# Chapter 12
# Molecular Similarity Between the 'Shared Epitope' of Rheumatoid Arthritis and Bacteria

## Contents

**Introduction: The Association of the 'Shared Epitope' EQR(K)RAA with Rheumatoid Arthritis** ..... 113

**Computer Analysis of the 'Shared Epitope' EQRRAA Sequence and the ESRRAL Sequence Found in the Haemolysins of *Proteus/Serratia* Microbes** ..... 114

**Implications for the Aetiology of Rheumatoid Arthritis** ..... 118

**Discussion and Conclusion** ..... 119

**References** ..... 119

## Introduction: The Association of the 'Shared Epitope' EQR(K)RAA with Rheumatoid Arthritis

The association between rheumatoid arthritis and some subtypes of HLA-DR4 is well established not only in white subjects but also in many other ethnic groups. HLA-DR4 can be subdivided into several subtypes, only some of which are associated with rheumatoid arthritis (Stastny 1978).

In white subjects, DR4/Dw4 and DR4/Dw14 subtypes are associated with rheumatoid arthritis, whereas in Japanese subjects, it is DR4/Dw15 that is the susceptibility factor (Maeda et al. 1981).

A. Ebringer, *Rheumatoid Arthritis and Proteus*,
DOI 10.1007/978-0-85729-950-5_12,

In Israel where DR4/Dw10 predominates, an association with HLA-DR1 has been reported in patients with rheumatoid arthritis (Schiff et al. 1982).

Analysis with synthetic oligonucleotides has shown that a particular region of the DRβ1 chain, from positions 70–74 coding for amino acids Gln-Arg-Arg-Ala-Ala (QRRAA), specific for DR1, Dw14 and Dw15, showed a strong association with rheumatoid arthritis compared to control subjects (Watanabe et al. 1989).

The sequence closely resembles that found in DRB1*0401 (DR4/Dw4) individuals, there being only one conservative substitution at position 71, from arginine to lysine (QKRAA). These two amino acids are positively charged, and thus the overall shape and charge configuration of these two sequences are similar.

This sequence EQ(K)RRAA has been described as the 'shared epitope' by the Winchester group from New York (Gregersen et al. 1987).

The glutamic acid (E) occupying position 69 is common to all DRβ1 molecules.

Furthermore, the EQRRAA sequence is also found in DRB1*1402 (DR6/DW16)-positive Yakima Indians affected by rheumatoid arthritis.

In contrast, the Dw10 haplotype has an aspartic acid at position 70 and glutamic acid at position 71, giving a net negative charge to the sequence, whereas Dw 13 has a glutamic acid at position 74.

The question arises as to whether there is 'molecular mimicry' between the EQR(K)RAA sequence and biological molecules found in the environment of rheumatoid arthritis patients.

## Computer Analysis of the 'Shared Epitope' EQRRAA Sequence and the ESRRAL Sequence Found in the Haemolysins of *Proteus/Serratia* Microbes

The rheumatoid arthritis susceptibility sequence spanning residues 69–74 (EQRRAA) was used to scan published sequences of molecules from *Proteus* microorganisms. No

Table 12.1 Comparison of the charge distribution (+ positive, – negative; *N* neutral) of the amino acid sequences of HLA-DR1 and DR4 molecules with the similarity sequence in *Proteus mirabilis* membrane haemolysin

| | Amino acid positions | | | | | | HLA |
|---|---|---|---|---|---|---|---|
| **Disease link** | **69** | **70** | **71** | **72** | **73** | **74** | **subtype** |
| Associated with RA | – | N | + | + | N | N | DR1, Dw14, Dw15 |
| | – | N | + | + | N | N | Dw4 |
| | – | N | + | + | N | N | *Proteus* haemolysin |
| Not associated with RA | – | – | – | + | N | N | Dw10 |
| | – | N | + | + | N | – | Dw13 |

*RA* rheumatoid arthritis

hexamer identity could be found in the Genbank database, but a closely related sequence (ESRRAL) spanning residues 32–37 of the surface membrane haemolysin of *Proteus mirabilis* (Hpm B polypeptide) was identified which had biochemical and charge similarity to the susceptibility sequence (Table 12.1).

A similar sequence was found in the membrane haemolysin of *Serratia marcescens* (Shl B polypeptide) but not in the haemolysins produced by ten other bacteria including *Escherichia coli* (Uphoff and Welch 1990).

The haemolysin molecule of *Proteus mirabilis* is composed of two polypeptides, Hpm B and Hpm A, with molecular weights of 63 and 165 kDa respectively. The Hpm B polypeptide is necessary for the extracellular secretion and activation of the structural haemolysin Hpm A and is thought to be located in the outer membrane where it could be involved in immune interactions. The ESRRAL sequence is hydrophilic and is an alpha helix (Fig. 12.1).

The secondary structure predictions were performed according to the published biochemical methods (Chou and Fasman 1978; Garnier et al. 1978).

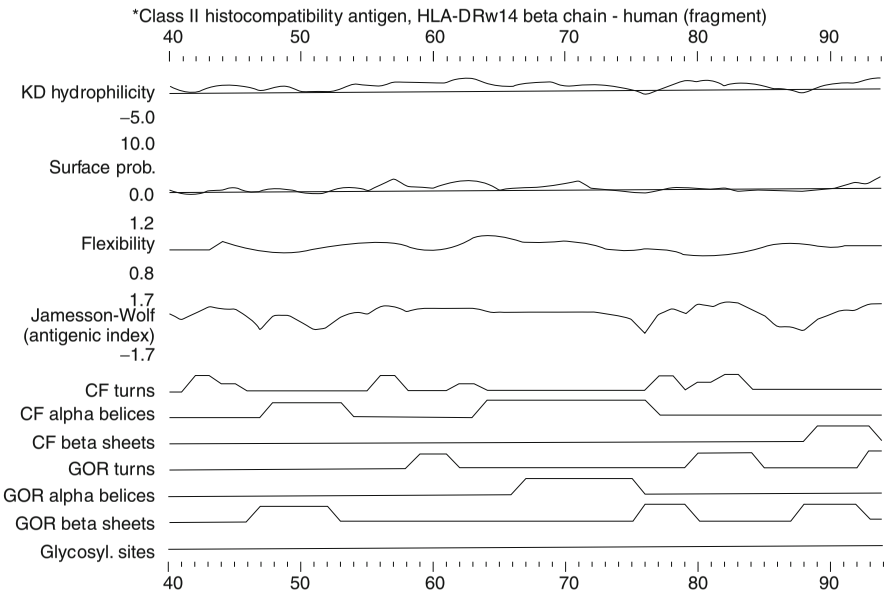


FIGURE 12.1 Computer plots generated from the amino acid sequences of HLA-DRw14 β chain and *Proteus haemolysin*. The hatched lines above and below the plots refer to the amino acid numbers of the protein sequence. Secondary structure predictions were performed according to published methods (Chou and Fasman 1978; Garnier et al. 1978)

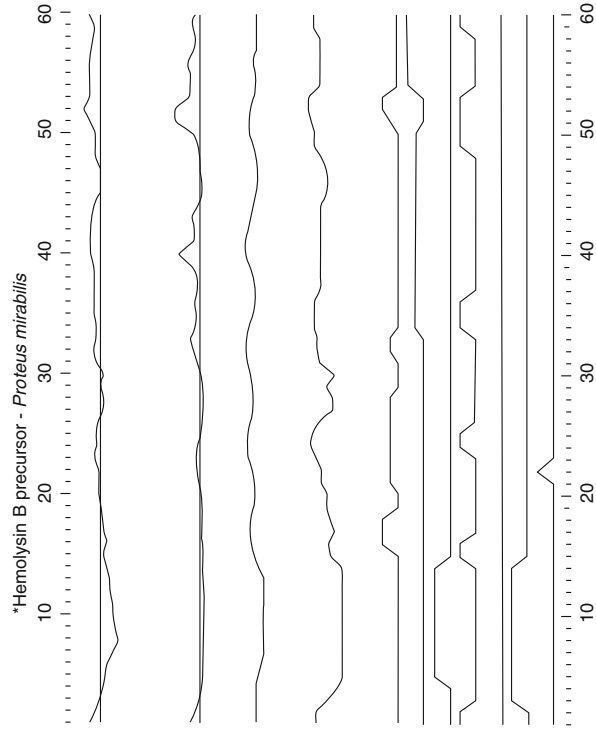


FIGURE 12.1 (continued)

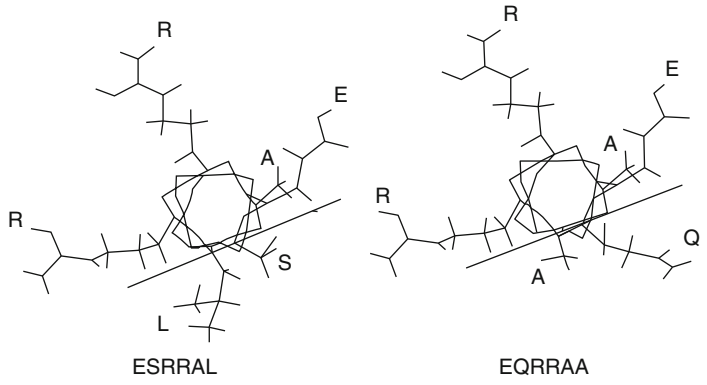


FIGURE 12.2 Helical models of ESRRAL and EQRRAA sequences. The predicted structures show a common charged surface of homologous residues available for immune interaction (*above line*). Modelling by DTMM (version 1.2) (Reprinted with permission from Springer Science+Business Media, Ebringer et al. (1992))

Molecular modelling of EQRRAA and ESRRAL motifs illustrates that the two sequences can be fitted to similar configurations producing almost identical antigenic epitopes (Fig. 12.2).

## Implications for the Aetiology of Rheumatoid Arthritis

The similarity sequence of *Proteus mirabilis* membrane haemolysin has the same charged amino acids arranged in the same order as the susceptibility sequence of DR1, Dw14 and Dw15 alleles, that is glutamic acid followed by a small neutral amino acid (serine) and then two positively charged arginines.

The three charged amino acids of the *Proteus mirabilis* sequence may provide an immunogenic group with a hydrophilicity value of +7.0, thereby indicating that it could be involved in immune interactions (Hopp and Woods 1981).

The important question is whether the *Proteus* peptide does indeed take up the predicted structure. If an antibody

were to be produced against the similarity sequence of the *Proteus* haemolysin membrane protein, it may bind more readily to the DR1, Dw14 and Dw15 sequences, in addition to the Dw4, but not to the Dw10 and Dw13, as there are significant charge differences between these sequences.

Thus, the *Proteus* haemolysin sequence discriminates between alleles of HLA-DR1 and HLA-DR4 that are and are not associated with rheumatoid arthritis.

If such antibodies activate the complement cascade and stimulate natural killer cells, then this could provide an explanation for the association of the 'shared epitope' with rheumatoid arthritis.

## Discussion and Conclusion

Computer analysis of the 'shared epitope' (EQRRAA) has identified a similarity sequence (ESRRAL) in both *Proteus* and *Serratia* microorganisms.

The question arises whether antibodies to these bacteria are relevant in the aetiology and pathogenesis of rheumatoid arthritis.

The results obtained by the Immunology group at King's College indicate that there is an amino acid homology between an outer membrane haemolysin protein of *Proteus haemolysin*, and *Serratia marcescens*, and the susceptibility sequence in HLA-DR1 and DR4 subtypes associated with rheumatoid arthritis (Ebringer et al. 1992).

## References

Chou PY, Fasman G. Prediction of the secondary structure of proteins from the amino acid sequence. Adv Enzymol. 1978;47: 145–7.

Ebringer A, Cunningham P, Ahmadi K, Wrigglesworth J, Hosseini R, Wilson C. Sequence similarity between HLA-DR1 and DR4 subtypes associated with rheumatoid arthritis and *Proteus/ Serratia* haemolysins. Ann Rheum Dis. 1992;51:1245–6.

Garnier J, Osguthorpe DJ, Robson BI. Analysis of the accuracy and implication of simple methods for predicting the secondary structure of globular proteins. J Mol Biol. 1978;120:97–120.

Gregersen PK, Silver J, Winchester RJ. The shared epitope hypothesis: an approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. Arthritis Rheum. 1987;30:1205–13.

Hopp TP, Woods KR. Prediction of protein antigenic determinants from amino acid sequences. Proc Natl Acad Sci USA. 1981;78: 3824–8.

Maeda H, Juji T, Mitsui H, Sonozaki H, Okitsu H. HLA-DR4 and rheumatoid arthritis in Japanese people. Ann Rheum Dis. 1981;40:299–302.

Schiff B, Mizrachi Y, Orgad S, Yaron M, Gazit E. Association of HLA-Aw31 and HLA-DR1 with adult rheumatoid arthritis. Ann Rheum Dis. 1982;41:403–4.

Stastny P. Association of the B-cell alloantigen DRW4 with rheumatoid arthritis. N Engl J Med. 1978;298:869–71.

Uphoff TS, Welch RA. Nucleotide sequencing of the *Proteus mirabilis* calcium independent haemolysin genes (hpm A and hpm B) reveals sequence similarity with the *Serratia marcescens* haemolysin genes (shl A and shl B). J Bacteriol. 1990;170:3177–88.

Watanabe Y, Tokunaga K, Matsuki K, Takeuchi F, Matsuta K, Maeda A, et al. Putative amino acid sequence of HLA DRβ chain contributing to rheumatoid arthritis susceptibility. J Exp Med. 1989;169:2263–8.