Immune complexome analysis

Kaname Ohyama†,‡, Naotaka Kuroda†,1

† Department of Environmental and Pharmaceutical Sciences, Graduate School of Biomedical Sciences, Nagasaki University; 1-14 Bunkyo-machi, Nagasaki, Nagasaki 852-8521, Japan

‡ Nagasaki University Research Center for Genomic Instability and Carcinogenesis (NRGIC); 1-12-4 Sakamoto-machi, Nagasaki, Nagasaki 852-8523, Japan

Running title: Immune complexome analysis

1 To whom correspondence should be addressed.

E-mail: n-kuro@nagasaki-u.ac.jp

Abbreviations: anti-CCP antibody, anti-citrulline-containing protein/peptide antibody; 2-DE, 2-dimensional gel electrophoresis; CIC, circulating immune complex; FDR, false discovery rate; IC, immune complex; LC-MS/MS, liquid chromatography tandem mass spectrometry; MALDI-TOF-MS, matrix laser desorption/ionization time-of-flight mass spectrometry; MS, mass spectrometry; OA, osteoarthritis; PF4, platelet factor 4; RA, rheumatoid arthritis; RF, rheumatoid factor; SLE, systemic lupus erythematosus; TSP-1, thrombospandin-1; VEGF, vascular endothelial growth factor
# Table of contents

1. Abstract 3

2. Immune complex 4

3. Identification of disease-associated antigens 5
   3.1 Immunoproteomics 6
   3.2 Immune complexome analysis 7

4. Immune complexome analysis of RA patients 9

5. Conclusion 11

References
1. Abstract

Immune complexes (ICs) are produced during an immune response and may reflect some aspects of an ongoing immune response. Therefore, the identity of antigens incorporated into ICs provides the information that in the future may aid in the development of diagnosis and treatment strategies for autoimmune diseases, infection, cancer, and transplantation therapy, and this information might be more relevant than information on free antigens. Because ICs may contain many antigens, comprehensive identification and profiling of such antigens is more effective than immunoblotting detection. Here, we introduced mass spectrometry (MS) -based two approaches (immunoproteomics and immune complexome analysis) to comprehensively identify the antigens. Immunoproteomics is a concept to identify disease-associated antigens that elicit immune responses by combining protein separation (2-dimensional electrophoresis, gel-free separation), immunological detection (Western blotting) and MS, or combining immunocapture and MS. Immune complexome analysis is designed for identifying antigens in circulating ICs and consists of ICs separation from serum and direct tryptic digestion followed by nano-liquid chromatography-tandem MS.
2. Immune complex

Our immune system resembles an intelligent security system, which continually monitors for foreign invaders; however, in some cases, this system recognizes the healthy parts as something harmful or foreign, and attacks them (autoimmune diseases). Immune complexes (ICs) are products of reactions that involve non-covalent interactions between foreign antigens or autoantigen and antibody molecules. IC formation constantly occurs even in healthy organisms; denatured proteins, antigens of gut bacteria, and dead cells undergo binding to antibodies followed by subsequent phagocytosis, mainly by macrophages located in the liver and spleen [1]. For a long time, ICs were thought to represent a common pathway for pathogenesis of several diseases (vasculitis, infections, and connective tissue autoimmune disorders). The pathology generally arises from insufficient clearance or deposition of ICs. Actually, ICs in the sera of the patients with rheumatoid arthritis (RA), systemic lupus erythematosus (SLE) and systemic sclerosis was significantly higher than that of healthy controls [2, 3]. Also, aberrant host proteins that occur during a certain state of diseases (e.g. cancer) can be recognized as an autoantigen and may form ICs [4, 5]. Reasons why the ICs underlie pathogenicity have been investigated by many researchers [6-8]. Clinical significance of deposited ICs in glomerulonephritis is well understood [9-11]; however, the relationship between ICs and many diseases are still unclear. Therefore, information on ICs is useful to reveal if ICs play an important role in each disease. Especially, identifying relevant antigens is informative because they trigger the humoral immune response. In future, such information may possibly provide new insights into pathophysiology and form the
basis for novel diagnostic and treatment strategies for autoimmune diseases, infection and cancer.

3. Identification of disease-associated antigens

ICs present in the human body are likely to contain many different antigens that may reflect underlying disease and/or differences between individuals. Therefore, comprehensive identification and profiling of such antigens may be more effective than immunoblotting detection of individual antigens. Microarrays are widely used to detect multiple antigens [12]. This approach depends on exposing serum samples from patients to an ordered array of putative antigens, capturing those antibodies that bind the antigens on the arrays. However, two major drawbacks of the array are that it is biased, given that antigens are selected based on the likelihood that they play a role in the disease and that the analytical comprehensiveness of this technique is fundamentally limited because only molecules represented on the arrays can be identified [13].

Contrastingly, with the progress in both mass spectrometry (MS) and bioinformatics, proteomic methods are fully developed and become common. Especially, a 2-dimensional liquid chromatography/tandem MS (2D-LC-MS/MS), or a combination of 2-dimensional gel electrophoresis (2-DE) and matrix laser desorption/ionization time-of-flight MS (MALDI-TOF-MS), which are followed by database-searching (sequence tag or peptide mass fingerprinting), is widely used. These methods drastically expand the comprehensiveness of protein identification; thus, the paradigm in biomarker research during the last decades has been shifted from hypothesis-driven to discovery-driven approaches. Based on those MS-based proteomic techniques, several attempts to
identify disease-associated antigens have been reported.

3.1 Immunoproteomics

Immunoproteomics is a concept to identify disease-associated antigens that elicit immune responses by combining protein separation (2-DE, gel-free separation), immunological detection (Western blotting) and MS (Fig. 1A), or combining immunocapture and MS (Fig. 1B) [14]. Proteins derived from cells or tissues (e.g. cancer cell) were separated by 2-DE. Antigenic proteins are detected by applying patient serum, which may include the antibodies specific to a disease, and subsequently using enzyme-labelled secondary antibodies. To identify immunogenic proteins, the corresponding spots are excised from gel and are in-gel digested. The digest is analyzed by MS or tandem MS, followed by peptide fingerprinting or sequence tag methods (Fig. 1A).

In contrast, antigen profiling by immunocapture MS begins with the immobilization of antibodies derived from patient serum (Fig. 1B). Almost all the antibodies are captured on Protein A or G which are bacterial-derived proteins and have specific affinity for the Fc domain of antibodies. Then, a protein mixture (cell or tissue lysate) was applied to a column or beads on which antibodies are immobilized, resulting in capturing antigens specific to antibodies present in patient serum. Finally, antigens eluted from the antibodies are applied to identify the proteins by MALDI-TOF-MS or surface-enhanced laser desorption/ionization MS. Using this immunocapture MS approach, Tjalsma et al. detected antigen profiles in colon cancer patients [15, 16]. Based on
the association of *Streptococcus bovis* with colon cancer, they prepared antigen candidate from *S. bovis* strain and identified bacterial antigens which were captured by antibodies derived from colon cancer patients [15]. Furthermore, the same authors also combined the immunocapture technique with an MS/MS analysis in a study aimed at identifying diagnostic antigens, which could not be identified in previous study [15], from *S. bovis* [16]. In this method, the antigens captured by antibodies in patient serum were eluted and are applied to in-solution tryptic digestion. All the digested peptides were separated on nano-LC column and subsequently sequenced by MS/MS spectra and protein database, which is called shotgun immunoproteomics [16].

3.2 **Immune complexome analysis**

Importantly, the identification of antigens in circulating immune complex (CIC) might be different from that of free antigens because CICs are the direct and real-time products of an immune response. CICs may sensitively reflect the pathophysiological change at the early stage or during the progression of a disease. Moore and colleagues reported a proteomic approach for profiling CIC in which CICs were isolated from the sera of juvenile idiopathic arthritis patients or healthy controls by Proceptor™ (ProGen Biologics), and subsequently were separated by 2-DE [17]. They found that thirty seven protein spots were overexpressed in CICs of the aggressive disease groups as compared to controls, 28 of which have been identified. However, most of protein spots identified in this study were derived from immunoglobulins and much information on antigens incorporated into CICs could not been obtained.
Ohyama et al. propose a novel proteomic strategy (immune complexome analysis, Fig. 2) that entails separation of ICs from serum, direct tryptic digestion, and nano-LC-MS/MS for the identification and profiling of antigens in CICs by using serum of patients and healthy donors [18]. CICs were purified by magnetic beads with immobilized Protein G (PureProteome®, Millipore). Beads (40 µl) were incubated with 10 µl of serum diluted with phosphate buffered saline (PBS, 1:9, v/v) for 30 min with gentle mixing. The beads with bound ICs were recovered with a magnet and washed three times with 500 µl of PBS. The beads were resuspended in 100 µl of 10 mM dithiothreitol and incubated at 56 ºC for 45 min; then, 100 µl of 55 mM iodoacetamide were added and the mixture was incubated at room temperature for 30 min in the dark. Subsequently, trypsin was added, and the mixture was incubated overnight at 37 ºC. Trifluoroacetic acid (10%) was added to stop the digestion, and the supernatant containing the peptide digests of antigens and antibodies was recovered. The peptide mixture (1 µl) was subjected to nano-LC-MS/MS. MS/MS data were searched against a human subdatabase from the public non-redundant protein database. Initial matching results were further evaluated by the filter criteria which were adjusted maintaining the empirically determined false discovery rate (FDR) at zero. FDR was calculated using the number of significant unique peptide in the reversed database divided by the number of those in the forward database. Ubiquitous keratins and trypsin were excluded as potential matches. In order to focus the aim on identifying antigens in CIC, the peptides derived from immunoglobulins were also excluded. More details of immune complexome analysis can be found in our recent literatures [18].
4. Immune complexome analysis of RA patients

RA is a progressive autoimmune disease characterized by chronic inflammation of the peripheral joints. The presence of ICs in serum and synovial fluid of RA patients is likely to contribute to the pathogenesis and to the articular damage, since they are responsible for the activation of complement, the stimulation of phagocytes and the release of cytokines, metalloproteinases and reactive oxygen intermediates [20-24].

RA is diagnosed based on classification criteria set by the American College of Rheumatology and serological assessment of two RA biomarkers, rheumatoid factor (RF) and anti-citrulline-containing protein/peptide (anti-CCP) antibody [25]. RF is not highly specific (81%) for RA; it is found in 4% of healthy persons [26, 27]. Anti-CCP antibody has a greater specificity (98%) for RA than RF; however, its sensitivity (68%) is lower than RF (78.6%) [23].

The proteins identified by immune complexome analysis in established RA patients and controls (healthy donors and osteoarthritis (OA) patients) are summarized in Table 1. The detection of several complements (C1, C3, C4) indicated that extraction of ICs from serum and peptide mapping by our method were successful. IgG and C3 were found together on the cartilage surface [28], and RA-associated autoantibodies were enriched in synovial fluid [29] relative to serum. Considering that CICs containing some of the antigens (clusterin, apolipoprotein E and vitronectin) identified in this study were also found in arthritis joints of RA and OA patients [30], our results suggest that the CICs possibly accumulate in lesion sites.

Of the antigens identified here, thrombospondin-1 (TSP-1) incorporated in CICs was 100%
specific for RA and appeared more sensitive (81%) than RF and anti-CCP antibodies (Table 1, Fig. 3). In RA patients, TSP-1 has been reported to be present in synovial tissues [31], and plasma concentrations of TSP-1 have been shown to be increased [32]. However, no information is available about whether TSP-1 is present in ICs. TSP-1 (M.W. 129,000) is a multifunctional glycoprotein and is produced in multiple cells including platelets, leukocytes and endothelial cells. Studies of animal models provided evidence that angiogenesis plays an important role in RA [33, 34]. In mice with collagen-induced arthritis, manifestations of inflammation develop in close correlation with the expansion of the synovial vasculature [35]. The administration of pro-angiogenic cytokines such as vascular endothelial growth factor (VEGF) [36] increases the severity of experimentally induced arthritis. Furthermore, serum angiogenesis markers, including VEGF and angiopoietin-1, are elevated very early in the course of the disease [37]. Because TSP-1 is known to be an endogenous inhibitor of angiogenesis [38], the formation of TSP-1-containing CICs may promote angiogenesis by eliminating active TSP-1 from circulation. However, the potent antiangiogenic properties of TSP-1 may be an effective therapeutic strategy for treatment of RA. Jou et al. demonstrated that direct administration of adenoviral vectors encoding TSP-1 significantly ameliorated the clinical course of collagen-induced arthritis in rats [39].

Although platelet factor 4 (PF4) was less sensitive (52.4%) than TSP-1, PF4 incorporated in CICs was also specifically detected in RA patients for the first time using our method (Table 1, Fig. 3). Recently, Xiao et al. demonstrated that the ICs following the binding of anti-PF4 antibodies to PF4 were present and stimulated human neutrophil activation and cell adhesion [40]. This adhesion mechanism enables leukocytes to migrate from blood and affect inflamed synovium. Furthermore, PF4 is known to be an angiostatic chemokine [41], and the production of anti-PF4
antibodies can accelerate angiogenesis.

Anti-CCP antibody is specific (94%-99%) for RA; however, 40% of early RA patients are negative for this marker [42]. Therefore, novel biomarkers are urgently required, especially for early RA and/or for RA lacking RF and anti-CCP antibody markers (i.e., seronegative RA). Already, several biomarkers, including autoantibodies and matrix metalloproteinases, have been developed for this purpose [43-45]. We evaluated the diagnostic potential of CIC-associated TSP-1 and CIC-associated PF4 in early RA patients who were divided into two groups: a seropositive group comprising RF-positive-CCP-positive and RF-positive-CCP-negative patients, and a seronegative group comprising only RF-negative-CCP-negative patients. CIC-associated TSP-1 was found only in early RA patients and was not found in controls (Sjögren’s syndrome (SS), SLE and healthy donors). Fifty-five% (22/40) of all the early RA patients, 56% (14/25) of seropositive and 53% (8/15) of seronegative patients had CIC-associated TSP-1 (Table 2). PF4-containing CICs were found in only three (8%) of the patients with early RA; in contrast, PF4-containing CICs were found in 52% of patients with established RA. Therefore, these PF4-containing CICs may promote disease progression. We demonstrated that CIC-associated TSP-1 has high potential as a novel biomarker for diagnosing early and/or seronegative RA. Further analyses using a large number of patients are warranted to determine the clinical benefit of using this novel biomarker.

5. Conclusion

ICs are source suitable for understanding nature and amount of humoral immune response. The relationship between ICs and diseases, particularly autoimmune diseases, are still unclear;
therefore, identifying the relevant antigens helps one to reveal if ICs play a role in a disease. Especially, such identification is at the heart of current approaches for discovering antibody biomarkers. Here, we introduced some approaches to identify the antigens in ICs. Among them, immune complexome analysis, using a simple protocol, is endowed with high throughputness and comprehensiveness. It may be generally applicable to the study of the relationship between ICs and immune response-related disease treatment in animals and humans.
REFERENCES


[9] B.F. Murphy, L. Kirszbaum, I.D. Walker, I.D et al., Sp-40, 40, a new identified normal human serum protein found in the SC5b-9 complex of complement and in the immune deposits in


[26] L. De Rycke, I. Peene, I.E. Hoffman et al., Rheumatoid factor and anti-citrullinated protein


[34] G. Clavel, N. Bessis, M.C. Boissier et al., Recent data on the role of angiogenesis in


[43] K. Somers, P. Geusens, D. Elewaut, et al., Novel autoantibody markers for early and


**FIGURE CAPTIONS**

**Fig. 1** Schemes of immunoproteomic approaches. A) 2-DE is based on protein separation based on $pI$ and molecular weight. Next, proteins are transferred from the gel and immobilized on a membrane by Western blotting. Antigenic proteins can be detected by applying patient serum to the blot, after which bound patient IgGs can be visualized by secondary labeled antibodies. Finally, the antigenic spot is excised and analyzed by MS. B) Antigen profiling by immunocapture MS is based on immobilization of patient IgGs, which are directly used to capture and isolate antigenic proteins from a complex mixture of proteins. Captured antigens are profiled by MS. Reprinted with permission from Ref. 14).

**Fig. 2** Immune complexome analysis to profile antigens included in circulating immune complexes. The first step of the protocol involves immune complex capture, using the magnetic beads with immobilization of Protein G. The second step is reduction/alkylation and tryptic digestion. The third step involves analysis of the captured antigens by nano-LC-MS/MS.

**Fig. 3** Fragmentation spectra of TSP-1 and PF4 in CIC obtained by nano-LC-MS/MS. Data on the masses of each tryptic peptide and of its fragmentation ions were compared to those predicted for all tryptic peptides from all proteins in the International Protein Index.
Table 1  Summary of proteins in ICs extracted from serum.

<table>
<thead>
<tr>
<th>Protein</th>
<th>RA (n = 21)</th>
<th>OA (n = 8)</th>
<th>Healthy donor (n = 13)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frequency</td>
<td>Frequency</td>
<td>Frequency</td>
</tr>
<tr>
<td><strong>Complement</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complement C1r subcomponent</td>
<td>19</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>Complement C1s subcomponent</td>
<td>12</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Complement component 1, q subcomponent, B chain precursor</td>
<td>11</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Complement C1q subcomponent subunit A</td>
<td>7</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Complement C1q subcomponent subunit C</td>
<td>12</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Complement component 3</td>
<td>16</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>Complement C4-A</td>
<td>21</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td>Complement component 4B preproprotein</td>
<td>7</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>C4b-binding protein alpha chain</td>
<td>21</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td>Isoform 1 of C4b-binding protein beta chain</td>
<td>3</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Complement factor H</td>
<td>12</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td><strong>Coagulation protein</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plasminogen</td>
<td>20</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>Prothrombin (fragment)</td>
<td>19</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>Platelet factor 4</td>
<td>12</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Adhesion protein</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isoform 1 of Fibronectin</td>
<td>21</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>Vitronectin</td>
<td>21</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>Isoform B of Fibulin-1</td>
<td>15</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td><strong>Others</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thrombospondin-1</td>
<td>17</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Histidine-rich glycoprotein</td>
<td>19</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td>Isoform HMW of Kininogen-1</td>
<td>18</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>Putative uncharacterized protein albumin</td>
<td>18</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>Galectin-3-binding protein</td>
<td>12</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Vitamin K-dependent protein S</td>
<td>8</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Rheumatoid factor RF-ET9 (fragment)</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Rheumatoid factor D5 light chain (fragment)</td>
<td>5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>VH3 protein (fragment)</td>
<td>4</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Anti-folate binding protein (fragment)</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 2  Number of patients with early RA carrying CIC-associated TSP-1 or CIC-associated PF4

<table>
<thead>
<tr>
<th></th>
<th>Early RA patients (n=40)</th>
<th>SS patients (n=16)</th>
<th>SLE patients (n=14)</th>
<th>Healthy donors (n=11)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Seropositive (n=25)</td>
<td>Seronegative (n=15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSP-1</td>
<td>14</td>
<td>8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PF4</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Fig. 3

**Thrombospondin-1**

![Thrombospondin-1 diagram](image)

**Platelet factor 4**

![Platelet factor 4 diagram](image)