

## Commentary

# Citrullination, a possible functional link between susceptibility genes and rheumatoid arthritis

Erik R Vossenaar, Albert JW Zendman and Walther J van Venrooij

Department of Biochemistry, University of Nijmegen, The Netherlands

Corresponding author: Erik R Vossenaar (e-mail: [e.vossenaar@ncmls.kun.nl](mailto:e.vossenaar@ncmls.kun.nl))

Received: 18 Sep 2003 Accepted: 23 Oct 2003 Published: 25 Nov 2003

*Arthritis Res Ther* 2004, **6**:1-5 (DOI 10.1186/ar1027)

© 2004 BioMed Central Ltd (Print ISSN 1478-6354; Online ISSN 1478-6362)

### Abstract

Antibodies directed to citrullinated proteins (anti-cyclic citrullinated peptide) are highly specific for rheumatoid arthritis (RA). Recent data suggest that the antibodies may be involved in the disease process of RA and that several RA-associated genetic factors might be functionally linked to RA via modulation of the production of anti-cyclic citrullinated peptide antibodies or citrullinated antigens.

**Keywords:** anti-cyclic citrullinated peptide autoantibodies, citrullination, genetic susceptibility, peptidylarginine deiminase, rheumatoid arthritis

### Introduction

The serum of rheumatoid arthritis (RA) patients contains a variety of antibodies directed against self-antigens. The most widely known of these autoantibodies is the rheumatoid factor; antibodies directed against the constant domain of IgG molecules (reviewed in [1]). The rheumatoid factor can not only be detected in roughly 75% of RA patients, but also in the serum of patients with other rheumatic or inflammatory diseases, and even in a substantial percentage of the healthy (elderly) population [2]. Its presence is therefore not very specific for RA.

Autoantibodies directed against citrullinated proteins have a much higher specificity for RA (reviewed in [3]). This family of autoantibodies includes the anti-perinuclear factor, the so-called anti-'keratin' antibodies, anti-filaggrin antibodies, anti-cyclic citrullinated peptide (anti-CCP) antibodies and probably also anti-Sa antibodies (for references see [3]). These autoantibodies all recognize epitopes containing citrulline (the naming of the antibody is simply determined by the substrate used to detect them).

Because citrulline is a nonstandard amino acid, it is not incorporated into proteins during translation. It can,

however, be generated by post-translational modification (citrullination) of protein-bound arginine by peptidylarginine deiminase (PAD) (EC 3.5.3.15; reviewed in [4]) enzymes (corresponding genes are annotated as *PADI*).

Anti-citrullinated protein antibodies can be detected (with the CCP2 assay) in up to 80% of RA sera with a specificity of 98%. Besides being very specific for RA, the antibodies can be detected very early in the disease and can predict clinical disease outcome. Furthermore, the antibodies are produced locally in the inflamed synovium, suggesting that they might play a role in the disease process (for references see [3]).

Because citrullinated proteins (e.g. fibrin) have been detected in the synovium of RA patients [5], PAD enzymes must also be present. At least five isotypes of PAD exist in mammals; two of these isotypes (PAD2 and PAD4) are known to be expressed in hemopoietic cells (for references see [4]) and are expressed in the RA synovium [6]. Of special interest is the PAD4 enzyme, which is normally present in the nucleus of granulocytes and CD14<sup>+</sup> monocytes, because genetic polymorphisms in the gene encoding this enzyme are associated with RA.

Figure 1

	SNP ID*	non-susceptible	susceptible	amino acid character <i>P</i> value
1	89	<b>S</b> AGC	<b>G</b> GGC	polar → non-polar <i>P</i> =0.07
2	90	<b>A</b> GCG	<b>V</b> GTG	non-polar → non-polar <i>P</i> =0.007
3	92	<b>A</b> GCG	<b>G</b> GGG	non-polar → non-polar <i>P</i> =0.0004
4	104	<b>L</b> CTG	<b>L</b> TTG	no substitution <i>P</i> =0.0005

Summary of the four exonal single nucleotide polymorphisms (SNPs) in *PADI4*. The actual SNP is indicated in bold. The amino acid that shows most conservation with other known peptidylarginine deiminases [4] is shaded gray. SNP ID\* according to Suzuki and colleagues as *padi4\_x* [7].

### PAD4 polymorphisms are associated with RA

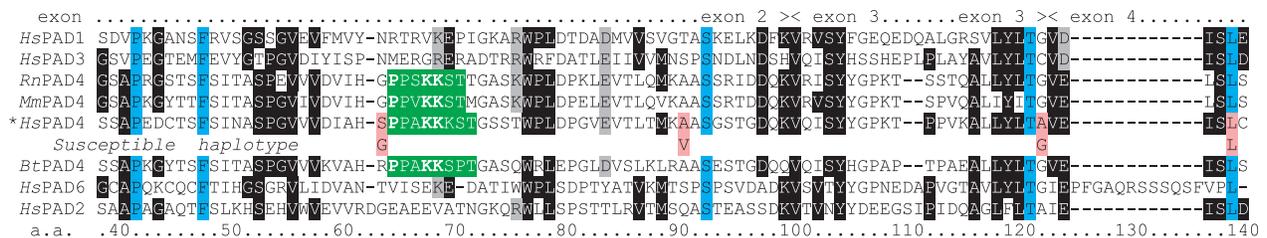
The existence of numerous single nucleotide polymorphisms (SNPs) in the *PADI* gene cluster (located on chromosome 1p36 [4]) was recently described by Suzuki and colleagues [7]. Eight of the 17 SNPs in *PADI4* were strongly associated ( $P < 0.001$ ) with RA, whereas SNPs in the other *PADI* genes were not. Because the SNPs within *PADI4* are in strong linkage disequilibrium, they segregate together in distinct haplotypes. The two most frequent haplotypes account for more than 85% of all individuals. One of these two haplotypes (referred to as the

susceptible haplotype) was more frequent in RA patients than in controls (case : control ratio = 1.28 versus 0.87 for the nonsusceptible haplotype).

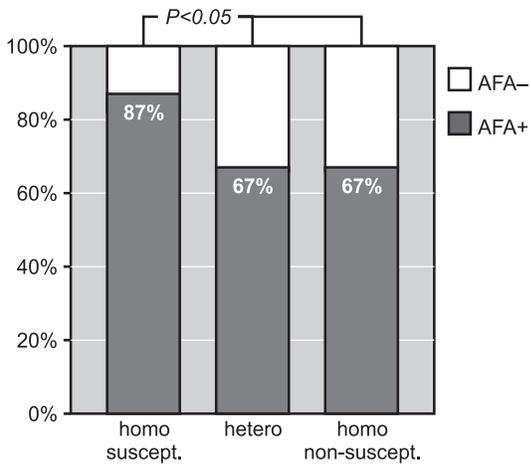
Four of the 17 SNPs in *PADI4* are located in exons of PAD4. Although three of them result in amino acid substitutions (Fig. 1), possible consequences for the function and activity of the PAD4 enzyme were not analyzed. The three SNPs leading to amino acid changes all appear at nonconserved places, as can be deduced from an alignment of PAD sequences (segment in Fig. 2; for complete alignment see [4]). The susceptible haplotype is more closely conserved to PAD4 sequences of other species (two of the three positions conserved) than the nonsusceptible haplotype (one of the three positions conserved). Interestingly, the fourth SNP, which does not lead to an amino acid substitution, is at a 100% conserved position. Only one of the three amino acid substitutions leads to a change in the electrostatic character of the residue. This SNP (*padi4\_89*) is located directly before the nuclear localization signal of PAD4 [8]. The nuclear localization signal was originally described in the nonsusceptible sequence [8]. The susceptible haplotype is conserved with the mouse sequence at this position and the mouse PAD4 also locates to the nucleus (our unpublished observations). Therefore, consequences for subcellular localization of the enzyme are not very likely. It would still be very interesting, however, to investigate possible effects of the amino acid substitutions on the functional properties of the enzyme (e.g. substrate specificity, calcium dependence, catalytic rate).

Eight of the 17 SNPs were significantly associated with RA ( $P < 0.001$ ); only two of these were exonal SNPs (*P* values presented in Fig. 1). Only one of these two SNPs

Figure 2



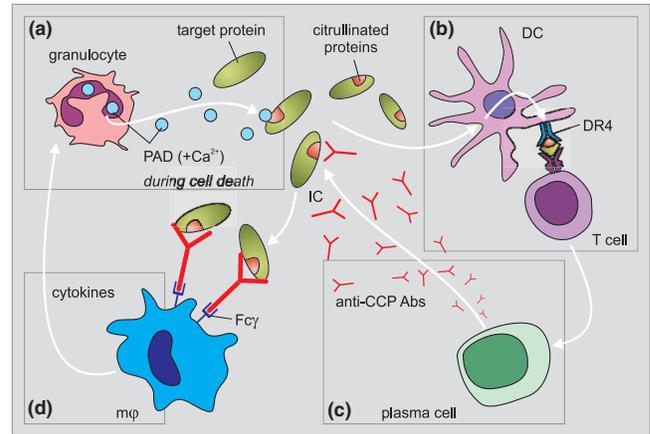
Multiple alignment of partial peptidylarginine deiminase (PAD) protein sequences based on a large full alignment described in [4] (available online: <http://www.mrw.interscience.wiley.com/suppmat/0265-9247/suppmat/2003/25/v25.1106.html>). Shown are segments of all five isotypes from the human (*Homo sapiens* [Hs], PAD1 NP\_037490, PAD2 NP\_031391, PAD3 NP\_057317, PAD4 NP\_036519 and PAD6 XP\_210118) and segments of PAD4 from the mouse (*Mus musculus* [Mm], NP\_035191), the rat (*Rattus norvegicus* [Rn], NP\_058923) and the cow (*Bos taurus* [Bt], based on BG364988). Conserved residues that are identical in more than 50% of all known PAD sequences are shaded black; fully conserved residues are shaded cyan. Conserved charged residues are also indicated (shaded light gray). Exon boundaries, based on PAD1 sequences, are annotated above the alignment. The monopartite nuclear localization signal (NLS) of PAD4 is shaded green, and conserved NLS residues are bold [8]. The four exonal single nucleotide polymorphisms are shaded pink. The nonsusceptible haplotype (S A A L) is shown in the alignment, and the susceptible (G V G L) haplotype is indicated below it. a.a., amino acid.

**Figure 3**

Correlation between the *PADI4* haplotype and autoantibodies to citrullinated proteins (anti-filaggrin antibodies [AFA]). Homozygous susceptible (homo suscept.) rheumatoid arthritis (RA) patients ( $n = 30$ ) are significantly more often AFA-positive than homozygous nonsusceptible (homo non-suscept.) RA patients ( $n = 33$ ) or heterozygous (hetero) RA patients ( $n = 66$ ) [7].

(*padi4\_92*) results in an amino acid substitution. Next to possible effects on 'protein character', the SNPs could influence mRNA stability or maturation (the SNPs most strongly associated with RA were located in introns of *PADI4*). Suzuki and colleagues measured the mRNA stability *in vitro* and showed that stability of susceptible transcripts is indeed higher (approximately threefold) than that of nonsusceptible transcripts [7]. They did not, however, investigate differences in PAD4 mRNA and protein levels between individuals with the susceptible haplotype versus those with the nonsusceptible haplotype. Nevertheless, Suzuki and colleagues hypothesize that the increased stability of the PAD4 mRNA may lead to more PAD4 enzyme being produced, and subsequently to an increased production of citrullinated proteins that serve as autoantigens. Their hypothesis is supported by the observation that RA patients homozygous for the susceptible haplotype frequently have significantly more antibodies to citrullinated proteins (87% versus 67%,  $P < 0.05$ ; Fig. 3) compared with heterozygous or homozygous nonsusceptible RA patients. Obviously, these PAD4 SNPs have functional effects *in vivo*.

The existence of polymorphisms in exons and in the 5' and 3' regions of PAD4 (designated in this reference with the old name PAD5) has also been reported by Caponi and colleagues [9]. One haplotype was more frequent in RA patients compared with controls (38% versus 17%,  $P < 0.007$ ) and appeared to be associated with the presence of antibodies to citrullinated proteins (anti-'keratin' antibodies) [9].

**Figure 4**

Possible links between rheumatoid arthritis (RA) specific anti-cyclic citrullinated peptide (anti-CCP) antibodies and RA-associated genetic factors (see text for details). (a) *PADI4* single nucleotide polymorphisms (SNPs) may lead to elevated PAD4 expression and to increased citrullination of proteins [7]. (b) RA-associated HLA-DR4 molecules (DR4) can bind and present citrullinated peptides much more efficiently than noncitrullinated peptides [17]. (c) IL-10 promoter SNPs are associated with increased anti-CCP antibody production and severity of the disease [19]. (d) Various cytokine polymorphisms are associated with RA and may lead to stronger effects of immune complex activated cells. Abs, antibodies; DC, dendritic cell; Fc $\gamma$ , Fc $\gamma$  receptor; IC, immune complex; m $\phi$ , macrophage; PAD, peptidylarginine deiminase.

### Genetic risk factors: A + B + C + D + ...

RA is a multifactorial disease and genetic risk factors are estimated to account for roughly 50% of the etiology [10]. The rest can be attributed to environmental factors, such as infectious agents, oral contraceptives and smoking [11]. Although many susceptibility loci have been found [12], well-defined functional effects of such RA-associated genetic factors have only very recently been described. The model in Fig. 4 shows how several independently described genetic risk factors for (severe) RA might be functionally linked to the production or effects of anti-CCP antibodies.

**A** SNPs in the gene for PAD4 cause increased mRNA stability of the susceptible transcript as described above. This might lead to increased levels of PAD4 enzyme (Fig. 4a).  $\text{Ca}^{2+}$  is needed for activity of PAD but, because normal intracellular  $\text{Ca}^{2+}$  levels are much too low for enzymatic activity (required concentration,  $>10^{-5}$  M; intracellular concentration,  $\sim 10^{-7}$  M), PAD enzymes are normally inactive. Only when control of calcium homeostasis is lost (e.g. during cell death or terminal differentiation) do the PAD enzymes become activated. Increased amounts of PAD may lead to increased citrullination of proteins [7]. When dying cells are not efficiently cleared (e.g. due to massive cell

death or defects in clearing machinery [13]) this could lead to exposure of the citrullinated proteins to the immune system. Citrullinated proteins may not be recognized as 'self' because they have been post-translationally modified, which has consequences for their charge and their structure [4,14]. Many known autoantigens become modified during cell death and, in particular, during apoptosis (for an overview see [15]).

**B** Correlation between RA and certain human leukocyte antigen haplotypes (e.g. HLA-DR4 [HLA-DRB1\*0401 and HLA-DRB1\*0404]) has been known for more than 25 years [16]. Recent molecular modeling data indicate that peptides containing citrulline, but not the corresponding arginine variant of the peptide, can efficiently be bound by HLA-DRB1\*0401 major histocompatibility complex molecules [17] (Fig. 4b). This citrulline-specific interaction might be the basis of a citrulline-specific immune response. T-cell proliferation assays with HLA-DRB1\*0401 transgenic mice showed that stimulation with citrullinated peptides, but not with the corresponding arginine peptides, induced proliferation and activation of T cells [17]. Although there is no absolute requirement for HLA-DR4 in order to develop anti-CCP antibodies, there is a strong correlation between HLA-DR4 status and anti-CCP positivity in RA patients [18].

**C** A specific SNP in the IL-10 promoter (-2849[AG/GG]) is associated with high IL-10 production [19]. IL-10 is a pleiotropic cytokine with many anti-inflammatory functions, but it can also stimulate inflammation by enhancing B-cell proliferation, differentiation and antibody production. Anti-CCP-positive RA patients with the 'high IL-10 haplotype' have significantly higher anti-CCP titers and more severe erosions than anti-CCP-positive patients with a 'low IL-10 haplotype' [19] (Fig. 4c). The anti-CCP antibodies that are locally produced in the inflamed synovium [20] will form immune complexes with locally produced citrullinated proteins [5]. Higher titers of the anti-CCP antibodies allow the formation of more immune complexes, which can be bound by inflammatory cells via their Fcγ receptors. This will activate these cells and cause the release of extra proinflammatory cytokines.

**D** Various polymorphisms in proinflammatory cytokines and their receptors (for references see [21,22]) are thought to be associated with RA (Fig. 4d). These genetic factors cause the release of larger amounts of cytokines upon stimulation or cause cells to be more sensitive towards these cytokines. The cytokines are the motor of the inflammation, causing influx and activation of more inflammatory cells. These cells will eventually die, allowing their PAD enzymes to become activated by influx of  $Ca^{2+}$ . With this the cycle is com-

plete and will continue if not stopped. The cycle will ultimately lead to the chronic inflammatory disease we call RA.

Besides these genetic factors, other susceptibility loci might also be involved. Their precise nature needs to be clarified in order to understand their possible role in the triggering or progression of RA.

## Concluding remarks

Recent literature on anti-CCP antibodies (reviewed in [3]) suggests that the antibodies might be involved in the disease process of RA. The antibodies are very specific for the disease, they are present very early in the disease and their presence is correlated with a more severe disease outcome. Anti-CCP antibodies and citrullinated antigens are also both produced at the site of inflammation. Furthermore, drops in anti-CCP titers during rituximab therapy or infliximab therapy are correlated with clinical improvement [23] (G Valesini, personal communication, 2003).

The very interesting study by Suzuki and colleagues [7], showing an association of *PADI4* genetic polymorphisms with RA underlines the relationship between citrullination and RA. Their study, however, leaves open some intriguing research questions. What are the effects of the amino acid substitutions on the enzymatic function of PAD? What are the effects on PAD enzyme levels *in vivo*? How are these *PADI4* SNPs distributed in a non-Japanese population? The answers to these and other questions will undoubtedly give a better insight in the etiology of this enigmatic disease.

## Competing interests

None declared.

## References

1. Mageed RA: **The RF antigen**. In *Manual of Biological Markers of Disease*. Edited by van Venrooij WJ, Maini RN. Dordrecht: Kluwer Academic; 1996:Section B1.1, 1-27.
2. Lisse JR: **Does rheumatoid factor always mean arthritis?** *Postgrad Med* 1993, **94**:133-134, 139.
3. Vossenaar ER, van Venrooij WJ: **Anti-CCP antibodies, a specific marker for (early) rheumatoid arthritis**. *Clin Appl Immunol Rev*, in press.
4. Vossenaar ER, Zendman AJW, van Venrooij WJ, Puijck G: **PAD, a growing family of citrullinating enzymes: genes, features and involvement in disease**. *BioEssays* 2003, **25**:1106-1118.
5. Masson-Bessière C, Sebbag M, Girbal-Neuhausser E, Nogueira L, Vincent C, Senshu T, Serre G: **The major synovial targets of the rheumatoid arthritis-specific antifilaggrin autoantibodies are deaminated forms of the alpha- and beta-chains of fibrin**. *J Immunol* 2001, **166**:4177-4184.
6. Chapuy-Ragaud S, Sebbag M, Nachat R, Baeten D, Foulquier V, Simon M, Senshu T, Yamada M, Takahara H, De Keyser F, Serre G: **Peptidylarginine deiminase isoforms expressed in the synovial membrane of rheumatoid arthritis patients [abstract]**. *Arthritis Res Ther* 2003, **5**(Suppl 1):S2.
7. Suzuki A, Yamada R, Chang X, Tokuhira S, Sawada T, Suzuki M, Nagasaki M, Nakayama-Hamada M, Kawaida R, Ono M, Ohtsuki M, Furukawa H, Yoshino S, Yukioka M, Tohma S, Matsubara T, Wakitani S, Teshima R, Nishioka Y, Sekine A, Iida A, Takahashi A, Tsunoda T, Nakamura Y, Yamamoto K: **Functional haplotypes of**

- PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis.** *Nat Genet* 2003, **34**:395-402.
8. Nakashima K, Hagiwara T, Yamada M: **Nuclear localization of peptidylarginine deiminase V and histone deimination in granulocytes.** *J Biol Chem* 2002, **277**:49562-49568.
  9. Caponi L, Petit-Teixeira E, Sebbag M, Bongioni F, Moscato S, Pratesi F, Osorio J, Guerrin-Weber M, Cornelis F, Serre G: **Analysis of the peptidylarginine deiminase V gene in rheumatoid arthritis [abstract].** *Arthritis Res Ther* 2003, **5(Suppl 1)**:S1-S2.
  10. MacGregor AJ, Snieder H, Rigby AS, Koskenvuo M, Kaprio J, Aho K, Silman AJ: **Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins.** *Arthritis Rheum* 2000, **43**:30-37.
  11. Gabriel SE: **The epidemiology of rheumatoid arthritis.** *Rheum Dis Clin North Am* 2001, **27**:269-281.
  12. Cornelis F, Faure S, Martinez M, Prud'homme JF, Fritz P, Dib C, Alves H, Barrera P, de Vries N, Balsa A, Pascual-Salcedo D, Maenaut K, Westhovens R, Migliorini P, Tran TH, Delaye A, Prince N, Lefevre C, Thomas G, Poirier M, Soubigou S, Alibert O, Lasbleiz S, Fouix S, Weissenbach J: **New susceptibility locus for rheumatoid arthritis suggested by a genome-wide linkage study.** *Proc Natl Acad Sci USA* 1998, **95**:10746-10750.
  13. Rodenburg RJ, Raats JM, Pruijn GJ, van Venrooij WJ: **Cell death: a trigger of autoimmunity?** *BioEssays* 2000, **22**:627-636.
  14. Tarcsa E, Marekov LN, Mei G, Melino G, Lee SC, Steinert PM: **Protein unfolding by peptidylarginine deiminase. Substrate specificity and structural relationships of the natural substrates trichohyalin and filaggrin.** *J Biol Chem* 1996, **271**:30709-30716.
  15. Utz PJ, Gensler TJ, Anderson P: **Death, autoantigen modifications, and tolerance.** *Arthritis Res* 2000, **2**:101-114.
  16. Silman AJ, Pearson JE: **Epidemiology and genetics of rheumatoid arthritis.** *Arthritis Res* 2002, **4(Suppl 3)**:S265-S272.
  17. Hill JA, Southwood S, Sette A, Jevnikar AM, Bell DA, Cairns E: **Cutting edge: the conversion of arginine to citrulline allows for a high-affinity peptide interaction with the rheumatoid arthritis-associated HLA-DRB1\*0401 MHC class II molecule.** *J Immunol* 2003, **171**:538-541.
  18. Goldbach-Mansky R, Lee J, McCoy A, Hoxworth J, Yarbboro C, Smolen JS, Steiner G, Rosen A, Zhang C, Menard HA, Zhou ZJ, Palosuo T, van Venrooij WJ, Wilder RL, Klippel JH, Schumacher HRJ, El-Gabalawy HS: **Rheumatoid arthritis associated autoantibodies in patients with synovitis of recent onset.** *Arthritis Res* 2000, **2**:236-243.
  19. Lard LR, van Gaalen FA, Schonkeren JJM, Pieterman EJ, Stoeken G, Vos K, Nelissen RGHH, Westendorp RGJ, Hoeben RC, Breedveld FC, Toes REM, Huizinga TWJ: **Association of the -2849 interleukin-10 promoter polymorphism with autoantibody production and joint destruction in rheumatoid arthritis.** *Arthritis Rheum* 2003, **47**:1841-1848.
  20. Masson-Bessière C, Sebbag M, Durieux JJ, Nogueira L, Vincent C, Girbal-Neuhauser E, Durroux R, Cantagrel A, Serre G: **In the rheumatoid pannus, anti-filaggrin autoantibodies are produced by local plasma cells and constitute a higher proportion of IgG than in synovial fluid and serum.** *Clin Exp Immunol* 2000, **119**:544-552.
  21. Klareskog L, Lorentzen J, Padyukov L, Alfredsson L: **Genes and environment in arthritis: can RA be prevented?** *Arthritis Res* 2002, **4(Suppl 3)**:S31-S36.
  22. Huizinga TW: **Genetics in rheumatoid arthritis.** *Best Pract Res Clin Rheumatol* 2003, **17**:703-716.
  23. Cambridge G, Leandro MJ, Edwards JC, Ehrenstein MR, Salden M, Bodman-Smith M, Webster AD: **Serologic changes following B lymphocyte depletion therapy for rheumatoid arthritis.** *Arthritis Rheum* 2003, **48**:2146-2154.

## Correspondence

Erik R Vossenaar, 161 Department of Biochemistry, PO Box 9101, 6500 HB Nijmegen, The Netherlands. Tel: +31 243613651; fax +31 243540525; e-mail: e.vossenaar@ncmls.kun.nl