

MARCH 1974

VOLUME 71

NUMBER 3

LIBRARY

APR 17 1974

UNIVERSITY OF
WASHINGTON



SCIENCE

PROCEEDINGS OF THE

National Academy of Sciences

S R R Apr 27 '74

OF THE UNITED STATES OF AMERICA

THE PROCEEDINGS OF THE
**National
Academy of
Sciences**
OF THE UNITED STATES OF AMERICA

*Officers
of the
Academy*

PHILIP HANDLER *President*
SAUNDERS MAC LANE *Vice President*
ALLEN V. ASTIN *Home Secretary*
HARRISON BROWN *Foreign Secretary*
E. R. PIORE *Treasurer*

*Editorial Board
of the
Proceedings*

ROBERT L. SINSHEIMER *Chairman*
ROBERT M. SOLOW *Vice Chairman*
MICHAEL KASHA *Vice Chairman*
ALLEN V. ASTIN *Home Secretary*
HARRISON BROWN *Foreign Secretary*
E. R. PIORE *Treasurer*

C. B. ANFINSEN	MARTIN D. KAMEN
ALEXANDER G. BEARN	HENRY S. KAPLAN
P. D. BOYER	SEYMOUR S. KEY
BERNARD D. DAVIS	MACLYN McCARTY
KINGSLEY DAVIS	EUGENE P. ODUM
HARRY EAGLE	ALEXANDER RICH
HERMAN EISEN	PAUL A. SAMUELSON
MARK KAC	

Managing Editor: PATRICIA ZEIS THOMAS
Senior Associate Managing Editor: MURRIE W. BURGAN
Assistant Managing Editor: GARY T. COCKS

Editorial correspondence: PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, 2101 Constitution Avenue, Washington, D.C. 20418.

Business correspondence: Printing and Publishing Office, National Academy of Sciences, 2101 Constitution Avenue, Washington, D.C. 20418.

Subscriptions are entered on a calendar-year basis only. The current subscription rate is \$35 (effective January 1, 1975, the subscription rate will be \$45). The price of a single issue is \$4. Volumes 1 through 61 are available from Walter J. Johnson, Inc., 111 Fifth Avenue, New York, N. Y. 10003; prices are available from that organization upon request.

Subscribers are requested to notify the Printing and Publishing Office of any change of address; also the local postmaster. The Academy is not responsible for nonreceipt of issues because of improper address unless a change of address is on file. The notice of address change should list both the old and new addresses.

Claims for replacement copies will not be honored more than 60 days after the mailing date for domestic subscribers and not more than 90 days after the mailing date for foreign subscribers.

Microfilms of complete volumes are available to regular subscribers only and may be obtained from University Microfilms, Xerox Corporation, Ann Arbor, Michigan 48103.

Second-class postage paid at Washington, D.C. and at additional mailing offices

PRINTED IN THE USA

THE PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES USA is published monthly by

THE NATIONAL ACADEMY OF SCIENCES, 2101 Constitution Avenue, Washington, D.C. 20418

Variable Region Sequences of Five Human Immunoglobulin Heavy Chains of the V_H III Subgroup: Definitive Identification of Four Heavy Chain Hypervariable Regions

(myeloma proteins/amino acid sequences/antibody combining site)

J. DONALD CAPRA AND J. MICHAEL KEHOE

Department of Microbiology, Mount Sinai School of Medicine of the City University of New York, 10 East 102 Street, New York, N.Y. 10029

Communicated by Henry G. Kunkel, November 5, 1973

ABSTRACT The variable regions of five human immunoglobulin heavy chains of the V_H III subgroup have been totally sequenced. Three of the heavy chains belonged to the IgG class and two to the IgA class. Examination of these sequences, and comparison with additional published heavy chain sequences, showed that a total of four hypervariable regions is characteristic of human heavy chain variable regions.

The relatively conserved character of large segments of the heavy chain variable region was very evident in these studies. The conserved segments, which are those sections located outside the hypervariable regions, comprise approximately 65% of the total heavy chain variable region. The following general structural pattern for antibody molecules emerges from this and related studies: an overall combining region superstructure is provided by the more conserved segments while the refinements of the active site specificity are a function of hypervariable regions.

The antibody combining site is now believed to reside exclusively in the variable regions of the heavy and light polypeptide chains of the immunoglobulin molecule. Evidence is accumulating from several laboratories which indicates that hypervariable regions within the variable region are directly involved in the antibody combining site as well as being responsible, at least in part, for the idiotypic determinants of myeloma proteins and specific antibodies (1-5).

The existence of three hypervariable regions in the variable region of human immunoglobulin heavy chains has been established by previous studies from this laboratory. Residues 31-37 were described as the first hypervariable region of the heavy chain (6), and, after fragmentation of IgG heavy chains with cyanogen bromide, two additional hypervariable regions were localized between residues 86-91 and 101-110 (7).

We have now completed the amino acid sequence from residues 41 to 84 of the three V_H III proteins originally reported (6, 7) as well as the complete V region sequence of two IgA proteins with V_H III variable regions. The data make apparent an additional area of sequence hypervariability between residues 51 and 68, thus supporting the observations of Cebra *et al.* made on pooled guinea-pig immunoglobulins (5). When these data on V_H III proteins are included with that available for proteins of the V_H I and V_H II subgroups and analyzed by the method of Wu and Kabat (3), four distinct areas of sequence hypervariability are observed.

MATERIALS AND METHODS

kappa), and Tur (IgA1 kappa) were isolated from serum or plasma by zone electrophoresis on polyvinyl copolymer ("Pevicon") (8). After further purification by gel filtration chromatography, they were reduced with 0.1 M β -mercaptoethanol and alkylated with iodoacetamide. The heavy and light chains were separated by gel filtration in propionic acid (9, 10).

Fragment Preparation: Heavy chains were treated with cyanogen bromide (11) and the resulting individual fragments purified by gel filtration chromatography in 5 M guanidine-HCl. Three proteins (Tei, Zap, and Tur) yielded a large N-terminal fragment comprising residues 1-85. Proteins Was and Jon, which contain a methionine residue at position 34, gave fragments comprising residues 1-34 and 35-85. Since all human IgG myelomas have a methionine at position 252, Tei and Was yielded a large fragment comprising residues 86-252. In protein Jon, however, an additional methionine was present at position 111. Consequently two distinct fragments comprising residues 86-111 and 112-252 were obtained from this protein. IgA1 proteins contain a methionine at residue 426 (12) so proteins Zap and Tur both yielded a very large fragment composed of residues 86-426.

Sequencing Procedure. Positions 1-85: On both the intact heavy chain as well as on the 1-85 fragment, proteins Tei, Zap, and Tur were sequenced 60 steps on the automated sequencer (13, 14). Tryptic peptides were prepared and separated on Dowex 50X4 with a pyridine-formate buffer system. In proteins Tei and Zap two invariant peptides were aligned by homology alone (70-74 and 75-78) while in protein Tur,

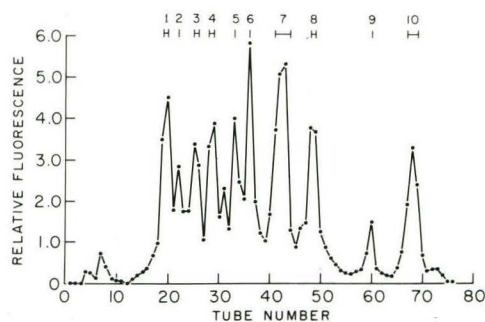


FIG. 1. Representative ion exchange chromatogram of tryptic hydrolysate of the amino terminal (1-85) cyanogen bromide fragment of the heavy chain variable region of human immunoglobulin heavy chains. The peptides corresponding to the amino acid sequence H I H H H I I H H I H are indicated above the peaks.

	10	20	
Tei	GLU VAL GLN LEU VAL GLU SER GLY GLY GLY LEU VAL GLN PRO GLY GLY SER LEU ARG LEU SER CYS ALA ALA SER		
Was	_____ LEU _____		
Jon	ASP _____ LYS _____		
Zap	_____ ALA _____ GLY _____		
Tur	_____ LEU _____		
	30	40	50
Tei	GLY PHE THR PHE SER THR SER ALA VAL TYR []	TRP VAL ARG GLN ALA PRO GLY LYS GLY LEU GLU TRP VAL	
Was	_____ SER _____ ASP _____ MET _____ []		
Jon	_____ ALA TRP MET LYS []		
Zap	_____ THR SER ARG PHE []		
Tur	_____ ARG VAL LEU SER SER []		
	60	70	
Tei	GLY TRP ARG TYR GLU GLY SER SER LEU THR HIS TYR ALA VAL SER VAL GLN GLY ARG PHE THR ILE SER ARG ASN		
Was	ALA _____ LYS _____ GLN GLU ALA _____ ASN SER _____ PHE _____ ASP THR _____ ASN _____		
Jon	VAL _____ VAL _____ GLN VAL VAL GLU LYS ALA PHE _____ ASN _____ ASN _____		
Zap	GLU PHE _____ VAL GLN _____ ALA ILE SER _____ ASP _____ ALA _____		
Tur	SER GLY _____ LEU ASN ALA _____ ASN LEU _____ PHE _____ ALA _____		
	80	90	100
Tei	ASP SER LYS ASN THR LEU TYR LEU GLN MET LEU SER LEU GLU PRO GLX ASX THR ALA VAL TYR TYR CYS ALA ARG		
Was	_____ _____ ASN ARG _____ ALA _____		
Jon	_____ ILE _____ VAL THR _____		
Zap	_____ ASN THR GLY _____ ALA _____		
Tur	_____ _____ GLN ALA _____ LEU _____		
	110	120	
Tei	VAL THR PRO ALA ALA ALA SER LEU THR PHE SER ALA VAL TRP GLY GLN GLY THR LEU VAL THR		
Was	PHE ARG GLN PRO PHE VAL GLN [] _____ PHE ASP _____ PHE _____		
Jon	_____ VAL VAL SER THR [] SER MET ASP _____ PRO _____		
Zap	THR ARG _____ GLY GLY TYR [] _____ ASP _____ SER		
Tur	LEU SER VAL THR _____ VAL [] ALA PHE ASP _____ LYS _____ SER		

FIG. 2. The amino-acid sequence of the variable regions of five human immunoglobulin heavy chains.

the isolation of chymotryptic peptides established the sequence unambiguously. In all cases, tryptic peptides were sequenced in the automated sequencer, often using 4-sulfo-phenylisothiocyanate (Pierce Chemical) on the lysine peptides (15). In proteins Was and Jon, which contained cyanogen bromide fragments 1-34 and 35-85, the first 60 residues were established by automated sequencing of the intact heavy chain. Thus, in these two proteins, sequencing cyanogen bromide fragment 1-34

Fragment 35-85 of proteins Was and Jon was sequenced 35 and 40 residues respectively; this, together with the C-terminal tryptic peptides mentioned above gave the complete sequence for this section. *Residues 86-121:* In proteins Tei, Was, Zap, and Tur the sequence was established by a continuous automated run of 45 steps from residue 86 into the C_{H1} domain. In both Zap and Tur, tryptic digestion and isolation of the resulting peptides confirmed a few question-

digestion. This sequence was obtained by difference since residues 1-34 were known from the initial study of the intact heavy chain. Ion fragment 112-253 was subjected to a long sequencer run which definitely established the sequence of residues 112-121 as well as providing sequence data into the C_H1 domain.

Ion Exchange Chromatography. An example of a Dowex 50×4 chromatogram is shown in Fig. 1 for a tryptic digest of the Tur 1-85 fragment; 6.5-ml fractions were collected and 0.1 ml of each fraction analyzed by the fluorescamine procedure initially described by Udenfriend *et al.* (16). Ninhydrin analysis was also performed after alkaline digestion of 0.5-ml aliquots. In most analyses, only the fluorescamine procedure was employed since it was much more sensitive. As shown in Fig. 1, 10 fractions were pooled. Each was subjected to amino-acid analysis and several useful peptides were isolated and sequenced. T-1 (Asn Thr Leu Tyr Leu Gln Hsr) (79-85), T-3 (Asn Asp Ser Lys) (75-78), T-7 (Gly Leu Gly Trp Val Ser Gly Arg) (46-53), and T-10 (Phe Thr Ile Ser Arg) (70-74).

RESULTS AND DISCUSSION

The amino-acid sequences of the variable regions of the five human myeloma proteins is displayed in Fig. 2. The variability-factor values determined by the method of Wu and Kabat (3) for these as well as all the other human V region sequences available is shown in Fig. 3. These calculations were based on 25 sequences from residues 1 to 34, 11 sequences from residues 35 to 85, and 14 sequences from residues 86 to 122. Previous to this study there were only six published complete V region sequences, all but one (Nie) of the V_HI (Eu) or V_HII (Daw, Cor, He, Ou) subgroup (for references see legend to Fig. 3). With five additional V_HIII sequences the variability within and between subgroups can now be compared more meaningfully. In addition, with the availability of 11 complete sequences and several fragments, the Wu-Kabat plot becomes more statistically significant.

A discussion of the sequences can be conveniently divided into those sections of the V region which are relatively constant (1-30, 38-50, 69-83, 92-100, and 111-121), and the hypervariable regions (31-37, 51-68, 84-91, and 101-110).

About 65% of the variable region of the heavy chain shows limited variation. In fact, there are 17 positions (14%) which have been absolutely invariant in all human heavy chains regardless of their V region subgroup assignment. Certain positions are subgroup specific since at these positions all, or nearly all, of the members of one subgroup have a particular amino acid, while members of the other subgroup contain a different amino acid. Utilizing the four available V_HII proteins, positions 3, 9, 16, 17, 19, 21, 23, 28, 29, 39, 42, 46, 50, 80, 81 and 82 appear to be subgroup specific. As noted previously, no subgroup specific residues are identifiable in the C terminal portion of the V region (7). There are thus 33 positions (27%) in the V region which are either invariant or subgroup specific. A comparison with the published sequences of myeloma proteins (17, 18), pooled immunoglobulins (5, 19, 20), and specifically purified antibodies (5, 21-23) from lower species, indicates that the particular amino acids found at these positions are characteristic of a wide variety of mammals and have been faithfully conserved during evolution. Such residues may have extremely im-

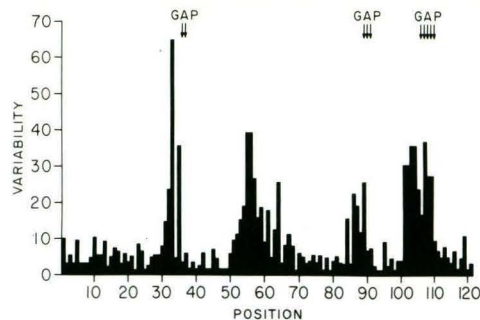


Fig. 3. Variability-factor values for the sequences shown in Fig. 1 as well as several other published sequences (36) determined according to the method of Wu and Kabat (3).

example, the provision of a distinct backbone structure which is crucial to antibody function.

As can be seen on inspection of Figs. 2 and 3, about a third of the heavy chain variable region can be considered "hypervariable." These regions deserve special consideration because of their specific implications for the formation of the antibody combining site, the nature of idiotypic determinants, and various theoretical conceptions of the origin of antibody diversity.

In light chains, affinity labels have been localized near or within hypervariable regions (23-25), thus providing direct support for the general concept that hypervariable regions participate directly in the antibody-combining site. For the heavy chain, recent work has also been consistent with this idea. For example, Ray and Cebra localized affinity labels to the first (31-37) and the fourth (101-110) heavy chain hypervariable regions (26), Haimovich *et al.* (27) localized an affinity label to residue 54 of the mouse myeloma protein 315 (which has anti-dinitrophenol activity), and Press and coworkers have localized affinity labels at or near the fourth hypervariable region in rabbit antibodies (28). Therefore, although the primary structure and affinity labeling studies of these proteins was being carried out independently, and even in different laboratories in many instances, there is a general implication from the experimental observations that the same regions of the molecule which show the highest degree of sequence variation are near or part of those particular regions of the heavy chain where affinity labels have been localized.

A second piece of evidence linking the antibody combining site to the hypervariable regions has come from comparisons of sequences obtained from pooled immunoglobulin heavy chains with those of specifically purified antibody heavy chains. Sequence analyses of rabbit (29), guinea pig (5), and other mammalian heavy chain pools (19), indicate that a definitive sequence cannot be obtained within those regions which have been identified as hypervariable on the basis of studies with myeloma proteins. However, when specifically purified antibodies are studied, a single major sequence can be determined, as has been shown most definitively by Cebra and his coworkers (5).

Additional support for the functional significance of hypervariable regions has been provided by current notions concerning the tertiary structure of the immunoglobulin molecule. Crystallographic analysis of human immunoglobulins has now advanced to the point where it has been possible

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts



Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research



With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips



Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.