

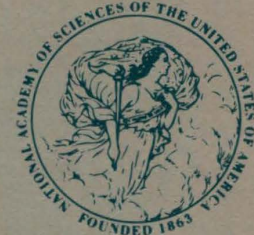
LIBRARY

DEI: 1974 NOVEMBER 1974

UNIVERSITY OF  
WASHINGTON

VOLUME 71

NUMBER 11



PROCEEDINGS OF THE

# National Academy of Sciences

OF THE UNITED STATES OF AMERICA

SCIENCE ROOM



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*  
GEORGE S. HAMMOND *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*  
GEORGE S. HAMMOND *Foreign Secretary*  
E. R. PIORE *Treasurer*

C. B. ANFINSEN	MARTIN D. KAMEN
ALEXANDER G. BEARN	HENRY S. KAPLAN
P. D. BOYER	SEYMOUR S. KETY
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:* Circulation Office of the PROCEEDINGS, 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 (effective January 1, 1975, the price of a single issue will be \$5). Volumes 1 through 61 are available from Walter J. Johnson, Inc., 355 Chestnut Street, Norwood, N.J. 07648; prices are available from that organization upon request.

*Subscribers* are requested to notify the Circulation Office of the PROCEEDINGS 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



## The Three-Dimensional Structure of a Phosphorylcholine-Binding Mouse Immunoglobulin Fab and the Nature of the Antigen Binding Site

(x-ray diffraction/domain structure/hypervariable cavity/hapten binding/antibody diversity)

DAVID M. SEGAL\*†, EDUARDO A. PADLAN\*, GERSON H. COHEN\*, STUART RUDIHOFF‡, MICHAEL POTTER‡, AND DAVID R. DAVIES\*

\* Laboratory of Molecular Biology, NIAMDD; † Immunology Branch, NCI; ‡ Laboratory of Cell Biology, NCI; National Institutes of Health, Bethesda, Maryland 20014

Communicated by Elvin A. Kabat, August 12, 1974

**ABSTRACT** The structure of the Fab of McPC 603, a mouse myeloma protein with phosphorylcholine binding activity, has been determined to 3.1-Å resolution. The four domains are found to be structurally similar with a well-defined double-layer structure. A large cavity exists at one end of the fragment, the walls of which are formed exclusively of hypervariable residues. Phosphorylcholine binds in this cavity and forms specific interactions with several well-defined amino-acid side chains of the protein. The hapten is bound asymmetrically and interacts more with the heavy chain than with the light chain.

We have been investigating the nature of antibody-antigen interactions by means of a crystallographic investigation of the Fab fragments of mouse myeloma proteins with known binding specificity (1, 2). The testing of myeloma proteins in mice for antigen-binding activity has revealed that many of them interact by precipitation, agglutination, or complement-fixation with common natural antigens from the environment of the mouse (3, 4). Immunochemical analysis of myeloma protein-antigen interactions has in many cases led to the identification of the chemical group (hapten) on the antigen molecule.

Antibodies, induced to dextran, levan, and phosphorylcholine in mice, share cross-specific antigenic (idiotypic) determinants with myeloma proteins of corresponding specificity (5, 6; M. Potter and R. Lieberman, unpublished observations). These findings strongly suggest the close relationship of antigen-binding myeloma proteins in mice to natural or induced antibodies. One of these myeloma proteins, from McPC 603, precipitates with antigens from several pathogenic organisms, and has been demonstrated to bind specifically to phosphorylcholine (3). We have previously reported crystallization of the Fab fragment of McPC 603 protein (1) and its structure at 4.5-Å resolution together with the location of its phosphorylcholine-binding site (2). Here we present the results of a 3.1-Å structure determination and the molecular details of phosphorylcholine binding. We compare these results with the crystallographic results on a human Fab fragment with hydroxyvitamin K<sub>1</sub> binding properties (7, 8), and with two human Bence-Jones proteins (9, 10).

### MATERIALS AND METHODS

The preparation and crystallization of McPC 603 pepsin Fab has been described (1). The crystals used were obtained

Abbreviations: C<sub>H1</sub> and V<sub>H</sub>; C<sub>L</sub> and V<sub>L</sub> signify the constant and variable domains of the heavy (H) and light (L) chains, respectively, of the Fab. H1, H2, and H3; L1, L2, and L3 are the first, second, and third hypervariable regions of the H and L chains, respectively.

from 42% saturated ammonium sulfate solutions, 0.1 M imidazole (pH 7.0). They belong to the space group P6<sub>3</sub> with  $a = b = 162.5$ ,  $c = 60.8$  Å. Heavy atom derivatives were prepared by soaking the crystals in solutions containing TmCl<sub>3</sub>, K<sub>2</sub>Pt(CNS)<sub>6</sub>, iodine, and combinations of these (E. A. Padlan, D. M. Segal, G. H. Cohen, T. Spande, and D. R. Davies, in preparation). Approximately 11,000 reflections (mean figure of merit 0.73) were used to calculate a 3.1-Å electron density map. Details of the crystallographic analysis are being published elsewhere. A Kendrew skeletal model of the entire Fab has been constructed with the aid of the amino-acid sequence, which is for the most part known. An optical comparator of the type proposed by Richards (11), but with the mirror parallel to the electron density sections, was used to facilitate the model building. Coordinates have been measured for the variable region of the molecule and improved by the Diamond model-building program (12). Hapten binding has been studied by difference-Fourier techniques using data from crystals of McPC 603 Fab soaked in saturating concentrations of phosphorylcholine as described (2), but here extended to 3.1-Å resolution.

The amino-acid sequence of the variable portion of the heavy chain (V<sub>H</sub>) is completely known (Table 1). Approximately half the C<sub>H1</sub> region sequence has been determined (Rudikoff, unpublished), and a substantial part of the remainder may be surmised from its expected homology with MOPC 315 (13). The V<sub>L</sub> sequence is mostly known and its determination is being actively pursued. The constant portion of the light chain (C<sub>L</sub>) is assumed to be identical with the sequence of the C region of other murine K light chains (14, 15). Light chain hypervariable regions were assigned according to Wu and Kabat (16). In the heavy chain these regions were identified by comparison with other heavy chain sequences, with deletions and insertions introduced to maximize homology (17).

### RESULTS

The electron density map is clear enough to permit tracing the course of the polypeptide chain with confidence. While many of the smaller side groups are lost in the background noise, particularly those on the exterior of the molecule, the larger aromatic side chains are for the most part clearly visible, and serve as markers in fitting the amino-acid sequence to the map. Although the four disulfide bridges were located, they do not exhibit the highest peak density in each domain.

*Overall Structure.* The overall appearance of the IgA(K) 603 Fab is very similar to that previously observed in the human



TABLE 1. Variable region sequence of McPC 603\*

1	10	20	30	40	50	58a 58b
V <sub>H</sub> : E-V-K-L-V-E-S-G-G-L-V-Q-P-G-G-S-L-R-L-S-C-A-T-S-G-F-T-F-S-B-F-Y-M-E-W-V-R-Q-P-P-G-K-R-L-E-W-I-A-A-S-R-B-K-G-B-K-Y-T--T-						
60	70	80	90	100	a b	110
Z-Y-S-A-S-V-K-G-R-F-I-V-S-R-B-T-S-Z-S-I-L-Y-L-Q-M-N-A-L-R-A-E-D-T-A-I-Y-Y-C-A-R-N-Y-Y-G-S-T-W-Y-F-D-V-W-G-A-G-T-T-V-T-V						
1	10	20	27 27a 27b 27c 27d 27e 27f	30	40	
V <sub>L</sub> †: D-I-V-M-T-Q-S-P-S-S-L-S-V-S-A-G-E-R-V-T-M-S-C-K-S-S-Z--S--L--L--B--S--G--B-Z-K-B-F-L-A-W-Y-Z-(Z)-K-P-G-Z-P-P-K-L-I-Y						

\* Heavy chain numbering is that previously reported (17), with gaps and insertions introduced to maximize homologies with other heavy chain sequences. Light chain numbering is according to Wu and Kabat (16).

† The first 35 residues of this light chain sequence have previously been reported (25).

IgG( $\lambda$ ) New Fab (7) and in the human Meg Bence-Jones dimer (9). The fragment has the overall dimensions  $40 \times 50 \times 75 \text{ \AA}$ . It consists of two globular regions, each approximately  $40 \times 40 \times 50 \text{ \AA}$ . Each globular region in turn is made up of two homologous domains; V<sub>H</sub>:V<sub>L</sub> in the variable region, and C<sub>H1</sub>:C<sub>L</sub> in the constant region. These pairs of domains are related to one another by an approximate 2-fold axis of symmetry. The pseudo 2-fold axes from the two regions are not colinear, but make an angle of approximately  $135^\circ$ .

**Domain Structure.** In both V and C regions the basic domain structure consists of straight segments of polypeptide chain arranged in two layers and forming a sandwich-like structure. The chains in each layer have the frequently observed left-handed twist.

Fig. 1 is a schematic representation of the folding for the variable and constant domains. Adjacent segments within each layer are antiparallel to each other and frequently assume a  $\beta$ -pleated sheet configuration. The interior of the sandwich contains principally hydrophobic residues. The intradomain disulfide and the invariant tryptophan are in close proximity and are located approximately in the middle of the domain. The extended segments are linked by bends with varying degrees of sharpness. The electron density in many of the tight bends can be fitted with  $\beta$ -bend configurations (18).

We have adopted the numbering scheme shown in Fig. 1, where we have denoted the stretches as S1, S2, S3, ... S9 and bends as B12, B23, B34, ... B89 with the bend numbers representing the stretches connected. Thus, we can think of the variable domains as consisting of nine stretches with eight bends and the constant domains as having seven stretches with six bends. Segments S4 and S5 and bend B45 are missing in the constant domains; this additional loop in the variable domains contains the second hypervariable regions of the light and heavy chains. Segments S1, S2, S6, and S7 form one layer of the sandwich and segments S3, S8, and S9 form the other. In the variable domains the additional loop occurs at the edge of the sandwich and is located approximately in the three-segment layer. The extended segments in the constant domains, are on the average, slightly longer than the homologous segments in the variable domains.

**Quaternary Structure.** It was clear even at low resolution that the two domains of each chain were oriented approximately at right angles to each other. In addition, the four domains of the fragment can be regarded as being located on the vertices of an elongated but distorted tetrahedron (2). The long axes of the domains of the light chain make an angle of approximately  $100^\circ$  with each other, while the corresponding angle for the heavy chain is approximately  $80^\circ$ . This results in a center to center distance of about  $40 \text{ \AA}$  between the intra-

domain disulfides of V<sub>L</sub> and C<sub>L</sub> compared with a distance of about  $30 \text{ \AA}$  between V<sub>H</sub> and C<sub>H1</sub>. Similar results have also been observed in the human Fab New (7) and in the human Meg Bence-Jones dimer (9). A possible explanation of this distortion of the symmetry of the McPC 603 fragment may lie in the more extensive interaction between the constant and variable domains of the heavy chain when compared with those of the light chain.

The close approach of the two heavy domains (V<sub>H</sub>, C<sub>H1</sub>) appears to be facilitated by the presence at the interface of amino acids with small side chains. For instance, involved in this contact are the stretch Gly-Gly-Gly (residues 8–10) in S1 and the stretch Gly-Ala-Gly (residues 109–111) in S9. In addition

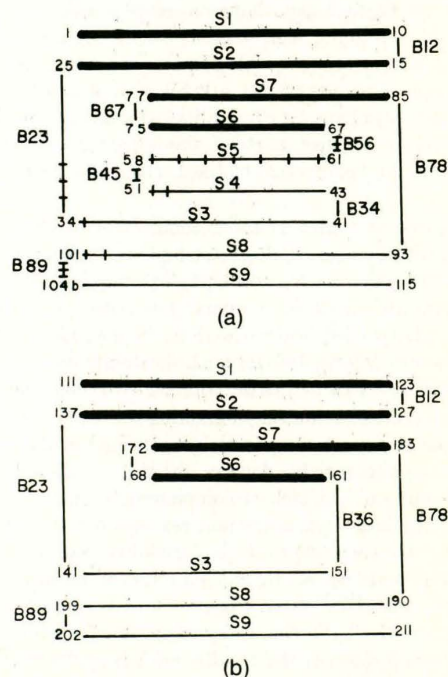


FIG. 1. (a) A schematic representation of the heavy chain variable domain. The relatively extended segments are denoted by S and are numbered sequentially from the NH<sub>2</sub> terminus. The bends B are numbered according to the segments they connect. The hypervariable regions (31–35, 49–66, 99–104b) are represented by the perpendicular bars. The segments drawn with thick lines belong to one layer of the sandwich structure, and the thin segments belong to the other layer. (b) A comparable schematic drawing of the C<sub>L</sub> domain. Note the absence of the loop S4, B45, S5, which is only found in the variable domains.



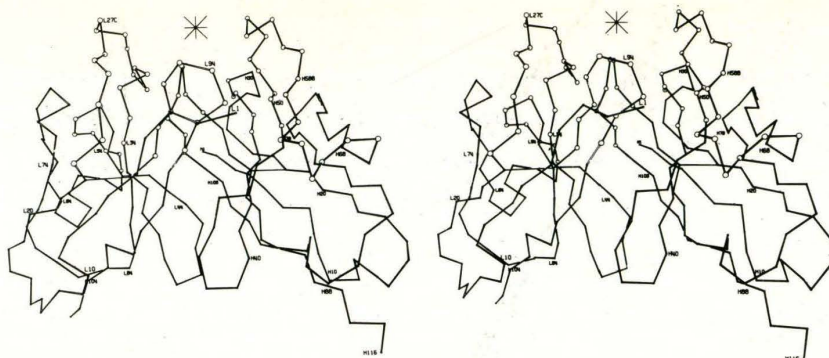


FIG. 2. Stereo drawing of the  $\alpha$ -carbon skeleton of the variable region. Open lines represent the light chain and solid lines the heavy chain. The hypervariable residues are represented by the large open circles. The star marks the location of the large hypervariable cavity.

tion, this heavy chain interdomain interaction involves no charged amino-acid side chains.

It is apparent that the dominant force maintaining the structural integrity of the McPC 603 Fab is due to the strong lateral interactions between homologous domains ( $V_H$ - $V_L$ ;  $C_{H1}$ - $C_L$ ) rather than the much weaker longitudinal interactions between domains of the same chain. The interactions between the two constant domains occurs principally between the layers containing segments S1, S2, S7, and S6. The interaction between the variable domains is quite different, and involves the segments S3, S8, and S9, as well as a major portion of the extra loops S4-S5. Additional intervariable domain contacts occur between parts of the hypervariable loops, especially in the regions B23 of the L chain and B89 of the H chain.

The constant region is more compact than the variable region, probably as a result of having fewer bulky residues in the interface between  $C_L$  and  $C_{H1}$ . The interacting layers in this region are about 10 Å apart; this results in a distance between the two S-S bridges of 18 Å. In the absence of complete sequence information for  $C_{H1}$ , the interactions across the C region interface cannot be determined with confidence.

In the variable region the interacting segments are 12-15 Å apart, and this greater separation is reflected in the distance between the intervariable domain disulfides of approximately 25 Å. In spite of our lack of complete light chain sequence, many of the interdomain contact residues can be identified. They span the complete range of variability. Some, for example, Leu 45 and Trp 47, have been found in all heavy chain sequences, and Gln 39 has been found in most (17, 19). Similarly in the light chain, Tyr 36, which makes contact across the interface, is the residue most frequently found in this position in kappa chains (19). In addition, many of the interacting residues are found in the hypervariable regions. Indeed, the large loop produced by the six-residue insertion in L1 is in intimate contact with most of the residues in H3. Moreover, some residues in L3 interact with residues in both hypervariable and nonhypervariable portions of the  $V_H$  domain.

*The Antigen-Binding Site.* The 4.5-Å investigation revealed the location of the hapten-binding site (2). The appearance of two peaks on the difference electron density map led to the

suggestion that, in the native crystals, the phosphate binding site was occupied by a sulfate ion. The present native electron density map at 3.1-Å resolution shows a large peak, tetrahedral in appearance, and making contact with side chains, in the phosphate position. The choline site is revealed by a 3.1-Å difference map between native and phosphorylcholine-soaked crystals.

The site of hapten binding is in a large wedge-shaped cavity, which exists at the amino-terminal end of the molecule (Fig. 2). The cavity is approximately 12-Å deep, 15-Å wide at the mouth, and 20-Å long, the walls of which are lined exclusively with hypervariable residues (16). Only five of the hypervariable regions contribute to the formation of the cavity; these are the L1, L3, H1, H2, and H3. The portion homologous to the fourth hypervariable region (residues 82-89), reported in human heavy chains (20), occurs predominantly in the bend B78 at the COOH-terminal end of the variable region and is quite removed from the cavity. L2 likewise does not form a part of the walls of the cavity, being screened from it by the large loop containing L1. The insertions in L1, H2, and H3 cause these hypervariable loops to extend farther out, thus increasing the cavity depth.

The phosphorylcholine occupies only a small part of the cavity and is bound asymmetrically (Fig. 3), being closer to the H chain than to the L chain. The choline moiety is bound in the interior of the cavity with the phosphate group more towards the exterior. The phosphate group appears to be exclusively bound by the H chain; in particular, specific interactions are formed between the phosphate and Tyr 33(H) and Arg 52(H). The hydroxyl group of Tyr 33(H) is apparently hydrogen bonded to one of the oxygens of the phosphate, as is one of the amino groups of the arginine side chain. Moreover, the close proximity of the positively charged guanidinium group of Arg 52(H) to the negatively charged phosphate should produce a large favorable electrostatic interaction. There is also in the immediate vicinity of the phosphate another positive group, that of Lys 54(H), also of the heavy chain which could help neutralize the negative charge of this portion of the hapten. The choline group appears to interact with both the L and H chains. The acidic side chain of Glu 35(H) is about 5 Å away from the positively charged nitrogen of the choline. In addition, there are van der Waals interactions between the choline and main chain atoms

# 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.