Chapter 1

Review of literature
1.1. Rheumatoid arthritis: an overview

Rheumatoid arthritis (RA) is a chronic, systemic and autoimmune disease with a prevalence of more than 1% worldwide. This disease is characterized by inflammation of the synovial membrane of diarthrodial joints. Early indications of RA are swelling and pain of the proximal inter-phalangeal and later, the larger joints become affected, especially those of the knee, elbow and ankle (Fig. 1.1). In normal healthy joints synovial membrane (SM) is mainly comprised of one or two layers of macrophages and fibroblast like synoviocytes embedded in a matrix composed of hyaluronan (HA) and collagen. Diarthrodial joints are covered by articular cartilage which is composed of a small number of chondrocytes. During such inflammatory arthritis, SM becomes thickens (synovitis) and swollen due to the infiltration of lymphocytes and plasma cells.

Fig. 1.1 Picture showing deformities in small joints of rheumatoid arthritis patient (Top). Schematic diagram showing the comparison among normal and the joints affected with osteoarthritis and rheumatoid arthritis (below).
from the synovial blood vessels. Hyperplasia or thickening of the membrane is promoted by cytokines and growth factors released from migrating cells. The SM becomes revascularised making it redder than normal. The cytokine enriched environment produced by pro-inflammatory cytokines (IL-1α, IL-6 and TNF-α) results in the aberrant growth of complex vessels known as pannus which invades the cartilage resulting in the irreparable degradation of the articular surfaces (Fig.1.2).

![Fig. 1.2 Several cytokine pathways involved in RA](image)

Synovial fluid (SF) not only lubricates the bearing surfaces of the joints, but also permits the transport of nutrients to joint cartilage and waste products to
the lymph. The classic studies showed SF as a dialysate of blood plasma (Ropes et al., 1954) but using sensitive techniques it was found that some of it components like cytokines, (Lettesjo et al., 1998) proteases, (Gaudin et al., 1997) rheumatoid factors (Smiley et al., 1968) are locally produced by the cells in the SM. The homogeneity of the SM infiltrate in an individual patient enables the use of synovial biopsies to evaluate disease severity and response to treatment (Rooney et al., 1988). Various pathological changes in joint structure like increased blood flow and vascular permeability brings about the changes in the protein constituent of SF in inflammatory joint diseases. The normal healthy joints, SF is only present in small amount (few drops to a maximum 4 ml) which is clear, pale, viscous and a total protein content of approximately 19 mg/mL while in RA, SF is high in amount, less in viscosity and cloudy due to susupended cells with the total protein content from 38.0 to 84.0 mg/mL. The high viscosity of SF is due to the presence of HA which is a high molecular weight (2x10^6 kDa) glycosaminoglycan (GAG) secreted by synovium cells and performs a protective role in the SF of articulating joints (Ghosh, 1994). Besides blood proteins, the concentrations of great numbers of pro-inflammatory cytokines like IL-1\(\beta\), IL-6, TNF-\(\alpha\), (Manicourt et al., 2000; Ribbens et al., 2000), myeloid related proteins (Frosch et al., 2000), serum amyloid A (SAA), C-reactive protein (CRP), S100 proteins (Foell et al., 2004), matrix metalloproteinase (MMPs) (Giannelli et al., 2004) derived from synovium and migrating leukocytes, degradation peptide fragments of proteins by MMPs (Belcher et al., 1997; Mansson et al., 1997) etc increases in SF. The entry of high molecular weight proteins like haptoglobin (85 kDa), IgG (150 kDa) and fibrinogen (340 kDa) were restricted which can enter now SF due to increased blood flow and vascular permeability for these high molecular weight proteins and others to equalise the plasma proteins concentrations.
1.1.1. Pathogenesis of rheumatoid arthritis

The precise etiology of the disease is unknown however; genetic and environmental factors seem to be involved in its pathogenesis (Kim-Howard et al., 2005).

1.1.1.1. Environmental Factors

Several infectious agents such as viruses, bacteria, and fungi have long been suspected as the cause for this disease. Certain infections or factors in the environment (Edwards and Cooper, 2006) as well as smoking tobacco has been reported to increase the risk of developing rheumatoid arthritis (Criswell et al., 2006). However, none of these environmental factors have been proven as the cause for this disease.

1.1.1.2. Genetic Factors

1.1.1.2.1. The major histocompatibility complex (MHC) in susceptibility to RA

The description of the human leukocyte antigen (HLA) associations with RA over two decades ago has been a major source of support for the hypothesis that genetic factors are important for susceptibility to RA. The contribution of HLA to the overall genetic risk has been variously estimated at between 30 and 50%. Since the late 1980s, a consensus has developed around susceptibility to RA being due to a closely related set of polymorphic sequences (the ‘shared epitope’) on several different DRB1 alleles, especially certain subtypes of the DR4 and DR1 allelic families (Weyand and Goronzy, 2000). It is still unclear how this set of class II alleles operates to confer susceptibility to RA, however, popular models invoke selective peptide presentation of autoantigens, biased thymic selection toward an autoreactive T-cell repertoire, direct effects on antigen processing, and a direct role for the shared epitope itself as a nominal peptide antigen, to name a
few. The unifying concept of the shared epitope has considerable appeal for understanding the MHC class II associations with RA, but it is clearly an oversimplification to consider the shared epitope as the only relevant MHC polymorphism for susceptibility to RA. There are important haplotypic influences on the degree of risk conferred by the shared epitope alleles. Certain shared epitope alleles, such as DRB1*0401, confer much greater risk than others, such as DRB1*0101. In addition, homozygosity for particular combinations of haplotypes, such as DRB1*0401/0404, appear to confer especially high risk or influence disease severity as well as risk (Suzuki et al., 2003). Differences in female versus male risk have been described for different alleles. Some investigators have proposed a role for DQ alleles on these haplotypes, although so far no convincing population data in humans supports this.

Another recent development has been the sequencing of most of the MHC, and the realization that the ‘central’ MHC contains many genes that could be directly involved in disease risk, or might interact with DRB1 alleles to modify risk. Tumor necrosis factor (TNF) is a particularly compelling candidate because of the obvious therapeutic importance of this cytokine. Several recent reports suggest that polymorphisms in the TNF region may interact with DR alleles to modify susceptibility to RA. These studies will need to be replicated and confirmed using large numbers of patients and well matched control individuals, but they emphasize that our understanding of the MHC influences on susceptibility to RA is far from complete.

1.1.1.3. Susceptibility genes outside of the MHC

One approach to identifying genetic susceptibility alleles outside the MHC is to analyze polymorphisms in genes that can reasonably be implicated in a pathway of pathogenesis. Obviously, success depends on having a model of
pathogenesis that at the very least involves relevant biochemical pathways, and on selecting the right genes in these pathways to study. It is highly uncertain whether either of these prerequisites can be fulfilled by our current knowledge. Most candidate genes that have been addressed to date involve immune recognition, cytokines and their receptors, or genes generally thought to be involved in inflammation. Of course ‘inflammation’ entails a large number of cellular processes, so that the list of candidates can get very broad, and could reasonably extend to a large fraction of the 100,000–150,000 genes now thought to reside in the human genome.

An additional problem is that the extent of genetic polymorphism is not well defined for many candidate genes of interest, and many polymorphisms exist in untranslated regulatory portions of genes, often in the form of single nucleotide polymorphisms (SNPs) or variable numbers of tandem repeats. Usually it is unclear whether these SNPs and variable numbers of tandem repeats have any functional significance, although there are examples where transcription does appear to correlate with these types of sequence changes. Clearly, the premise underlying a candidate gene approach dictates that functional changes should be of most interest. Recent surveys indicate that SNPs within coding regions are biased towards silent, and presumably functionally irrelevant, substitutions.

This is consistent with evolutionary selection against deleterious mutations. Nevertheless, functionally irrelevant polymorphisms may be in linkage disequilibrium with other sequence differences that do affect function. This implies that a general screen of SNPs or other types of polymorphisms within and around a candidate gene can be a rational approach, regardless of their direct functional effects.
Positive associations between RA and a number of candidate genes have been reported, and a few have been replicated (Gregersen, 1999). Variable results have been reported for the T-cell receptor loci. Among the cytokines (other than TNF-α), a recent report of an association with interleukin (IL-4) is provocative. Studies of IL-1, IL-1 receptor antagonist, and IL-10 show only weak or negative associations. IL-10 is of interest because promoter polymorphisms in this gene are associated with differences in levels of transcription, and low IL-10 expressing haplotypes may influence the asthma phenotype. Other candidate genes that do not fall strictly into the immunological category include corticotrophin releasing hormone, glutathione S-transferase, and Nramp1. Overall, the evidence for associations with most of these candidates is only suggestive at best.

1.1.2. Diagnosis of rheumatoid arthritis

There is no single test which clearly diagnoses early RA. When a patient first develops joint pains, it may be difficult for a doctor to say that he definitely has RA. This is because there are many other causes of joint pains. The rheumatologist examines the patients’ joints for inflammation and deformity, the skin for rheumatoid nodules, and other parts of the body for inflammation. Disease activity score (DAS) (calculated using DAS-28 disease score calculator, (Prevoo et al., 1995) is generally used for the diagnosis of the disease on the basis of the number of tender joints involved, number of swollen joints, erythrocyte sedimentation rate (ESR) and visual analogue score (VAS) for general health as subjectively estimated by the patients. The commonly used indices to depict clinical response to therapy in RA include the American College of Rheumatology (ACR) response, Health Assessment Questionnaire (HAQ) score and Disease Activity Score (Sato et al., 2006). In time, X-rays of joints may begin
to show typical erosions (early damage) and other features of RA, makes the diagnosis more certain (Scott et al., 1984). The diagnosis will be based on the pattern of symptoms, the distribution of the inflamed joints, and the blood and x-ray findings. The distribution of joint inflammation is important to the doctor in making a diagnosis (Arnett et al., 1988). In RA, the small joints of the hands, wrists, feet, and knees are typically inflamed in a symmetrical distribution (affecting both sides of the body). When only one or two joints are inflamed, the diagnosis of RA becomes more difficult. The doctor may then perform other tests to exclude arthritis due to infection or gout. The detection of rheumatoid nodules, most often around the elbows and fingers, can help the diagnosis.

A blood test called the erythrocyte sedimentation rate (ESR rate) is a measure of how fast red blood cells fall to the bottom of a test tube. ESR is used as a crude measure of the inflammation of the joints which is usually faster during disease flares, and slower during remissions. Another blood test that is used to measure the degree of inflammation present in the body is the C-reactive protein (CRP). The RF factor, ANA, ESR and C-reactive protein tests can also be abnormal in other systemic autoimmune conditions. Therefore, abnormalities in these blood tests alone are not sufficient for a firm diagnosis of rheumatoid arthritis (Skogh et al., 2003).

Analysis of the joint fluid, in the laboratory (arthrocentesis), can help to exclude other causes of arthritis, such as infection and gout (Norberg et al., 1983). Blood tests can detect inflammation, characteristic antibodies, and anaemia which may suggest the presence of RA. These tests include those for IgM and IgG RF, anti-CCP antibodies. Presence of these characteristic autoantibodies like RF factors and anti-citrullinataed protein antibodies (ACPA) has been implicated in diseases diagnosis and are widely studied in patients with RA. ACPA against
various citrullinated autoantigens like keratin (antiperinuclear antibodies, APA), filaggrin (antifilaggarin antibodies, AFA), vimentin (antivimentin antibodies, anti SA) can be detected very early even before the onset of clinical symptoms. Protein citrullination is a post translational modification of arginine to citrulline in the presence of PAD (peptidyl arginine deaminase) enzyme (Cantaert et al., 2006; van Boekel et al., 2002). A test for anti-CCP antibodies is most helpful in looking for the cause of previously undiagnosed inflammatory arthritis when the traditional blood test for RA, i.e. RF, is not present. Citrulline antibodies have been felt to represent the earlier stages of rheumatoid arthritis in this setting (Nielen et al., 2004). Another antibody called “the antinuclear antibody” (ANA) is also frequently found in patients with RA (Chellingworth et al., 1984). We had recently reported a significant increased level of anti-MBL-antibodies in the RA patients as compared to healthy controls (Gupta et al., 2006).

1.1.3. Treatment of rheumatoid arthritis

There is no cure for RA. However, much can be done to help. The aims of treatment are to reduce pain and stiffness in affected joints, to prevent joint damage as much as possible, to minimize any disability caused by pain, joint damage, or deformity and to reduce the risk of developing associated conditions such as heart disease. During a flare-up of inflammation, if the patient rests the affected joint(s) it helps to ease pain. Special wrist splints, footwear, gentle massage, or applying heat may also help. Medicines which may be advised to ease pain and stiffness like non-steroidal anti-inflammatory drugs (NSAIDs) or 'anti-inflammatories' like aspirin and paracetamol, a course of steroid tablets such as prednisolone are good at easing pain and stiffness.

There are a number of drugs called 'disease-modifying anti-rheumatic drugs' (DMARDs) that ease symptoms and reduce the damaging effect of the
disease on the joints. It includes sulfasalazine, methotrexate, gold injections, gold tablets, penicillamine, leflunomide and hydroxychloroquine etc. These drugs work by blocking the effects of chemicals involved in causing joint inflammation. It is these drugs which have improved the prognosis in recent years for many people with RA. It is usual to start a DMARD as soon as possible after RA has been diagnosed to try and limit the disease process as much as possible. DMARDs have no immediate effect on pains or inflammation. Another line of drugs which have recently been developed and found very effective are biologic response modifiers drugs (BRMDs) which include TNF blockers like etanercept, infliximab, adalimumab and IL-1 blocker anakinra (Sautner and Leeb, 2003; Weaver, 2004). They show promise but their long-term benefits are still being evaluated. One problem with these drugs is that they need to be given by injection.

To minimize disability as much as possible the patient needs to remain active thus regular exercise may help to reduce pain and improve joint function. A physiotherapist can advise on exercises to keep muscles around joints as mobile and strong as possible. To reduce the risk of developing associated diseases such as heart disease, stroke, osteoporosis, and certain cancers the patient is advised to eat a good healthy diet and exercise regularly, lose weight if overweight, not to smoke (in addition to increasing the risk of cancer, heart disease and stroke, smoking may also make symptoms of RA worse).

Some people try complementary therapies such as special diets, bracelets, acupuncture, etc. There is little research evidence to say how effective such treatments are for RA.
1.2. Plasma proteomics

Blood plasma is an exceptional and the most complex human-derived proteome, containing other tissue proteomes as subsets. It is most difficult to characterize on account of the large proportion of albumin (55%), the wide dynamic range in abundance of other proteins, and the tremendous heterogeneity of its predominant glycoproteins. There are about 300 different proteins found in human plasma (Fig. 1.3) (Anderson and Anderson, 2002). The dynamic range of proteins at high end can be 35-50 mg/ml for serum albumin and 0-5 pg/ml for cytokines at low end. Classical definition of plasma proteins by Putnam et al. includes those proteins that carry out their function in circulation, thus excluding those protein that act as messengers between the tissues or that leak into blood plasma during tissue damage (Putnam, 1984). Serum, the protein solution remaining after plasma (or whole blood) is allowed to clot, is very similar to plasma: prothrombin is cleaved to thrombin, fibrinogen is removed (to form the clot), and a limited series of other protein changes (mainly proteolytic cleavages) take place. A further important feature of the plasma proteome is that it is the furthest removed, among tissue proteomes, from the mRNA level. While many of the major plasma proteins are synthesized in the liver and comprise many of its most abundant mRNAs (Anderson and Seilhamer, 1997), it is known that their plasma levels correlate only poorly with message abundance in liver (Kawamoto et al., 1996) and presumably even more poorly for proteins synthesized in smaller organs (individually or collectively). For these reasons, plasma is a biological system that can only be approached with protein methods and thus remains beyond the scope of DNA or RNA-based diagnostics.
The attraction of plasma for disease diagnosis lies in two characteristics: the ease by which it can be safely obtained and it reflects the patho-physiological state of the body at given time. For e.g. the concentration of high abundance protein, albumin, which is daily synthesized by liver (~12g) and has a half life of 21 days, is measured clinically as an indication of severe liver disease or malnutrition. Similarly, the level of IL-6 is measured as an indicator of infection and inflammation. Thus, proteins at all abundance levels prove to be clinically useful. Tissue leakage proteins are important because a serious pathology can be detected in a small volume of tissue by measuring release into plasma of a high abundance tissue protein for e.g. cardiac myoglobin (Mb) is present in plasma from normal subjects at 1–85 ng/ml but is increased to 200–1,100 ng/ml by a myocardial infarction (Cloonan et al., 1979). Realising the essentiality of plasma proteins as a source of information for various diseases human proteome organisation (HUPO) initiated the human plasma proteome project with the objective to facilitate major disease-related studies of candidate biomarkers for earlier diagnosis, better stratification of
newly diagnosed patients, pathways-based monitoring of targeted therapies, and design of preventive interventions (www.hupo.org/research/hppp/).

Plasma peptidome or fragmentome, the low-molecular-weight range of the circulatory proteome in the plasma, offers an important area for the discovery of prognostic or diagnostic marker for many diseases (Hortin, 2006). It is a ‘recording’ of the cellular and extracellular enzymatic events that take place at the level of the tissue microenvironment which can provide important information in diseases like cancer (Petricoin et al., 2006). Peptidome profiling allowed the discovery and validation of platelet factor 4 as a new discriminating marker in pancreatic cancer (Fiedler et al., 2009). Peptidome analysis of plasma in hepatitis patients has led to the identification of few signature peptides which has helped in understanding the disease mechanism (Taneja et al., 2011).

Post-translational modifications of proteins are important in many biological processes, and post-translational changes have relevance to diseases. Protein glycosylation is an important post translational modification (PTM) with wide degree of heterogeneity. Since most of the plasma proteins are glycoproteins in nature, they hold an immense potential as biomarker for diseases. This has been described in details in next few pages of this chapter.
1.3. **Protein glycosylation**

Protein glycosylation is the most common post translational modification of proteins. It configures most of the proteins in biologically active forms for diverse physiological functions (Varki et al., 2009). Glycans provide stability to the proteins which are exceptionally prone to destabilization due to alteration in chemical and physical environment of the cell (Cromwell et al., 2006; Wang, 2005). Besides stability, some other important functions performed by glycans in the proteins are protection from proteolytic enzymes and viruses, host pathogen interaction, immune system, protein folding etc (Bergeron et al., 1998; Ricardo and Sola, 2008; Sasisekharan and Myette, 2003). In addition, oligosaccharide sequences are themselves recognised as antigens and as ligands for carbohydrate binding proteins (lectins). Deficient or erroneous modifications of these proteins often lead to defects in the function of the protein which may often lead to a

![Fig. 1.4. Bio-synthesis of the N-linked core oligosaccharide](image-url)
disease. These incorrectly modified proteins are detected as non-self molecules by the immune system and thus consequently results in autoimmunity. The immune system represents the diversity of oligosaccharide structure and function, particularly, the relevance of oligosaccharide heterogeneity, antigenicity, and immune regulatory activity in microbial pathogenesis and autoimmunity (Hounsell and Davies, 1993).

1.3.1. **Biosynthesis of N and O-linked glycans**

Biosynthesis of glycans is an enzymatically regulated process taking place in endoplasmic reticulum (ER) and golgi bodies (Fig. 1.4). The glycans are covalently attached to the growing polypeptide chains while the later is being synthesized in the ER. For N-linked glycans, the precursor oligosaccharide structure which consists of GlcNAc2Glc3Man9, is synthesized as a lipid linked dolichol pyrophosphate precursor on the ER membrane which is transferred to

![Fig.1.4. Different types of N–linked glycan structures](image-url)
the growing polypeptide chain by oligosaccharyl transferase enzyme complex (OST) to the asparagine residue in the sequence Asn-X-Thr or Ser (where X is any residue other than proline) of the peptide (Helenius and Aebi, 2001; Varki et al., 2009). Further, the processing and trimming of the precursor molecule by various glucosidases and mannosidases present in the lumen of the ER remove three glucose (Glc) and one mannose (Man) molecule and is transferred to the golgi complex where the addition and deletion of various monosaccharide residues by variety of enzymes like glycosidases and glycosyltransferases, give rise to the structural heterogeneity in the oligosaccharide structures.

Depending on the presence of monosaccharide residues, N-linked glycans are broadly classified into three categories i.e. high mannose, complex, and hybrid (Fig. 1.5). All the three types of N-glycans have a common core pentasaccharide structure, consists of two N-acetylglucosamine (GlcNAc) and three mannose molecules. High-mannose oligosaccharides have two to six additional Man residues linked to the pentasaccharide core while complex-type oligosaccharides have two or more branches, each containing at least one

![Fig. 1.6. O-linked Core 1 structures](image-url)
GlcNAc, one galactose (Gal) and eventually one sialic acid (SA) residues. Hybrid type contains both high mannose and complex type of structures. There are various types of O-linked glycan structures known in mammals but the most commonly studied is of mucin type in which the N-acetylgalactosamine (GalNAc) is covalently linked to the hydroxyl group of serine or threonine (GalNAc-α1-ser/thr, Tn antigen) residue through O-glycosidic bond by polypeptide α-N-acetyl galactosaminyl transferases (ppGalNAcTs) localized to the Golgi apparatus (Fig 1.6). It is presumed that ppGalNAcTs acts on the correctly folded proteins after successfully existing from the ER (Rottger et al., 1998). This is a common step for the biosynthesis of all the mucin type O-glycans. After existing from ER, Tn antigen can have access to many glycosyltransferases present in the golgi for further addition of sugars forming different core structures (Core 1-9) (Table 1). The most common is core 1 β3 galactosyl transferase or T-synthase which adds Gal to GalNAc forming core 1 (Galβ1-

<table>
<thead>
<tr>
<th>O-glycans</th>
<th>Structures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tn antigen</td>
<td>GalNAcαSer/Thr</td>
</tr>
<tr>
<td>Sialyl-Tn antigen</td>
<td>Sia α2-6GalNAcαSer/Thr</td>
</tr>
<tr>
<td>Core 1 (T antigen)</td>
<td>Galβ1-3GalNAcαSer/Thr</td>
</tr>
<tr>
<td>Core 2</td>
<td>GlcNAcβ1-6(Galβ1-3)GalNAcαSer/Thr</td>
</tr>
<tr>
<td>Core 3</td>
<td>GlcNAc β1-3GalNAcαSer/Thr</td>
</tr>
<tr>
<td>Core 4</td>
<td>GlcNAcβ1-6(GlcNAcβ1-3)GalNAc αSer/Thr</td>
</tr>
<tr>
<td>Core 5</td>
<td>GalNAcα1-3GalNAcαSer/Thr</td>
</tr>
<tr>
<td>Core 6</td>
<td>GlcNAcβ1-6GalNAcαSer/Thr</td>
</tr>
<tr>
<td>Core 7</td>
<td>GalNAcα1-6GalNAcαSer/Thr</td>
</tr>
<tr>
<td>Core 8</td>
<td>Galα1-3GalNAcαSer/Thr</td>
</tr>
</tbody>
</table>
3 GalNAc-\(\alpha\)-ser/thr) or T-antigen (Thomsen–Friedenreich antigen) which may or may not be further sialyated (NeuAc\(\alpha\)2-3, Gal\(\beta\)1-3GalNAc-ser/thr).

Besides that, other non mucin types of O-glycan structures are also known where O-linked fucose (Fuc) and glucose (Glc) are found like in epidermal growth factor (EGF) (Harris and Spellman, 1993), O-linked GlcNAc in nuclear proteins like c-Myc (Chou et al., 1995), Gal on hydroxyl lysines of collagen in the sequence Gly-Xaa-Hyl-Gly (Michaelsson et al., 1994) and the arabinosylation of hydroxyprolines in some vegetal proteins (Kieliszewski et al., 1995).

1.3.2. Protein glycosylation and immunity

Within the immune system, various classes of glycan-binding receptors (lectins) exist that recognize specific glycan structures presented on a protein backbone or lipid structure. Galectins, C-type lectin receptor (CLR) and siglecs are the three most common carbohydrate binding receptor present on immune cells in which galectins are secretory while rest of the two are membrane bound (Crocker et al., 2007; Leffler et al., 2004; Weis et al., 1998). Galectins have a conserved carbohydrate recognition domain (CRD) that recognizes glycans containing the disaccharide N-acetyllactosamine (Gal-\(\beta\) (1-4)-GlcNAc) structure. Galectins are mostly expressed by activated but not resting T and B cells, and they are significantly upregulated in activated macrophages and regulatory T cells (Rabinovich et al., 2007).

C-type lectin receptors (CLRs) are a large family of receptors containing calcium-dependent carbohydrate binding proteins and most of the CLRs contain one or two CRD domains. The mannose-specific CLRs contain an EPN (Glu-Pro-Asn) amino acid motif and have specificity for mannose- and/or fucose-terminated glycans while the Gal specific CLRs contain the QPD (Gln-Pro-Asp) sequence in the CRD and recognize Gal terminated or GalNAc terminated glycan
structures. CLR\textsuperscript{s} are mainly restricted to antigen presenting cells (APCs) such as macrophages and dendritic cells (DCs) while some have also been identified on NK cells or endothelial cell (Weis \textit{et al.}, 1998). Selectins are important members of C-type of lectins whose role are well studied in lymphocyte adhesion, activation, extravastation and homing. L-selectin is expressed by almost all the leukocytes whose ligand is present on the high endothelial venules (HEVs) of vascular endothelium. The lymphocytes tether and role over the HEVs by the interaction of L-selectin and its ligand. P-selectin is constitutively found in Weibel-Palade bodies of endothelial cells and in alpha granules of platelets. Within minutes after activation by inflammatory mediators, P-selectin is mobilized to the cell surface. The cell surface expression of P-selectin is short lived so they play a role in early events of leucoctyes rolling. E-selectin production is strongly and rapidly induced by a variety of inflammatory mediators on the surface of human vein endothelial cells. Since all the selectins contain CRD domain, the interaction of selectins with their ligands are mediated through fucosylated tetrasaccharide sialyl lewis\textsuperscript{x} (sLe\textsuperscript{x}) or sLe\textsuperscript{a} glycan present on N-linked or O-linked structures of ligand glycoprotein (Tedder \textit{et al.}, 1995; Varki \textit{et al.}, 2009). Siglecs (sialic acid-binding immunoglobulin-like lectins) are a family of membrane proteins with variable numbers of immunoglobulin domains having specificity for sialic acid containing glycans. Siglecs are widely expressed on immune cells, including neutrophils, monocytes, B cells, DCs, NK cells, eosinophils and basophils. All of these receptors interacts with the glycan structures presents on the surface of pathogens and instructs dendritic cells and subsequently T and B Cells. Thus, the information regarding various immunologically important functions like activation, differentiation, maturation, recognition and homing are
included in the form of glycans which are decoded by the above mentioned receptors whose expression and functions are regulated at sites of inflammation.

1.3.3. Aberrant glycosylation and diseases

As mentioned above, glycosylation of proteins is mediated through various exoglycosidases and glycosyl transferases in ER and glogi bodies. Their expression depends on tissue specific manner and external stimulus by various regulatory factors like cytokines, hormones etc (Alavi et al., 1998). Alteration in external stimulus, increases the stress inside the cell which subsequently affects glycosylation of proteins. These aberrant glycosylation patterns can be used as marker for the disease. Disease specific glycosylation changes in the plasma proteins have been identified as potential biomarkers for prostate cancer (Kyselova et al., 2007) where increased fucosylation of serum ribonuclease 1 and altered glycosylation of other proteins have been reported in pancreatic cancer (Barrabes et al., 2007). Studies by Mondal et al have shown the increased fucosylation and sialylation in plasma alpha-feto protein (AFP) (Mondal et al., 2011) and increased fucosylation in plasma AGP (Mondal et al., 2009) in patients of hepatitis induced hepatocellular carcinoma (HCC). An abnormal microheterogeneity of plasma Hp, alpha1-antitrypsin (AAT) and transferring was detected in the serum of alcoholic patients (Gravel et al., 1996). Alteration in glycosylation of serum AGP has been extensively studied in various inflammatory conditions. Within the sera of patients with acute inflammation, increase in bi-antennary and decrease in tri- and tetra-antennary structures were observed, as well as increase in alpha1, 3-fucosylation, at most glycosylation sites. In the sera of patients with chronic inflammation, increased rates of tri-antennary alpha1,3-fucosylation at sites 3 and 4 and tetra-antennary alpha1,3-fucosylation at sites 3, 4 and 5 were detected (Higai et al., 2005). Quantitative and
qualitative differences in glycan masses of Hp from plasma of psoriatic patients revealed the increase in fucosylated, tetraantennary and triantennary glycopeptides in patients compared to controls (Maresca et al., 2010). The glycosylation study of plasma AAT in congenital disorder of glycosylation type I (CDG-I) has shown that it is under glycosylated (Mills et al., 2003).

An interesting study by Chui et al has shown the role of N-linked glycosylation in development of autoimmune response. The author showed that the mice mutated in α-mannosidase II gene showed the symptoms of systemic autoimmune disease similar to human systemic lupus erythematosus (SLE) like the presence of anti-nuclear and anti-Sm antibodies in the serum and progressive decrease in kidney functions. α-mannosidase II regulates the hybrid to complex branching pattern of extracellular asparagine (N)-linked oligosaccharide chains (N-glycans). Glycoproteins having complex type glycosylation were enriched from various tissues of mannosidase-II deficient mice and found that the glycosylation of proteins in kidneys are mostly affected due to the mutation in mannosidase II enzyme while in other organs glycoproteins compensated the deficiency using alternate pathways (Chui et al., 2001). Similarly, another study by Demetriou et al found that mice deficient in the locus encoding N-acetylglucosaminyltransferase-5 (Mgat5), results in lack of β (1-6) branching of tri- and tetra-antennary complex N-glycans, ligands for galectin-3, enhanced T cell receptor (TCR)-mediated signaling and induced ‘hyper-TH1’ responses.

**Table 2.** Glycosylation changes in some autoimmune diseases

<table>
<thead>
<tr>
<th>Disease</th>
<th>Aberrant Glycosylation of proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>IgA nephropathy (IgAN)</td>
<td>agalactosylation of O-linked glycan in IgA.</td>
</tr>
<tr>
<td>Sjogren’s syndrome (pSS)</td>
<td>Over sialylation of N-linked glycan in IgA, agalactosylation of serum IgG</td>
</tr>
</tbody>
</table>
and greater susceptibility to autoimmune disease, which was the result of a restricted TCR aggregation by binding to galectins in the immunological synapse (Demetriou et al., 2001). Table 2 summarises the list of some of the autoimmune diseases undergoing glycosylation alteration. Alteration in glycosylation of IgG has also been widely studied in pathogenesis of disease and the concentration of terminally agalactosylated IgG (IgG0) was found to be increased and correlated with the disease severity in patients of RA. Rheumatoid factor (RF factor), the characteristic auto-antibodies against the Fc region of IgG0, are also increased in the sera of patients with RA (Carson et al., 1987; Liljeblad et al., 2000). The altered glycosylation of IgG activates the complement system via mannose binding lectin (MBL) through the interaction with penultimate GlcNAc residue exposed after degalactosylation (Malhotra et al., 1995). Besides IgG, other serum proteins also undergo glycosylation changes in RA patients (Elliott et al., 1997; Goodarzi et al., 1998; Nakagawa et al., 2007; Przybysz et al., 2007; Raghav et al., 2006; Ryden et al., 2002).

A study by Dzhambazov et al, showed the role of glycosylation and T cell activation by Collagen type II (CII)-derived glycopeptide in RA where T cell

<table>
<thead>
<tr>
<th>Disease</th>
<th>Presence/Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Systemic lupus erythematosus (SLE)</td>
<td>Presence of agalactosylated IgG (IgG0) in serum</td>
</tr>
<tr>
<td>Tn-syndrome</td>
<td>Presence of Tn-antigen (GalNAcα1-Thr/Ser-peptide) on blood cells. Sometimes presence of sialyl Tn-antigen (Neu5Acα 2-6GalNAcα1-Thr/Ser peptide) on blood cells.</td>
</tr>
<tr>
<td>Rheumatoid arthritis (RA)</td>
<td>Presence of agalactosylated IgG (IgG0) in serum</td>
</tr>
<tr>
<td>Multiple sclerosis</td>
<td>Reduce con-A reactivity of CSF glycoproteins</td>
</tr>
</tbody>
</table>
hybridoma lines specific for nonmodified collagen peptide (residues 260–270), hydroxylated peptide (the precursor to O-linked glycosylation), O-linked β-galactosylated peptide, and O-linked β-galactosylated-glucosylated peptide were used to compare CII from normal and arthritic joints (Dzhambazov et al., 2005). Both normal and arthritic samples activate glycosylated-CII-specific hybridomas; however, the arthritic samples also activate the non modified-CII-specific hybridoma indicating that normal tissue primarily contains glycosylated CII, whereas arthritic tissue contains a mixture of glycosylated and non-glycosylated CII, suggesting that a change in glycosylation or, more specifically, a lack of glycosylation in cartilage is a key component of arthritis progression glycosylation. To investigate whether the observed patho-physiological similarities that develop in both the collagen induced experimental model of arthritis (CIA) and RA are associated with similar glycosylation changes, relative activity of the beta, 1-4 galactosyltransferase (GTase) was assessed within various tissues. Tissue-specific changes were assessed by comparison of GTase activity in peripheral (P.GTase) and paired splenic lymphocytes. A marked reduction in P.GTase activity was found in CIA, compared with age-matched unimmunised mice and the adjuvant controls while splenic GTase activity was found to be similar in all tissue (Alavi et al., 1998). Advanced glycation end products (AGEs) are a heterogeneous group of molecules formed from the non-enzymatic reaction of reducing sugars with free amino groups of proteins, lipids, and nucleic acids under the oxidative stress environment like in RA. Although the presence of several AGE modified products like pentosidine, carboxy-methyl lysine have been identified in RA patients, but the presence of IgM antibodies against AGE modified IgG has been reported in RA patients (Drinda et al., 2002; Ligier et al., 1998; Newkirk et al., 1998). These AGE modified proteins activate immune system via its receptor (RAGE), resulting in signal pathways linked to pro-
inflammatory responses and differentiation of T-cells (Chen et al., 2008; Hofmann et al., 1999). The receptor for AGEs in synovial tissue has been identified in RA (Drinda et al., 2002).
1.4. **Animal model of RA**

An ongoing and challenging issue for most disease-based studies centers on study design and samples, and for human samples, the availability of human tissues will always remain difficult. In addition to the difficulties in obtaining human samples, another inherent problem is that human samples are naturally heterogeneous. This large degree of human heterogeneity largely stems from differences in genetic background, environment and medical history. This makes ‘discovery’ in a highly variable system rather difficult (Bousette et al., 2008). Conversely, many of these variable factors complicating analysis of human derived samples can be decreased using animal models, which are genetically identical. Disease and control animals are housed under identical conditions, restricting environmental effects, and can be examined at identical times/disease points, even at both pre- and post-symptomatic times. Tissue procurement is often hindered by the invasiveness of the procedure in humans and when possible is limited in the amount that can be harvested (i.e. only small amounts are ethically reasonable). This technical hindrance can be avoided with the use of animal models. Furthermore, the use of animal models permits in vivo sampling analysis (Huang et al., 2006).

1.4.1. **Collagen induced arthritis (CIA)**

CIA is most widely studied model for RA that is induced in susceptible mouse strains following an intradermal immunization with collagen-II (CII) emulsified in Freund’s adjuvant (Trentham et al., 1977). CIA can be induced using native autologous or heterologous CII and is specific to CII, since immunization with types I or III collagen failed to induce disease CIA. CII is a major protein constituent of joint cartilage and the immunization provokes an autoimmune response that attacks the joints. Like RA, CIA is characterized by
the presence of fibrin deposition, hyperplasia of synovial cells, periosteal bone formation, mononuclear infiltrates, pannus formation and eventual ankylosis of one or more articular joints (Courtenay et al., 1980). In addition, the presences of rheumatoid factor and systemic manifestations have been reported in animals with CIA. Moreover, susceptibility to both CIA and RA is strongly associated with the expression of specific MHC class II Ab gene (coding for the A beta chain) (Gregersen et al., 1987). In mice, susceptibility to CIA is mediated predominantly by I-A\(^d\), an MHC class II molecule which binds the same immunodominant CII peptide region located between positions 256 and 270 of CII as the human RA-associated allele HLA-DR4 (DRB1*0401) (Andersson et al., 1998). This is a post translationally modified glycopeptide with lysine residues at position 264 which can be hydroxylated, galactosylated and glucogalactosylated. This modified peptide is recognised by different subset of T cells.

T and B cells play a central role in the pathogenesis of CIA. The major role of B cells is production of arthritogenic anti-CII antibodies, which is clearly shown by the fact that antibodies reactive with CII can bind to cartilage and induce arthritis (Terato et al., 1992). The role of T cells in CIA is more complex and can be divided into two main pathways. Firstly, T cells provide help to B cells for the production of arthritogenic anti-CII antibodies (Corthay et al., 1999). Secondly, T cells themselves are believed to play a role in joint inflammation through activation of other cells, e.g. synovial macrophages. Transfer of T cells can induce synovitis but not clinically evident arthritis and blockage of T cells or T cell function ameliorate development of arthritis (Taylor et al., 1996).