIG, H., WEBER, E., AND SCHRAMM, G., NATURE,  
VOL. 186, NO.4729, PP.922-925, JUNE 18, 1960

FUNATSU, G., TSUGITA, A., AND FRAENKEL-CONRAT, H., ARCH.  
BIOCHEM. BIOPHYS., VOL. 105, NO.1, PP.25-41, APR. 1964

TOBACCO MOSAIC VIRUS STRAIN DAHLMENSE

ACETYL- AT AMINO END

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 S O S I T S P S Q F V F L S S V W A D P I E L L N V C T S S  
 31 L G N Q F Q T Q Q A R T T Q V Q Q F S E V W K P F P Q S T V  
 61 R F P G D V O K V O R O N A V L D P L I T A L L G T F D T R  
 91 N R I I E V E N Q Q S P T T A E T L D A T R R V D D A T V A  
 121 I R S A N I N L V N E L V R G T G L O N Q N T F E S M S G L  
 151 V W T S A P A S •

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 SER TYR SER ILU THR SER PRO SER GLN PHE VAL PHE LEU SER SER  
 VAL TRP ALA ASP PRO ILU GLU LEU LEU ASN VAL CYS THR SER SER  
 31 LEU GLY ASN GLN PHE GLN THR GLN GLN ALA ARG THR THR GLN VAL  
 GLN GLN PHE SER GLU VAL TRP LYS PRO PHE PRO GLN SER THR VAL  
 61 ARG PHE PRO GLY ASP VAL TYR LYS VAL TYR ARG TYR ASN ALA VAL  
 LEU ASP PRO LEU ILU THR ALA LEU LEU GLY THR PHE ASP THR ARG  
 91 ASN ARG ILU ILU GLU VAL GLU ASN GLN GLN SER PRO THR THR ALA  
 GLU THR LEU ASP ALA THR ARG ARG VAL ASP ASP ALA THR VAL ALA  
 121 ILU ARG SER ALA ASN ILU ASN LEU VAL ASN GLU LEU VAL ARG GLY  
 THR GLY LEU TYR ASN GLN ASN THR PHE GLU SER MET SER GLY LEU  
 151 VAL TRP THR SER ALA PRO ALA SER \*\*\*

COMPOSITION

11 ALA A	12 GLN Q	13 LEU L	16 SER S
9 ARG R	7 GLU E	2 LYS K	17 THR T
10 ASN N	6 GLY G	1 MET M	3 TRP W
7 ASP D	0 HIS H	8 PHE F	5 TYR O
1 CYS C	7 ILU I	8 PRO P	15 VAL V

TOTAL NO. OF ACIDS = 158

- WITTMANN-LIEBOLD, B. AND WITTMANN, H. G., Z. VERERBUNGS.,  
VOL. 94, PP. 427-435, 1963

## CHYMOTRYPSINOGEN-A - BOVINE

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 C G V P A I Q P V L S G L S R I V G D E E A V P G S W P W Q  
 31 V S L Q D K T G F H F C G G S L I N E N W V V T A A H C G V  
 61 T T S D V V V A G E F D Q G S S S E K I Q K L K I A K V F K  
 91 N S K O N S L T I N N N I T L L K L S T A A S F S Q T V S A  
 121 V C L P S A S D D F A A G T T C V T T G W G L T R O T N A N  
 151 T P D R L Q Q A S L P L L S N T N C K K O W G T K I K D A M  
 181 I C A G A S G V S S C M G D S G G P L V C K K N G A W T L V  
 211 G I V S W G S S T C S T S T P G V O A R V T A L V N W V Q Q  
 241 T L A A N •

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 CYS GLY VAL PRO ALA ILU GLN PRO VAL LEU SER GLY LEU SER ARG  
 ILU VAL GLY ASP GLU GLU ALA VAL PRO GLY SER TRP PRO TRP GLN  
 31 VAL SER LEU GLN ASP LYS THR GLY PHE HIS PHE CYS GLY GLY SER  
 LEU ILU ASN GLU ASN TRP VAL VAL THR ALA ALA HIS CYS GLY VAL  
 61 THR THR SER ASP VAL VAL VAL ALA GLY GLU PHE ASP GLN GLY SER  
 SER SER GLU LYS ILU GLN LYS LEU LYS ILU ALA LYS VAL PHE LYS  
 91 ASN SER LYS TYR ASN SER LEU THR ILU ASN ASN ASN ILU THR LEU  
 LEU LYS LEU SER THR ALA ALA SER PHE SER GLN THR VAL SER ALA  
 121 VAL CYS LEU PRO SER ALA SER ASP ASP PHE ALA ALA GLY THR THR  
 CYS VAL THR THR GLY TRP GLY LEU THR ARG TYR THR ASN ALA ASN  
 151 THR PRO ASP ARG LEU GLN GLN ALA SER LEU PRO LEU LEU SER ASN  
 THR ASN CYS LYS LYS TYR TRP GLY THR LYS ILU LYS ASP ALA MET  
 181 ILU CYS ALA GLY ALA SER GLY VAL SER SER CYS MET GLY ASP SER  
 GLY GLY PRO LEU VAL CYS LYS LYS ASN GLY ALA TRP THR LEU VAL  
 211 GLY ILU VAL SER TRP GLY SER SER THR CYS SER THR SER THR PRO  
 GLY VAL TYR ALA ARG VAL THR ALA LEU VAL ASN TRP VAL GLN GLN  
 241 THR LEU ALA ALA ASN \*\*\*

## COMPOSITION

22 ALA A	10 GLN Q	19 LEU L	28 SER S
4 ARG R	5 GLU E	14 LYS K	23 THR T
14 ASN N	23 GLY G	2 MET M	8 TRP W
9 ASP D	2 HIS H	6 PHE F	4 TYR O
10 CYS C	10 ILU I	9 PRO P	23 VAL V

TOTAL NO. OF ACIDS = 245

- HARTLEY, B.S., BROWN, J.R., KAUFFMAN, D.L., AND SMILLIE, L.R.,  
NATURE, VOL.207, NO.5002, PP.1157-1159, SEPT.11, 1965

THIS SEQUENCE HAS BEEN CORRECTED BY DELETING SER (S)  
WHICH WAS AT POSITION 215.

BROWN, J.R., AND HARTLEY, B. S., BIOCHEM J., VOL. 89, 59P, 1963

THE ACTIVE SITE SERINE IS AT POSITION 195

KEIL, B., PRUSIK, Z., AND SORM, F., BIOCHIM. BIOPHYS. ACTA,  
VOL. 78, P. 559-561, 1963

DISULPHIDE BRIDGES LINK POSITIONS 1-122, 42-58, 136-201,  
168-182 AND 191-220.

KOSTKA, V., MELOUN, B., AND SORM, F., COLLECTION CZECH.  
CHEM. COMMUN., VOL. 28, PP.2779-2805, 1963 .

HARTLEY, B.S., NATURE, VOL. 201, NO. 4962, PP.1284-1287, MARCH 28, 1964

## TRYPSINOGEN

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 V D D D D K I V G G O T C G A N T V P D Q V S L N S G O H F  
 31 C G G S L I N S Q W V V S A A H C O K S G I Q V R L G E D N  
 61 I N V V E G D E Q F I S A S K S I V H P S O N(P,L,T,N)N N D  
 91 I M L I K L K S A A S L N S R V A S I S L P T S C A S A G T  
 121 Q C L I S G W G N T K S S G T S O P D V L K C L K A P I L S  
 151 D S S C K S A O P G Q I T S N M F C A G O L E G G K N S C Q  
 181 G D S G G P V V C S G K L Q G I V S W G S G C A Q K N K P G  
 211 V O T K V C N O V S W I K Q T I A S N \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 VAL ASP ASP ASP ASP LYS ILU VAL GLY GLY TYR THR CYS GLY ALA  
 ASN THR VAL PRO TYR GLN VAL SER LEU ASN SER GLY TYR HIS PHE  
 31 CYS GLY GLY SER LEU ILU ASN SER GLN TRP VAL VAL SER ALA ALA  
 HIS CYS TYR LYS SER GLY ILU GLN VAL ARG LEU GLY GLU ASP ASN  
 61 ILU ASN VAL VAL GLU GLY ASP GLU GLN PHE ILU SER ALA SER LYS  
 SER ILU VAL HIS PRO SER TYR ASN(PRO,LEU,THR,ASN)ASN ASN ASP  
 91 ILU MET LEU ILU LYS LEU LYS SER ALA ALA SER LEU ASN SER ARG  
 VAL ALA SER ILU SER LEU PRO THR SER CYS ALA SER ALA GLY THR  
 121 GLN CYS LEU ILU SER GLY TRP GLY ASN THR LYS SER SER GLY THR  
 SER TYR PRO ASP VAL LEU LYS CYS LEU LYS ALA PRO ILU LEU SER  
 151 ASP SER SER CYS LYS SER ALA TYR PRO GLY GLN ILU THR SER ASN  
 MET PHE CYS ALA GLY TYR LEU GLU GLY GLY LYS ASN SER CYS GLN  
 181 GLY ASP SER GLY GLY PRO VAL VAL CYS SER GLY LYS LEU GLN GLY  
 ILU VAL SER TRP GLY SER GLY CYS ALA GLN LYS ASN LYS PRO GLY  
 211 VAL TYR THR LYS VAL CYS ASN TYR VAL SER TRP ILU LYS GLN THR  
 ILU ALA SER ASN \*\*\*



## COMPOSITION

14 ALA A	10 GLN Q	14 LEU L	33 SER S
2 ARG R	4 GLU E	15 LYS K	10 THR T
16 ASN N	25 GLY G	2 MET M	4 TRP W
10 ASP D	3 HIS H	3 PHE F	10 TYR O
12 CYS C	15 ILU I	9 PRO P	18 VAL V

TOTAL NO. OF ACIDS = 229

- WALSH, K., AND NEURATH, H., PROC. NATL. ACAD. SCI. U.S., VOL. 52, NO.4, PP.884-889, 1964

KAUFFMAN, D. L., J. MOL. BIOL., VOL.12, PP.929-932, 1965

DISULPHIDE BRIDGES WERE FOUND BETWEEN LINKS 13-143, 31-47  
115-216, 122-189, 154-168, AND 179-203.  
THE ACTIVE SERINE IS AT LINK 183.

## PAPAIN

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1/I P E O V D W R Q K G A V T P V K N Q G S C G S C W/A F/I I/  
 31 R N T P O O E G V Q R O C R S R E K G P D A A K T D G V R Q  
 61 V Q P O N Q G A L L O S I A N Q P S V V L Q A A G K D F Q L  
 91 O R G G I F V G P C G N K V D H A V A A V G D N P G O I L I  
 121 K N S W G T G W G E N G O I R I K T G N L N Q O S E Q E L L  
 151 D C D R R S O G C O P G D G W/S A L/V A Q O G I H O R G T G  
 181 N S O G V C G L O T S S F O P V K N •

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1/ILU PRO GLU TYR VAL ASP TRP ARG GLN LYS GLY ALA VAL THR PRO  
 VAL LYS ASN GLN GLY SER CYS GLY SER CYS TRP/ALA PHE/ILU ILU/  
 31 ARG ASN THR PRO TYR TYR GLU GLY VAL GLN ARG TYR CYS ARG SER  
 ARG GLU LYS GLY PRO TYR ALA ALA LYS THR ASP GLY VAL ARG GLN  
 61 VAL GLN PRO TYR ASN GLN GLY ALA LEU LEU TYR SER ILU ALA ASN  
 GLN PRO SER VAL VAL LEU GLN ALA ALA GLY LYS ASP PHE GLN LEU  
 91 TYR ARG GLY GLY ILU PHE VAL GLY PRO CYS GLY ASN LYS VAL ASP  
 HIS ALA VAL ALA ALA VAL GLY TYR ASN PRO GLY TYR ILU LEU ILU  
 121 LYS ASN SER TRP GLY THR GLY TRP GLY GLU ASN GLY TYR ILU ARG  
 ILU LYS THR GLY ASN LEU ASN GLN TYR SER GLU GLN GLU LEU LEU  
 151 ASP CYS ASP ARG ARG SER TYR GLY CYS TYR PRO GLY ASP GLY TRP/  
 SER ALA LEU/VAL ALA GLN TYR GLY ILU HIS TYR ARG GLY THR GLY  
 181 ASN SER TYR GLY VAL CYS GLY LEU TYR THR SER SER PHE TYR PRO  
 VAL LYS ASN \*\*\*

## COMPOSITION

13 ALA A	12 GLN Q	10 LEU L	12 SER S
11 ARG R	6 GLU E	9 LYS K	7 THR T
12 ASN N	27 GLY G	0 MET M	5 TRP W
7 ASP D	2 HIS H	4 PHE F	19 TYR O
7 CYS C	10 ILU I	10 PRO P	15 VAL V

TOTAL NO. OF ACIDS = 198

- LIGHT, A., FRATER, R., KIMMEL, J., AND SMITH, E.L., PROC. NATL. ACAD. SCI. U.S., VOL.52, NO.5, PP.1276-1283, NOV. 1964

DISULPHIDE BRIDGES ARE FORMED BETWEEN CYSTEINES AT POSITIONS 43 AND 152, 100 AND 186, AND 22 AND 159.

THE ACTIVE SULFHYDRYL GROUP IS AT POSITION 25.

## LYSOZYME - CHICKEN

LYSOZYME HAS A BETA (1-4) GLUCOSAMINIDASE ACTIVITY WITH THE ABILITY TO HYDROLYSE A MUCOPOLYSACCHARIDE COMPONENT OF SOME BACTERIAL CELL WALLS RELEASING N-ACETYL AMINO SUGARS DERIVED FROM GLUCOSAMINE AND MURAMIC ACID.

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 K V F G R C E L A A A M K R H G L D N O R G O S L G N W V C  
 31 A A K F E S N F N T Q A T N R N T D G S T D O G I L Q I N S  
 61 R W W C N D G R T P G S R N L C N I P C S A L L S S D I T A  
 91 S V N C A K K I V S D G D G M N A W V A W R N R C K G T D V  
 121 Q A W I R G C R L \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 LYS VAL PHE GLY ARG CYS GLU LEU ALA ALA ALA MET LYS ARG HIS  
 GLY LEU ASP ASN TYR ARG GLY TYR SER LEU GLY ASN TRP VAL CYS  
 31 ALA ALA LYS PHE GLU SER ASN PHE ASN THR GLN ALA THR ASN ARG  
 ASN THR ASP GLY SER THR ASP TYR GLY ILU LEU GLN ILU ASN SER  
 61 ARG TRP TRP CYS ASN ASP GLY ARG THR PRO GLY SER ARG ASN LEU  
 CYS ASN ILU PRO CYS SER ALA LEU LEU SER SER ASP ILU THR ALA  
 91 SER VAL ASN CYS ALA LYS LYS ILU VAL SER ASP GLY ASP GLY MET  
 ASN ALA TRP VAL ALA TRP ARG ASN ARG CYS LYS GLY THR ASP VAL  
 121 GLN ALA TRP ILU ARG GLY CYS ARG LEU \*\*\*

## COMPOSITION

12 ALA	A	3 GLN	Q	8 LEU	L	10 SER	S
11 ARG	R	2 GLU	E	6 LYS	K	7 THR	T
13 ASN	N	12 GLY	G	2 MET	M	6 TRP	W
8 ASP	D	1 HIS	H	3 PHE	F	3 TYR	O
8 CYS	C	6 ILU	I	2 PRO	P	6 VAL	V

TOTAL NO. OF ACIDS = 129

- CANFIELD, R., J. BIOL. CHEM., VOL.238, NO.8, PP.2698-2707, AUG., 1963

CANFIELD, R., LIU, A.K., J. BIOL. CHEM., VOL.240, NO.5, PP. 1997-2002, MAY 1965

ABOVE SEQUENCE CONFIRMED IN THIS WORK.  
DISULPHIDE BONDS ARE FOUND BETWEEN 6 AND 127, 30 AND 115, 64 AND 80, AND 76 AND 94.

JOLLES, J., JAUREGUI-ADELL, J., BERNIER, I., AND JOLLES, P., BIOCHIM. BIOPHYS. ACTA, VOL.78, PP.668-689, 1963

THIS SEQUENCE DIFFERS FROM THE ABOVE AS FOLLOWS, 40-GLN, 41-ALA, 42-THR, 43-THR, 46-ASP, 58-ASN, 59-ILU, 92-ASN, AND 93-VAL.

BLAKE, C.C.F., KOENING, D.F., HAIR, G.A., NORTH, A.C.T., PHILLIPS, D.C., AND SARMA, V.R., NATURE, NO. 4986, PP. 757-761, MAY 22, 1965

A 2 ANGSTROM RESOLUTION FOURIER SYNTHESIS HAS BEEN PERFORMED BY X-RAY CRYSTALLOGRAPHIC METHODS. THE LOCATION OF THE FOUR DISULPHIDE BRIDGES HAS BEEN CONFIRMED. THE BINDING SITE OF THE INHIBITOR N-ACETYL-GLUCOSAMINE AND ITS DIMER HAS BEEN FOUND TO BE VERY EXTENSIVE INVOLVING RESIDUES AT POSITIONS 44, 46, 47, 48, 50, 52, 57, 59, 61-63, 72, 73, 97, 99-101, 103, 107-110, 113, AND 114.

## GLUCAGON - BOVINE

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9  
 1 H S Q G T F T S D O S K O L D S R R A Q D F V Q W L M N T \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 HIS SER GLN GLY THR PHE THR SER ASP TYR SER LYS TYR LEU ASP  
 SER ARG ARG ALA GLN ASP PHE VAL GLN TRP LEU MET ASN THR \*\*\*

## COMPOSITION

1 ALA A	3 GLN Q	2 LEU L	4 SER S
2 ARG R	0 GLU E	1 LYS K	3 THR T
1 ASN N	1 GLY G	1 MET M	1 TRP W
3 ASP D	1 HIS H	2 PHE F	2 TYR O
0 CYS C	0 ILU I	0 PRO P	1 VAL V

TOTAL NO. OF ACIDS = 29

- BROMER, W.W., SINN, L.G., AND BEHRENS, O.K., J. AM. CHEM. SOC., VOL. 79, PP. 2807-2810, JUNE 5, 1957

## ARGININE VASOPRESSIN - BOVINE

THE C TERMINAL GLYCINE IS PRESENT AS THE AMIDE.  
THE TWO CYSTEINES ARE LINKED BY A DISULPHIDE BOND.

1 2 3 4 5 6 7 8 9

1 C O F Q N C P R G \*

1 2 3 4 5 6 7 8 9

1 CYS TYR PHE GLN ASN CYS PRO ARG GLY \*\*\*

## COMPOSITION

0 ALA	A	1 GLN	Q	0 LEU	L	0 SER	S
1 ARG	R	0 GLU	E	0 LYS	K	0 THR	T
1 ASN	N	1 GLY	G	0 MET	M	0 TRP	W
0 ASP	D	0 HIS	H	1 PHE	F	1 TYR	O
2 CYS	C	0 ILU	I	1 PRO	P	0 VAL	V

TOTAL NO. OF ACIDS = 9

• DU VIGNEAUD, V., LAWLER, H. C., AND POPENOE, E. A., J. AM. CHEM. SOC., VOL.75, PP.4880-4881, OCT. 5, 1953

ACHER, R., AND CHAUVET, J., BIOCHIM. BIOPHYS. ACTA, VOL. 12, PP.487-488, 1953

THIS WORK CONFIRMED THE SEQUENCE ABOVE, HOWEVER GLU (E) AND ASP (D) WERE NOT DISTINGUISHED FROM GLN (Q) AND ASN (N).

## LYSINE VASOPRESSIN - PIG

THE C TERMINAL GLYCINE IS PRESENT AS THE AMIDE.  
THE TWO CYSTEINES ARE LINKED BY A DISULPHIDE BOND.

1 2 3 4 5 6 7 8 9

1 C.O.F.Q.N.C.P.K.G.\*

1 2 3 4 5 6 7 8 9

1 CYS.TYR.PHE.GLN.ASN.CYS.PRO.LYS.GLY \*\*\*

## COMPOSITION

0 ALA A	1 GLN Q	0 LEU L	0 SER S
0 ARG R	0 GLU E	1 LYS K	0 THR T
1 ASN N	1 GLY G	0 MET M	0 TRP W
0 ASP D	0 HIS H	1 PHE F	1 TYR O
2 CYS C	0 ILU I	1 PRO P	0 VAL V

TOTAL NO. OF ACIDS = 9

\* POPENOE, E. A., LAWLER, H. C., AND DU VIGNEAUD, V.,  
J. AM. CHEM. SOC., VOL.74, P.3713, JULY 20, 1952



## OXYTOCIN - BOVINE

THE C TERMINAL GLYCINE IS PRESENT AS THE AMIDE.  
 THE TWO CYSTEINES ARE LINKED BY A DISULPHIDE BOND.  
 OXYTOCIN IS THE PRINCIPAL UTERINE CONTRACTING AND MILK EJECTING  
 HORMONE OF THE POSTERIOR PITUITARY.

1 2 3 4 5 6 7 8 9

1 C O I Q N C P L G \*

1 2 3 4 5 6 7 8 9

1 CYS TYR ILU GLN ASN CYS PRO LEU GLY \*\*\*

## COMPOSITION

0 ALA	A	1 GLN	Q	1 LEU	L	0 SER	S
0 ARG	R	0 GLU	E	0 LYS	K	0 THR	T
1 ASN	N	1 GLY	G	0 MET	M	0 TRP	W
0 ASP	D	0 HIS	H	0 PHE	F	1 TYR	O
2 CYS	C	1 ILU	I	1 PRO	P	0 VAL	V

TOTAL NO. OF ACIDS = 9

• DU VIGNEAUD, V., RESSLER, C., TRIPPETT, S., J. BIOL. CHEM.,  
 VOL.205, PP.949-957, 1953

TUPPY, H. AND MICHL, H., MONATSH. CHEM., VOL.84,  
 PP.1011-1020, 1953

## HYPERTENSIN - BOVINE

1 2 3 4 5 6 7 8 9 0

1 D R V D V H P F H L \*

1 2 3 4 5 6 7 8 9 10

1 ASP ARG VAL TYR VAL HIS PRO PHE HIS LEU \*\*\*

## COMPOSITION

0 ALA	A	0 GLN	Q	1 LEU	L	0 SER	S
1 ARG	R	0 GLU	E	0 LYS	K	0 THR	T
0 ASN	N	0 GLY	G	0 MET	M	0 TRP	W
1 ASP	D	2 HIS	H	1 PHE	F	1 TYR	D
0 CYS	C	0 ILE	I	1 PRO	P	2 VAL	V

TOTAL NO. OF ACIDS = 10

• ELLIOT, D. F., AND PEART, W. S., BIOCHEM. J., VOL.65,  
PP.246-254, 1957

## ALPHA MELANOCYTE-STIMULATING HORMONE - BOVINE, PIG, AND HORSE

ACETYL AT AMINO END.  
C-TERMINAL VALINE IS AMINATED.

1 2 3 4 5 6 7 8 9 0 1 2 3  
S O S M E H F R W G K P V \*

1 2 3 4 5 6 7 8 9 10 11 12 13  
SER TYR SER MET GLU HIS PHE ARG TRP GLY LYS PRO VAL \*\*\*

## COMPOSITION

0 ALA	A	0 GLN	Q	0 LEU	L	2 SER	S
1 ARG	R	1 GLU	E	1 LYS	K	0 THR	T
0 ASN	N	1 GLY	G	1 MET	M	1 TRP	W
0 ASP	D	1 HIS	H	1 PHE	F	1 TYR	O
0 CYS	C	0 ILU	I	1 PRO	P	1 VAL	V

TOTAL NO. OF ACIDS = 13

• HARRIS, J. I. AND LERNER, A. B., NATURE, VOL.179, NO.4574,  
PP.1346-1347, JUNE 29, 1957 (PIG)

LI, C. H., LABORATORY INVESTIGATION, VOL. 8, NO.2,  
PP.574-587, 1959 (BOVINE)

DIXON, J. S. AND LI, C. H., J. AM. CHEM. SOC., VOL.82,  
PP.4568-4572, SEPT. 5, 1960 (HORSE)

## BETA MELANOCYTE-STIMULATING HORMONE - BOVINE

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8

1 D S G P O K M E H F R W G S P P K D \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

1 ASP SER GLY PRO TYR LYS MET GLU HIS PHE ARG TRP GLY SER PRO

PRO LYS ASP \*\*\*

## COMPOSITION

0 ALA	A	0 GLN	Q	0 LEU	L	2 SER	S
1 ARG	R	1 GLU	E	2 LYS	K	0 THR	T
0 ASN	N	2 GLY	G	1 MET	M	1 TRP	W
2 ASP	D	1 HIS	H	1 PHE	F	1 TYR	O
0 CYS	C	0 ILE	I	3 PRO	P	0 VAL	V

TOTAL NO. OF ACIDS = 18

\* GESCHWIND, I.I., LI, C. H., AND BARNAFI, L., J. AM. CHEM. SOC., VOL. 79, PP.1003-1004, FEB. 20, 1957

BETA MELANOCYTE-STIMULATING HORMONE - PIG

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8  
 1 D E G P O K M E H F R W G S P P K D \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 ASP GLU GLY PRO TYR LYS MET GLU HIS PHE ARG TRP GLY SER PRO  
 PRO LYS ASP \*\*\*

COMPOSITION

0 ALA A	0 GLN Q	0 LEU L	1 SER S
1 ARG R	2 GLU E	2 LYS K	0 THR T
0 ASN N	2 GLY G	1 MET M	1 TRP W
2 ASP D	1 HIS H	1 PHE F	1 TYR O
0 CYS C	0 ILU I	3 PRO P	0 VAL V

TOTAL NO. OF ACIDS = 18

• HARRIS, J. I. AND ROOS, P., NATURE, VOL.178, NO.4524, P. 90,  
 JULY 14, 1956

GESCHWIND, I.I., LI, C. H., AND BARNAFI, L., J. AM. CHEM.  
 SOC., VOL. 79, PP.620-625, FEB. 5, 1957

BETA MELANOCYTE-STIMULATING HORMONE - HORSE

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8  
 1 D E G P O K M E H F R W G S P R K D \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 ASP GLU GLY PRO TYR LYS MET GLU HIS PHE ARG TRP GLY SER PRO  
 ARG LYS ASP \*\*\*

COMPOSITION

0 ALA A	0 GLN Q	0 LEU L	1 SER S
2 ARG R	2 GLU E	2 LYS K	0 THR T
0 ASN N	2 GLY G	1 MET M	1 TRP W
2 ASP D	1 HIS H	1 PHE F	1 TYR O
0 CYS C	0 ILU I	2 PRO P	0 VAL V

TOTAL NO. OF ACIDS = 18

\* DIXON, J. S. AND LI, C. H., GEN. COMP. ENDOCRINOL.,  
 VOL.1, PP.161-169, 1961

## BETA MELANOCYTE-STIMULATING HORMONE - HUMAN

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2  
 1 A E K K D E G P O R M E H F R W G S P P K D \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 ALA GLU LYS LYS ASP GLU GLY PRO TYR ARG MET GLU HIS PHE ARG  
 TRP GLY SER PRO PRO LYS ASP \*\*\*

## COMPOSITION

1 ALA	A	0 GLN	Q	0 LEU	L	1 SER	S
2 ARG	R	3 GLU	E	3 LYS	K	0 THR	T
0 ASN	N	2 GLY	G	1 MET	M	1 TRP	W
2 ASP	D	1 HIS	H	1 PHE	F	1 TYR	O
0 CYS	C	0 ILU	I	3 PRO	P	0 VAL	V

TOTAL NO. OF ACIDS = 22

- HARRIS, J. I., NATURE, VOL. 184, NO. 4681, PP.167-169,  
 JULY 18, 1959

## BETA CORTICOTROPIN - PIG

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 S D S M E H F R W G K P V G K K R R P V K V D P G A E D D Q  
 31 L A E A F P L E F •

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 SER TYR SER MET GLU HIS PHE ARG TRP GLY LYS PRO VAL GLY LYS  
 LYS ARG ARG PRO VAL LYS VAL TYR PRO GLY ALA GLU ASP ASP GLN  
 31 LEU ALA GLU ALA PHE PRO LEU GLU PHE \*\*\*

## COMPOSITION

3 ALA A	1 GLN Q	2 LEU L	2 SER S
3 ARG R	4 GLU E	4 LYS K	0 THR T
0 ASN N	3 GLY G	1 MET M	1 TRP W
2 ASP D	1 HIS H	3 PHE F	2 TYR O
0 CYS C	0 ILE I	4 PRO P	3 VAL V

TOTAL NO. OF ACIDS = 39

• WHITE, W. F., AND LANDMANN, W. A., J. AM. CHEM. SOC.,  
 VOL. 77, PP.1711-1712, MARCH 20, 1955

HOWARD, K. S., SHEPHERD, R. G., EIGNER, E. A., DAVIS, D. S.,  
 AND BELL, P. H., J. AM. CHEM. SOC., VOL.77, PP.3419-3420,  
 JUNE 20, 1955

BELL, P. H., J. AM. CHEM. SOC., VOL.76, PP.5565-5567, NOV. 1954

THIS SEQUENCE DIFFERS FROM THAT SHOWN ABOVE BY REMOVING THE ASP (D)  
 FROM POSITION 29 AND INSERTING IT BETWEEN POSITIONS 24 AND 25.



## ALPHA CORTICOTROPIN - SHEEP AND BOVINE

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 S D S M E H F R W G K P V G K K R R P V K V O P D G E A E D  
 31 S A Q A F P L E F •

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 SER TYR SER MET GLU HIS PHE ARG TRP GLY LYS PRO VAL GLY LYS  
 LYS ARG ARG PRO VAL LYS VAL TYR PRO ASP GLY GLU ALA GLU ASP  
 31 SER ALA GLN ALA PHE PRO LEU GLU PHE \*\*\*

## COMPOSITION

3 ALA	A	1 GLN	Q	1 LEU	L	3 SER	S
3 ARG	R	4 GLU	E	4 LYS	K	0 THR	T
0 ASN	N	3 GLY	G	1 MET	M	1 TRP	W
2 ASP	D	1 HIS	H	3 PHE	F	2 TYR	O
0 CYS	C	0 ILU	I	4 PRO	P	3 VAL	V

TOTAL NO. OF ACIDS = 39

- LI, C.H., GESCHWIND, I. I., COLE, D., RAACK, I. D., HARRIS, J.I., AND DIXON, J. S., NATURE, VOL.176, NO.4484, PP.687-689, OCT. 8, 1955 (SHEEP)

LI, C. H., DIXON, J. S., AND CHUNG, D., J. AM. CHEM. SOC., VOL. 80, P.2587, 1958 (BOVINE)

## INSULIN A - BOVINE

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1  
 1 G I V E Q C C A S V C S L D Q L E N D C N \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 GLY ILU VAL GLU GLN CYS CYS ALA SER VAL CYS SER LEU TYR GLN  
 LEU GLU ASN TYR CYS ASN \*\*\*

## COMPOSITION

1 ALA A	2 GLN Q	2 LEU L	2 SER S
0 ARG R	2 GLU E	0 LYS K	0 THR T
2 ASN N	1 GLY G	0 MET M	0 TRP W
0 ASP D	0 HIS H	0 PHE F	2 TYR D
4 CYS C	1 ILU I	0 PRO P	2 VAL V

TOTAL NO. OF ACIDS = 21

- SANGER, F. AND THOMPSON, E. O. P., BIOCHEM J., VOL.53, PP. 353-374, 1953

THE AMIDE GROUPS WERE SUBSEQUENTLY DETERMINED.

RYLE, A. P., SANGER, F., SMITH, L.F., AND KITAI, R.,  
 BIOCHEM. J., VOL. 60, PP. 541-556, 1955

## INSULIN A - BONITO

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1  
 1 G I(H,E,E,C(C,K,P,H)C,D,L)F E L E D D C N •

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 GLY ILU(HIS,GLU,GLU,CYS(CYS,LYS,PRO,HIS)CYS,ASP,LEU)PHE GLU  
 LEU GLU ASP TYR CYS ASN \*\*\*

## COMPOSITION

0 ALA	A	0 GLN	Q	2 LEU	L	0 SER	S
0 ARG	R	4 GLU	E	1 LYS	K	0 THR	T
1 ASN	N	1 GLY	G	0 MET	M	0 TRP	W
2 ASP	D	2 HIS	H	1 PHE	F	1 TYR	O
4 CYS	C	1 ILU	I	1 PRO	P	0 VAL	V

TOTAL NO. OF ACIDS = 21

\* KOTAKI, A., J. BIOCHEM. (TOKYO), VOL.53, NO.1, PP.61-70, 1963

## INSULIN A - HORSE

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1  
 1 G I V E Q C C T G I C S L O Q L E N O C N \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 GLY ILU VAL GLU GLN CYS CYS THR GLY ILU CYS SER LEU TYR GLN  
 LEU GLU ASN TYR CYS ASN \*\*\*

## COMPOSITION

0 ALA	A	2 GLN	Q	2 LEU	L	1 SER	S
0 ARG	R	2 GLU	E	0 LYS	K	1 THR	T
2 ASN	N	2 GLY	G	0 MET	M	0 TRP	W
0 ASP	D	0 HIS	H	0 PHE	F	2 TYR	O
4 CYS	C	2 ILU	I	0 PRO	P	1 VAL	V

TOTAL NO. OF ACIDS = 21

- HARRIS, J. I., SANGER, F., AND NAUGHTON, M. A., ARCH. BIOCHEM. BIOPHYS., VOL.65, PP.427-438, 1956

SOME EVIDENCE FOR THE SEQUENCE WAS DERIVED FROM HOMOLOGY WITH BOVINE INSULIN.

## INSULIN A - SHEEP

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1  
 1 G I V E Q C C A G V C S L O Q L E N D C N \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 GLY ILU VAL GLU GLN CYS CYS ALA GLY VAL CYS SER LEU TYR GLN  
 LEU GLU ASN TYR CYS ASN \*\*\*

## COMPOSITION

1 ALA	A	2 GLN	Q	2 LEU	L	1 SER	S
0 ARG	R	2 GLU	E	0 LYS	K	0 THR	T
2 ASN	N	2 GLY	G	0 MET	M	0 TRP	W
0 ASP	D	0 HIS	H	0 PHE	F	2 TYR	O
4 CYS	C	1 ILU	I	0 PRO	P	2 VAL	V

TOTAL NO. OF ACIDS = 21

\* BROWN, H., SANGER, F., AND KITAI, R., BIOCHEM. J., VOL.60,  
 PP.556-565, 1955

SOME EVIDENCE FOR THE SEQUENCE WAS DERIVED FROM HOMOLOGY  
 WITH BOVINE INSULIN.

## INSULIN A - SPERM WHALE, FIN-WHALE, PIG, AND HUMAN

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1  
 1 G I V E Q C C T S I C S L O Q L E N D C N \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 GLY ILU VAL GLU GLN CYS CYS THR SER ILU CYS SER LEU TYR GLN  
 LEU GLU ASN TYR CYS ASN \*\*\*

## COMPOSITION

0 ALA	A	2 GLN	Q	2 LEU	L	2 SER	S
0 ARG	R	2 GLU	E	0 LYS	K	1 THR	T
2 ASN	N	1 GLY	G	0 MET	M	0 TRP	W
0 ASP	D	0 HIS	H	0 PHE	F	2 TYR	O
4 CYS	C	2 ILU	I	0 PRO	P	1 VAL	V

TOTAL NO. OF ACIDS = 21

\* BROWN, H., SANGER, F., AND KITAI, R., BIOCHEM. J., VOL.60,  
 PP.556-565, 1955 (PIG)

SOME EVIDENCE FOR THE SEQUENCE WAS DERIVED FROM HOMOLOGY  
 WITH BOVINE INSULIN.

HARRIS, J. I., SANGER, F., AND NAUGHTON, M. A., ARCH. BIOCHEM.  
 BIOPHYS., VOL.65, PP.427-438, 1956 (SPERM WHALE)

HAMA, H., TITANI, K., SAKAKI, S., AND NARITA, K., J. BIOCHEM.  
 (TOKYO), VOL.56, NO.3, PP.285-293, 1964 (FIN-WHALE)

THIS WORK CONFIRMED THE SEQUENCE ABOVE, EXCEPT GLU (E) AND GLN (Q)  
 WERE INTERCHANGED AT POSITIONS 15 AND 17.

NICOL, D. S. H. AND SMITH, L. F., NATURE, VOL.187, NO.4736,  
 PP.483-485, AUG. 6, 1960 (HUMAN)

## INSULIN A - SEI-WHALE

I

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1

1 G I V E Q C C A S T C S L O Q L E N D C N \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

1 GLY ILU VAL GLU GLN CYS CYS ALA SER THR CYS SER LEU TYR GLN

LEU GLU ASN TYR CYS ASN \*\*\*

## COMPOSITION

1 ALA	A	2 GLN	Q	2 LEU	L	2 SER	S
0 ARG	R	2 GLU	E	0 LYS	K	1 THR	T
2 ASN	N	1 GLY	G	0 MET	M	0 TRP	W
0 ASP	D	0 HIS	H	0 PHE	F	2 TYR	O
4 CYS	C	1 ILU	I	0 PRO	P	1 VAL	V

TOTAL NO. OF ACIDS = 21

ISHIHARA, Y., SAITO, T., ITO, Y., AND FUJINO, M., NATURE,  
VOL.181, NO.4621, PP.1461-1469, MAY 24, 1958 (SEI-WHALE)

INSULIN B - BOVINE, SHEEP, HORSE, HUMAN, PIG, AND SPERM WHALE

TWO DISULPHIDE BONDS CONNECT THE A AND B CHAINS.

A7 IS BONDED TO B7 AND A20 IS BONDED TO B19. IN ADDITION THERE IS A BOND FROM A6 TO A11.

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 F V N Q H L C G S H L V E A L O L V C G E R G F F O T P K A

•

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 PHE VAL ASN GLN HIS LEU CYS GLY SER HIS LEU VAL GLU ALA LEU  
 TYR LEU VAL CYS GLY GLU ARG GLY PHE PHE TYR THR PRO LYS ALA

\*\*\*

#### COMPOSITION

2 ALA	A	1 GLN	Q	4 LEU	L	1 SER	S
1 ARG	R	2 GLU	E	1 LYS	K	1 THR	T
1 ASN	N	3 GLY	G	0 MET	M	0 TRP	W
0 ASP	D	2 HIS	H	3 PHE	F	2 TYR	O
2 CYS	C	0 ILU	I	1 PRO	P	3 VAL	V

TOTAL NO. OF ACIDS = 30

- RYLE, A. P., SANGER, F., SMITH, L.F., AND KITAI, R., BIOCHEM. J., VOL. 60, PP.541-556, 1955 (BOVINE, SHEEP, AND PIG)

SANGER, F. AND TUPPY, H., BIOCHEM. J., VOL.49, PP.481-490, 1951 (BOVINE)

THE AMIDE GROUPS WERE SUBSEQUENTLY DETERMINED.

HARRIS, J. I., SANGER, F., AND NAUGHTON, M. A., ARCH. BIOCHEM. BIOPHYS., VOL.65, PP.427-438, 1956 (SPERM WHALE AND HORSE)

ISHIHARA, Y., SAITO, T., ITO, Y., AND FUJINO, M., NATURE, VOL.181, NO.4621, PP.1461-1469, MAY 24, 1958 (SPERM AND SEI-WHALE)

NICOL, D. S. H. AND SMITH, L. F., NATURE, VOL.187, NO.4736, PP.483-485, AUG. 6, 1960 (HUMAN)

HUMAN INSULIN B CHAIN IS IDENTICAL WITH ABOVE EXCEPT THAT POSITION 30 IS THR (T).



## INSULIN B - BONITO

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 A A N(P,H,L)C(G,S,H,L,V,E,A,L)D L(V,C,G,E)R G F F O Q P K •

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 ALA ALA ASN(PRO,HIS,LEU)CYS(GLY,SER,HIS,LEU,VAL,GLU,ALA,LEU)  
 TYR LEU(VAL,CYS,GLY,GLU)ARG GLY PHE PHE TYR GLN PRO LYS \*\*\*

## COMPOSITION

3 ALA	A	1 GLN	Q	4 LEU	L	1 SER	S
1 ARG	R	2 GLU	E	1 LYS	K	0 THR	T
1 ASN	N	3 GLY	G	0 MET	M	0 TRP	W
0 ASP	D	2 HIS	H	2 PHE	F	2 TYR	O
2 CYS	C	0 ILU	I	2 PRO	P	2 VAL	V

TOTAL NO. OF ACIDS = 29

• KOTAKI, A., J. BIOCHEM.(TOKYO), VOL.51, NO.4, PP.301-309, 1962

FIBRINOPEPTIDE A - BOVINE

FIBRINOPEPTIDES ARE THOSE PORTIONS OF VERTEBRATE FIBRINOGEN MOLECULES WHICH ARE PROTEOLYTICALLY REMOVED BY THE ENZYME THROMBIN. THEIR REMOVAL PERMITS SPONTANEOUS POLYMERIZATION OF THE PARENT MOLECULES TO FORM AN INSOLUBLE FIBRINOGENEL. SINCE THE FUNCTION OF THE FIBRINOPEPTIDES IS RATHER NON-SPECIFIC, LARGE SEQUENCE CHANGES ARE OBSERVED AMONG CLOSELY RELATED SPECIES.

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9  
 1 E D G S D P P S G D F L T E G G G V R /

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 GLU ASP GLY SER ASP PRO PRO SER GLY ASP PHE LEU THR GLU GLY  
 GLY GLY VAL ARG ///

COMPOSITION

0 ALA A	0 GLN Q	1 LEU L	2 SER S
1 ARG R	2 GLU E	0 LYS K	1 THR T
0 ASN N	5 GLY G	0 MET M	0 TRP W
3 ASP D	0 HIS H	1 PHE F	0 TYR O
0 CYS C	0 ILU I	2 PRO P	1 VAL V

TOTAL NO. OF ACIDS = 19

• DOOLITTLE, R. F. AND BLOMBACK, B., NATURE, VOL. 202, NO. 4928, PP. 147-152, APRIL 11, 1964

FIBRINOPEPTIDE A - SHEEP

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9  
 1 A D D S D P V G G E F L A E G G G V R /

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 ALA ASP ASP SER ASP PRO VAL GLY GLY GLU PHE LEU ALA GLU GLY  
 GLY GLY VAL ARG ///

COMPOSITION

2 ALA	A	0 GLN	Q	1 LEU	L	1 SER	S
1 ARG	R	2 GLU	E	0 LYS	K	0 THR	T
0 ASN	N	5 GLY	G	0 MET	M	0 TRP	W
3 ASP	D	0 HIS	H	1 PHE	F	0 TYR	O
0 CYS	C	0 ILU	I	1 PRO	P	2 VAL	V

TOTAL NO. OF ACIDS = 19

- DOOLITTLE, R. F. AND BLOMBACK, B., NATURE, VOL. 202, NO. 4928, PP. 147-152, APRIL 11, 1964

## FIBRINOPEPTIDE A - GOAT

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9

1 A D D S D P V G G E F L A E G G G V R /

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

1 ALA ASP ASP SER ASP PRO VAL GLY GLY GLU PHE LEU ALA GLU GLY  
GLY GLY VAL ARG ///

## COMPOSITION

2 ALA A	0 GLN Q	1 LEU L	1 SER S
1 ARG R	2 GLU E	0 LYS K	0 THR T
0 ASN N	5 GLY G	0 MET M	0 TRP W
3 ASP D	0 HIS H	1 PHE F	0 TYR O
0 CYS C	0 ILU I	1 PRO P	2 VAL V

TOTAL NO. OF ACIDS = 19

- DOOLITTLE, R. F. AND BLOMBACK, B., NATURE, VOL. 202, NO. 4928, PP. 147-152, APRIL 11, 1964

FIBRINOPEPTIDE A - REINDEER

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9  
 1 A D G S D P A G G E F(L,A,E,G,G,G,V)R /

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 ALA ASP GLY SER ASP PRO ALA GLY GLY GLU PHE(LAU,ALA,GLU,GLY,  
 GLY,GLY,VAL)ARG ///

COMPOSITION

3 ALA A	0 GLN Q	1 LEU L	1 SER S
1 ARG R	2 GLU E	0 LYS K	0 THR T
0 ASN N	6 GLY G	0 MET M	0 TRP W
2 ASP D	0 HIS H	1 PHE F	0 TYR O
0 CYS C	0 ILU I	1 PRO P	1 VAL V

TOTAL NO. OF ACIDS = 19

- DOOLITTLE, R. F. AND BLOMBACK, B., NATURE, VOL. 202, NO. 4928, PP. 147-152, APRIL 11, 1964

## PIBRINOPEPTIDE A - PIG

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7  
 1 A E V Q D K G E F L A E G G G V R /

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 ALA GLU VAL GLN ASP LYS GLY GLU PHE LEU ALA GLU GLY GLY GLY  
 VAL ARG ///

## COMPOSITION

2 ALA	A	1 GLN	Q	1 LEU	L	0 SER	S
1 ARG	R	3 GLU	E	1 LYS	K	0 THR	T
0 ASN	N	4 GLY	G	0 MET	M	0 TRP	W
1 ASP	D	0 HIS	H	1 PHE	F	0 TYR	O
0 CYS	C	0 ILU	I	0 PRO	P	2 VAL	V

TOTAL NO. OF ACIDS = 17

\* DOOLITTLE, R. F. AND BLOMBACK, B., NATURE, VOL. 202,  
 NO. 4928, PP. 147-152, APRIL 11, 1964

## FIBRINOPEPTIDE A - HUMAN

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6  
 1 A D S G E G D F L A E G G G V R /

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 ALA ASP SER GLY GLU GLY ASP PHE LEU ALA GLU GLY GLY GLY VAL  
 ARG ///

## COMPOSITION

2 ALA A	0 GLN Q	1 LEU L	1 SER S
1 ARG R	2 GLU E	0 LYS K	0 THR T
0 ASN N	5 GLY G	0 MET M	0 TRP W
2 ASP D	0 HIS H	1 PHE F	0 TYR O
0 CYS C	0 ILU I	0 PRO P	1 VAL V

TOTAL NO. OF ACIDS = 16

- DOOLITTLE, R. F. AND BLOMBACK, B., NATURE,\* VOL. 202,  
 NO. 4928, PP. 147-152, APRIL 11, 1964

PHOSPHO-SERINE OCCURS AT POSITION 3 IN ABOUT HALF THE MOLECULES.  
 A MINOR COMPONENT FRAGMENT, WITH THE N TERMINAL ALANINE MISSING,  
 HAS BEEN DETECTED IN ALL INDIVIDUALS.

FIBRINOPEPTIDE A - RABBIT

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6  
 1 V D P G E T S F L(T,E,G,G)D A R /

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 VAL ASP PRO GLY GLU THR SER PHE LEU(THR,GLU,GLY,GLY)ASP ALA  
 ARG ///

COMPOSITION

1 ALA	A	0 GLN	Q	1 LEU	L	1 SER	S
1 ARG	R	2 GLU	E	0 LYS	K	2 THR	T
0 ASN	N	3 GLY	G	0 MET	M	0 TRP	W
2 ASP	D	0 HIS	H	1 PHE	F	0 TYR	O
0 CYS	C	0 ILU	I	1 PRO	P	1 VAL	V

TOTAL NO. OF ACIDS = 16

- DOOLITTLE, R. F. AND BLOMBACK, B., NATURE, VOL. 202, NO. 4928, PP. 147-152, APRIL 11, 1964



FIBRINOPEPTIDE B - BOVINE

PYRROLIDONE CARBOXYLIC ACID - AT AMINO END  
 SO4 ATTACHED TO TYROSINE AT POSITION 5

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 I F P T D D D E G Q D D R P K V G L G A R /

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 PHE PRO THR ASP TYR ASP GLU GLY GLN ASP ASP ARG PRO LYS VAL  
 GLY LEU GLY ALA ARG ///

COMPOSITION

1 ALA A	1 GLN Q	1 LEU L	0 SER S
2 ARG R	1 GLU E	1 LYS K	1 THR T
0 ASN N	3 GLY G	0 MET M	0 TRP W
4 ASP D	0 HIS H	1 PHE F	1 TYR O
0 CYS C	0 ILU I	2 PRO P	1 VAL V

TOTAL NO. OF ACIDS = 20

\* DOOLITTLE, R. F. AND BLOMBACK, B., NATURE, VOL. 202,  
 NO. 4928, PP. 147-152, APRIL 11, 1964

FIBRINOPEPTIDE B - SHEEP

SO4 ATTACHED TO TYROSINE AT POSITION 5

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 G O L D O D E V D D N R A K L P L D A R /

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 GLY TYR LEU ASP TYR ASP GLU VAL ASP ASP ASN ARG ALA LYS LEU  
 PRO LEU ASP ALA ARG ///

COMPOSITION

2 ALA A	0 GLN Q	3 LEU L	0 SER S
2 ARG R	1 GLU E	1 LYS K	0 THR T
1 ASN N	1 GLY G	0 MET M	0 TRP W
5 ASP D	0 HIS H	0 PHE F	2 TYR O
0 CYS C	0 ILU I	1 PRO P	1 VAL V

TOTAL NO. OF ACIDS = 20

\* DOOLITTLE, R. F. AND BLOMBACK, B., NATURE, VOL. 202,  
 NO. 4928, PP. 147-152, APRIL 11, 1964

FIBRINOPEPTIDE B - GOAT

SO4 ATTACHED TO TYROSINE AT POSITION 5

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 G O L D D D E V D D N R A K L P L D A R /

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 GLY TYR LEU ASP TYR ASP GLU VAL ASP ASP ASN ARG ALA LYS LEU  
 PRO LEU ASP ALA ARG ///

COMPOSITION

2 ALA A	0 GLN Q	3 LEU L	0 SER S
2 ARG R	1 GLU E	1 LYS K	0 THR T
1 ASN N	1 GLY G	0 MET M	0 TRP W
5 ASP D	0 HIS H	0 PHE F	2 TYR O
0 CYS C	0 ILU I	1 PRO P	1 VAL V

TOTAL NO. OF ACIDS = 20

- DOOLITTLE, R. F. AND BLOMBACK, B., NATURE, VOL. 202, NO. 4928, PP.147-152, APRIL 11, 1964

FIBRINOPEPTIDE B - REINDEER

PYRROLIDONE CARBOXYLIC ACID - AT AMINO END  
 SO<sub>4</sub> ATTACHED TO TYROSINE AT POSITION 4  
 A MUTANT HAS BEEN FOUND WHERE GLYCINE REPLACES HISTIDINE  
 IN POSITION 9.

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9  
 1 L A D O D E V(E,H,D)R A K L H L D A R /

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15'  
 1 LEU ALA ASP TYR ASP GLU VAL(GLU,HIS,ASP)ARG ALA LYS LEU HIS  
 LEU ASP ALA ARG ///

COMPOSITION

3 ALA A	0 GLN Q	3 LEU L	0 SER S
2 ARG R	2 GLU E	1 LYS K	0 THR T
0 ASN N	0 GLY G	0 MET M	0 TRP W
4 ASP D	2 HIS H	0 PHE F	1 TYR O
0 CYS C	0 ILU I	0 PRO P	1 VAL V

TOTAL NO. OF ACIDS = 19

\* DOOLITTLE, R. F. AND BLOMBACK, B., NATURE, VOL. 202,  
 NO. 4928, PP. 147-152, APRIL 11, 1964

FIBRINOPEPTIDE B - PIG

SO4 ATTACHED TO TYROSINE AT POSITION 4

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9  
 1 A I D O D E D E D G R P K V H V D A R /

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 ALA ILU ASP TYR ASP GLU ASP GLU ASP GLY ARG PRO LYS VAL HIS  
 VAL ASP ALA ARG ///

COMPOSITION

2 ALA	A	0 GLN	Q	0 LEU	L	0 SER	S
2 ARG	R	2 GLU	E	1 LYS	K	0 THR	T
0 ASN	N	1 GLY	G	0 MET	M	0 TRP	W
5 ASP	D	1 HIS	H	0 PHE	F	1 TYR	O
0 CYS	C	1 ILU	I	1 PRO	P	2 VAL	V

TOTAL NO. OF ACIDS = 19

• DOOLITTLE, R. F. AND BLOMBACK, B., NATURE, VOL. 202,  
 NO. 4928, PP. 147-152, APRIL 11, 1964

## FIBRINOPEPTIDE B - HUMAN

PYRROLIDONE CARBOXYLIC ACID - AT AMINO END  
PHOSPHO-SERINE OCCURS IN POSITION 11.

1 2 3 4 5 6 7 8 9 0 1 2 3  
1 G V N D N E E G F F S A R /

1 2 3 4 5 6 7 8 9 10 11 12 13  
1 GLY VAL ASN ASP ASN GLU GLU GLY PHE PHE SER ALA ARG ///

## COMPOSITION

1 ALA	A	0 GLN	Q	0 LEU	L	1 SER	S
1 ARG	R	2 GLU	E	0 LYS	K	0 THR	T
2 ASN	N	2 GLY	G	0 MET	M	0 TRP	W
1 ASP	D	0 HIS	H	2 PHE	F	0 TYR	O
0 CYS	C	0 ILE	I	0 PRO	P	1 VAL	V

TOTAL NO. OF ACIDS = 13

- DOOLITTLE, R. F. AND BLOMBACK, B., NATURE, VOL. 202,  
NO. 4928, PP. 147-152, APRIL 11, 1964

FIBRINOPEPTIDE B - RABBIT

SD4 ATTACHED TO TYROSINE AT POSITION 4

1 2 3 4 5 6 7 8 9 0 1 2 3

1 A D D O(D,E,P,L,D,V)D A R /

1 2 3 4 5 6 7 8 9 10 11 12 13

1 ALA ASP ASP TYR(ASP,GLU,PRO,LEU,ASP,VAL)ASP ALA ARG ///

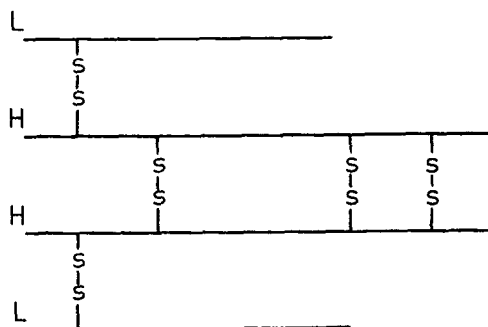
COMPOSITION

2 ALA	A	0 GLN	Q	1 LEU	L	0 SER	S
1 ARG	R	1 GLU	E	0 LYS	K	0 THR	T
0 ASN	N	0 GLY	G	0 MET	M	0 TRP	W
5 ASP	D	0 HIS	H	0 PHE	F	1 TYR	O
0 CYS	C	0 ILU	I	1 PRO	P	1 VAL	V

TOTAL NO. OF ACIDS = 13

- DOOLITTLE, R. F. AND BLOMBACK, B., NATURE, VOL. 202, NO. 4928, PP. 147-152, APRIL 11, 1964

Immunoglobulins are serum proteins distinguishable by electrophoretic mobilities, sedimentation coefficients and differential solubilities in variable ethanol-salt solutions. Of these, the gamma globulins are associated with normal antibody function. A proposal for the structure of gamma globulin has been made by Porter (1959) and Fleishman *et al.*, (1963).



Gamma globulin is thought to be a tetramer consisting of two pairs of identical polypeptide chains held in a particular configuration by disulfide bonds. There are two L (m.w. 20-25,000 each) and two H chains (m.w. 50,000-55,000 each). Because of the chemical problems associated with elucidation of gamma globulin structure, attention has turned to the abundantly produced, structurally similar globulins found in multiple myeloma.

Bence-Jones proteins are found exclusively in the urine of all multiple myeloma patients, and probably represent abnormal protein synthesized by the multiple myeloma tumor cell. They are thought to be made exclusively of L chains, related to gamma globulins, (Edelman and Gally, 1962, S. Cohen, 1963, Putnam 1962). It is thought that determination of the amino acid sequence of a particular individual's Bence-Jones protein would reflect a homologous sequence in that individual's antibody structure, thereby partially elucidating the structure of gamma globulin.

- Cohen, S. *Biochem. J.* Vol. 89, p. 334 (1963)  
 Edelman, G. M. and Gally, J. A. *J. Exp. Med.* Vol. 116, p. 207 (1962)  
 Fleishman, J. B. *et al.* *Biochem. J.* Vol. 88, p. 220 (1963)  
 Porter, R. R. *Biochem. J.* Vol. 73, p. 119 (1959)  
 Putnam, F. W. *Biochim. Biophys. Acta.* Vol. 63, p. 539 (1962)



## BENCE-JONES PROTEIN

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0  
 1 D(T,S,S,S,E,E,P,M,I)L S(S,G,A,V)D R(D,T,T,S,S,E,E,A,V,I,I,I,  
 31 F,C)L(O,D,W,E,E,P,G)K K A P K L L I D D A S K L E(S,P,G,A,V)  
 61 R F S(D,T,T,S,G,G,G)F T(D,S,S,E,E,P,I,L)I A T O(D,D,T,E,E,P,  
 91 L,L,C,O,F,F)G(T,G,G)K V D F K R T(S,P,A,A,V)V F I(D,S,E,E,P,  
 121 P,F)L K S(T,S,G,A)V(V,C)L L D(D,P,F)O R E A K V E W K V(D,D,  
 151 D,S,S,E,E,G,A,L)E S(D,T,S,E,E,V)K D(T,S)O S S S T L L T L S  
 181 K A D O E K H K L O A C E V(T,E,G,H)L S(T,S,P,V)K S F D R G  
 211 E C \*

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
 1 ASP(THR,SER,SER,SER,GLU,GLU,PRO,MET,ILU)LEU SER(SER,GLY,ALA,  
 VAL)ASP ARG(ASP,THR,THR,SER,SER,GLU,GLU,ALA,VAL,ILU,ILU,ILU,  
 31 PHE,CYS)LEU(TYR,ASP,TRP,GLU,GLU,PRO,GLY)LYS LYS ALA PRO LYS  
 LEU LEU ILU TYR ASP ALA SER LYS LEU GLU(SER,PRO,GLY,ALA,VAL)  
 61 ARG PHE SER(ASP,THR,THR,SER,GLY,GLY,GLY)PHE THR(ASP,SER,SER,  
 GLU,GLU,PRO,ILU,LEU)ILU ALA THR TYR(ASP,ASP,THR,GLU,GLU,PRO,  
 91 LEU,LEU,CYS,TYR,PHE,PHE)GLY(THR,GLY,GLY)LYS VAL ASP PHE LYS  
 ARG THR(SER,PRO,ALA,ALA,VAL)VAL PHE ILU(ASP,SER,GLU,GLU,PRO,  
 121 PRO,PHE)LEU LYS SER(THR,SER,GLY,ALA)VAL(VAL,CYS)LEU LEU ASP  
 (ASP,PRO,PHE)TYR ARG GLU ALA LYS VAL GLU TRP LYS VAL(ASP,ASP,  
 151 ASP,SER,SER,GLU,GLU,GLY,ALA,LEU)GLU SER(ASP,THR,SER,GLU,GLU,  
 VAL)LYS ASP(THR,SER)TYR SER SER SER THR LEU LEU THR LEU SER  
 181 LYS ALA ASP TYR GLU LYS HIS LYS LEU TYR ALA CYS GLU VAL(THR,  
 GLU,GLY,HIS)LEU SER(THR,SER,PRO,VAL)LYS SER PHE ASP ARG GLY  
 211 GLU CYS \*\*\*

## COMPOSITION

13	ALA	A	0	GLN	Q	17	LEU	L	29	SER	S
5	ARG	R	24	GLU	E	14	LYS	K	17	THR	T
0	ASN	N	13	GLY	G	1	MET	M	2	TRP	W
20	ASP	D	2	HIS	H	10	PHE	F	8	TYR	Y
5	CYS	C	8	ILE	I	11	PRO	P	13	VAL	V

TOTAL NO. OF ACIDS = 212

- HILSCHMANN, N. AND CRAIG, L.C., PROC. NATL. ACAD. SCI. U.S., VOL.53, NO.6, PP.1403-1409, 1965

AUTHOR INDEX

100,000

	PAGE
ACHER, R.	TI BOPA 5.001
	PR BOAR 8.101
AMBLER, R. P.	CY PS 1.006
	AZ PS 3.003
ANDERER, F. A.	TM TM 6.001
ANSFIELD, M. J.	TI BOPA 5.001
BAHL, O. P.	CY RS 1.009
BAGLIONI, C.	GL HUH 2.020
BARNAFI, L.	TN BOBM 8.102
	TN PGBM 8.203
BARTSCH, R. G.	DH CH 3.001
BEALE, D.	GL HUH 2.020
BEHRENS, O. K.	GN BO 8.001
BELL, P. H.	TN PGAC 8.206
BENSON, A.	CY PG 1.005
	FE CP 3.002
BERNIER, I.	LS CH 7.201
BLAKE, C. C. F.	LS CH 7.201
BLOMBACK, B.	FB BOA 9.001
	FB SHA 9.002
	FB GTA 9.003
	FB RDA 9.004
	FB PGA 9.005
	FB HUA 9.006
	FB RTA 9.007
	FB BOB 9.101
	FB SHB 9.102
	FB GTB 9.103
	FB RDB 9.104
	FB PGB 9.105
	FB HUB 9.106
	FB RTB 9.107
	GL HUHA 2.001
	GL HUHB 2.002
	GL HOHA 2.006
	GN BO 8.001
	IS SHA 8.304
	IS WPA 8.305
	TR BOCH 7.001
	AZ PS 3.003
	GL LEHB 2.007
	LS CH 7.201
	TI BOPA 5.001
	PR BOAR 8.101
	TN SBAC 8.207
	10.000
	TN SBAC 8.207
	GL HUHA 2.001
	GL HUHG 2.003
	BJ HU 10.001
	TN PGAC 8.206
	TN BPAM 8.201
	TN HOBM 8.204
	TN SBAC 8.207
	TI BOPA 5.001
BRAUNITZER, G.	
BROMER, W. W.	
BROWN, H.	
BROWN, J. R.	
BROWN, L. H.	
BUETTNER-JANUSCH, J.	
CANFIELD, R.	
CHAUVET, J.	
CHUNG, D.	
COHEN, S.	
COLE, D.	
CORMICK, J.	
CRAIG, L. C.	
DAVIS, D. S.	
DIXON, J. S.	
DLOUHA, V.	

DOOLITTLE, R. F.

FB BOA 9.001  
 FB SHA 9.002  
 FB GTA 9.003  
 FB RDA 9.004  
 FB PGA 9.005  
 FB HUA 9.006  
 FB RTA 9.007  
 FB BOB 9.101  
 FB SHB 9.102  
 FB GTB 9.103  
 FB RDB 9.104  
 FB PGB 9.105  
 FB HUB 9.106  
 FB RTB 9.107  
 DH CH 3.001  
 PR BOAR 8.101  
 PR PGLS 8.102  
 PR BOOX 8.103

DUS, K.

DU VIGNEAUD, V.

EDELMAN, G. M.  
 EDMUNDSON, A. B.  
 EFRON, M. L.  
 EIGNER, E. A.  
 ELLIOT, D. F.  
 FLEISHMAN, J. B.  
 FRAENKEL-CONRAT, H.  
 FRATER, R.  
 FUJINO, M.

10.000  
 GL WHMY 2.101  
 GL HUH 2.020  
 TN PGAC 8.206  
 PR BOHY 8.104  
 10.000  
 TM TM 6.001  
 PA PA 7.101  
 IS WHA 8.306  
 IS BOB 8.321  
 TM TM 6.001

FUNATSU, G.  
 GALLY, J. A.  
 GEHRING-MULLER, R.

10.000  
 GL HUHA 2.001  
 GL HUHB 2.002  
 GL HUH 2.020  
 TN BOBM 8.202  
 TN PGBM 8.203  
 TN SBAC 8.207  
 IS WPA 8.305  
 TN BPAM 8.201  
 TN PGBM 8.203  
 TN HUBM 8.205  
 TN SBAC 8.207  
 IS HOA 8.303  
 IS WPA 8.305  
 IS BOB 8.321

GERALD, P. S.  
 GESCHWIND, I. I.

TR BOCH 7.001  
 CY PG 1.005  
 GL HUHA 2.001  
 GL LEHB 2.007  
 GL HUH 2.020  
 GL HUHA 2.001  
 GL HUHB 2.002  
 BJ HU 10.001  
 GL HUHA 2.001  
 GL HUHB 2.002

HAMA, H.  
 HARRIS, J. I.

HARTLEY, B. S.  
 HIGA, H.  
 HILL, R. J.  
 HILL, R. L.

HILSCHMANN, N.

HILSE, K.

HOBOM, G.	GL HUHA 2.001
HOWARD, K. S.	GL HUHB 2.002
HUNT, J. A.	TN PGAC 8.206
INGRAM, V. M.	GL HUH 2.020
ISHIHARA, Y.	GL HUH 2.020
ITO, Y.	IS WHA 8.306
JAUREGUI-ADELL, J.	IS BOB 8.321
JOLLES, J.	IS WHA 8.306
JOLLES, P.	IS BOB 8.321
JONES, R. T.	LS CH 7.201
KAMEN, M. D.	LS CH 7.201
KASSELL, B.	LS CH 7.201
KAUFFMAN, D. L.	GL HUHG 2.003
KEIL, B.	DH CH 3.001
KIMMEL, J.	TI BOPA 5.001
KINGMA, S.	TR BOCH 7.001
KITAI, R.	TR BOTR 7.002
KOENING, D. F.	TR BOCH 7.001
KONIGSBERG, W.	PA PA 7.101
KOSTKA, V.	GL HUH 2.020
KOTAKI, A.	IS BOA 8.301
KREIL, G.	IS SHA 8.304
LANDMANN, W. A.	IS WPA 8.305
LASKOWSKI, M.	IS BOB 8.321
LAWLER, H. C.	LS CH 7.201
LEHMANN, H.	GL HUHA 2.001
LERNER, A. B.	TR BOCH 7.001
LI, C. H.	IS BNA 8.302
LIGHT, A.	IS BNB 8.322
LIU, A. K.	CY HO 1.003
MAIR, G. A.	CY TF 1.007
MARGOLIASH, E.	TN PGAC 8.206
MARTIN, N.	TI BOPA 5.001
MATSUBARA, H.	PR BOAR 8.101
MATSUDA, G.	PR PGLS 8.102
MELOUN, B.	GL HUH 2.020
MICHL, H.	TN BPAM 8.201
	TN BPAM 8.201
	TN BOBM 8.202
	TN PGBM 8.203
	TN HOBM 8.204
	TN SBAC 8.207
	PA PA 7.101
	LS CH 7.201
	LS CH 7.201
	CY CH 1.002
	CY HO 1.003
	CY PG 1.005
	GL HUH 2.020
	CY HU 1.004
	CY PG 1.005
	GL HOHA 2.006
	TI BOPA 5.001
	TR BOCH 7.001
	PR BOOX 8.103

MOORE, S.	RN BO	4.001
MOWER, H. F.	FE CP	3.002
MULLER, C. J.	GL HUH	2.020
MURAKAMI, H.	CY BY	1.001
MURAYAMA, M.	GL HUH	2.020
NAKASHIMA, T.	CY PG	1.005
	FE CP	3.002
NARITA, K.	CY BY	1.001
	IS WPA	8.305
NAUGHTON, M. A.	IS HOA	8.303
	IS WPA	8.305
	IS BOB	8.321
NEEDLEMAN, S. B.	CY CH	1.002
	CY PG	1.005
NEURATH, H.	TR BOTR	7.002
NICOL, D. S. H.	IS WPA	8.305
	IS BOB	8.321
NORTH, A. C. T.	LS CH	7.201
NOUVEL, G.	TI BOPA	5.001
PALEUS, S.	CY PG	1.005
	CY RR	1.010
	CY SM	1.011
	GL HUH	2.020
	LS CH	7.201
	PR BOHY	8.104
		10.000
PAULING, L.	PR BOAR	8.101
PHILLIPS, D. C.	PR PGLS	8.102
PEART, W. S.	TI BOPA	5.001
PORTER, R. R.	TR BOCH	7.001
POPENOE, E. A.		10.000
POSPISILOVA, D.	TN SBAC	8.207
PRUSIK, Z.	TI BOPA	5.001
PUTNAM, F. W.	PR BOOX	8.103
RAACK, I. D.	GL HUH	2.020
RADICEVIC, M.	TN PGBM	8.203
RESSLER, C.	GL HUHA	2.001
RHINESMITH, H. W.	GL HUHB	2.002
ROOS, P.	IS BOA	8.301
RUDLOFF, V.	IS BOB	8.321
	IS WHA	8.306
RYLE, A. P.	IS BOB	8.321
	IS WPA	8.305
SAITO, T.	IS BOA	8.301
	IS HOA	8.303
SAKAKI, S.	IS SHA	8.304
SANGER, F.	IS WPA	8.305
	IS BOB	8.321
	LS CH	7.201
SARMA, V. R.	GL HUH	2.020
SCHNEIDER, R. G.	TM TM	6.001
SCHRAMM, G.	GL HUHA	2.001
SCHROEDER, W. A.	GL HUHG	2.003
	GL HUH	2.020
	GL HUH	2.020
SCHWARTZ, H. C.		

SHELTON, J. B.  
 SHELTON, J. R.  
 SHEPHERD, R. G.  
 SINN, L. G.  
 SMILLIE, L. B.  
 SMITH, D. B.  
 SMITH, E. L.  
  
 SMITH, L. F.  
  
 SMYTH, D. G.  
 SORM, F.  
  
 STEIN, W. H.  
 STEWART, J. W.  
  
 SWENSON, R. T.  
 TANAKA, M.  
 THOMPSON, E. O. P.  
 TITANI, K.  
  
 TRIPPETT, S.  
 TSUGITA, A.  
 TUPPY, H.  
  
  
  
 UHLIG, H.  
 WALSH, K.  
 WATSON-WILLIAMS, E. J.  
 WEBER, E.  
 WHITE, W. F.  
 WITTMANN, H. G.  
 WITTMANN-LIEBOLD, B.  
  
  
 YAOI, Y.  
 YASUNOBU, K. T.  
  
 ZUCKERKANDL, E.

GL HUHA 2.001  
 GL HUHG 2.003  
 GL HUHA 2.001  
 GL HUHG 2.003  
 TN PGAC 8.206  
 GN BO 8.001  
 TR BOCH 7.001  
 GL HOHB 2.005  
 CY HO 1.003  
 CY HU 1.004  
 CY RS 1.009  
 PA PA 7.101  
 IS BOA 8.301  
 IS WPA 8.305  
 IS BOB 8.321  
 RN BO 4.001  
 TI BOPA 5.001  
 TR BOCH 7.001  
 RN BO 4.001  
 CY CH 1.002  
 CY PG 1.005  
 GL HUH 2.020  
 FE CP 3.002  
 IS BOA 8.301  
 CY BY 1.001  
 IS WPA 8.305  
 PR BOOX 8.103  
 TM TM 6.001  
 CY HO 1.003  
 CY PG 1.005  
 CY SW 1.008  
 CY RR 1.010  
 CY SM 1.011  
 PR BOOX 8.103  
 IS BOB 8.321  
 TM TM 6.001  
 TR BOTR 7.002  
 GL HUH 2.020  
 TM TM 6.001  
 TN PGAC 8.206  
 TM TMD 6.002  
 GL HUHA 2.001  
 GL HUHB 2.002  
 TM TMD 6.002  
 CY BY 1.001  
 CY PG 1.005  
 FE CP 3.002  
 GL GOHB 2.004