**PATHOGENESIS OF RHEUMATOID ARTHRITIS**

Jonathan CW Edwards  
Bloomsbury Rheumatology Unit  
University College London

- The epidemiology of rheumatoid arthritis (RA) implicates genes, smoking and chance

- Rheumatoid factor (RF), major histocompatibility complex (MHC) class II and tumour necrosis factor alpha (TNFα) remain the key clues to pathogenesis

- RF production requires class II and T-cells but does not require T-cell autoreactivity

- Current methods do not reveal which T-cells are involved

- Immune complex/Fc gamma receptor type IIIa (FcγRIIIa)-induced TNFα release offers a unifying effector mechanism

- ‘Disease memory’ may be carried by T-, B- or plasma cells

**WHAT CAUSES RHEUMATOID ARTHRITIS?**

**Genes**

The monozygotic twin of a subject with rheumatoid arthritis (RA) has a 25% chance of sharing disease. Two X chromosomes and major histocompatibility complex (MHC) class II account for part of this risk, but other genes contribute, some of which may soon be identified. Within class II there may be more than one factor. Part of the risk seems to relate to a specific amino acid sequence – the ‘shared epitope’ on DR0404, 0401 and 0101 – but part does not.

Female predominance may suggest that antibodies are important in pathogenesis. Those other autoimmune diseases for which there is strong evidence for antibody mediation tend to have a female bias. The class II association implicates antigen presentation to T-cells, but the specific basis for the link is unclear. The shared epitope might favour binding of certain peptides, but it is not obvious how this would promote autoimmunity. Alternatively, different class II allotypes may be expressed at different levels on different antigen-presenting cells, or otherwise vary in functional potency. Information in this area is emerging, but difficult to interpret.

**Smoking**

Evidence for an environmental factor in RA was for a long time tenuous, based largely on studies of migration from primitive to western social systems. Smoking has now emerged as a clear external risk.

Environmental factors are often seen as ‘triggers’ which determine disease onset. However, monozygotic twins rarely acquire RA at the same time and, certainly, smoking is unlikely to determine time of onset. The trigger concept may be misconceived. It may also be that it takes years for clinical disease to evolve.

**Chance**

The slow evolution of many autoimmune processes both in man and mice presents a significant mystery. The usual model of immune dynamics cannot cope. An immune re-
response takes 20 days, not 2 years. It is commonly assumed that non-genetic factors in aetiology are environmental. However, Stastny, who identified the human leucocyte antigen (HLA)-DR4 association, pointed out that the remainder of causation might be random. This idea, familiar in cancer, is often forgotten in autoimmunity, yet over 10 billion purely random events occur in the immune system every day. Neither antibodies nor T-cell receptors can be made to order; they all arise by chance gene recombination or mutation.

The lack of a link with external events of the sort seen in Reiter’s syndrome points strongly to a chance element in RA, as does the continued rise in incidence with age. Another factor pointing to random events is the timeframe of RA. Alterations in immunoglobulins, including rheumatoid factor (RF), may precede symptoms by many years. Remissions after gold or cytolytic therapies may last for a year or more before relapse occurs. It seems that, whether or not external factors are involved, the immune response in RA can take years to evolve.

The mathematical model that may fit best is that used in genetically determined cancer. Disease arises through a sequence of very infrequent chance events with early events (in cancer a deoxyribonucleic acid (DNA) repair defect) facilitating further events. What such random events might be in RA is not known but one of the obvious candidates in adults is physiological immunoglobulin gene mutation leading to new antibody species. Such mutations are commonplace, so there must be something very unusual about antibodies involved in disease evolution. An emerging possibility is that they are unusual in being both antibodies and immunomodulators, thereby entraining a cascade of pathogenesis.

**WHAT GENERATES INFLAMMATION?**

**Macrophage cytokines**

Therapeutic neutralisation of tumour necrosis factor alpha (TNFα) abrogates many features of RA. Blockade of interleukins (IL)-1 and 6 also appears to provide benefit. It seems likely that inflammation in RA is mediated primarily by macrophage cytokine production. The potency of TNFα neutralising agents supports the concept that TNFα is the crucial cytokine, but others may play both parallel and amplifying roles.

What has been more difficult to establish is the cause of macrophage activation. T-cell cytokines have been considered likely candidates on the basis that T-cells are present in RA synovium and that several animal models of RA can be transferred by T-cells. However, there is still doubt over a role for T-cells in initiation of synovitis. Moreover, despite being T-cell-transferable, there is now evidence that in many animal models inflammation is generated via antibody.

Macrophage immigration and activation (once misnamed synovial hyperplasia) is present at the earliest stages of both synovitis and nodules. Although definitive evidence is hard to acquire, repeated observations suggest that it precedes T-cell infiltration. When T-cells do arrive they settle about 100 microns away from the synovial intima and nodule palisade. As highlighted by Firestein and Zvaifler, γ-interferon levels are rather low and the pattern of intimal macrophage activation, with low levels of CD64, is not that expected from T-cell signals. Within subintima, where T-cells accumulate, macrophages do show high levels of CD64, and contact-dependent signals from T-cells to macrophages in this region may well provide important amplification once cells have accumulated. However, they do not explain the initial activation of the intimal layer.

In the absence of T-cells, the likely signal would seem to be antibody, but antibody with the ability to selectively activate macrophages in the intimal layer. This could conceivably be a tissue-domain-specific autoantibody, but this idea is problematic, particularly for extra-articular
lesions like pericarditis. The answer remains open but a likely candidate has recently re-emerged, having been first brought to light in 1983 by Nardella et al – the small immune complex.27

**Immune complex/Fc gamma receptor interactions**

In the 1970s immune complex-induced inflammation was seen in terms of deposition of large complexes in basement membranes and complement activation. However, Nardella et al demonstrated an alternative mechanism whereby soluble complexes activated macrophages through immunoglobulin G (IgG) Fc receptors (FcγR) to produce monokines.27 At that time Fc receptors were poorly characterised and, for reasons that are not entirely clear, the idea was not pursued for some time.

New information on the distribution and function of specific human FcγR rekindled interest in immune complex/FcγR interactions in the mid-1990s.28-30 It has also subsequently emerged that many animal models of RA are dependent on FcγR for the genesis of inflammation.20,21 Even in the T-cell receptor transgenic mouse developed by Matsumoto et al,21 which appears to be a pure case of T-cell autoreactivity, inflammation is mediated by antibody and FcγR.

Of the three types of FcγR capable of generating pro-inflammatory cytokines, there are strong indications that FcγRIIIa is most likely to be involved in the generation of inflammation in RA. It is selectively expressed in those tissues in which macrophage activation is seen in RA; synovial intima, serosae, alveoli, sclera, salivary gland, lymphoid tissue, bone marrow, Kupffer cells and subcutaneous tissues at sites of nodules.28 It is the only receptor apparently capable of generating TNFα when only two or three receptors are cross-linked, indicating that it is specifically involved in responses to very small complexes.29

**Rheumatoid factor**

RF is neither specific for, nor always present in, RA.31 This has led to doubts about its relevance to pathogenesis. However, the same applies to many other antibodies likely to have pathogenic roles. The reason probably relates to the difficulty of devising an in vitro assay which directly reflects pathogenic activity in vivo.

What seems reasonably clear is that pentameric IgM-RF, at least in the circulation, is not pro-inflammatory. When complexed with IgG, pentameric IgM-RF will fix complement and be cleared via complement receptors on erythrocytes. Circulating IgM-RF is quite frequently seen in healthy people following immunisation, or those with chronic infection, without any sign of joint disease. RA subjects differ from normals not so much in having RF, but in having RF of all isotypes. IgG- and IgA-RF are largely restricted to RA.35 Small complexes of self-associated IgG-RF are measurable in the circulation of approximately 50% of RA subjects.30 These complexes are theoretically ideally suited to the generation of cytokines via FcγRIIIa.29,50 They are too small to fix complement52 and will therefore tend to evade clearance. They are also small enough to pass through endothelium and access tissue macrophages. IgA-RF may also form small complexes with IgG molecules, potentially capable of cross-linking FcγR. Intriguingly, Xu et al have shown that up to 20% of IgM in RA subjects is in a monomeric form which might allow IgM-RF to form complexes with similar properties.33

There have been major difficulties in correlating immune complex levels to clinical disease in RA. This may in part reflect problems with identifying those complexes capable of generating cytokine release. Much of the work on immune complexes used complement in methods of detection and it is now clear that this is inappropriate. A further difficulty with trying to correlate complex levels with disease is that complexes may be sequestered in synovium. RF-secreting plasma cells are known to colonise RA synovium. Complex-forming IgG-RF secreted locally may drive macrophage activation without significant levels of antibody appearing in the circulation, which may partly explain seronegative RA.

**The synovial milieu**

Although expression of FcγRIIIa provides a mechanism for targeting synovium in immune complex disease, other factors almost certainly make synovium susceptible to inflammation. The arthritis of TNFα transgenic mice illustrates the sensitivity of synovial fibroblasts to cytokines in terms of expression of a range of molecules including vascular cell adhesion molecule (VCAM)-1, metalloproteinases and prostanoids.34,35 VCAM-1 is of key importance in B-cell and plasma cell survival.35 Many of the features of synovial fibroblasts are shared with bone marrow cells, both arising from perichondrial cells in early development. Synovium precedes bone marrow in evolution, suggesting that bone marrow has ‘borrowed’ synovial features and used them for haemopoietic purposes.

The specialised nature of synovial fibroblasts may explain the tendency for lymphoid follicles to develop in chronic synovitis, allowing the immune response to be amplified locally.36 This ectopic lymphoid tissue may also take part in responses to exogenous agents such as viruses. A significant part of the problem in established RA may simply be the presence of active lymphoid
tissue, with concomitant cytokine production, within a moving joint.

Although it has been suggested that synovial fibroblasts in RA may take on unusual properties and become autonomous in some way, most of the features seen in these cells (such as VCAM-1 expression) are consistent with the known hyper-responsiveness to cytokines of synovial cells in general.

WHAT PERPETUATES DISEASE?

Although RA is sometimes seen as the conversion of an acute event into a chronic process, the weight of evidence is that RA is chronic from the outset. Histological studies indicate that the elements of early and late synovitis are much the same. Disturbances of antibody production gather momentum for years before symptoms appear. The underlying process appears to involve an immune response with antigen presentation to T-cells via class II, often production of RF of all isotypes and macrophage cytokine generation. Understanding of this process is crucially dependent on knowledge of the T-cell populations involved.

Which T-cells?

For 20 years attempts have been made to identify T-cell specificities that might drive the immune response in RA, with limited success. It may be that assessment of T-cell responses in vitro with current techniques is unreliable. Responses to almost any antigen, foreign or self, can be found under the right conditions, and controls are fraught with sampling problems. Nevertheless, it may be that looking for specific T-cell autoreactivities is the wrong approach.

A different approach is to focus on T-cells that might fuel chronic RF production (accepting that it might or might not be legitimate to extrapolate to seronegative disease). A key factor is that there do not appear to be anti-IgG T-cells in RA. This has the important implication that T-cell help for RF B-cells, by definition, breaks the rules.

The likely mechanism is illustrated by experiments by Roosnek and Lanzavecchia in which RF B-cells were shown to be able to present any antigen to T-cells by taking up antigens prebound to IgG through their surface RF. In this way RF B-cells can obtain help from any T-cell in the presence of its antigen. This can be interpreted in two ways. It might imply that RF production could be driven by T-cells responding to any self antigen. It could also imply that RF is not driven by autoreactive T-cells at all. As in chronic infection, RF production in RA may be driven by T-cells recognising foreign antigens, perhaps derived from endogenous viruses such as Epstein–Barr.

If help comes from autoreactive T-cells a number of self antigens might be implicated, including the chaperone protein BiP, the spliceosome protein RA33, collagen type II and proteins bearing citrulline residues. Antibodies to all these antigens have been described in RA. Each might represent an alternative ‘primary antigen’, to which a secondary RF response is linked. If so responses to these antigens might be expected to be mutually exclusive, as seen for nuclear antigens in systemic sclerosis. Relationships between these responses are not well known and further work would be valuable.

Some of these antibody responses may be secondary to joint inflammation and damage. This would seem quite likely for antibodies to collagen type II. It has also been noted that transgenic mice which develop arthritis as a consequence of constitutive TNFα production develop antibodies to RA33, presumably as a consequence of inflammation.

A further interpretation is that T-cell autoreactivities may be driven by signals from RF B-cells, such that if these signals are removed self-tolerance would be restored. This comes close to the situation in which T-cell help for RF B-cells is derived purely from T-cells recognising foreign antigens. In both cases the underlying question becomes the reasons for persistence of RF B-cells, other than T-cell help per se.

Self-perpetuating B-cells and plasma cells

In order to survive, most B-cells require a combination of T-cell help and a supply of antigen in the form of immune complexes bound to the complement fragment C3d. RF B-cells are unusual in that they can acquire help from T-cells not recognising IgG. They are also unusual in that any IgG-containing immune complex can potentially provide them with their antigen as part of a complex, since their antigen is IgG. This should mean that any RF B-cell can survive whenever IgG-based immune complexes are present. This would explain why IgM-RF production occurs briefly after immunisation and persistently during chronic infection.

The situation for IgG-RF secreting B-cells may be different from IgM-RF secreting B-cells in an important respect. IgG-RF secreting B-cells, or their daughter plasma cells, can generate their own IgG-based immune complexes. It is known that such complexes are present in RA joints and polymerise sufficiently to activate complement and acquire C3d. In normal individuals the absence of IgG-RF implies a powerful block to class switching from IgM-RF. However, the above argument suggests that once an IgG-RF B-cell has survived long enough to give rise to plasma cells secreting significant amounts of self-complexing IgG-RF then its survival may be secured.
indefinitely. Moreover, the signals generated by IgG-RF complexes should promote the long-term survival of B-cells committed to RF of all isotypes – as seen in RA.

This brings the argument back to rare antibody species as immunomodulators, as suggested in the first section. IgG-RF is the perfect candidate because it is its own antigen and can perturb the rules of immune complex formation. The details cannot be covered here and we are a long way from knowing precise pathways, but plausible mechanisms are readily available. Seronegative RA may be based on ‘hidden’ RF, but a variety of other complex-perturbing autoantibodies could play a similar role, including anti-C1q. Essentially these are autoantibodies capable of encouraging more autoantibodies by devious means.

An unresolved question

Current debate about RA is beginning to focus in on the cell type responsible for disease persistence. If IgG-RF B-cells, arising at random by Ig gene mutation, can self-perpetuate without requiring autoreactive T-cells, ‘disease memory’ may be seen as residing in B-cells. One possibility is that, either prior to or as a consequence of, RF B-cell expansion, abnormal T-cell responses may be established which ensure continuation of the process; the memory may be in T-cells. A further possibility is that persistence of disease is dependent on long-lived plasma cells secreting self-complexing antibody capable of supporting the survival of relatively transient B-cell clones. The cell type carrying disease memory becomes crucial with the advent of cell-specific depletion therapies.

Evidence from immunoablative therapies currently leaves this key question unanswered. Remissions of 2–3 years have been seen following either broad spectrum high dose chemotherapy (cyclophosphamide > 10 gm) or selective B-cell depletion, but lifelong remission has not reproducibly been achieved. Partial T-cell depletion has produced little clinical effect, but total T-cell ablation is not an option. High-dose chemotherapy does not kill all cells of any of the three types. Immunoglobulin levels fall rather variably and B- and T-lymphocyte levels do not fall to zero. B-cell depletion reduces circulating B-cell counts to zero but solid tissue levels are not known.

The crucial experiment has yet to be done and the model proposed remains unproven and alternative models exist. Nevertheless the achievement of substantial periods of remission following ablative therapies gives encouragement that with technical developments not only may the experiment be possible but it may also be successful for our patients.

REFERENCES


