PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

A Possible Basis for Major Histocompatibility Complex-Restricted T-Cell Recognition

M. M. Davis, Yueh-Hsiu Chien, Pamela J. Bjorkman, J. F. Elliott, M. Iwashima, E. P. Rock and P. A. Patten

Phil. Trans. R. Soc. Lond. B 1989 **323**, 521-524
doi: 10.1098/rstb.1989.0030

Phil. Trans. R. Soc. Lond. B **323**, 521–524 (1989)
Printed in Great Britain

A possible basis for major histocompatibility complex-restricted T-cell recognition

By M. M. Davis, Yueh-Hsiu Chien, Pamela J. Bjorkman, J. F. Elliott†, M. Iwashima, E. P. Rock and P. A. Patten

The Howard Hughes Medical Institute and Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, California 94305, *U.S.A.*

Four distinct T-cell antigen-receptor gene loci have now been identified and partly characterized: α, β, γ and δ. All of these loci can rearrange in an immunoglobulin-like fashion and express polypeptides that contribute to either α:β or γ:δ T-cell receptor–CD3 complexes. Surprisingly, the T-cell receptor (TCR) δ coding regions are located entirely, or almost entirely, within the TCR α locus and share at least some of the V region gene segments, thus at least partly linking the two different types of receptor heterodimers. Analysis of potential T-cell receptor diversity, particularly that of the δ chain, indicates a striking concentration of somatic polymorphism in the V–J junctional region of the two heterodimers, four to six orders of magnitude higher than similar calculations for immunoglobulin light- and heavy-chain combinations. In contrast, the number of possible V region combinations in T-cell receptors is one hundredth to one thousandth that of immunoglobulins. TCR α: β heterodimers are known to recognize many possible fragments of antigens embedded in the peptide-binding clefts of a relatively small number of major histocompatibility complex (MHC) molecules. Thus it is attractive to speculate that the V–J junctional portions of both types of T-cell receptor contact peptide antigens, whereas the remaining diversity regions contact the MHC. This contention is supported by molecular modelling studies and has interesting implications for the evolution of antigen-receptor genes.

Introduction

For some time it has been known that T-cell recognition of antigen occurs in a major histocompatibility complex- (MHC-)restricted fashion (Katz *et al.* 1973; Rosenthal & Shevach 1973; Zinkernagel & Doherty 1974). Much recent evidence suggests that the antigens 'seen' by T-cell receptors (TCRs) are fragments of proteins bound to MHC molecules at a single site (Benacerraf 1978; Shimonkevitz *et al.* 1983; Babbitt *et al.* 1985; Buus *et al.* 1986; Townsend *et al.* 1986; Guillet *et al.* 1986; Bjorkman *et al.* 1987*a*). By contrast, the immunoglobulin (Ig) B-cell receptor binds to native antigen alone. Structurally and genetically, however, both immunoglobulins and T-cell receptors seem very similar. Both are derived from the relatively random juxtaposition of different coding segments (V, D and J) of DNA to product proteins that differ in their *N*-terminal domains (V domains), but are the same elsewhere (C domains) (Tonegawa 1983; Kronenberg *et al.* 1986; Davis & Bjorkman 1988). Ig V region domains from the heavy and light chain polypeptides (V_H and V_L) pair to form the ligand-binding region (Eisen 1980). By analogy, it seems likely that the binding site for antigen and MHC is formed by pairs of TCR V domains (either V_α:B_β or V_γ:V_δ). In the Ig variable regions, sequence diversity is concentrated in three distinct 'hypervariable regions' (Wu & Kabat 1970; Kabat

† Present address: DNAX Research Institute, Palo Alto, California 94304, U.S.A.

et al. 1987). These amino acids form the principal points of contact with antigens and are thus referred to as complementarity determining regions (CDRs) (Amzel & Poljak 1979; Davies & Metzger 1983).

Inferred similarity between TCR and Ig structures

Sequence data suggest that TCR variable regions fold into a β-sheet tertiary structure similar to Ig variable regions (Patten *et al.* 1984; Barth *et al.* 1985; Arden *et al.* 1985; Becker *et al.* 1985; Hedrick *et al.* 1984; Novotny *et al.* 1986). In antibodies, the variable regions from the heavy and light chains (V_H and V_L) are paired such that the three complementarity regions (CDR1, 2 and 3) from each domain form the antigen-binding site (Amzel & Poljak 1979; Davies & Metzger 1983; Chothia & Lesk 1987). The overall geometry of V_L–V_H pairing is conserved in the Fabs whose structures are known (Novotny & Haber 1985; Chothia *et al.* 1985) resulting in a similar arrangement of CDRs in these binding sites. Most of the amino acids involved in the interface between Ig V_H and V_L are identically placed in TCR V region sequences (Barth *et al.* 1985; Arden *et al.* 1985; Becker *et al.* 1985; Novotny *et al.* 1986), suggesting that the overall geometry of the TCR V_α:V_β and V_γ:V_δ combining sites will be similar to that of V_H:V_L.

The three-dimensional structures of a number of proteolytic fragments (Fabs) of antibodies complexed to antigens have been described (Amit *et al.* 1986; Colman *et al.* 1987; Sheriff *et al.* 1987; Segal *et al.* 1974). In contrast, the mode of association between TCRs and their more complex ligand (antigen/MHC) is not well understood. Nor is it clear why two completely independent recognition systems are necessary in an organism. Here we compare the patterns of diversity of Igs and TCRs and present a model for TCR interaction with a complex of MHC plus antigen. The model assumes structural similarity between the Ig and TCR combining sites to align the TCR V domains over the known structure of an MHC molecule (Bjorkman *et al.* 1987*a*, *b*).

Potential diversity of TCRs compared with Igs

Compared with Igs, the generation of diversity in TCR heterodimers indicates a striking concentration of sequence polymorphism in the CDR3-equivalent region (Davis & Bjorkman 1988). As indicated in table 1, there are significantly fewer TCR V gene segments than IgV gene segments, and much less combinatorial diversity ($V_H \times V_\kappa$ against $V_\alpha \times V_\beta$, for example). By contrast, the diversity at the V–J junction of TCRs (CDR3) greatly exceeds that of Igs. Particularly striking is the case of the adult TCR δ chain that seems to express only a few V_δ sequences but has been estimated to have up to 10^{13} possible amino acid sequences in its V–J region (Elliott *et al.* 1988). The unique features of TCR genes that contribute to diversity in this region are listed in table 1 and include *N*-region addition in all four TCR chains (as opposed to only V_H of immunoglobulins), the large number of J_α and J_β gene segments and the use of two different D regions simultaneously in TCR δ (Davis & Bjorkman 1988; Elliott *et al.* 1988; Chien *et al.* 1987).

Conclusions

What this skewing of diversity towards the CDR3 equivalent region of TCRs might mean is suggested by the fact that in an antibody-combining site, the CDR3 residues are located in the middle, in between the CDR1 and CDR2 contributed residues of each V region (Chothia & Lesk

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

TABLE 1. SEQUENCE DIVERSITY IN T-CELL RECEPTOR AND IMMUNOGLOBULIN GENES

(Calculated potential amino acid sequence diversity in T-cell receptor and immunoglobulin genes without allowance for somatic mutation. The approximate number of V gene segments are listed for TCR α, β, γ and δ contrasted with V_H and V_L. The first two hypervariable regions of immunoglobulins (CDR1 and CDR2) and their equivalents in TCRs are encoded within the V gene segments. The pairing of random V regions ($V_H \times V_L$, $V_\alpha \times V_\beta$ or $V_\gamma \times V_\delta$) generates the combinatorial diversity listed as 'variable region combinations.' The magnitude of combinational diversity in TCRs is lower than in immunoglobulins because of the decreased number of TCR V gene segments. Estimates for the number of amino acid sequences that might result from diversity within the junctional region are contrasted for TCRs and immunoglobulins. The third immunoglobulin hypervariable region (CDR3) and its TCR equivalent are encoded almost entirely within the D and/or J region gene segments. The last few amino acids encoded by a TCR V-gene segment can contribute to the TCR CDR3-equivalent region, but the effects of these residues on junctional diversity are not included in these calculations. The mechanisms for diversity generation within the junctional region used for this calculation include usage of different D and J gene segments, *N*-region addition up to six nucleotides at each junction, variability in the 3′ joining position in V and J gene segments, and translation of D regions in different reading frames. Numbers are corrected for out-of-frame joining codon redundancy and *n*-region mimicry of germline sequences as detailed in Elliott *et al.* (1988).)

	Ig		TCR I			TCR II
	H	κ	α	b	γ	δ
variable segments	250–1000	250	100	25	7	10
diversity segments	10	0	0	2	0	2
Ds read in all frames	rarely	—	—	often	—	often
N-region addition	V–D, D–J	none	V–J	V–D, D–J	V–J	V–D_1, D_1–D_2, D_1–J
joining segments	4	4	50	12	2	2
variable region combinations	62 500–250 000		2500			70
junctional combinations	*ca.* 10^{11}		*ca.* 10^{15}			*ca.* 10^{18}

1987). In the HLA-A2 molecule, the putative peptide binding region is located in between two nearly parallel α-helices on the surface of the structure. The distance between these two α-helices (*ca.* 18 Å†) is almost identical to the distance between the CDR1 and CDR2 of one IgV domain (in a heterodimer or homodimer) and its partner domain. Thus, an Ig molecule (or TCR) can be 'fit' over an MHC structure such that the CDR3-equivalent residues are spanning a significant portion of the apparent antigen-binding site (Davis & Bjorkman 1988). This may explain the peculiar patterns of diversity that we see in TCRs as opposed to Igs and has interesting evolutionary predictions as well (see Davis & Bjorkman 1988).

M. M. D. is a scholar of the PEW Foundation, P. J. B. is a Postdoctoral Fellow of the American Cancer Society and J. F. E. is a Centennial Fellow of the Medical Research Council of Canada. We thank the NIH and Howard Hughes Medical Institute for grant support.

REFERENCES

Amit, A. G., Mariuzza, R. A., Phillips, S. E. V. & Poljak, R. J. 1986 Three-dimensional structure of an antigen–antibody complex at 2.8 Å resolution. *Science, Wash.* **233**, 747–753.

Amzel, L. M. & Poljak, R. J. 1979 Three-dimensional structure of immunoglobulins. *A. Rev. Biochem.* **48**, 961–997.

Arden, B., Klotz., J., Sui, G. & Hood, L. 1985 Diversity and structure of genes of the alpha family of mouse T-cell antigen receptor. *Nature, Lond.* **316**, 783–787.

Babbitt, B. P., Allen, P. M., Matsueda, G., Haber, E. & Unanue, E. R. 1985 Binding of immunogenic peptides to Ia histocompatibility molecules. *Nature, Lond.* **317**, 359–361.

Barth, R., Kim, B., Lan, N., Hunkapiller, T., Sobieck, N., Winoto, A., Gershenfeld, H., Okada, C., Hansburg, D., Weissman, I. & Hood, L. 1985 The murine T-cell receptor uses a limited repertoire of expressed V_β gene segments. *Nature, Lond.* **316**, 517–523.

† 1 Å = 10^{-10} m = 10^{-1} nm.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Becker, D. M., Patten, P., Chien, Y., Yokota, T., Eshhar, Z., Giedlin, M., Gascoigne, N. R. J., Goodnow, C., Wolf, R., Arai, K. & Davis, M. M. 1985 Variability and repertoire size in T-cell receptor V_α and V_β gene segments. *Nature, Lond.* **317**, 430–434.

Benacerraf, B. 1978 A hypothesis to relate the specificity of T lymphocytes and the activity of I region-specific Ir genes in macrophages and B lymphocytes. *J. Immunol.* **120**, 1809–1812.

Bjorkman, P. J., Saper, M. A., Samraoui, B., Bennett, W. S., Strominger, J. L. & Wiley, D. C. 1987*a* Structure of the human class I histocompatibility antigen, HLAL-2. *Nature, Lond.* **329**, 506–512.

Bjorkman, P. J., Saper, M. A., Samraoui, B., Bennett, W. S., Strominger, J. L. & Wiley, D. C. 1987*b* The foreign antigen binding site and T-cell recognition regions of Class I histocompatibility antigens. *Nature, Lond.* **329**, 512–518.

Buus, S., Colon, S., Smith, C., Freed, J. H., Miles, C. & Grey, H. M. 1986 Interaction between a 'processed' ovalbumin peptide and Ia molecules. *Proc. natn. Acad. Sci. U.S.A.* **83**, 3968–3971.

Chien, Y., Iwashima, M., Wettstein, D. A., Kaplan, K. B., Elliott, J. F., Born, W. & Davis, M. M. 1987 T-cell receptor δ gene rearrangements in early thymocytes. *Nature, Lond.* **330**, 722–727.

Chothia, C., Novotny, J., Bruccoleri, R. & Karplus, M. 1985 Domain association in immunoglobulin molecules. The packing of variable domains. *J. molec. Biol.* **186**, 651–663.

Chothia, C. & Lesk, A. M. 1987 Canonical structures for the hypervariable regions of immunoglobulins. *J. molec. Biol.* **196**, 901–917.

Colman, P. M., Laver, W. G., Varghese, J. N., Baker, A. T., Tulloch, P. A., Air, G. M. & Webster, R. G. 1987 Three-dimensional structures of a complex of antibody with influenza virus neuraminidase. *Nature, Lond.* **326**, 358–363.

Davies, D. R. & Metzger, H. 1983 Structural basis of antibody function. *A. Rev. Immunol.* **1**, 87–117.

Davis, M. M. & Bjorkman, P. J. 1988 T-cell antigen receptor genes and T-cell recognition. *Nature, Lond.* **334**, 395–402.

Elliott, J. F., Rock, E. P., Patten, P. A., Davis, M. M. & Chien, Y. 1988 The adult T-cell receptor δ-chain is diverse and distinct from that of fetal thymocytes. *Nature, Lond.* **331**, 627–631.

Eisen, H. N. 1980 *Immunology*. New York: Harper & Row.

Guillet, J.-G., Lai, M.-Z., Briner, T. J., Smith, J. A. & Gefter, M. L. 1986 Interaction of peptide antigens and class II major histocompatibility complex antigens. *Nature, Lond.* **324**, 260–262.

Hedrick, S. M., Nielsen, E. A., Kavaler, J., Cohen, D. I. & Davis, M. M. 1984 Sequence relationships between putative T-cell receptors polypeptides and immunoglobulins. *Nature, Lond.* **308**, 153–158.

Kabat, E. A., Wu, T. T., Reid-Miller, M., Perry, H. M. & Gottesman, K. S. 1987 *Sequences of proteins of immunological interest*. Bethesda, Maryland: Public Health Service.

Katz, D. H., Hamaoka, T. & Benacerraf, B. 1973 Cell interactions between histoincompatible T and B lymphocytes. II. Failure of physiologic cooperative interactions between T and B lymphocytes from allogeneic donor strains in humoral response to hapten–protein conjugates. *J. exp. Med.* **137**, 1405–1418.

Kronenberg, M., Sui, G., Hood, L. E. & Shastri, N. 1986 The molecular genetics of the T-cell antigen receptor and T-cell antigen recognition. *A. Rev. Immunol.* **4**, 529–591.

Novotny, J., Tonegawa, S., Saito, H., Kranz, D. M. & Eisen, H. N. 1986 Secondary, tertiary, and quaternary structure of T-cell-sepdivid immunoglobin-like polypeptide chains. *Proc. natn. Acad. Sci. U.S.A.* **83**, 742–746.

Novotny, J. & Haber, E. 1985 Structural invariants of antigen binding: comparisons of immunoglobulin V_L–V_L domain dimers. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4592–4596.

Patten, P., Yokota, T., Rothbard, J., Chien, Y., Arai, K. & Davis, M. M. 1984 Structure, expression and divergence of T-cell receptor beta-chain variable regions. *Nature, Lond.* **312**, 40–46.

Rosenthal, A. S. & Shevach, E. M. 1973 Function of macrophages in antigen recognition by guinea pig T lymphocytes. *J. exp. Med.* **138**, 1194–1212.

Segal, D. M., Padlan, E. A., Cohen, G. H., Rudikoff, S., Potter, M. & Davies, D. R. 1974 The three-dimensional structure of a phosphorylcholine-binding mouse immunoglobulin Fab and the nature of the antigen binding site. *Proc. natn. Acad. Sci. U.S.A.* **71**, 4298–4302.

Sheriff, S., Silverton, E. W., Padlan, E. A., Cohen, G. H. Smith-Gill, S. J., Finzel, B. C. & Davies, D. R. 1987 Three-dimensional structure of an antibody–antigen complex. *Proc. natn. Acad. Sci. U.S.A.* **84**, 8075–8079.

Shimonkevitz, R., Kappler, J. W., Marrack, P. & Grey, H. M. 1983 Antigen recognition by H-2 restricted T cells. I. Cell free antigen processing. *J. exp. Med.* **158**, 303–316.

Tonegawa, S. 1983 Somatic generation of antibody diversity. *Nature, Lond.* **302**, 575–581.

Townsend, A. R. M., Rothbard, J., Gotch, G. M., Bahadur, G., Wraith, D. & McMichael, A. J. 1986 The epitopes of influenza nucleoprotein recognized by cytotoxic T lymphocytes can be defined with short synthetic peptides. *Cell* **44**, 959–968.

Wu, T. T. & Kabat, E. A. 1970 Analysis of the sequences of the variable regions of Bence-Jones proteins and myeloma light chains and their implications for antibody complementarity. *J. exp. Med.* **132**, 211–250.

Zinkernagel, R. M. & Doherty, P. C. 1974 Restriction of *in vitro* T-cell mediated cytotoxicity in lymphocytic choriomeningitis within a syngeneic or semi-allogeneic system. *Nature, Lond.* **248**, 701–702.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES