Structure–function analyses of a stereotypic rheumatoid factor unravel the structural basis for germline-encoded antibody autoreactivity

Rheumatoid factors (RFs) are autoantibodies against the fragment-crystallizable (Fc) region of IgG. In individuals with hematological diseases such as cryoglobulinemia and certain B cell lymphoma forms, the RFs derived from specific heavy- and light-chain germline pairs, so-called “stereotypic RFs,” are frequently produced in copious amounts and form immune complexes with IgG in serum. Of note, many structural details of the antigen recognition mechanisms in RFs are unclear. Here we report the crystal structure of the RF YES8c derived from the IGHV1-69/IGKV3-20 germline pair, the most common of the stereotypic RFs, in complex with human IgG1-Fc at 2.8 Å resolution. We observed that YES8c binds to the CH2–CH3 elbow in the canonical antigen-binding manner involving a large antigen–antibody interface. On the basis of this observation, combined with mutational analyses, we propose a recognition mechanism common to IGHV1-69/IGKV3-20 RFs: (1) the interaction of the Leu432–His335 region of Fc enables the highly variable complementarity-determining region (CDR)-H3 to participate in the binding, (2) the hydrophobic tip in the CDR-H2 typical of IGHV1-69 antibodies recognizes the hydrophobic patch on Fc, and (3) the interaction of the highly conserved RF light chain with Fc is important for RF activity. These features may determine the putative epitope common to the IGHV1-69/IGKV3-20 RFs. We also showed that some mutations in the binding site of RF increase the affinity to Fc, which may aggravate hematological diseases. Our findings unravel the structural basis for germline-encoded antibody autoreactivity.

This work was supported by Japan Society for the Promotion of Science (JSPS) KAKENHI grants 23655160 and 15K15184 (to T. U.) and the Platform Project for Supporting Drug Discovery and Life Science Research (Platform for Drug Discovery, Informatics, and Structural Life Science) from AMED. The authors declare that they have no conflicts of interest with the contents of this article.

This article contains Figs. S1–S11, Tables S1 and S2, and Experimental procedures.

The atomic coordinates and structure factors (code 5XMH) have been deposited in the Protein Data Bank (http://wwpdb.org/).

1 Both authors contributed equally to this work.

2 To whom correspondence may be addressed: Laboratory of Protein Structure, Function, and Design, Graduate School of Pharmaceutical Sciences, Kyushu University, Fukuoka 812-8582, Japan. Fax: 81-92-642-6667; E-mail: shiroya.phar.kyushu-u.ac.jp.

3 To whom correspondence may be addressed: Laboratory of Protein Structure, Function, and Design, Graduate School of Pharmaceutical Sciences, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan. Fax: 81-92-642-6667; E-mail: ueda.phar.kyushu-u.ac.jp.

4 The abbreviations used are: RF, rheumatoid factor; Fc, fragment crystallizable (Fc) region of IgG; RA, rheumatoid arthritis; MC, mixed cryoglobulinemia; WM, Waldenstrom’s macroglobulinemia; HCV, hepatitis C virus; BSA, buried surface area; SPR, surface plasmon resonance; SHM, somatic hypermutation; CDR, complementarity-determining region; VH, variable domain of the heavy chain; VL, variable domain of the light chain; CH, constant domain of the heavy chain.
RFs derived from IGHV1-69/IGKV3-20 are highly related to HCV-associated lymphoma (24). Of note, more than 90% of patients with MC are HCV carriers (25). In patients with HCV-associated MC, IgM antibodies from IGHV1-69/IGKV3-20 have RF activity with or without somatic hypermutations (26). These findings indicate that IGHV1-69/IGKV3-20-derived antibodies have some characteristics reflective of RF activity.

Despite physiological importance, the structural details of the antigen recognition mechanism have not been clarified for the IGHV1-69/IGKV3-20 RFs. Research has shown that IGHV1-69/IGKV3-20 RFs from healthy immunized donors, RA patients, and MC patients have long third complementarity-determining regions of the heavy chain (CDR-H3s) with a highly diverse amino acid composition, despite being restricted in length (12–15 amino acids) (14, 26, 27). On the other hand, it has been shown that the epitope of this type of RF is mostly located at the CH2–CH3 elbow in the Fc region (9). These findings lead to the following questions. Do the IGHV1-69/IGKV3-20 RFs recognize the same epitope with the same binding mechanism? If so, how does the highly diverse CDR-H3 participate in RF binding? Why does the CH2–CH3 cleft region of Fc tend to be the epitope of these RFs? And last, why is this germline combination prone to having RF activity? So far, the crystal structures of two RFs have been determined in complex with Fc. One of these is a RA patient–derived IgM rheumatoid factor, RF-AN, that recognizes the CH2–CH3 cleft and the CH3 domain close to the C terminus (29). However, these RFs are not stereotypic RFs and therefore cannot provide answers to the aforementioned questions.

Here we report the 2.8 Å resolution crystal structure of an IGHV1-69/IGKV3-20 stereotypic monoclonal human RF, YES8c, in complex with the Fc fragment of human IgG1 (IgG1-Fc). YES8c is a monoreactive IgM RF obtained from bone marrow B cells from an RA patient with macroglobulinemia (30). Crystal structure and mutagenesis studies revealed a common mechanism used by a wide range of IGHV1-69/IGKV3-20 stereotypic RFs to interact with the same position in the Fc fragment.

Results

Sequence and binding characteristics of YES8c-Fab

Amino acid sequence analysis of YES8c showed that the germline of the V, D, and J regions in the heavy chain are IGHV1-69, IGHD2-2, and IGHJ4, respectively (Fig. S1A). The germline of the V and J regions in the light chain are IGKV3-20 and IGKJ2, respectively (Fig. S1B). Therefore, YES8c is an IGHV1-69/IGKV3-20 stereotypic RF. According to previous research, YES8c is a monoreactive IgM RF that does not show cross-reactivity to self-antigens other than IgG-Fc; however, the precise binding characteristics are unknown (30). There are two large groups of the IgG subclass-specific RFs, Ga (binds to IgG1, 2 and 4) and Pan (binds to all subclasses) (10). We investigated the subclass specificity of YES8c. Although YES8c showed a preference for IgG1, it interacted with IgG1 and IgG4 based on ELISA results (Fig. 1A). This specificity was also confirmed by surface plasmon resonance (SPR) (Fig. S2). Based on these results, YES8c was categorized as another subclass specificity group. We measured the affinity of YES8c Fab for IgG1 by SPR (Fig. 1, B and C). The dissociation constant (Kd) was 160 ± 30 μM.

YES8c–IgG1-Fc complex structure and implications for low-affinity interactions

The structure of YES8c–Fab in complex with human IgG1-Fc was determined at 2.8 Å resolution (Table 1). Two YES8c–IgG1-Fc complexes were observed in the asymmetric crystal unit (Fig. 2A). Each YES8c–Fc complex was designated complex 1 or complex 2. In complex 2, the electron density of the constant region of YES8c was too weak to build atomic models, which might be due to crystal packing (Fig. S3). The heavy chain of YES8c recognizes the CH2–CH3 cleft and the light chain mainly recognizes the CH3 domain. The binding site of the heavy chain overlaps with that of S. aureus protein A, which competitively inhibits the binding of most RFs to Fc regions (Fig. 2B) (31, 32).

In the case of the rheumatoid factor RF-AN, the Fc region was bound to the edge of the conventional antigen-binding site of Fab, deviating from the pseudo-axis between VH and VL
Structure of the stereotypic rheumatoid factor

Table 1
Crystallographic data collection and refinement statistics

| YES8c-Fab–IgG1-Fc | YES8c-Fab–IgG1–IgG1-Fc |
|--------------------|------------------------|
| **Data collection** |                        |
| Space group        | P6_3                   |
| Number of crystals | 1                      |
| Cell dimensions    |                         |
| a, b, c (Å)        | 189.9, 189.9, 79.6      |
| α, β, γ (°)        | 90.0, 90.0, 120.0       |
| Resolution (Å)     | 20.0–2.80 (2.85–2.80)* |
| Rmerge (%)         | 0.074 (0.355)*          |
| CC50 of the highest-resolution shell (%) | 0.917 |
| rmerge (%)         | 19.1 (2.6)*             |
| Completeness (%)   | 98.7 (95.1)*            |
| Redundancy (%)     | 4.6 (3.8)*              |

| **Refinement** |                        |
| Resolution (Å) | 20.0–2.80               |
| No. of reflections | 40,124                 |
| Rwork (%)        | 0.221                   |
| Rfree (%)        | 0.270                   |
| No. of atoms     | 8,175                   |
| Protein          | 54                      |
| Water            | 52.0                    |
| Average B-factor | 64.32                   |
| RMSDs            | Bond lengths (Å)        | 0.003                   |
|                  | Angles (°)              | 0.632                   |
| Ramachandran plot (%) |                          |
| Favored         | 95.12                   |
| Allowed         | 40.4                    |
| Disallowed      | 0.85                    |

*The highest shell statistics are in parentheses.

(Fig. 2C) (28). In contrast, in the case of YES8c, the Fc binds near the center of the axis (Fig. 2C). Consequently, 22 residues of YES8c participate in binding to Fc. This number is larger than those of two structurally determined RFs, RF-AN (nine residues) and RF61 (14 residues), whose affinities are higher than YES8c (Table 2). YES8c has a large buried surface area (BSA) within the complex, 2101 Å² for complex 1 and 1996 Å² for the YES8c–Fc complex. The protrusion at Leu432–His435 (0.72 and 0.73) between the YES8c heavy chain and Fc are larger than those of two structurally determined RFs, RF-AN (nine residues) and RF61 (five residues and 255 Å²) (Table 2). CDR-L1 and CDR-L3 form a flat binding face and recognize the flat β-sheet region of the CH3 domain (Fig. 4A). The shape complementarity score between the light chain and the Fc region was low (0.57 and 0.58) (Table 2). A slight difference in the binding orientation between the light chain and Fc was observed when comparing complex 1 and complex 2 (Fig. S7). This caused differences in the interfacial area at Asn384 and Lys439 in Fc, and L-Ser30 in YES8c (Fig. S5, A and B). To investigate the role of the residues on the light chain involved in Fc recognition, the binding of the Ala mutants was assessed (Fig. 4B). In the ELISA, the binding of the L-S29A, L-S31A, L-Y32A, and L-S94A mutant was significantly lower than that of the WT, and the binding decreased by one-half for L-Y96A, whereas the L-S93A mutant had no effect on the interaction. SPR analysis confirmed the large decreases in binding of L-S94A and L-Y96A (Fig. S8).

Fc recognition by the heavy chain of YES8c

The residues in the light chain were highly conserved among the IGHV1-69/IGKV3-20 RFs throughout the chain (Fig. S4B). For the light chain of YES8c, the number of residues participating in Fc binding (10–11 residues), and the binding interface area (430–450 Å²) is much larger than that of RF-AN (two residues and 196 Å²) and RF61 (five residues and 255 Å²) (Table 2). CDR-L1 and CDR-L3 form a flat binding face and recognize the flat β-sheet region of the CH3 domain (Fig. 4A). The shape complementarity score between the light chain and the Fc region was low (0.57 and 0.58) (Table 2). A slight difference in the binding orientation between the light chain and Fc was observed when comparing complex 1 and complex 2 (Fig. S7). This caused differences in the interfacial area at Asn384 and Lys439 in Fc, and L-Ser30 in YES8c (Fig. S5, A and B). To investigate the role of the residues on the light chain involved in Fc recognition, the binding of the Ala mutants was assessed (Fig. 4B). In the ELISA, the binding of the L-S29A, L-S31A, L-Y32A, and L-S94A mutant was significantly lower than that of the WT, and the binding decreased by one-half for L-Y96A, whereas the L-S93A mutant had no effect on the interaction. SPR analysis confirmed the large decreases in binding of L-S94A and L-Y96A (Fig. S8).

The effects of somatic hypermutations on YES8c

According to previous reports, some antibodies acquire RF activity as a result of somatic hypermutations (SHMs), and others possess RF activity without SHMs (26, 35, 36). Although the structures of RF-AN and RF61 indicated the importance of SHMs for RF activity, binding analyses using mutants have not been performed. In this study, we investigated the effects of SHMs on YES8c by mutation analysis. In the light chain, there are three amino acid substitutions in the V region (H-A33P, H-M48V, H-I53L) and one substitution in the J region (H-D100EF) (Fig. S1A). In the heavy chain, there are seven amino acid substitutions in the V region (H-A33P, H-M48V, H-I53L, H-K62R, H-S76R, H-L82V, and H-S82BV) and one substitution in the J region (H-D100EF) (Fig. S1A). Among them, the residues near the Fc binding site in YES8c (Leu28, His33, and His53) were substituted with the germline residues (L-I28V, H-P33A, and CH3 (Fig. 3C). As pointed out in previous reports, the amino acid composition of CDR-H3 in IGHV1-69/IGKV3-20 RFs is diverse, although the length is restricted to 12–15 residues (Fig. 3D) (14, 17, 27).

The hydrophobicity of the CDR-H2 tip is a characteristic of the heavy chain derived from the IGHV1-69 germline. In YES8c, the hydrophobic residues (H-Leu53 and H-Phe54) at the CDR-H2 tip interact with the hydrophobic moiety in the CH2–CH3 cleft comprised of Ile253 and Leu214 (Fig. 3E). The residues in CDR-H2 form a large interaction interface (Fig. S5B). Comparing the sequence of the IGHV1-69/IGKV3-20 RFs, the hydrophobicity in CDR-H2 is highly conserved (Fig. 3F). To investigate the role of the residues in CDR-H2, the binding of the Ala mutants of YES8c was measured by ELISA (Fig. 3G). The bindings were lower in H-I52A and H-L53A mutants compared with the WT. The binding was significantly lower in the H-F54A mutant. SPR analysis confirmed the large decrease in binding of the H-F54A mutant (Fig. S6).
and H-L53I), and the bindings were analyzed by ELISA (Fig. 5A). In the L-I28V mutant, the binding was similar to that of the WT. The binding was lower in the H-P33A mutant than in the WT. In contrast, binding increased in the H-L53I mutant compared with WT. Binding of the mutants was also analyzed by SPR (Fig. 5B and Fig. S9). A binding response similar to that of the WT was observed for L-I28V. A response was not observed for H-P33A, even at the highest concentration. The affinity of H-L53I (\(K_d = 84 \mu M\)) was 2-fold higher than that of WT.

Table 2

|                  | YES8c (Complex 1) | YES8c (Complex 2) | RF-AN | RF61 |
|------------------|-------------------|-------------------|-------|------|
| Contact residues |                   |                   |       |      |
| L chain          | 10                | 11                | 2     | 5    |
| H chain          | 14                | 10                | 7     | 9    |
| Buried surface area (Å) |       |                   |       |      |
| Total            | 2,101             | 1,996             | 1,458 | 1,689 |
| L chain          | 450               | 427               | 196   | 255  |
| H chain          | 593               | 577               | 552   | 578  |
| Fc               | 1,058             | 992               | 710   | 856  |
| Salt bridges*    | 0                 | 0                 | 1     | 2    |
| Hydrogen bonds*  | 7                 | 8                 | 5     | 12   |
| van der Waals contacts* |       |                   |       |      |
| Shape complementarity (Sc score) |       |                   |       |      |
| LH-Fc            | 0.65              | 0.67              | 0.69  | 0.69 |
| L-Fc             | 0.57              | 0.58              | 0.44  | 0.65 |
| H-Fc             | 0.73              | 0.72              | 0.75  | 0.69 |

* The criteria of the atomic contacts are as follows: C·C, 4.1 Å; C·N, 3.8 Å; C·O, 3.7 Å; O·O, 3.3 Å; O·N, 3.4 Å; N·N, 3.4 Å; C·S, 4.1 Å; O·S, 3.7 Å; N·S, 3.8 Å.

Discussion

The RFs derived from the IGHV1-69/IGKV3-20 germline combination have long been studied as paraprotein-IgM RFs observed in MC and WM. IGHV1-69/IGKV3-20 RFs are detected at a high level as a stereotypic RF in patients with B cell lymphoma caused by HCV infection (13–15). The germline combinations for RF-AN and RF61 are IGHV3-9/IGLV1-47 and IGHV4-39/IGLV1-47, respectively, and the crystal structures of both RFs could not explain the binding properties of stereotypic RFs (28, 29). Here we determined the crystal structure of the IGHV1-69/IGKV3-20 RF, designated YES8c, in complex with IgG1-Fc.

![Figure 2. Crystal structure of YES8c in complex with IgG1-Fc.](image-url)
Binding analysis showed that YES8c interacted with IgG1 and IgG4 despite some preference for IgG1. The epitope residues were different for IgG3 and the other subclasses at the 384, 422, 435, and 436 positions (Table S2). In particular, position 435, which is considered the key residue for the Ga specificity (9), is histidine in IgG1, 2, and 4 but arginine in IgG3. The structure of the complex suggested that the side chain at this position collides with H-Phe54 containing an H435R substitu-
Figure 5. Binding of the germline-reverted mutant RFs to IgG1-Fc. A, binding of the germline-reverted mutants to IgG1-Fc was measured by ELISA. Data show mean ± S.D. (error bars) from triplicates from one representative experiment of two independent experiments with similar results. B, plot of equilibrium binding responses against concentration of YES8c Fab mutants for the sensorgrams shown in Fig. S10. The solid line represents the nonlinear fit of the 1:1 binding model for the WT and the H-L53I mutant that gave $K_d$ values of 130 $\mu M$ and 84 $\mu M$, respectively. The responses for L-I28V and H-P33A could not be fit with the 1:1 binding model. RU, resonance unit.

With a larger interfacial area and complementarity score than the light chain, the heavy chain likely plays a central role in the YES8c–Fc interaction. We propose a model that allows for the diversity of the CDR-H3 loop of IGHV1-69/IGKV3-20 RFs in the RF-Fc binding site. In the YES8c–IgG1–Fc structure, the protrusion at Leu$^{432}$–His$^{435}$ in Fc sticks into the VL–VH pocket in YES8c, and Asn$^{434}$ interacts with the main chain of CDR-H3 from inside. This may prevent the CDR-H3 loop from inserting into the binding interface. Consequently, a large variation in CDR-H3 could be accommodated in the space between CH2 and CH3 with only a slight change in the loop structure. However, it has been suggested that stereotypic RFs are restricted with regards to the CDR-H3 length (14, 27). Most of the IGHV1-69/IGKV3-20 RFs have a CDR-H3 composed of 12–15 amino acids. Considering the YES8c–Fc structure, it is likely that the interaction between CDR-H3 and Fc is not possible when CDR-H3 is shorter than 12 amino acids. In contrast, if CDR-H3 is longer than 15 amino acids, then the loop may not be accepted in the space between CH2 and CH3, or it may interrupt the binding of the protrusion at Leu$^{432}$–His$^{435}$ by covering the VL–VH pocket. To validate this model, structure determination of the other IGHV1-69/IGKV3-20 RFs with different CDR-H3 will be required.

Hydrophobic residues in CDR-H2 are considered to participate in the pattern recognition–like interaction of naive antibodies derived from the IGHV1-69 germline. According to the crystal structures of the anti-influenza hemagglutinin antibody CR6261 (37), anti-HCV E2 envelope antibody AR3C (18), and anti-HIV-1 gp41 antibody D5 (20), the CDR-H2 of the IGHV1-69–derived antibody recognizes a hydrophobic patch in the antigen. Therefore, the naïve antibodies derived from IGHV1-69 are considered pattern recognition antibodies with specificity for CDR-H2. Our mutational analysis showed that the binding of H-F54A mutants to IgG1-Fc largely decreased, indicating that the hydrophobic tip of CDR-H2 (H-Phe54) contributes to the RF activity of YES8c.

The YES8c–Fc structure showed that the interface between the light chain and Fc was flat and the binding was loose. However, the number of YES8c light chain residues interacting with IgG1-Fc and the interfacial area between the light chain and Fc are much larger than found for RF-AN and RF61. The residues involved in the YES8c–Fc interaction were highly conserved among the IGHV1-69/IGKV3-20 RFs (Fig. S4B). Additionally, a large decrease in binding was observed for substitution of L-Ser29, L-Tyr32, and L-Ser94 to Ala. These findings suggest that the recognition of the highly conserved hydrophilic patch by the light chain is important for the recognition of IGHV1-69/IGKV3-20 RFs despite its loose and flat interface.

In this study, the binding of the germline-reverted mutants with IgG1-Fc was explored. Binding of the H-P33A mutant was largely decreased, indicating that the SHM at this position is important for conferring RF activity to YES8c. In contrast, the affinity of the H-L53I mutant to IgG1-Fc doubled, with no changes observed for the L-I28V mutant. It remains unclear whether a combined germline-reverted mutant has RF activity. In the case of RF-TS1, an IGHV1-69/IGKV3-20 RF, the naïve antibody without SHMs had a higher affinity for Fc than RF-TS1 (36). On the other hand, in mixed cryoglobulinemia related to HCV, the IGHV1-69/IGKV3-20 IgM acquired RF activity with SHMs (26). Furthermore, we showed that the affinity of WT YES8c for IgG1–Fc was further increased by a single-base substitution in the antigen-binding site (H-L53I, H-T56K, H-N58K, L-Q27E, and L-S27AN), indicating that further SHMs could make YES8c an aggravated pathogenic factor.

YES8c binds to the CH2–CH3 elbow in the canonical antigen-binding manner. Combined with mutational analyses, we propose a recognition mechanism common to the IGHV1-69/IGKV3-20 stereotypic RFs that mediates RF activity. Additionally, amino acid substitutions increased the affinity of YES8c to Fc, showing the potential of this RF for aggravating the viru-
Structure of the stereotypic rheumatoid factor

Figure 6. The YES8c sites relative to the increase in binding affinity for IgG1-Fc. A, the residues (H-Thr56 and H-Asn58) in CDR-H2 located near the acidic residues on Fc (Glu445 and Glu449), B, the residues (L-Gln27 and L-Ser27A) in CDR-L1. C, plot of equilibrium binding responses against concentration of YES8c-Fab mutants for the sensorgrams shown in Fig. S11, B–E. The solid line represents a nonlinear fit of the 1:1 binding model for H-T56K, H-N58K, L-Q27E, and L-S27AN. The K_d values are indicated. RU, resonance unit.

Discussion

The increased binding affinity of YES8c-Fab for IgG1-Fc is influenced by the residues (H-Thr56 and H-Asn58) in CDR-H2 and the residues (L-Gln27 and L-Ser27A) in CDR-L1. These findings indicate that the heavy chain (VH) from YES8c IgM and the light chain composed of variable domain of the light chain (VL) from YES8c IgM and the CH1 from IgG1 are important for the increased binding affinity of YES8c-Fab for IgG1-Fc. Furthermore, the residues (L-Gln27 and L-Ser27A) in CDR-L1 may also contribute to the increased binding affinity of YES8c-Fab for IgG1-Fc.

Experimental procedures

Preparation of YES8c-Fab expressed in Escherichia coli

YES8c-Fab was prepared for crystallization as described in our previous report (40). In brief, the heavy chain composed of variable domain of the heavy chain (VH) from YES8c IgM and CH1 from IgG1 and the light chain composed of variable domain of the light chain (VL) from YES8c IgM and the κ light chain (CLκ) were separately expressed in E. coli as inclusion bodies. The inclusion bodies were solubilized in a buffer containing 6 M guanidine hydrochloride, and then the heavy chain and light chain were mixed and refolded using a stepwise dialysis method. The refolded Fab was purified by Resource S cation exchange chromatography (GE Healthcare).

Preparation of YES8c-Fab mutants in the silkworm baculovirus expression system

The precise methods are described in the Supplemental Experimental Procedures. In brief, the recombinant bacmid DNAs of the heavy and light chains were generated separately by transfection of bacmid DNA into Bombyx mori–derived cells. The recombinant virus mixture of the heavy and light chains were mixed and refolded using a stepwise dialysis method. The refolded Fab was purified by Resource S cation exchange chromatography (GE Healthcare).

Preparation of the IgG1-Fc fragment

The humanized IgG1 mAb rituximab (Chugai Pharma, Tokyo, Japan) was treated with 1/10 (w/w) of papain activated with 10 mM L-cysteine for 3 h at 37 °C. The papain-digested IgG1 was applied to COSMOGEL® Ig-Accept protein A resin (Nakalai, Kyoto, Japan). The resin was washed with 5 column volumes of wash buffer (50 mM Tris-HCl and 100 mM NaCl (pH 7.0)). Protein was eluted with elution buffer (0.1 M Gly-HCl (pH 3.0)), and the eluent was mixed with 1 M Tris-HCl (pH 8.5). The Fc fragment was further purified by gel filtration using a Superdex200 10/300 column with running buffer (20 mM Tris-HCl and 150 mM NaCl (pH 7.5)).

ELISA

Human IgG2 (ab90284), IgG3 (ab118426), and IgG4 (ab90286) were purchased from Abcam. IgG was immobilized on a 96-well Maxisorb immunoplate (Thermo Fisher Scientific) by incubation with 50 μl/well of 20 μg/ml IgG in immobilization buffer (50 mM Na_2CO_3-NaHCO_3 and 10% NaN_3 (pH 9.6)) for 1 h at 25 °C. The plate was washed three times with PBS-T (PBS with 0.05% Tween 20) and then blocked with 5% skim milk for 1 h at 25 °C. After washing with PBS-T, the wells were treated with Strept-Tactin® alkaline phosphatase conjugate (Iba Lifesciences) diluted with PBS-T and incubated for 1 h at 37 °C. After washing with PBS-T, 50 μl/well of 1 mg/ml p-nitrophenyl phosphate was added, and then absorbance at 405 nm was measured using a Scientific Multiscan plate reader (Thermo Fisher Scientific).

Surface plasmon resonance

The affinity of YES8c-Fab for IgG1-Fc was measured by SPR using a Biacore X100 or Biacore 3000 instrument (GE Healthcare). HBS-EP buffer (10 mM HEPES (pH 7.4), 150 mM NaCl, 3 mM EDTA, and 0.005% (v/v) surfactant P20) was used as the running buffer. YES8c-Fab was dialyzed overnight into HBS-EP buffer. Approximately 1000 resonance units of human IgG were immobilized on a CM5 sensor chip by amine coupling. The reference flow cell was prepared by blocking with ethanolamine after activation with 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide and N-hydroxysuccinimide. Sequentially diluted YES8c-Fab solutions were injected serially over the flow cells for 60 s at a flow rate of 30 μl/min. The maximum concentration of YES8c-Fab was limited to 40 μM because of its tendency for nonspecific binding at high concentration. The reference response was subtracted from the sensorgram to obtain the actual binding response. The equilibrium dissociation constant (K_d) was obtained by fitting a plot of the responses at equilibrium against the concentration with one-site-specific binding model using Prism 6.0 software.
Crystallization, data collection, and structure determination

YES8c–Fab expressed in E. coli and IgG1–Fc were mixed at a molar ratio of 2:1 and concentrated to 10 mg/mL. The crystals were obtained by sitting drop vapor diffusion in a reservoir solution (0.1 M Tris-HCl (pH 8.5), 0.2 M sodium acetate, and 18% (w/v) PEG4000) and incubated for 1 week at 20 °C. The crystals were soaked in soaking buffer (reservoir solution with 20% glycerol) and flash-frozen in liquid nitrogen. X-ray diffraction data were collected at the beamline AR-NE3A at the Photon Factory (Tsukuba, Japan). The diffraction data were processed using HKL2000 (HKL Research Inc.). The structure was determined by molecular replacement using Phaser-MR (42) (PHENIX software package (43)) with the Fab structure (PDB code 1VGE) and human Fc structure (PDB code 3DO3) as search models. Model rebuilding and refinement were performed using Coot (44) and phenix.refine (45), respectively. The final Rwork/Rfree factors were 0.221/0.270. The crystallographic statistics are shown in Table S1. The intermolecular atomic contacts were determined using the program CONTACT in the CCP4 suite (46). The buried surface area was calculated with a default probe radius of 1.4 Å using the program AREAIMOL (47) in the CCP4 suite. The shape complementarity was calculated using the program SC in CCP4 with a default probe radius of 1.7 Å.

Author contributions—M. S. and T. U. designed and supervised the research. M. S., Y. I., K. S., and J. M. L. prepared proteins. M. S. and K. S. performed crystallization. M. S. performed diffraction data collection and structure determination. Y. I. and K. S. performed the binding study. Data were analyzed by M. S., Y. I., J. M. L., T. K., and T. U. M. S., J. M. L., T. K., and T. U. wrote the manuscript.

Acknowledgments—We thank the Research Support Center, Research Center for Human Disease Modeling, Kyushu University Graduate School of Medical Sciences for technical assistance. We thank Prof. Kouhei Tsumoto at the University of Tokyo and Dr. Jose Caaveiro at the School of Medical Sciences for technical assistance. We thank Prof. T. U. M. S., J. M. L., T. K., and T. U. designed and supervised the research. M. S., Y. I., K. S., and J. M. L. prepared proteins. M. S. and K. S. performed crystallization. M. S. performed diffraction data collection and structure determination. Y. I. and K. S. performed the binding study. Data were analyzed by M. S., Y. I., J. M. L., T. K., and T. U. M. S., J. M. L., T. K., and T. U. wrote the manuscript.

References
1. Randen, I., Thompson, K. M., Pascual, V., Victor, K., Beale, D., Cordwell, J., Forre, O., Capra, J. D., and Natvig, J. B. (1992) Rheumatoid factor V genes from patients with rheumatoid arthritis are diverse and show evidence of an antigen-driven response. *Immunol. Rev.* **128**, 49–71CrossRef Medline
2. Pascual, V., Randen, I., Thompson, K., Sioud, M., Forre, O., Natvig, J., and Capra, J. D. (1990) The complete nucleotide sequences of the heavy chain variable regions of six monospecific rheumatoid factors derived from Epstein-Barr virus-transformed B cells isolated from the synovial tissue of patients with rheumatoid arthritis: further evidence that some autoantibodies are unmutated copies of germ line genes. *J. Clin. Invest.* **86**, 1320–1328CrossRef Medline
3. Victor, K. D., Randen, I., Thompson, K., Forre, O., Natvig, J. B., Fu, S. M., and Capra, J. D. (1991) Rheumatoid factors isolated from patients with autoimmune disorders are derived from germ line genes distinct from those encoding the Wa, Po, and B1a cross-reactive idiotypes. *J. Clin. Invest.* **87**, 1603–1613CrossRef Medline
4. Pascual, V., Victor, K., Randen, I., Thompson, K., Steinitz, M., Forre, O., Fu, S. M., Natvig, J. B., and Capra, J. D. (1992) Nucleotide sequence analysis of rheumatoid factors and polyreactive antibodies derived from patients with rheumatoid arthritis reveals diverse use of VH and VL gene segments and extensive variability in CDR-3. *Scand. J. Immunol.* **36**, 349–362CrossRef Medline
5. Ermel, R. W., Kenny, T. P., Chen, P. N., and Robbins, D. L. (1993) Molecular analysis of rheumatoid factors derived from rheumatoid synovium suggests an antigen-driven response in inflamed joints. *Arthritis Rheum.* **36**, 380–388CrossRef Medline
6. Mantovani, L., Wilder, R. L., and Casali, P. (1993) Human rheumatoid B-1a (CD5+ B) cells make somatically hypermutated high affinity IgM rheumatoid factors. *J. Immunol.* **151**, 473–488Medline
7. Youngblood, K., Fruchter, L., Ding, G., Lopez, I., Bonagura, V., and Davidson, A. (1994) Rheumatoid factors from the peripheral blood of two patients with rheumatoid arthritis are genetically heterogeneous and somatically mutated. *J. Clin. Invest.* **93**, 852–861CrossRef Medline
8. Artandi, S. E., Canfield, S. M., Tao, M. H., Calame, K. L., Morrison, S. L., and Bonagura, V. R. (1991) Molecular analysis of IgM rheumatoid factor binding to chimeric IgG. *J. Immunol.* **146**, 603–610Medline
9. Artandi, S. E., Calame, K. L., Morrison, S. L., and Bonagura, V. R. (1992) Monoclonal IgM rheumatoid factors bind IgG at a discontinuous epitope comprised of amino acid loops from heavy-chain constant-region domains 2 and 3. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 94–98CrossRef Medline
10. Bonagura, V. R., Artandi, S. E., Davidson, A., Randen, I., Agostino, N., Thompson, K., Natvig, J. B., and Morrison, S. L. (1993) Mapping studies reveal unique epitopes on IgG recognized by rheumatoid arthritis-derived monoclonal rheumatoid factors. *J. Immunol.* **151**, 3840–3852Medline
11. Silverman, G. J., Goldfien, R. D., Chen, P., Mageed, R. A., Jefferis, R., Gotli, F., Frangione, B., Fong, S., and Carson, D. A. (1988) Idiotypic and subgroup analysis of human monoclonal rheumatoid factors: implications for structural and genetic basis of autoantibodies in humans. *J. Clin. Invest.* **82**, 469–475CrossRef Medline
12. Thompson, K. M., Randen, I., Berretzen, M., Forre, O., and Natvig, J. B. (1994) Variable region gene usage of human monoclonal rheumatoid factors derived from healthy donors following immunization. *Eur. J. Immunol.* **24**, 1771–1778CrossRef Medline
13. Charles, E. D., Green, R. M., Marukian, S., Talal, A. H., Lake-Bakaar, G. V., Jacobson, I. M., Rice, C. M., and Dustin, L. B. (2008) Clonal expansion of immunoglobulin M + CD27+ B cells in HCV-associated mixed cryoglobulinaemia. *Blood* **111**, 1344–1356Medline
14. Bende, R. J., Aarts, W. M., Riedl, R. G., de Jong, D., Pals, S. T., and van Noesel, C. J. (2005) Among B cell non-Hodgkin’s lymphomas, MALT lymphomas express a unique antibody repertoire with frequent rheumatoid factor reactivity. *J. Exp. Med.* **201**, 1229–1241CrossRef Medline
15. De Re, V., De Vita, S., Gasparotto, D., Marzotto, A., Carbone, A., Ferraccioli, G., and Boiocchi, M. (2002) Salivary gland B cell lymphoproliferative disorders in Sjogren’s syndrome present a restricted use of antigen receptor gene segments similar to those used by hepatitis C virus-associated non-Hodgkin’s lymphomas. *Eur. J. Immunol.* **32**, 903–910CrossRef Medline
16. Stamatopoulos, K., Belessi, C., Moreno, C., Boudjograh, M., Guida, G., Smilevska, T., Belhou, L., Stella, S., Stavroyianni, N., Crespo, M., Hadzimitiou, A., Sutton, L., Bosch, F., Lautaris, N., Anagnostopoulos, A., et al. (2007) Over 20% of patients with chronic lymphocytic leukemia carry stereotyped receptors: pathogenetic implications and clinical correlations. *Blood* **109**, 259–270CrossRef Medline
17. Bende, R. J., Slot, L. M., Hoogeboom, R., Wormhoudt, T. A., Adeoye, O. A., Guikema, J. E., and van Noesel, C. J. (2015) Stereotypic rheumatoid factors that are frequently expressed in mucosa-associated lymphoid tissue-type lymphomas are rare in the labial salivary glands of patients with Sjogren’s syndrome. *Arthritis Rheumatol.* **67**, 1074–1083CrossRef Medline
18. Kong, L., Giao, E., Nieuwja, T., Kadam, R. U., Cogburn, K. E., Hua, Y., Dai, X., Stanfield, R. L., Burton, D. R., Ward, A. B., Wilson, I. A., and Law, M. (2013) Hepatitis C virus E2 envelope glycoprotein core structure. *Science* **342**, 1090–1094CrossRef Medline
19. Chen, C. H., Hadlock, K. G., Foug, S. K., and Levy, S. (2001) V(H)1-69 gene is preferentially used by hepatitis C virus-associated B cell lympho-
Structure of the stereotypic rheumatoid factor

mas and by normal B cells responding to the E2 viral antigen. Blood 97, 1023–1026 CrossRef Medline

20. Luftig, M. A., Mattu, M., Di Giovinone, P., Geleziunas, R., Hrin, R., Barbato, G., Bianchi, E., Miller, M. D., Pessi, A., and Carfi, A. (2006) Structural basis for HIV-1 neutralization by a gp41 fusion intermediate-directed antibody. Nat. Struct. Mol. Biol. 13, 740–747 CrossRef Medline

21. Ekiert, D. C., Bhabha, G., Elsliger, M. A., Friesen, R. H., Jongeneelen, M., Dustin, L. B. (2013) Somatic hypermutations confer rheumatoid factor RF61 and IgG1 Fc reveals a novel epitope and evidence for affinity maturation. J. Mol. Biol. 368, 620–637 CrossRef Medline

22. Ying, T., Prabakaran, P., Du, L., Shi, W., Feng, Y., Wang, Y., Wang, L., Li, J., Sasso, E. H., Barber, C. V., Nardella, F. A., Yount, W. J., and Mannik, M. (2016) Antigen specificities of human monoclonal and polyclonal IgM rheumatoid factors: the C γ2-C γ3 interface region contains the major determinants. J. Immunol. 140, 3098–3107 Medline

23. Deisenhofer, J. (1981) Crystallographic refinement and atomic models of a human Fc fragment and its complex with fragment B of protein A from Staphylococcus aureus at 2.9- and 2.8-A resolution. Biochemistry 20, 2381–2370 CrossRef Medline

24. Sasso, E. H., Barber, C. V., Nardella, F. A., Yount, W. J., and Mannik, M. (1988) Antigenic specificities of human monoclonal and polyclonal IgM rheumatoid factors: the C γ2-C γ3 interface contains the major determinants. J. Immunol. 140, 3098–3107 Medline

25. Carayannopoulou, M. O., Potter, K. N., Li, Y., Natvig, J. B., and Capra, J. D. (2000) Evidence that human immunoglobulin M rheumatoid factors can be derived from the natural autoantibody pool and undergo an antigen driven immune response in which somatically mutated rheumatoid factors have lower affinities for immunoglobulin G Fc than their germline counterparts. Scand. J. Immunol. 51, 327–336 CrossRef Medline

26. Charles, E. D., Bhabha, G., Elsliger, M. A., Friesen, R. H., Jongeneelen, M., Dustin, L. B., and Thompson, K. M. (1995) Structural basis for the functional analysis of BmNPV he65. Biochemistry 34, 343–350 CrossRef Medline

27. Afonine, P. V., Grosse-Kunstleve, R. W., Headd, J. J., Moriarty, N. W., Oeffner, R., Read, R. J., and McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 66, 213–221 CrossRef Medline

28. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132 CrossRef Medline

29. Afonine, P. V., Grosse-Kunstleve, R. W., Emsley, P., Evans, P. R., Keegan, R. M., Kriissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Muskhot, G. D., Pannu, N. S., Potterton, E. A., Powell, H. R., et al. (2011) Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. 67, 252–252 CrossRef Medline

30. Saif, E. B., and Kullaars, A. B. J. (1997) Distributing many points on a sphere. Math. In-tell. 19, 5–11

31. Eisenberg, D., Schwartz, E., Komaromy, M., and Wall, R. (1984) Analysis of membrane and surface protein sequences with the hydrophobic moment plot. J. Mol. Biol. 179, 125–142 CrossRef Medline

32. Afonine, P. V., Grosse-Kunstleve, R. W., Emsley, P., Evans, P. R., Keegan, R. M., Kriissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Muskhot, G. D., Pannu, N. S., Potterton, E. A., Powell, H. R., et al. (2011) Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. 67, 252–252 CrossRef Medline

33. Sasso, E. H., Barber, C. V., Nardella, F. A., Yount, W. J., and Mannik, M. (1988) Antigenic specificities of human monoclonal and polyclonal IgM rheumatoid factors: the C γ2-C γ3 interface contains the major determinants. J. Immunol. 140, 3098–3107 Medline

34. Lawrence, M. C., and Colman, P. M. (1993) Shape complementarity at protein/protein interfaces. J. Mol. Biol. 234, 946–950 CrossRef Medline

35. Randen, I., Brown, D., Thompson, K. M., Hughes-Intones, N., Pascal, V., Victor, K., Capra, J. D., Farre, O., and Natvig, J. B. (1992) Clonally related IgM rheumatoid factors undergo affinity maturation in the rheumatoid synovial tissue. Immunol. 148, 3296–3301 Medline

36. Dustin, L. B. (2013) Somatic hypermutations confer rheumatoid factor RF61 and IgG1 Fc reveals a novel epitope and evidence for affinity maturation. J. Mol. Biol. 368, 620–637 CrossRef Medline