COLD SPRING HARBOR SYMPOSIA ON QUANTITATIVE BIOLOGY

VOLUME XLI

Origins of Lymphocyte Diversity

COLD SPRING HARBOR LABORATORY

1977

COLD SPRING HARBOR SYMPOSIA ON QUANTITATIVE BIOLOGY
VOLUME XLI

International Standard Book Number 0-87969-040-2 (clothbound)
Library of Congress Catalog Card Number 34-8174

Printed in the United States of America

COLD SPRING HARBOR SYMPOSIA ON QUANTITATIVE BIOLOGY

Founded in 1933 by

REGINALD G. HARRIS

Director of the Biological Laboratory 1924 to 1936

Previous Symposia Volumes

I (1933) Surface Phenomena
II (1934) Aspects of Growth
III (1935) Photochemical Reactions
IV (1936) Excitation Phenomena
V (1937) Internal Secretions
VI (1938) Protein Chemistry
VII (1939) Biological Oxidations
VIII (1940) Permeability and the Nature of Cell Membranes
IX (1941) Genes and Chromosomes: Structure and Organization
X (1942) The Relation of Hormones to Development
XI (1946) Heredity and Variation in Microorganisms
XII (1947) Nucleic Acids and Nucleoproteins
XIII (1948) Biological Applications of Tracer Elements
XIV (1949) Amino Acids and Proteins
XV (1950) Origin and Evolution of Man
XVI (1951) Genes and Mutations
XVII (1952) The Neuron
XVIII (1953) Viruses
XIX (1954) The Mammalian Fetus: Physiological Aspects of Development
XX (1955) Population Genetics: The Nature and Causes of Genetic Variability in Population
XXI (1956) Genetic Mechanisms: Structure and Function
XXII (1957) Population Studies: Animal Ecology and Demography
XXIII (1958) Exchange of Genetic Material: Mechanism and Consequences
XXIV (1959) Genetics and Twentieth Century Darwinism
XXV (1960) Biological Clocks
XXVI (1961) Cellular Regulatory Mechanisms
XXVII (1962) Basic Mechanisms in Animal Virus Biology
XXVIII (1963) Synthesis and Structure of Macromolecules
XXIX (1964) Human Genetics
XXX (1965) Sensory Receptors
XXXI (1966) The Genetic Code
XXXII (1967) Antibodies
XXXIII (1968) Replication of DNA in Microorganisms
XXXIV (1969) The Mechanism of Protein Synthesis
XXXV (1970) Transcription of Genetic Material
XXXVI (1971) Structure and Function of Proteins at the Three-dimensional Level
XXXVII (1972) The Mechanism of Muscle Contraction
XXXVIII (1973) Chromosome Structure and Function
XXXIX (1974) Tumor Viruses
XL (1975) The Synapse

The Symposium Volumes are published by the Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, and may be purchased directly from the Laboratory or through booksellers. Price of Volume XLI—2-part set $60.00 (inc. postage). May be purchased only as a complete set. Price subject to change without notice.

Model-building Studies of Antigen-binding Sites: The Hapten-binding Site of MOPC-315

E. A. PADLAN, D. R. DAVIES, I. PECHT,* D. GIVOL* AND C. WRIGHT†

*Laboratory of Molecular Biology, National Institute of Arthritis, Metabolism and Digestive Diseases, National Institutes of Health, Bethesda, Maryland 20014; * Department of Chemical Immunology, The Weizmann Institute of Science, Rehovot, Israel; † Laboratory of Molecular Biophysics, Department of Zoology, Oxford University, Oxford, England*

The molecular basis for the structural diversity of antibody combining sites has become apparent through the recent X-ray diffraction studies on several immunoglobulin (Ig) fragments (see Davies et al. 1975a,b for a review). These structures reveal that the combining site is formed by bringing together the three hypervariable regions (Wu and Kabat 1970) of V_L and of V_H to form a continuous complementarity-providing surface. A quantitative comparison of the tertiary structures of a number of variable domains from both light and heavy chains has demonstrated that their nonhypervariable or framework regions are very similar, with the principal differences occurring in the hypervariable loops (Padlan and Davies 1975).

In view of this structural invariance, the immunoglobulin variable region can be regarded as consisting of a rigid framework to which are attached the hypervariable loops. These loops are not large, generally consisting of at most 17 residues; in addition, the configurations of a number of different hypervariable loop regions are already known from X-ray diffraction. Thus it might be possible to construct by model building new loop regions (and hence the Ig binding sites) from a knowledge of the sequences alone. Since it is clearly impractical to determine by X-ray diffraction the structures of all the interesting antibodies, this model-building method offers an attractive alternative approach when sequence data are available. We have examined the potential usefulness of this possibility by attempting to build the V regions of the rabbit anti-type-III polysaccharide antibody (Davies and Padlan 1976) and of protein MOPC-315, a mouse immunoglobulin with dinitrophenyl (DNP)-binding specificity (Eisen et al. 1968), based on the known three-dimensional structure of McPC-603 Fab′ (Padlan et al. 1973, 1974; Segal et al. 1974a,b) together with the amino acid sequences of these immunoglobulin Fab's.

In this paper, we describe the results of two independent attempts to construct the model of MOPC-315, one at the National Institutes of Health and the other at the Weizmann Institute. The two models were ultimately merged to give a final model that is presented and is discussed in terms of the known binding specificity, the kinetic mapping of the binding site (Pecht et al. 1972a,b; Haselkorn et al. 1974), and the results of affinity labeling with various reagents (Goetzl and Metzger 1970a,b; Givol et al. 1971; Haimovich et al. 1972). We discuss some of the limitations and the potential of this method of model building, the assumptions that have to be made, and some of the additional information that will be needed to put this tentative approach on a firmer footing.

MATERIALS AND METHODS

Sequence Alignment

The amino acid sequences of the V_H domains of proteins McPC-603 (Rudikoff and Potter 1974) and MOPC-315 (Francis et al. 1974; L. Hood, M. Margolies, D. Givol and R. Zakut, unpubl.) were aligned for maximum sequence homology and structural analogy, as presented in Table 1a. The sequence of New V_H (Poljak et al. 1974) is included for comparison. A similar alignment of the 603 (Segal et al. 1974a) and 315 V_L (Dugan et al. 1973) sequences is given in Table 1b. However, only the sequence of the first 49 residues is known for 603, and the residues given for the rest of the sequence are those which most frequently occur in mouse kappa chains (McKean et al. 1973). Included in Table 1b are the sequences of the V_L domains of REI (Epp et al. 1974), Mcg (Fett and Deutsch 1974), and New (Chen and Poljak 1974) proteins.

The sequence numbering schemes used in these tables are those obtained from the original references. Whenever a residue position is referred to in the text, however, the quoted number corresponds to the 315 sequence. The hypervariable regions of the light chain are: L1, residues 23–36; L2, residues 52–58; and L3, residues 91–99; and those of the heavy chain are: H1, residues 31–36; H2, residues 49–66; and H3, residues 99–104.

Model Building

Model 1 was constructed at the Weizmann Institute from Nicholson molecular models (Labquip, Reading, England), which have a scale of 1 cm = 1 Å, and CPK space-filling models (1.25 cm = 1 Å). Model 2 and the final model were constructed at the National Institutes of Health from Kendrew

(2 cm = 1 Å) molecular models (Repetition Engineers, Cambridge, England).

The general principle used for model building was to construct first the framework part of the variable region, based on the structure of 603 Fab'. The hypervariable loops were then constructed, leaving the structure as little changed as possible except when forced by amino acid insertions and deletions.

An attempt was made to maximize the structural stability within each loop by forming hydrogen bonds whenever possible and maintaining the phi and psi peptide angles within reasonable limits (Ramakrishnan and Ramachandran 1965). The interactions between loops were then maximized, leaving no large holes in the domain interior, while minimizing steric hindrance between groups.

The L1 regions in the kappa chains 603 and REI are simple loops, whereas the corresponding regions are helical in New (Poljak et al. 1974) and Mcg (Schiffer et al. 1973), which are both lambda chains. Since the light chain of protein 315 is of the lambda type and since this region has the same number of residues as New and Mcg with no gross differences in the nature of the amino acid side chains, it is most likely that this region in 315 will have a similar configuration to that found in New and Mcg. Accordingly, the L1 loop in 315 was built to conform as closely as possible to the corresponding region in the lambda chains with the aid of the atomic coordinates of the Mcg backbone kindly provided by M. Schiffer and coworkers (pers. comm.).

L2 was built by assigning to it the same backbone conformation as in protein 603. In this region, 603 and REI are not significantly different (Padlan and Davies 1975). Moreover, kappa and lambda light chains have the same number of residues in this part of the molecule (Dayhoff 1972); an exception is New (Table 1b), which has a seven-residue deletion in this part of the molecule. The L2 region in Mcg appears to have the same structure as 603 (Edmundson et al. 1974). Thus it is reasonable to assume that 315 L2 will have the same configuration as that of 603.

L3 was built essentially as in 603, the main difference being the possibility of forming more hydrogen bonds in 315. The L3 loop in 315 was built as two, fully antiparallel, segments, the loop in 603 being less regular. Although the sequence of 603 is not known in this region, the six mouse kappa chains sequenced so far (McKean et al. 1973) all have a proline at position 94. The possible occurrence of this residue in 603 could explain the fewer hydrogen bonds formed in 603 L3.

Protein 315 has one more residue than 603 in H1 (Table 1a). Originally, the sequence of 315 H1 contained a lysine in position 35 (Francis et al. 1974). Structurally, this residue is analogous to Met 34 of 603 since both occur two positions from the structure the domain interior. Its replacement by a lysine side chain would cause a significant rearrangement of the H1 backbone in order to expose the charged amino group to solvent. A reexamination of the sequence of this part of 315 led to the assignment of a Trp instead of the Lys in position 35 (L. Hood, M. Margolies, D. Givol and R. Zakut, unpubl.). A Trp side chain can easily be accommodated in the domain interior without significantly changing the configuration from that observed in 603.

The alignment shown in Table 1a leads us to conclude that the additional residue in H1 of 315 can be best accommodated in the exposed loop at the beginning of H1. The glycine at position 32 permits the construction of a sharp turn in this part of 315.

H2 has three more residues in 603 than in 315, whereas 315 and New are of the same length in this region. An initial model of 315 H2 was built by simply excising three residues from the end of the 603 H2 loop. Minor adjustments were then made to make the 315 H2 loop resemble as closely as possible the corresponding region in New (Poljak et al. 1974).

Basically the same procedure was employed in building the 315 H3 region, which has the same length as New and which is two residues shorter than 603. Here, however, the C-terminal segment of 315 H3 was made to approximate the corresponding region in 603, rather than in New. The difference between 603 and New, aside from the two-residue insertion in 603, is the configuration of the segment immediately preceding the phenylalanine at position 105, which almost always contains a large hydrophobic residue (Dayhoff 1972). In 603, the side group of Trp 104a, which is structurally analogous to Leu 103 of 315, projects into the hapten-binding cavity. On the other hand, the backbone of this segment in New appears to be different (Poljak et al. 1974). In view of the greater structural similarity of residues 103–105 of 315 (Leu-Tyr-Phe) to those of 603 (Trp-Tyr-Phe) rather than to those of New (Gly-Cys-Ile), the configuration of 603 was followed in building this segment of 315. However, it should be kept in mind that an alternative configuration might be that observed in New.

As soon as a tentative model of the binding site was completed, a likely site for DNP-binding was located using the criteria that the nitro groups of the DNP moiety must be hydrogen-bonded to the protein and that the stacking van der Waals interaction between the DNP ring and the side group of Trp 93 (L) must be maximal. The model was then adjusted to accommodate the DNP-hapten while maintaining the general stability of the various loop structures. Further adjustments were then made to ensure the feasibility of labeling Tyr 33 (L) and Lys 52 (H) by specific affinity reagents (Givol et al. 1971; Haimovich et al. 1972). The adjustments made to the tentative model to accommodate the

Table 1. Amino Acid Sequences of V_H Domains of Proteins McPC-603 and MOPC-315

(a) Alignment of V_H Sequences

		0	**2**	0	4	0	2	0	**4**	**0**	0
McPC-603	1	Glu	Val	Lys	Leu	Val	Glu	Ser	Gly	Gly	Gly
MOPC-315	1	Asp	Val	Gln	Leu	Gln	Glu	Ser	Gly	Pro	Gly
New	1	Pca	Val	Gln	Leu	Pro	Glu	Ser	Gly	Pro	Glu
		0	4	0	0	0	0	0	4	0	4
	11	Leu	Val	Gln	Pro	Gly	Gly	Ser	Leu	Arg	Leu
	11	Leu	Val	Lys	Pro	Ser	Gln	Ser	Leu	Ser	Leu
	11	Leu	Val	Ser	Pro	Gly	Glx	Thr	Leu	Ser	Leu
		0	4	0	4	0	1	4	0	4	0
	21	Ser	Cys	Ala	Thr	Ser	Gly	Phe	Thr	Phe	Ser
	21	Thr	Cys	Ser	Val	Thr	Gly	Tyr	Ser	Ile	Thr
	21	Thr	Cys	Thr	Gly	Ser	Thr	Val	Ser	Thr	Phe
		0		4	2	4	2	4	C	3	C
	31	Asp		Phe	Tyr	Met	Glu	Trp	Val	Arg	Gln
	31	Ser	Gly	Tyr	Phe	Trp	Asn	Trp	Ile	Arg	Gln
	31	Ala		Val	Tyr	Ile	Val	Trp	Val	Arg	Gln
		2	0	0	0	0	C	2	C	4	4
	40	Pro	Pro	Gly	Lys	Arg	Leu	Glu	Trp	Ile	Ala
	40	Phe	Pro	Gly	Asn	Lys	Leu	Glu	Trp	Leu	Gly
	40	Pro	Pro	Gly	Arg	Gly	Leu	Glu	Trp	Ile	Ala
		1	4	2	0	0	0	0	0	2	0
	50	Ala	Ser	Arg	Asn	Lys	Gly	Asn	Lys	Tyr	Thr
	50	Phe	Ile	Lys	Tyr	Asp	Gly	–	–	–	Ser
	50	Tyr	Val	Phe	Tyr	His	Gly	–	–	–	Thr
		2	0	3	1	0	0	4	0	0	1
	58b	Thr	Glu	Tyr	Ser	Ala	Ser	Val	Lys	Gly	Arg
	57	Asx	(Tyr , Gly)		Asx	Pro	Ser	Leu	Lys	Asn	Arg
	57	Ser	Asp	Thr	Asp	Thr	Pro	Leu	Arg	Ser	Arg
		4	0	4	0	2	0	0	0	0	3
	68	Phe	Ile	Val	Ser	Arg	Asp	Thr	Ser	Gln	Ser
	68	Val	Ser	Ile	Thr	Arg	Asp	Thr	Ser	Glu	Asn
	68	Val	Thr	Met	Leu	Val	Asn	Thr	Ser	Lys	Asn
		0	4	1	4	0	4	0	0	4	0
	78	Ile	Leu	Tyr	Leu	Gln	Met	Asn	Ala	Leu	Arg
	78	Gln	Phe	Phe	Leu	Lys	Leu	Asp	Ser	Val	Thr
	78	Gln	Phe	Ser	Leu	Arg	Leu	Ser	Ser	Val	Thr
		0	0	2	3	4	c	4	C	4	c
	88	Ala	Glu	Asp	Thr	Ala	Ile	Tyr	Tyr	Cys	Ala
	88	Thr	Glx	Asx	Thr	Ala	Thr	Tyr	Tyr	Cys	Ala
	88	Ala	Ala	Asp	Thr	Ala	Val	Tyr	Tyr	Cys	Ala
		3	3	2	C	0	0	C	C	C	C
	98	Arg	Asn	Tyr	Tyr	Gly	Ser	Thr	Trp	Tyr	Phe
	98	Gly	Asp	Asn	Asp	His	–	–	Leu	Tyr	Phe
	98	Arg	Asx	Leu	Ile	Ala	–	–	Gly	Cys	Ile
		1	4	C	4	0	4	2	0	4	0
	106	Asp	Val	Trp	Gly	Ala	Gly	Thr	Thr	Val	Thr
	106	Asp	Tyr	Trp	Gly	Gln	Gly	Thr	Thr	Leu	Thr
	106	Asx	Val	Trp	Gly	Gln	Gly	Ser	Leu	Val	Thr
		4	0	0							
	116	Val	Ser	Ser							
	116	Val	Ser	Ser							
	116	Val	Ser	Ser							

(a) Above each residue in the McPC-603 sequence is its structural location designated by: 0, completely exposed to solvent; 1, mainly exposed; 2, partly exposed, partly buried in the domain interior; 3, mainly buried; 4, completely buried; or C, in contact with the homologous domain. The numbers at the left alongside the sequences correspond to the sequence number of the first residue in each row as obtained from the original publica-

DOCKET ALARM

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.