CONCISE REPORTS

Shared amino acid sequences between major histocompatibility complex class II glycoproteins, type XI collagen and *Proteus mirabilis* in rheumatoid arthritis

C Wilson, A Ebringer, K Ahmadi, J Wigglesworth, H Tiwana, M Fielder, A Binder, C Ettelaie, P Cunningham, C Joannou, S Bansal

Division of Biomolecular Sciences, Immunology Section, King's College, London W8 7AH, United Kingdom

C Wilson
A Ebringer
K Ahmadi
J Wigglesworth
H Tiwana
M Fielder

Department of Computing
P Cunningham

Department of Pharmacy
S Bansal

Department of Rheumatology, Middlesex Hospital, UCL School of Medicine, London, United Kingdom
A Ebringer

Department of Rheumatology, Lister Hospital, Stevenage, Herts, United Kingdom
A Binder

Department of Chemistry and Biochemistry, The Royal Free Hospital, London, United Kingdom
C Ettelaie

Department of Biochemistry, UMDS School of Medicine, London, United Kingdom
C Joannou

Correspondence to:
Dr Alan Ebringer.

Accepted for publication
17 October 1994

Abstract

Objectives—To show molecular similarity between two sequences of *Proteus mirabilis* (haemolysin—ESRRAL; urease—IRRET) with HLA-DR antigens (EQRRAA) which are associated with rheumatoid arthritis (RA) and type XI collagen (LRREI), respectively; and, in patients with RA, to measure levels of antibody against a 16-mer synthetic peptide containing the ESRRAL sequence, and the haemolysin and urease proteins of *Proteus mirabilis*.

Methods—The homologous sequences EQRRAA and ESRRAL were modelled withALCHEMY III, using the crystalline structure of DRB1*0101 (HLA-DR1). Sera from 40 patients with RA, 30 with ankylosing spondylitis (AS), and 30 controls were tested against synthetic ESRRAL peptide and the haemolysin of *Proteus mirabilis* by enzyme linked immunosorbent assay. Similar tests were also carried out on sera from 20 patients with RA, 40 with AS, and 15 controls, against *Proteus mirabilis* urease.

Results—Molecular modelling of the homologous sequences ESRRAL/EQRRAA and IRRET/LRREI showed stereochemical similarities. Antibodies to the 16-mer synthetic peptide containing the ESRRAL sequence, the haemolysin, and urease proteins were significantly increased in RA patients compared with AS patients ($p < 0.001$) and healthy controls ($p < 0.001$). No such increases were observed with three control peptides including the EDERAA sequence of DRB1*0402 (HLA-DR4/Dw10), the haemolysin proteins of *Streptococcus pyogenes* and *Vibrio parahaemolyticus*, and the urease of *Bacillus pasteurii*.

Conclusion—The additive effect of the immune responses to the two *Proteus mirabilis* antigens, haemolysin (ESRRAL) and urease (IRRET), could be relevant in the aetiopathogenesis of RA.

(Ann Rheum Dis 1995; 54: 216–220)

The association between rheumatoid arthritis (RA) and some subtypes of HLA-DR4 is well

established:¹ a particular region of the DRβ1 chain, from position 70–74, encoding amino acids Gln-Arg-Arg-Ala-Ala (QRRAA) and found in DRB1*0101 (DR1) and DR4 (subtypes DRB1*0404 (Dw14) and DRB1*0405 (Dw15)), has been identified as the molecular sequence responsible for susceptibility to RA.² 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); 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 with RA. Our group has reported an amino acid homology between an outer membrane haemolysin protein of *Proteus mirabilis* and the susceptibility sequence in HLA-DR1 and DR4 subtypes associated with RA.³ Because RA is a systemic disorder and cartilage destruction is a major feature of the disease, a search of the protein database was made for any sequence homology between *Proteus mirabilis* and collagen. An amino acid homology was identified between *Proteus mirabilis* urease (IRRET), amino acid residues 337–341⁴ and α2(XI) collagen (LRREI) residues 421–425.⁵ Type XI collagen is a component of hyaline cartilage which is composed of three different polypeptide subunits α1(XI), α2(XI), α3(XI). Early studies from our laboratory showed that anti-HLA-DR4 tissue typing sera bound significantly to *Proteus mirabilis* compared with non-DR4 tissue typing sera; no such interaction was found with *Escherichia coli*.⁶ In addition, we have reported an increase in *Proteus* antibodies in RA patients⁷ which has been confirmed by two independent groups, from Dublin⁸ and Newcastle.⁹ The hexamer sequence ESRRAL and the pentamer IRRET of *Proteus mirabilis* are both hydrophilic, which suggests that they may be immunogenic.

The present study was undertaken to compare spatial configurations of EQRRAA/ESRRAL, and LRREI/IRRET sequences for similarity using molecular modelling, and to determine if antibodies to ESRRAL and to the haemolysin and urease proteins were present in RA patients.

Patients and methods

PATIENTS

Sera were collected from active (erythrocyte sedimentation rate (ESR) > 15 mm/1st h) RA patients attending the Rheumatology Department at the Lister Hospital, Stevenage, Herts, and active AS patients attending the AS Research Clinic at the Middlesex Hospital. The diagnosis of RA was according to the American Rheumatism Association criteria and that of AS by the New York criteria.

In the first study (antibodies against ESRRAL and the haemolysin protein), sera were from 40 patients with RA (10 male/30 female; mean age 59 (range 37–77) years; mean ESR 46.2 (SE 4.0) mm/1st h), 30 AS patients (23 males/seven females; mean age 49 (27–76) years; mean ESR 30.1 (5.2) mm/1st h), and 30 healthy control subjects (15 male/15 female; mean age 23 (21–27) years).

In the second study (antibodies against the urease protein), patients and controls were different from those in the first study. They comprised 20 patients with RA (five male/15 female; mean age 56 (20–75) years; mean ESR 59.8 (7.7) mm/1st h), 40 AS patients (29 males/11 females; mean age 45 (21–76) years; mean ESR 25.9 (3.6) mm/1st h), and 15 healthy control subjects (mean age 29 (21–52) years).

MOLECULAR MODELLING

Comparison of space filling models of a predicted ESRRAL sequence of *Proteus* haemolysin and EQRRAA sequence within DRB1*0101 (HLA-DR1) from known crystallographic structure was carried out to study spatial configuration. The ESRRAL model was constructed as a helix with torsional angles corresponding to those observed in the known EQRRAA sequence and the two sequences were superimposed (root mean square (RMS) = 0.046). The ESRRAL and EQRRAA sequences were also compared with the sequence EDERAA, present in DRB1*0402 (HLA-DR4/Dw10), which is not associated with RA.

Comparison of space filling models was made of the predicted IRRET of *Proteus mirabilis* urease and LRREI of $\alpha 2$ (XI) collagen. Both sequences were constructed as helices and the two sequences were superimposed (RMS = 0.011). Structures were modelled using Alchemy III (Tripos ASSOC Inc, St Louis, USA).

ELISA STUDIES WITH SYNTHETIC PEPTIDES

Peptides were prepared by solid phase synthesis and analysed for purity by high performance liquid chromatography. The test peptide was LGSISESRRALQDSQR (16-mer), which represents amino acid residues 27–42 of *Proteus mirabilis* haemolysin (hpm B);¹⁰ three control peptides, SQKDILEDERAADVTY (16-mer) DRB1*0402 (HLA-DR4/Dw10), YASGASGAS (9-mer), and DAHKSEVAHRFLDLGEENFKALVL (24-mer), were used to exclude non-specific

binding. Sera were tested against the *Proteus mirabilis* haemolysin peptide and control peptides by enzyme linked immunosorbent assay (ELISA). Briefly, polystyrene microtitre plates (Dynatech, Billingshurst, UK) were coated with the synthetic peptide (5.0 µg/well) overnight at 4°C. After absorption and washing with phosphate buffered saline (PBS)-Tween, the plates were saturated with 1% bovine serum albumin (BSA)-PBS-Tween and incubated for one hour at 37°C, followed by further washing with PBS-Tween. Serum samples (200 µl at 1/50 dilution in PBS-Tween) were added and the plates incubated for 90 minutes at 37°C, followed by washing with PBS-Tween. Peroxidase conjugated rabbit anti-human class specific IgG, IgA or IgM (DAKO Ltd, Bucks, UK) diluted 1/500 in PBS-Tween was added and the plates incubated for 90 minutes at 37°C then washed. After exposure of the enzyme to substrate (2,2'-azinobis (3-ethylbenz-thiazoline-6-sulphonic acid), Sigma Chemical Company Ltd, Dorset, UK) at room temperature for 20 minutes, the reaction was stopped with sodium fluoride and the optical density (OD) measured at wavelength 630 nm. All assays were carried out in duplicate and under code, so that the status of each serum sample under investigation was not known to the tester.

PURIFICATION AND ASSAY OF THE 63 KDA OUTER MEMBRANE HAEMOLYSIN PROTEIN

Purification of the 63 kDa outer membrane haemolysin protein was carried out as originally described for *E. coli*.¹¹ Sera obtained from the same subjects as above, were tested against *Proteus mirabilis* haemolysin (63 kDa) protein, and against two control haemolysin proteins, obtained from *Streptococcus pyogenes* and *Vibrio parahaemolyticus* (Sigma Chemical Company Ltd) by ELISA. Briefly, polystyrene microtitre plates (Dynatech) were coated with the haemolysin protein (2.0 µg/well) overnight at 4°C. Serum samples were diluted 1/200 in PBS-Tween and the plates saturated with 0.1% BSA-BPS-Tween. The remainder of the assay procedure was as previously described for the synthetic peptides.

PURIFICATION AND ASSAY OF 280 KDA UREASE PROTEIN

Purification of the 280 kDa urease protein was by the method of Jones and Mobley.¹² Sera obtained from different subjects than in the first study were tested against *Proteus mirabilis* urease protein and *Bacillus pasteurii* urease (Sigma Chemical Company) by ELISA. Polystyrene microtitre plates (Dynatech), were coated with the urease protein (2 µg/well) and assays carried out as previously described.

Results

Molecular modelling of the two structures ESRRAL and EQRRAA showed a common surface of homologous residues. The positions

32, 34, 35 and 69, 71, 72 form a motif common to HLA-DR and *Proteus mirabilis* occupying the same stereochemical space. However, the EDERAA sequence of DRB1*0402 (HLA-DR4/Dw10) which is not associated with RA, differed from both the ESRRAL and EQRRAA motifs (fig 1). Molecular modelling of the two sequences IRRET and LRREI also showed a common surface of homologous residues. The positions 338, 339, 340 and 421, 422, 423, form a motif common to type XI collagen and *Proteus mirabilis* urease occupying the same stereochemical space (fig 1). The common surface of homologous residues observed in the sequences ESRRAL/EQRRAA

and IRRET/LRREI, which may be involved in immune interactions, consists of two positively charged arginines and one negatively charged glutamic acid.

Antibodies to *Proteus mirabilis* haemolysin peptide LGSISESRRALQDSQR of the IgG class were significantly increased in RA patients (mean 0.263 (SE 0.011) OD units) compared with AS patients (0.162 (0.012) OD units) ($t = 5.98$, $p < 0.001$) and healthy controls (0.158 (0.014) OD units (95% confidence limits (CL) of mean ± 0.026) ($t = 5.90$, $p < 0.001$) (fig 2). The difference between the titres in AS patients and healthy controls tested against the haemolysin peptide was not

Figure 1 Comparison of space filling models. A: ESRRAL sequence of *Proteus mirabilis* haemolysin predicted from known crystallographic structure; B: EQRRAA sequence within DRB1*0101 (HLA-DR1), predicted from known crystallographic structure; C: EDERAA sequence of DRB1*0402 (HLA-DR4/Dw10); D: predicted LRREI sequence of $\alpha 2$ (XI) collagen; E: predicted IRRET sequence of *Proteus mirabilis* urease.

Figure 2 Antibody titres for IgG immunoglobulin in controls (C) and patients with ankylosing spondylitis (AS) or rheumatoid arthritis (RA) when tested by ELISA against the synthetic peptides, EDERAA, ESRRAL and the haemolysin protein of *Proteus mirabilis*. OD = Optical density. Bars indicate means. NS = Not significant compared with controls.

significant. There was no increase in IgA and IgM antibodies against the haemolysin peptide, and no significant reactivity by the RA patients against the DRB1*0402 (HLA-DR4/Dw10) EDERAA peptide (fig 2) or the other two control peptides used in this study (data not shown).

Titres of antibody to *Proteus mirabilis* (63 kDa) protein of the IgG class were significantly increased in RA patients (0.304 (0.013) OD units) compared with AS patients (0.190 (0.075) OD units) ($t = 7.11$, $p < 0.001$) and healthy controls (0.201 (0.014) OD units) (95% CL ± 0.028) ($t = 5.52$, $p < 0.001$), but there was no significant difference between the titres in AS patients and controls (fig 2). In contrast, there was no increase in IgA and IgM antibodies against the haemolysin protein, and no significant reactivity by RA patients against the two control *Streptococcus pyogenes* and *Vibrio parahaemolyticus* haemolysin proteins (data not shown).

Antibody titres against *Proteus mirabilis* urease protein of the IgG class were significantly increased in patients with RA (0.68 (0.07) OD units) compared with AS patients (0.39 (0.02) OD units) ($t = 5.09$, $p < 0.001$) and healthy controls (0.32 (0.04) OD units) ($t = 4.06$, $p < 0.001$), but there was no significant difference between the titres in AS patients and controls, and no significant reactivity by RA patients against *Bacillus pasteurii* urease (fig 3).

Figure 3 Antibody titres for IgG immunoglobulin in controls (C) and patients with ankylosing spondylitis (AS) or rheumatoid arthritis (RA) when tested by ELISA against the urease proteins of *Proteus mirabilis* and *Bacillus pasteurii*. OD = Optical density. Bars indicate means. NS = Not significant compared with controls.

Discussion

In this study, we have demonstrated for the first time that patients with RA have an increased titre of antibodies against a synthetic peptide derived from *Proteus mirabilis* membrane haemolysin (hpm B), containing the hexamer sequence ESRRAL which is homologous with the RA susceptibility sequence EQRRRA of the DRβ1 chain of class II major histocompatibility complex molecules and with the native haemolysin protein (fig 2). The ESRRAL motif was found in three of 67 000 sequences: *Proteus mirabilis*, *Serratia marcescens*, and *Vibrio cholerae*. However, we were unable to find any significant increase in antibodies to *Serratia* in patients with RA (data not shown). This demonstration of a significant increase in IgG, but not IgA, antibodies to *Proteus mirabilis* membrane haemolysin sequence and native protein suggests that the site of infection in RA patients is not at a mucosal surface involving IgA, but in some interstitial tissue, leading to IgG production. In a related study, auto-antibodies against a 16-mer synthetic peptide of DRB1*0405 (HLA-DR4/Dw15) β1 chain, containing the EQRRRA sequence, were reported to be increased in Japanese patients with RA, when tested by ELISA.¹³ We have also demonstrated that RA patients have an increased titre of antibodies to the urease protein derived from *Proteus mirabilis*, containing a sequence IRRET which is homologous with a sequence LRREI present in type XI collagen. The IRRET motif was found in six microbial proteins (*Bacillus sphaericus*, murine leukaemia virus, *Escherichia coli*, *Klebsiella aerogenes*, *Proteus vulgaris*, *Proteus mirabilis*) out of 30 000

sequences examined in the protein database. The high level of *Proteus mirabilis* antibodies was not attributable to non-specific effects of inflammation, because the AS patients also had increased ESR values, yet their levels of *Proteus mirabilis* antibodies were similar to those found in healthy control subjects.

The ESRRAL and IRRET sequences are both hydrophilic, and in helical configurations where they may be involved in immune interactions. It has been reported that *Proteus mirabilis* may be isolated more frequently from RA patients than from non-RA patients or healthy controls,¹⁴ although some groups could not find such increased *Proteus mirabilis* isolation.¹⁵ The results of this study suggest that RA patients have a significant increase in antibodies to a *Proteus* membrane haemolysin sequence containing a hexamer which is homologous with the RA susceptibility sequence, and this could provide an explanation for the 'shared epitope' hypothesis. Moreover, our RA patients also had a significant increase in antibodies to *Proteus mirabilis* urease containing a pentamer sequence which is homologous with a sequence in type XI collagen. Further investigations are required to determine if antibodies to ESRRAL/IRRET and haemolysin/urease will react with HLA alleles and collagen to produce tissue damage, thereby contributing to the pathogenesis of this disease.

We thank the Trustees of the Middlesex Hospital for their support. MF Acknowledges the support of the SERC.

- 1 Nepom G T, Byers P, Seyfried C E, et al. HLA genes associated with rheumatoid arthritis. Identification of susceptibility alleles using oligonucleotide probes. *Arthritis Rheum* 1989; 33: 15-21.

- 2 Watanabe Y, Tokunaga K, Matsuki K, et al. Putative amino acid sequence of HLA DRB chain contributing to rheumatoid arthritis susceptibility. *J Exp Med* 1989; 169: 2263-8.
- 3 Ebringer A, Cunningham P, Ahmadi K, Wigglesworth J, Hosseini R, Wilson C. Sequence similarity between HLA-DR1 and DR4 subtypes associated with rheumatoid arthritis and *Proteus/Serratia* membrane haemolysins. *Ann Rheum Dis* 1992; 51: 1245-6.
- 4 Jones B D, Mobley H L T. *Proteus mirabilis* urease: Nucleotide sequence determination and comparison with jack bean urease. *J Bacteriol* 1989; 171: 6414-22.
- 5 Kimura T, Cheah K S E, Chan S D H, et al. The human $\alpha 2$ (XI) collagen (COL1A2) chain molecular cloning of cDNA and genomic DNA reveals characteristics of a fibrillar collagen with differences in genomic organisation. *J Biol Chem* 1989; 264: 13910-6.
- 6 Ebringer A, Khalafpour S, Wilson C. Rheumatoid arthritis and *Proteus*: a possible aetiological association. *Rheumatol Int* 1989; 9: 223-8.
- 7 Ebringer A, Ptaszynska T, Corbett M, et al. Antibodies to *Proteus* in rheumatoid arthritis. *Lancet* 1985; ii: 305-7.
- 8 Rogers P, Hassan J, Bresnihan B, Feighery C, Whelan A. Antibodies to *Proteus* in rheumatoid arthritis. *Br J Rheumatol* 1988; 27 (suppl 2): 90-4.
- 9 Deighton C M, Gray J W, Bint A J, Walker D J. Specificity of the *Proteus* antibody response in rheumatoid arthritis. *Ann Rheum Dis* 1992; 51: 1242-4.
- 10 Uphoff T S, Welch R A. Nucleotide sequencing of the *Proteus mirabilis* calcium independent hemolysin genes (*hpm A* and *hpm B*) reveals sequence similarity with the *Serratia marcescens* hemolysin genes (*shl A* and *shl B*). *J Bacteriol* 1990; 172: 1206-15.
- 11 Schnaitman C. Outer membrane proteins of *Escherichia coli* O111. Evidence that the major protein of *Escherichia coli* O111 outer membrane consists of four distinct polypeptide species. *J Bacteriol* 1974; 118: 442-53.
- 12 Jones B D, Mobley H L T. Genetic and biochemical diversity of ureases of *Proteus*, *Providencia*, and *Morganella* species isolated from urinary tract infection. *Infect Immun* 1987; 55: 2198-203.
- 13 Takeuchi F, Kosuge E, Matsuta K, et al. Antibody to a specific HLA DR $\beta 1$ sequence in Japanese patients with rheumatoid arthritis. *Arthritis Rheum* 1990; 33: 1867-8.
- 14 Ebringer A, Wilson C, Ahmadi K, Corbett M, Rashid T, Shipley M. Rheumatoid arthritis as a reactive arthritis to *Proteus* infection: prospects for therapy. In: Machtey I, ed. *Progress in Rheumatology*, vol 5. Petach Tiqva, Israel: Rheumatology Service, Hasharon Hospital, Golda Medical Centre, 1993; 77-83.
- 15 McDonagh J, Gray J, Sykes H, Walker D J, Bint A S, Deighton C M. Anti-*Proteus* antibodies and *Proteus* organisms in rheumatoid arthritis: A clinical study. *Br J Rheumatol* 1994; 33: 32-5.