Direct Identification of Naturally Processed Autoantigen-derived Peptides Bound to HLA-DR15*

(Received for publication, January 11, 1996, and in revised form, May 7, 1996)

R. G. Phelps, A. N. Turner, and A. J. Rees
From the Department of Medicine and Therapeutics, University of Aberdeen, Aberdeen, AB9 22D Scotland, United Kingdom

Biochemical analysis of HLA class II-associated peptides from antigen-pulsed cells is a potentially useful approach to the analysis of antigen processing and presentation because it examines directly which antigen-derived peptides are presented. This is especially advantageous in the analysis of self-antigen presentation where conventional approaches utilizing antigen-specific T cells may be biased by the presence of self-tolerance. However, successful biochemical analysis has been reported for only one exogenous antigen and no autoantigens. We have used a novel analytical approach coupling biochemical data with the reported properties of class II-associated peptides to characterize the peptides derived from a clinically relevant autoantigen presented on the disease-associated class II type. Incubation of the target of autoimmune attack in patients with Goodpasture's disease, the 230-amino acid NC1 domain of the α3 chain of type IV collagen (Goodpasture antigen, α3(IV)NC1), with human B cells homozygous for HLA-DR15, the allele carried by 80% of patients, we find that α3(IV)NC1 is presented as at least two sets of three to five peptides centered on common core sequences (nested sets). Synthetic peptides containing these core sequences bind to HLA-DR15 with intermediate affinity (IC₅₀, 1.1–6 μM).

Extracellular antigens are recognized by T helper cells in the form of processed peptides bound to HLA class II molecules on the surface of antigen presenting cells (APC). However, the mechanisms determining which antigen-derived peptides are displayed are poorly understood. Extracellular antigens are uptaken by APC into endosomes where, in an acidic reducing environment, proteolytic cleavage generates antigen-derived peptides (processing). Those peptides able to form stable complexes with HLA class II molecules are protected from further proteolysis and transported to the cell surface for T cell recognition. Which peptides are made available and in what relative proportion is believed to profoundly influence immune responses, but direct biochemical analysis of antigen-derived class II-associated peptides has been reported for only one antigen.

HLA class II-associated peptides have been purified and characterized from several in vitro propagated APC types bearing a variety of class II types (1–5) and cells extracted from spleen and thymus (6). They comprise complex mixtures of short (12–25 residues) peptides and derive mainly from membrane-associated and endosomal proteins. Peptides derived from exogenous antigens, which presumably are present at substantially lower concentrations within endosomes, comprise a small proportion of identified peptides (3, 4, 7). Nevertheless, Nelson et al. were able to identify peptides derived from hen egg lysozyme (HEL) among the IA*-associated peptides of a murine APC pulsed with high concentrations of HEL (8). The peptides contained a common core sequence and variable degrees of NH₂- and COOH-terminal overhang, a nested set, now recognized as characteristic of class II-associated peptides and to have a structural basis in the "open at both ends" class II peptide binding groove (9). Importantly this core sequence was the immunodominant epitope of HEL on IA*, showing that the approach can identify T cell epitopes and supporting the concept that immune responses to exogenous antigens are directed toward the predominant antigen-derived peptides available.

In contrast to exogenous antigens, the relationship between the presentation of self-antigens and the specificity of autoimmune responses has not been established in any experimental or human autoimmune disease. There may be important differences. Because some degree of tolerance to self-antigens is likely to exist and be most complete to the predominant self-antigen derived peptides presented, autoactivity may be directed at other epitopes (10–12). This may be the case in an animal model of autoimmunity in which autoactivity was found to be directed at a peptide with undetectable binding affinity for class II (13). If the methods used to predict T cell epitopes within exogenous proteins (on the basis of predicted high binding affinity for class II (14, 15)) are to be extended to the prediction of the peptide targets of autoimmunity (16), it will be important to understand the relationship between how an autoantigen is presented as peptides and the autoimmune response to it.

We are investigating Goodpasture's disease as a model of antigen presentation in human autoimmune nephritis. Here we describe a novel approach to the analysis of biochemical data that has enabled us to characterize antigen-derived peptides presented by APC pulsed with modest quantities of the Goodpasture antigen. Goodpasture's disease is an uncommon form of glomerulonephritis caused by autoimmunity to a component of glomerular and certain other basement membranes (17–20). It is a good model of autoimmunity because its pathogenesis is known, the target antigen (α3(IV)NC1) is the same in all patients (21), and over 80% of patients carry the same class II allele, HLA-DR15 (22, 23). In this work we have sought to identify biochemically the predominant peptides derived from...
**Experimental Procedures**

Preparation of Recombinant α3(IV)NC1—cDNA for α3(IV)NC1 (19) was expressed in Escherichia coli as fusion proteins with either glutathione S-transferase (expression vector pGEX (24)) or a 6-histidine residue tag (pET14b from Novagen). In each case a linker containing a threonine residue tag (pET14b from Novagen) was inserted in case of small run to run variations in retention time.

α3(IV)NC1 present by the disease-associated class II molecule HLA-DR15.

**Amino Acid Residue Positions**

| Position | α3(IV)NC1-derived peptides with the indicated core sequences. |
|----------|--------------------------------------------------------------|
| α3(IV)NC1 | Calculating the mass of all possible antigen-derived peptides assuming full reduction of disulfide bonds and absence of amino acid modifications. Peptides with calculated mass within the 95% confidence intervals of a measured extra mass were considered to match. In each experiment occurrences of each of the possible (226) 9-amino acid cores within peptides matched to extra masses were counted, and cores were ranked according to the number of different extra masses matched. Numbering of amino acid residues is relative to the sequence SPAT (expression vector pGEX (24)) or a 6-histidine residue tag (pET14b from Novagen). In each case a linker containing a threonine residue tag (pET14b from Novagen) was inserted in case of small run to run variations in retention time.

α3(IV)NC1 presented by the disease-associated class II molecule HLA-DR15.

**Occurrence within peptides matched to extra masses identified in two experiments**

| Core α3(IV)NC1-derived sequence | Occurrence within peptides matched to extra masses |
|--------------------------------|---------------------------------------------------|
| TLGSCLQR8FT | 53 |
| TVCEPAAIAIV | 113 |
| LEEFRASPF | 169 |
| FCNNDVVCN | 68 |
| AIAAXHSHQTT | 119 |
| HSOQTD1PPCC | 125 |
| GSIFIMFTS | 144 |
| TSAGSEGTTG | 151 |
| ASPGSCEELE | 163 |
| LEEFRASPF or FCNNDVVCN | 11 |

**Retention Time Prediction**

Retention time prediction is well validated for short peptides but may be less reliable when extrapolated to peptides longer than 12 residues because of secondary structure considerations.

**Peptide Binding Assays**

The myelin basic protein peptide MBP was purified by HPLC, and their composition was confirmed by mass spectrometry and amino acid analysis. Truncation analogues were generated by manual Edman degradation and partial carboxypeptidase digestion employing carboxypeptidase P and Y (Boehringer Mannheim). HPLC conditions were identical to those used for class II-associated peptide pools. Retention time predictions were made as described (27). The amino acid retention coefficients reported in Ref. 27 were used to seed regression analysis of the measured retention times of 26 synthetic peptides. Measured and predicted retention times exhibited good fit to a linear model ($R^2 = 0.93$). Retention time prediction was well validated for short peptides but may be less reliable when extrapolated to peptides longer than 12 residues because of secondary structure considerations. This source of error was minimized by attempting retention time prediction only for peptides of similar length and sequence to the peptides used to calculate the amino acid coefficients.
TABLE II

Two nested sets of α3(IV)NC1-derived peptides matched to extra masses identified in first experiment

| Extra massa | Massb | Position | Length | Matched peptides | Sequence |
|-------------|-------|----------|--------|------------------|----------|
| 1910.9      | 1910.1| 160      | 18     | QALASPGSC LEEFRASPFP |          |
| 2024.1      | 2023.3| 160      | 19     | QALASPGSC LEEFRASPFP L |          |
| 2024.3      | 2024.3| 161      | 19     | ALAGSC LEEFRASPFP LE |          |
| 1823.1      | 1824.1| 162      | 17     | LASPSC LEEFRASPFP L |          |
| 1944.0      | 1943.2| 163      | 18     | ASPSC LEEFRASPFP L |          |
| 2197.6      | 2196.4| 166      | 20     | GSC LEEFRASPFP LECHGRGT |          |
| 1849.0      | 1848.0| 169      | 16     | LEEFRASPFP LECHGRG |          |
| 2509.1      | 2509.9| 55       | 22     | GSCLTFTMPFL FCNVDVCC |          |
| 2024.3      | 2023.1| 62       | 18     | TTMPFL FCNVDVCC FASR |          |
| 2178.7      | 2179.5| 62       | 19     | TTMPFL FCNVDVCC FASR |          |
| 2068.0      | 2060.2| 65       | 17     | PFL FCNVDVCC FASRN |          |
| 2024.3      | 2075.3| 65       | 18     | PFL FCNVDVCC FASRN |          |

a Mean of two determinations of mass; 95% confidence limits ± 0.07%.
b Calculated from sequence assuming Cys reduced using the following atomic weights of the elements: C = 12.011, H = 1.00794, N = 14.0067, O = 15.9994, S = 32.06.

Fig. 2. A “degenerate” HLA-DR15 peptide binding motif based on published data designed to identify peptides with the minimal requirements for binding to the gene products of either of DRB1*1501 or DRB5*0101, the two DR molecules expressed in the DR15 type. An X signifies an unspecified amino acid.

RESULTS

MGAR, a human Epstein-Barr virus transformed B cell homoygous for HLA-DR15, was incubated with (treated) or without (control) recombinant human α3(IV)NC1, and the purified HLA-DR15-associated peptide pools were separated by reverse phase HPLC. When the chromatograms were compared (Fig. 1a), minor but distinct differences suggested the presence of extra peptides in the peptide pool obtained from antigen-treated APC. In order to confirm the presence of extra peptides and to precisely measure their mass, each HPLC fraction was analyzed by matrix-assisted laser-desorption ionization time-of-flight mass spectrometry. Most fractions were found to contain complex mixtures of peptides with molecular masses between 1300 and 3000 Da, as is typical of HPLC-separated class II-associated peptides (3). Similar spectra were obtained for the majority of corresponding control and antigen-treated fractions, indicating that antigen treatment had not greatly perturbed baseline class II peptide presentation. In the first experiment 17 masses present in spectra from antigen-treated fractions could not be found in spectra obtained for any adjacent control fraction (Fig. 1b). The extra, putatively α3(IV)NC1-derived, masses ranged between 1389.8 and 2509.1 Da (95% confidence limits ± 0.07%) and in general occurred in fractions collected where the control and treated chromatograms differed. Attempts to further purify and obtain the sequences of the “extra masses” were unsuccessful because each was present at low levels (estimated to be 0.1–10 pmol) within complex peptide mixtures (comprising <5% of total peptide).

In order to identify which antigen-derived peptides could account for the extra masses, the sequence of α3(IV)NC1 was searched for peptides with matching calculated mass. 1–12 (median 5) peptides could be matched to each extra mass. Because HLA class II-associated peptides characteristically comprise nested sets enclosing common core sequences (1), we examined the sequences of the matched peptides for the recurring presence of a 9 (or more) amino acid core. Three putative nested sets each comprising peptides matched to six different extra masses were identified. Each set contained a 9–11-amino acid core sequence (Table I) and one (Table II, upper set) was centered on a core sequence containing a DR15 binding motif. The motif, deduced from published data, specifies a 9-amino acid core sequence with hydrophobic residues at positions 1 and 4 (Fig. 2). A further six putative nested sets each containing extra masses matched to five extra masses were identified, one with a core sequence matching the DR15 motif (Table II, lower set).

In order to extend these observations a second experiment was performed with 4-fold more antigen. Seventeen extra masses were identified (1568.5–2449.13 Da ± 0.05%) and matched to 1–10 (median 4) α3(IV)NC1-derived sequences. The number of extra masses that could be matched by mass to peptides containing the core sequences identified in the first experiment are shown in Table I. Eleven of the 17 extra masses could again be matched to peptides containing the two previously identified DR15 motif-containing core sequences (Table III). The probability that this recurrent observation could arise by chance was investigated by matching 8,000 random numbers in the range of the observed extra masses to α3(IV)NC1-derived peptides. Random masses could be matched to peptides containing either (or both) core sequence in 2651/8000 (p = 0.33, σ = 0.0526). The expected number of chance matches is therefore 5.5–5.8, and the probability of observing 11/17 matches lies between 0.0058 and 0.0099 (Binomial distribution using p = 1.96σ). This result suggested that at least a proportion of the sequences assigned to 11 extra masses (shown in Table II) were likely to be correct.

Although the APC had been pulsed with a higher concentration of antigen, again insufficient peptide could be purified to obtain confirmatory sequence data. Instead the sequence assignments shown in Table III were tested by examining the properties of synthetic peptides with some of the proposed sequences. Three out of 6 (marked with asterisks in Table III) exhibited consistent HPLC retention times. Chromatograms for two of these are shown in Fig. 3. Five other extra masses were in fractions with retention times close to those predicted
for their respective proposed sequences (marked with ‡ in Table III). Retention time prediction is likely to be acceptably accurate here because it is attempted only for peptides very similar to those used to calibrate the HPLC equipment; 17 of the 26 synthetic peptides used in the calibration were truncation variants of peptides containing one of the two 9-amino acid core sequences common to these sequences in Table III. Therefore analysis of the retention times of synthetic peptides with the sequences proposed in Table III supports sequence assignment in 8/11 and refutes assignment in 3/11.

Because class II-associated peptides are known to be capable of binding to affinity-purified class II molecules (1), presumably by displacing other peptides, the binding capacity of synthetic peptides was examined in an inhibition binding assay (Fig. 4). Peptides containing either core sequence bound to affinity-purified HLA-DR15 with affinity (IC50 1.1–6 μM) intermediate within the range reported for class II-associated peptides (1, 29) and similar (less than 5 fold difference) to that measured for the well described (26) DR15-binding myelin basic protein peptide 86–98 (98A). Similar results were obtained for longer (20 residue) peptides encoding these core sequences, suggesting that all the peptides in Table III are likely to bind to HLA-DR15 with high affinity.

Takentogetherthedatasuggestthatofthe17extramasstobiochemically detectable among the DR15-associated peptides of a3(IV)NC1-pulsed human B cells (in the second experiment), five correspond to peptides encoding the core sequence LEEFRASPF, three correspond to peptides encoding FCNVNDV, and nine remain unidentified.  

### DISCUSSION

Class II-associated peptides have generally been studied by analyzing the responses of T cells to antigen or peptide-pulsed antigen presenting cells. Although this approach has the advantage of great sensitivity, it does not indicate the range and relative proportion of antigen-derived peptides presented nor necessarily the predominant peptides (30), because it can only detect peptides recognized by available T cells. These could be very different, especially in the case of autoantigens, both because of self-tolerance and because of practical limitations in the raising of panels of T cell clones with widely varying specificities and restrictions. In comparison, direct biochemical analysis of class II-associated peptides purified from antigen-pulsed cells potentially indicates both the range and relative proportion of antigen-derived peptides available for T cell recognition. This is advantageous because both parameters are believed important in determining the initiation of immune (and autoimmune) responses (8, 31, 32) and the development of self-tolerance (11, 33).
Natural Processing of Goodpasture Antigen

The use of a biochemical approach in the characterization of antigen-derived major histocompatibility complex class II-associated peptides was first described by Nelson et al. (8). However, it is striking that the only reported successes with this approach have used HEL as antigen (5, 8). This reflects the difficulty in analyzing the small quantity of antigen-derived peptides that can be isolated from even large numbers of antigen-pulsed cells (5, 34, 35) and perhaps the difficulty in preparing large quantities of antigen. Goodpasture disease may be particularly suitable for this approach. Unlike many autoimmune diseases, the target antigen is known and can be made in large quantities as a recombinant molecule. Also α3(IV)NC1, like HEL, is both a small molecule and cationic, factors that may promote uptake into APC (36). However, success depended upon the recognition that α3(IV)NC1, like HEL, was likely to be presented in the form of nested sets of peptides. A strategy could then be devised by which candidate sequences for extra masses could be distinguished from other mass-matched sequences. However the numerical analysis utilized can only be presented and that these are generally among those with high-affinity for class I molecules would suggest that HLA-determined differences in the predominant antigen-derived peptides are presented and that these are generally among those with highest class II binding affinity, data for autoantigens is lacking. Indeed there are theoretical arguments and experimental results that suggest that other peptides presented at lower level are important. In particular it may be expected that self-tolerance will be most securely established to the predominant self-peptides presented (11). If the rapidly accumulating knowledge on the peptide binding characteristics of class II molecules is to be utilized in the prediction of autoimmune disease-associated epitopes, it is important to clarify the relation between self-antigen-derived peptide presentation, class II type, and autoreactivity. Our data identify for the first time some of the predominant peptides derived from a human autoantigen presented on a disease-associated class II type, laying a foundation for investigating this relationship in human autoimmunity.

Acknowledgments—We are grateful to Prof. J. Fothergill and I. Davidson in the Department of Molecular and Cell Biology for access to and assistance in the use of a mass spectrometer.

REFERENCES

1. Chicz, R., Urban, R., Lane, W., Gorga, J., Stern, L., Vignali, D., and Strominger, J. (1992) Nature 358, 764–768
2. Falk, K., Roubitschke, O., Steinwich, S., Jung, G., and Ramhans, H.-G. (1994) Immunogenetics 39, 230–242
3. Hunt, F., Michel, H., Dickenson, T., Shabanowitz, J., Cox, A., Sakkajogi, K., Appella, E., Grey, H., and Sette, A. (1992) Science 256, 1817–1820
4. Rudensky, A., Preston-Hurlbut, P., Hong, S., Barlow, A., and J aneway, C. A., Jr. (1991) Nature 353, 622–627
5. Vignali, D., Urban, R., Chicz, R., and Strominger, J. (1993) Eur. J. Immunol. 23, 1602–1607
6. Marrack, P., Ignatowicz, L., Kappler, J. W., Boyml, J., and Freed, J. H. (1993) J. Exp. Med. 178, 2173–2183
7. Chicz, R. M., Urban, R. G., Gorga, J. C., Vignali, D. A., Lane, W. S., and Strominger, J. L. (1992) J. Exp. Med. 176, 27–47
8. Nelson, C. A., Roof, R. W., McCourt, D. W., and Unanue, E. R. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 7380–7383
9. Brown, J. H., Jaredetzky, T. S., Garga, J. C., Stern, L. J., Urban, R. G., Strominger, J. L. and, and Wiley, D. C. (1993) Nature 364, 33–39
10. Sercarz, E. E., Lehmann, P. V., Ametani, A., Benichou, G., Miller, A., and Kourilsky, P. (1993) Annu. Rev. Immunol. 11, 729–766
11. Cribbs, R., Kabelopoulos, J. M., Cabaniols, J. P., and Kanellopoulos, J. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 1816–20
12. Lehmann, P. V., Forsthuber, T., Miller, A., and Sercarz, E. E. (1992) Nature 358, 155–157
13. Fairchild, P., Wildgoose, R., Atherton, E., Webb, S., and Wraith, D. (1993) Int. Immunol. 5, 1151–1158
14. Hammer, J., Bono, E., Gallazzi, F., Belusin, C., Nagy, Z., and Sinigaglia, F. (1994) J. Exp. Med. 180, 2353–2358
15. Marshall, K., Wilson, K., Lian, J., Woods, A., Zoller, D., and Rothbard, J. (1995) Biomed. Pept. Prot. Nucleic Acids. 1, 157–162
16. Hammer, J., Gallazzi, F., Bono, E., Karr, R. W., Guenot, J., Valsasini, P., Nagy, Z. A., and Sinigaglia, F. (1995) J. Exp. Med. 183, 1847–1853
17. Sauls, J., Wieslander, J., Lang, R., Guinones, S., and Hudson, B. (1988) J. Biol. Chem. 263, 13374–13380
18. Morrison, K., Marimaya, M., Yang-Feng, T., and Reeder, S. (1991) Am. J. Hum. Genet. 49, 540–554
19. Turner, N., Mason, P. J., Brown, R. Fox, M., Powey, S., Rees, A., and Pusey, C. D. (1992) J. Clin. Invest. 89, 592–601
20. Kalluri, R., Wilson, C. B., Weber, M., Gunwar, S., Chonko, A. M., Nielsen, E. G., and Hudson, B. G. (1995) J. Am. Soc. Nephrol. 6, 1178–1185
21. Pusey, C. D., Dash, A., Kershaw, M. J., Morgan, A., Reilly, A., Rees, A. J., and Lockwood, C. M. (1987) Lab. Invest. 56, 23–31
22. Burns, A. P., Fisher, M., Li, P., Pusey, C. D., and Rees, A. J. (1995) J. Med. 88, 93–100
23. Huey, B., McCormack, K., Capper, J., Ratcliff, C., Colombe, B., and Sinigaglia, F. (1992) J. Biol. Chem. 267, 31–40
24. Smith, D., and Johnson, K. (1988) Gene (Amst.) 67, 31–40
25. Vogt, A., Krohnopher, H., Kalkbacher, H., Kalluri, R., and Sinigaglia, F. (1994) J. Immunol. 153, 1665–1673
26. Wucherpfennig, K. W., Sette, A., Appella, E., Grey, H., and Sette, A. (1992) Kidney Int. 41, 307–313
27. Smith, D., and Johnson, K. (1988) Annu. Rev. Immunol. 6, 93–100
28. Quinton, P. E. (1991) J. Exp. Med. 174, 1111–1120
29. Buus, S., Sette, A., Colon, S., Jenis, D., and Grey, H. (1986) Cell 47, 1071–1077
30. Viner, N., Nelson, C. A., and Unanue, E. R. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 2212–2218
31. Buus, S., Sette, A., Colon, S. M., Miles, C., and Grey, H. M. (1987) Science 235, 1353–1358
32. Theofilopoulos, A. (1995) Immunol. Today 16, 90–97
33. Cabaniols, J. P., Cribbs, R., Kourilsky, P., Sinigaglia, F., and Kalluri, R. (1994) Eur. J. Immunol. 24, 1743–1749
34. Doria, S., Grey, H. M., and Sette, A. (1990) Science 249, 1028–1030
35. Moore, S., Schaffold, J., and Watts, C. (1995) Immunology 85, 523–530
36. Apple, R. J., Dorn, P. L., Muckerheide, A., and Michael, J. B. (1988) J. Immunol. 140, 3290–3295