Universiteit
Leiden
The Netherlands

Dysregulation of autoreactive B cell responses in autoimmune diseases: from initial triggering to persistent activation

Neppelenbroek, S.

Citation

Neppelenbroek, S. (2026, July 9). *Dysregulation of autoreactive B cell responses in autoimmune diseases: from initial triggering to persistent activation*. Retrieved from https://hdl.handle.net/1887/4307822

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/4307822

Note: To cite this publication please use the final published version (if applicable).

5

CHAPTER 5

Clonal autoantibodies identify fungal antigens as triggers of autoreactive B cells in a human autoimmune disease

Sam Neppelenbroek[1], Sophie I.E. Liem[1], Theo van Laar[2], Eva M. Hoekstra[1], Corrie M. Wortel[1], E.W. Nivine Levarht[1], Cynthia M. Fehres[1], Nynke H. Dekker[2,3], Jeska K. de Vries-Bouwstra[1], René E.M. Toes[1], Hans U. Scherer[1]

[1] Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands.
[2] Department of BioNanoScience, Kavli Institute of Nanoscience, Delft University of Technology, Delft, The Netherlands.
[3] Department of Physics and Kavli Institute of Nanoscience Discovery, University of Oxford, Oxford, United Kingdom.

Abstract

Transformative observations indicate unprecedented success of B cell-depleting interventions in many human autoimmune diseases, calling for a deeper understanding of the triggers leading to B cell-mediated autoimmunity and its maintenance in human disease. Here, we show that the autoimmune response to human topoisomerase 1 (TOP1), a hallmark of the severe autoimmune disease systemic sclerosis characterized by vasculopathy and fibrosis, reacts to TOP1 of microbial origin. More specifically, taking TOP1 from *S. cerevisiae* as prototype, fungal TOP1 was specifically recognized at the polyclonal and monoclonal level by patient IgG. Importantly, it also activated B cells expressing a patient-derived, human TOP1-reactive B cell receptor, providing functional evidence for a breach of B cell tolerance. Notably, patients affected by interstitial lung disease, the most deleterious disease manifestation, most frequently showed recognition of fungal TOP1. These findings identify fungi as potential drivers of immune dysregulation in human autoimmunity and make microbial antigen-reactive cells important therapeutic targets.

One sentence summary

Recognition of microbial antigen by autoantibodies mediates autoreactive B cell activation and associates with severe disease in systemic sclerosis.

Available as preprint on *BioRxiv, 2026.03.17.712484*

Introduction

Systemic sclerosis (SSc) is a chronic autoimmune disease (AID) in which vasculopathy and fibrosis affect the function of internal organs and the skin. The clinical disease course varies strongly between patients, particularly in terms of organ involvement and rate of disease progression. This heterogeneity, together with a lack of effective and curative treatments, makes the disease unpredictable and complicates its clinical management. Therefore, it is crucial to understand the mechanisms driving the disease to reduce SSc-related morbidity and mortality.

Anti-nuclear autoantibodies hallmark SSc and define its clinical phenotype and course of disease[1]. As is the case for other AIDs, striking clinical improvements have been observed in severely ill patients upon B cell depletion, for example with rituximab or anti-CD19 chimeric antigen-receptor (CAR)-T cells[2-4]. These observations indicate a pivotal role of (autoreactive) B cells in disease initiation and progression. How such B cells are generated, activated and maintained, however, remains unclear. The human nuclear protein topoisomerase 1 (TOP1) is the most prevalent target of autoreactive B cells in patients with diffuse cutaneous SSc[5,6]. Anti-TOP1 autoantibodies (ATAs) associate with fibrosis of skin and lungs as well as mortality[6-8]. Recently, we also found an association between the activity of the ATA B cell response with disease progression and pulmonary fibrosis[9,10]. Hence, it is conceivable that active, TOP1-reactive B cells contribute directly or indirectly to the inflammatory and fibrotic processes in affected organs. Despite these indications, the triggers driving the initiation and activation of TOP1-reactive B cells are undefined.

Microbes have been implicated in the break of B cell tolerance towards self-antigens in several autoimmune diseases[11-13]. Antigens derived from microbes can trigger the B cell receptor (BCR) of autoreactive B cells by structurally resembling the autoantigen, a concept known as molecular mimicry[13]. For example, commensal bacteria expressing orthologs of the human Ro60 protein were identified as targets of autoreactive B cells in systemic lupus erythematosus[14]. By analogy, we considered that antibodies and BCRs reactive to human TOP1 could also recognize TOP1 from microbial species. SSc patients could be exposed to such microbial TOP1 antigens in affected organs, especially the lungs and the skin[15]. Also, the majority of ATA-IgG$^+$ patients harbor ATA-IgA, the predominant isotype in mucosal immune responses[9,16]. Nevertheless, there are only limited studies on the presence and abundance of microbes in affected SSc tissues[17-19]. Moreover, it is unknown whether ATAs can recognize microbial antigens. Elucidating this possibility would be relevant as it may indicate that microbial TOP1 contributes to the breakdown of immune tolerance to human TOP1 in systemic sclerosis, as well as the activation of TOP1-reactive B cells in progressive disease.

Based on the above, we searched for structural homologies between human and microbial TOP1. We studied the recognition of microbial TOP1 by patient-derived ATAs and the ability of microbial TOP1 to activate ATA-expressing B cells. We found that TOP1 from *Saccharomyces cerevisiae* (*S. cerevisiae*), a microbial TOP1 sharing structural similarities with human TOP1, is recognized by selected human ATAs and capable of activating ATA-expressing human B cells. The recognition of this microbial TOP1 was enhanced in severe ATA$^+$ SSc patients and in patients with clinically relevant interstitial lung disease (ILD).

Together, these results provide evidence that microbial TOP1 could be involved in the initiation and activation of TOP1-reactive B cells in patients and suggest that this process could drive severe SSc.

Methods

Study design

Biological samples from SSc patients were obtained at the Department of Rheumatology at Leiden University Medical Center (LUMC, The Netherlands). All SSc patients fulfilled the 2013 American College of Rheumatology (ACR)/European League Against Rheumatism (EULAR) classification criteria for SSc[20]. Patients on B cell-depleting therapies and patients with a history of hematopoietic stem cell transplantation were excluded.

Cohort 1 - Plasma samples from ATA$^+$ SSc patients (n=22), ACA$^+$ SSc patients (n=20) and HDs (n=16) included in a previous study (Table 1)[10]. Patients were recruited irrespective of disease state or organ involvement.

Table 1. Characteristics of SSc patients and healthy donors included in the cohort 1.

	ATA$^+$ SSc patients (n=22)	ACA$^+$ SSc patients (n=20)	Healthy donors (n=16)
Age, years, mean (SD)	52 (16)	62 (16)	37 (12)
Female, n (%)	12 (55%)	18 (90%)	12 (75%)
Duration since first NRP symptom, months, median, (IQR)	63 (39-102)	120 (95-221)	
Diffuse cutaneous SSC, n (%)	7 (32%)	0 (0%)	
mRSS, median (IQR)	5 (2-7)	3 (0-8)	
ILD on HRCT, n (%)	14 (64%)	0 (0%)	
Clinically relevant ILD, n (%)	2 (9%)	0 (0%)	
Pulmonary hypertension, n (%)	0 (0%)	2 (10%)	
Digital ulcers, n (%)	2 (9%)	0 (0%)	
Current immunosuppressive use, n (%)	12 (55%)	6 (30%)	

Patients and controls were included in the context of a previous study in which patient selection and characteristics are described in detail (Wortel *et al*. RMD Open, 2023)[10]. Clinical characteristics represent data collected at moment of sampling. SD = standard deviation, IQR = interquartile range, NRP = non-Raynaud's phenomenon, mRSS = modified Rodnan skin score, ILD = interstitial lung disease, HRCT = high-resolution computer tomography scan, Clinically relevant ILD = presence of ILD on HRCT combined with a forced vital capacity below 80% of the predicted value.

Cohort 2 - To replicate findings in cohort 1 and to explore possible associations with disease severity, serum samples from ATA$^+$ SSc patients (n=39) were selected from the CCISS cohort[21]. Selection was performed by two independent physicians (JKdVB, SIEL), experienced in the field of SSc, with the aim to identify patients at both ends of the SSc disease severity spectrum: severe ATA$^+$ SSc (n=18) versus mild ATA$^+$ SSc (n=21) (Table 2). This approach was chosen because currently available disease activity indices have significant limitations for hypothesis-generating studies. For example, they often fail to distinguish between disease activity and irreversible damage, and inadequately capture involvement of specific organ systems, such as the gastrointestinal tract[22]. Patients with severe SSc were selected by chart review and clinical expertise taking into account all available data

over time including involvement of skin, lung, heart, kidney, gastrointestinal tract and vascular system. Next, patients with mild disease matched with severe patients for age and disease duration. Both physicians had to agree on classification of a patient as being mild or severe; if not, the patient was excluded from cohort 2. Samples collected at time of inclusion in the CCISS cohort were used. Serum samples from age- and sex-matched ACA$^+$ SSc patients (n=10) and HDs (n=10) obtained from the CCISS cohort and Leiden University Medical Center Voluntary Donor Service, respectively, were used as controls (Table 2).

Table 2. Characteristics of SSc patients and healthy donors included in the cohort 2.

	ATA$^+$ SSc patients (n=39)				ACA$^+$ SSc patients (n=10)	Healthy donors (n=10)
	Baseline		Follow-up			
	Mild (n=21)	Severe (n=18)	Mild (n=21)	Severe (n=18)		
Age, years, mean (SD)	46 (16)	50 (14)	-	-	49	51 (3)
Female, n (%)	16(76%)	10 (56%)	-	-	7 (70%)	7 (70%)
Duration since first NRP symptom, months, median, (IQR)	16 (6-64)	12 (4-38)	-	-	109 (76-193)	
Baseline to follow-up, months, median (IQR)	-	-	40 (20-79)	30 (13-55)	-	
Diffuse cutaneous SSC, n (%)	1 (5%)	9 (50%)	1 (5%)	12 (67%)	1 (10%)	
mRSS, median (IQR)	4 (2-7)	24 (3-23)	4 (2-6)	9 (4-20)	4 (2-5)	
ILD on HRCT, n (%)	10 (48%)	8 (44%)	12 (57%)	14 (78%)	0 (0%)	
Clinically relevant ILD, n (%)	1 (5%)	3 (17%)	1 (5%)	5 (28%)	0 (0%)	
Pulmonary hypertension, n (%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	
Digital ulcers, n (%)	3 (14%)	1 (6%)	1 (5%)	5 (28%)	2 (20%)	
ASCT between baseline & follow-up, n (%)	-	-	0	3 (17%)	-	
Current immunosuppressive use, n (%)	6 (29%)	6 (33%)	8 (38%)	11 (61%)	2 (20%)	
Azathioprine, n	1	0	1	0	0	
Hydroxychloroquine, n	1	0	1	0	1	
Corticosteroids > 5mg/day, n	0	4	0	2	0	
Methotrexate, n	4	1	6	3	2	
Cyclophosphamide, n	0	0	0	3	0	
Mycophenolic acid, n	0	1	0	3	0	
Other (infliximab, tocilizumab), n	0	1	0	2	0	

Baseline characteristics represent clinical data collected at cohort entry. SD = standard deviation, IQR = interquartile range, NRP = non-Raynaud's phenomenon, mRSS = modified Rodnan skin score, ILD = interstitial lung disease, HRCT = high-resolution computer tomography scan, clinically relevant ILD = presence of ILD on HRCT combined with a forced vital capacity below 80% of the predicted value, ASCT = autologous stem cell transplantation.

Cohort 3 - To replicate the clinical associations found in cohort 2, serum samples from all ATA$^+$ SSc patients included in the CCISS cohort (n=163) until 31st of December, 2023, which fulfilled the ACR 2013 criteria for SSc were used[20,21]. Samples collected at time of inclusion in the CCISS cohort were used. Serum samples from age- and sex-matched ACA$^+$ SSc patients (n=50) and HDs (n=30) obtained from the CCISS cohort and Leiden University Medical Center Voluntary Donor Service (LuVDS), respectively, were used as controls (Table 3).

Clinical data acquisition

All included patients participate in the ongoing prospective CCISS cohort in which clinical variables of SSc patients are monitored in a structural follow-up program which included an annual comprehensive assessment of organ involvement[21]. ILD was diagnosed based on the presence of ground glass opacities or interstitial fibrosis on high-resolution

Table 2. Characteristics of SSc patients and healthy donors included in the cohort 2.

	ATA+ SSc patients (n=163)	ACA+ SSc patients (n=50)	Healthy donors (n=30)
Age, years, mean (SD)	51 (15)	53 (12)	47 (6)
Female, n (%)	107 (66%)	31 (62%)	19 (63%)
Duration since first NRP symptom, months, median, (IQR)	17 (6-81)	38 (10-80)	
Diffuse cutaneous SSC, n (%)	64 (39%)	3 (06)	
mRSS, median (IQR)	6 (2-12)	2 (0-4)	
ILD on HRCT, n (%)	99 (61%)	4 (8%)	
Clinically relevant ILD, n (%)	42 (26%)	0 (0%)	
Pulmonary hypertension, n (%)	1 (1%)	3 (7%)	
Digital ulcers, n (%)	21 (13%)	5 (10%)	
Current immunosuppressive use, n (%)	66 (41%)	9 (18%)	
Azathioprine, n	7	2	
Hydroxychloroquine, n	4	2	
Corticosteroids > 5mg/day, n	31	2	
Methotrexate, n	26	3	
Cyclophosphamide, n	4	0	
Mycophenolic acid, n	10	1	
Other (infliximab, tocilizumab), n	5	0	

Baseline characteristics represent clinical data collected at cohort entry. SD = standard deviation, IQR = interquartile range, NRP = non-Raynaud's phenomenon, mRSS = modified Rodnan skin score, ILD = interstitial lung disease, HRCT = high-resolution computer tomography scan, clinically relevant ILD = presence of ILD on HRCT combined with a forced vital capacity below 80% of the predicted value.

computed tomography (HRCT) scan of the thorax. Clinically relevant ILD was defined as presence of ILD on HRCT combined with a FVC below 80% of predicted. Skin involvement was evaluated by assessing the modified Rodnan Skin Score (MRSS)[23]. Pattern of skin involvement was used to categorize patients into diffuse cutaneous, limited cutaneous or non-cutaneous disease subsets.

Study approval

The study was approved by the ethical committee of the Leiden University Medical Center (protocol number P17.151) and the Leiden University Biobank Toetsing Commissie (CME no. B16.037, REU 043/SH/sh, R09.003/SH/s, LuVDS23.039 and LuVDS25.011). All patients and HDs gave written informed consent for study participation.

Analysis of structural homology between human and microbial TOP1 using FoldSeek

Structural homology between human TOP1 and microbial proteins was analyzed using Foldseek[24]. A crystal structure of human TOP1 (1EJ9, Protein Data Bank) was used as input[25]. Foldseek was operated in default mode (3Di/AA) with an E-value cut-off of 0.001 to search databases AlphaFold/Uniprot[50] (v4), AlphaFold/Swiss-Prot (v4), AlphaFold/Proteome (v4), CATH50 (v3.4.0) and PDB100 (version 20240101). Taxonomic filters were applied to specifically identify proteins from fungi, bacteria and viruses.

Recombinant production and purification of human TOP1

Human TOP1 was recombinantly expressed in insect cells (Protein Facility, Leiden University Medical Center) and subsequently purified as described in detail before[26].

In short, a recombinant bacmid encoding full-length human TOP1 with a N-terminal polyhistidine tag (6x histidine) was generated and transfected into sedentary Sf9 insect cells (ThermoFisher). Culture supernatant containing baculovirus was used to transduce suspension Sf9 cells to recombinantly express human TOP1. Upon culturing, TOP1 was purified from cell lysates using a HisTrap HP column (Cytiva) followed by cation exchange chromatography (HiTrap SP FF, Cytiva). Purity and integrity of purified human TOP1 was validated by gel electrophoresis.

Recombinant production and purification of S. cerevisiae TOP1

S. cerevisiae TOP1 expression was induced using galactose in a *S. cerevisiae* cell line with a galactose-inducible *S. cerevisiae* TOP1 gene containing a N-terminal calmodulin purification tag derived from the Calmodulin Binding Protein (CPB) followed by a TEV protease recognition sequence (yAE42 cells, provided by Dr. J.F.X. Diffley[27]). *S. cerevisiae* TOP1 expression was induced by addition of 2% galactose to yAE42 cells cultured in YP medium (1% yeast extract, 2% peptone) supplemented with 2% raffinose. Upon culturing for three hours, cells were harvested and resuspended in TOP1 lysis buffer (25mM Tris-HCl (pH=7.5), 0.02% NP-40 substitute, 10% glycerol, 300mM NaCl, and 1mM DTT) supplemented with protease inhibitors (cOmplete™, EDTA-free Protease Inhibitor Cocktail, Sigma) and 0.3mM phenylmethylsulphonyl fluoride. Cell suspension was dropped into liquid nitrogen and frozen droplets were ground in a freezer mill (6875, SPEX SamplePrep) for six cycles (2 minute run time, 1 minute cool time, rate of 15 cycles per second). Resulting powder was resuspended in TOP1 lysis buffer supplemented with protease inhibitors and lysate was cleared by centrifugation. Cleared lysate was supplemented with $CaCl_2$ to a final concentration of 2mM and incubated with Sepharose 4B Calmodulin Beads (GE Healthcare) for 1 hour at 4°C. Beads were washed 20 times with TOP1 binding buffer (25mM Tris-HCl (pH=7.5), 0.02% NP-40 substitute, 10% glycerol, 300mM NaCl, 2 mM $CaCl_2$ and 1mM DTT) and protein was eluted using TOP1 elution buffer (25mM Tris-HCl (pH=7.5), 0.02% NP-40 substitute, 10% glycerol, 300mM NaCl, 2 mM EDTA, 2mM EGTA, and 1 mM DTT). *S. cerevisiae* TOP1 was further purified by size exclusion chromatography using a Superose 6 increase 10/300 GL column (GE Healthcare) equilibrated in TOP1 GF buffer (25mM Tris-HCl pH 7.5, 0.02% NP-40 substitute, 10% glycerol, 150mM NaCl and 1mM DTT). Purity and integrity of purified *S. cerevisiae* TOP1 was validated by gel electrophoresis.

Generation of patient-derived ATA mAbs

Patient-derived ATA mAbs were generated as described in detail previously[26]. In short, BCR sequences obtained from single-cell sorted human TOP1-reactive B cells of an ATA^+ SSc patient were cloned into plasmids encoding antibody constant domains to recombinantly produce IgG1 antibodies in Freestyle™ 293-F cells (Gibco). Produced IgG was purified from culture supernatants using a Protein G column (GE Healthcare) and purity and integrity of the generated mAbs was validated by gel electrophoresis. An anti-TT-IgG mAb, generated in a similar procedure as described above, was used as negative control.

Gel electrophoresis

The purity and integrity of human and *S. cerevisiae* TOP1 and the mAbs were assessed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Human and *S. cerevisiae* TOP1 was loaded on a polyacrylamide gel with a 4-15% gradient upon addition

of Laemmli buffer (Bio-Rad). The mAbs were loaded on a polyacrylamide gel with a 4-15% gradient upon incubation for five minutes at 95°C in Laemmli buffer in presence (reduced) or absence (non-reduced) of 2% β-mercaptoethanol (Merck). PageRuler™ Plus Prestained Protein ladder (ThermoFisher Scientific) was used as reference. Proteins were separated for 60-90 minutes at 100-120 volt. Upon running the gels, gels were washed in Milli-Q water for 5 minutes and protein was subsequently stained using InstantBlue® Coomassie Protein Stain (Abcam) for one hour. Upon overnight destaining in Milli-Q water, gels were visualized using a ChemiDoc™ Touch Imaging system (Bio-Rad).

Anti-human TOP1-IgG and anti-S. cerevisiae TOP1-IgG ELISA

High binding 384-well microplates (Corning) were coated with 2.5µg/ml human or *S. cerevisiae* TOP1 in PBS (pH=8). Plates were blocked with 1% BSA in PBS and samples were subsequently applied. Serum samples were diluted 3200x or 800x for the ELISA coated with human and *S. cerevisiae* TOP1, respectively. Saturated serum samples were diluted further. The mAbs were applied in a four-fold serial dilution with a starting concentration of 30µg/ml. Pooled serum samples of ATA^+ SSc patients were used as standards. IgG was detected using 1:5000 diluted goat anti-human IgG-HRP (DAKO, P0214) with ABTS as substrate. Plates were measured with a Spectramax I3x Multi-Mode Microplate Reader (Molecular Devices) using SoftMax® Pro 7 (version 7.0.2, Molecular Devices). Samples of HDs were used to determine the cut-off (mean plus two times the standard deviation) for anti-human TOP1-IgG and anti-*S. cerevisiae* TOP1-IgG seropositivity.

Generation of Ramos B cell lines expressing TOP1-reactive IgG BCRs

Ramos B cell lines expressing TOP1-reactive IgG BCRs were generated based on the BCR sequence of B cell clones 2F8, 7G6 and 9D11. To this end, vectors encoding the variable domain of these BCRs as membrane-bound IgG were retrovirally transduced into a Ramos B cell line in which IgD, IgM and AID was knocked out (MDL-AID KO Ramos cells), as described in detail before[28,29]. To determine GFP and surface IgG expression on the transduced cells, Ramos cells were stained with anti-human IgG-Fc BV786 (564230, BD Biosciences) for 30 minutes on ice and subsequently measured using a BD LSRFortessa 4L (BD Biosciences). Data was analyzed using FlowJo™ (version 10.9.0). Ramos cells were sorted repeatedly using a CytoFLEX SRT sorter (Beckman Coulter) based on GFP and surface IgG expression.

Stimulation of ATA-expressing Ramos B cell lines with human and S. cerevisiae TOP1

ATA-expressing Ramos B cell lines were stimulated with 5µg/ml human TOP1, *S. cerevisiae* TOP1 or polyclonal AffiniPure™ $F(ab')_2$ fragment goat anti-human IgG (109-006-097, Jackson ImmunoResearch) for 5 minutes at 37°C in RPMI 1640 supplemented with 100 units/ml penicillin/streptomycin, glutamax, 10mM HEPES (all Gibco) and 1% FCS (Bodinco). Upon stimulation, cells were fixed in 4% paraformaldehyde (BioLegend) for 20 minutes at 37°C and subsequently permeabilized in True-Phos™ Perm Buffer (BioLegend) for 75 minutes at -20°C. Afterwards, cells were stained with anti-Syk(pY352)-Alexa Fluor™ 647 (557817, BD Biosciences) diluted tenfold in PBS, 0.5% BSA, 0.02% NaN_3. Cells were measured on a BD LSRFortessa 4L flow cytometer (BD Biosciences). Data were analyzed using FlowJo (version 10.9.0).

Statistical analysis

The statistical analysis plan with defined primary and secondary research questions (indicated for separate analyses below) was formulated before performance of the statistical analysis. The researcher performing the laboratory studies (SN) was blinded for clinical parameters of SSc patients, and statistical analyses were performed by other researchers (SIEL/EMH). Use of statistical tests are indicated in the figure legends. A *p*-value < 0.05 was considered statistically significant. Statistical analyses were performed using Graphpad Prism (version 9.3.1) and IBM SPSS Statistics (version 29.0.0.0).

Cohort 1-3 - Kruskal Wallis test combined with Dunn's multiple comparisons test was used to statistically compare anti-human TOP1-IgG and anti-*S. cerevisiae* TOP1-IgG levels between ATA^+ SSc patients, ACA^+ SSc patients and HDs. The non-parametric Spearman's rank correlation coefficient was used to describe correlations between anti-human TOP1-IgG and anti-*S. cerevisiae* TOP1-IgG levels.

Cohort 2 - Primary research question: Does the recognition of *S. cerevisiae* TOP1 differ between patients with mild versus severe SSc? Mann-Whitney U test was used to test for statistically significant differences in anti-human TOP1-IgG and anti-*S. cerevisiae* TOP1-IgG levels between patients with mild versus severe SSc. Relation between other patient characteristics (clinically relevant ILD, radiographic ILD, disease subset, sex, MRSS and digital ulcers) and anti-human TOP1-IgG or anti-*S. cerevisiae* TOP1-IgG levels were explored using the Mann-Whitney U test or described using the non-parametric Spearman's rank correlation coefficient for dichotomous or continuous variables, respectively.

Cohort 3 - Primary research question: Does the recognition of *S. cerevisiae* TOP1 associate with clinically relevant ILD? Clinically relevant ILD was defined as primary research outcome based on the relation between anti-*S. cerevisiae* TOP1 and clinically relevant ILD observed in cohort 2, the strong association between ATA and ILD, and clinically relevant ILD being one of the main and most prevalent cause of morbidity and mortality in SSc[30,31]. To evaluate the independent association between the presence of anti-*S. cerevisiae* TOP1-IgG and clinically relevant ILD, a multivariable logistic regression was performed, adjusted for sex, disease subset and elevated C-reactive protein (CRP > 5mg/ml). These variables are known risk factors for interstitial lung disease[30,32-34]. Mann-Whitney U test was used to test for statistically significant differences in anti-human TOP1-IgG and anti-*S. cerevisiae* TOP1-IgG levels between patients with versus without clinically relevant ILD. Relation between other patient characteristics (radiographic ILD, disease subset, sex, MRSS and digital ulcers) and anti-human TOP1-IgG or anti-*S. cerevisiae* TOP1-IgG levels were analyzed using the Mann-Whitney U test or the non-parametric Spearman's rank correlation coefficient for dichotomous or continuous variables, respectively.

Results

TOP1 is structurally conserved in microbes, particularly in fungi

To identify microbial proteins that could be recognized by human ATAs, we performed a homology search for proteins sharing structural similarities with human TOP1. We used Foldseek with a crystal structure of human TOP1 (1EJ9, Protein Data Bank) as input[24,25].

Foldseek searches databases for proteins or protein fragments with resolved or predicted tertiary protein structures that show structural similarities with a protein of interest[24]. Like other crystal structures of human TOP1, 1EJ9 does not contain the poorly conserved N-terminal and linker domains[25]. A taxonomic filter was applied to only identify proteins of fungal, bacterial or viral origin. The search identified 145 fungal, 273 bacterial and 3 viral hits (Fig. 1A). Most hits were TOP1 proteins in the respective microbial compartments (Suppl. Table S1-S3). Fungal hits showed the highest template modeling scores and lowest expect values (E-values) (Fig. 1A). Interestingly, strong structural similarities were observed between human TOP1 and TOP1 from, for example, *S. cerevisiae*, *Pseudomonas aeruginosa* and *Variola virus*, despite relatively low sequence identities (41.2%, 17.3% and 17.2%, respectively) (Fig. 1B-D). Together, these findings provided rational to explore cross-recognition of microbial TOP1 molecules by human ATAs.

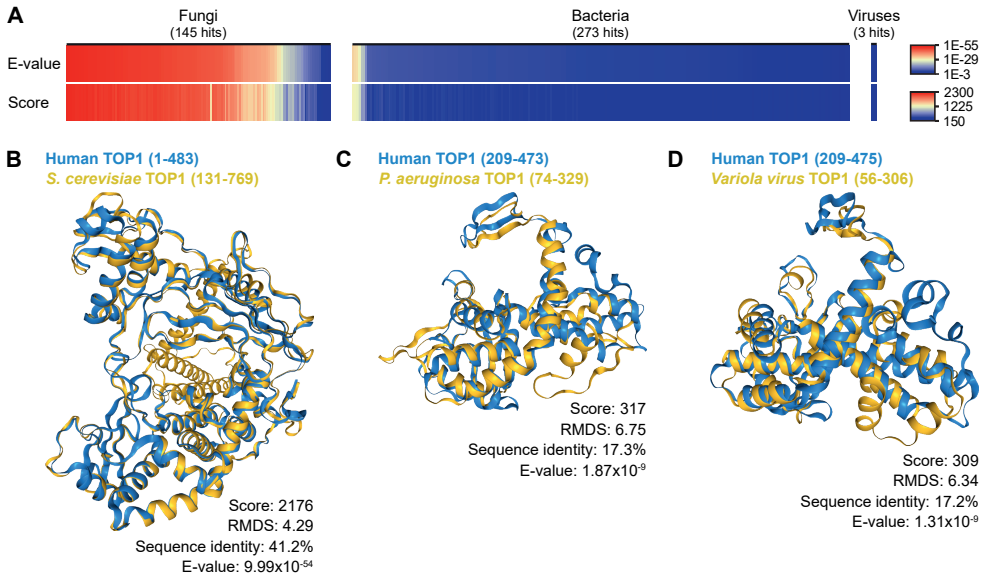


Figure 1. Homology search for the identification of microbial proteins sharing structural similarities with human TOP1. (A) Overview of microbial proteins having structural homology with human TOP1 as predicted by Foldseek. Homology search was based on a crystal structure of human TOP1 (1EJ9, Protein Data Bank). A taxonomic filter was used to specifically identify fungal, bacterial and viral proteins. Hits were ranked based on E-values. **(B-D)** Structural homology as predicted by Foldseek between human TOP1 (in blue) and TOP1 from *S. cerevisiae* (AlphaFold/Proteome (v4) database) **(B)**, *P. aeruginosa* (AlphaFold/Proteome (v4) database) **(C)** and *Variola virus* (PDB100 database (20240101)) **(D)** (all in yellow). Amino acid positions of predicted structurally overlapping parts are indicated between brackets. Score = structural bit score. RMDS = root mean square deviation. Expect values = E-values.

Circulating antibodies from ATA⁺ SSc patients recognize TOP1 from S. cerevisiae

We selected TOP1 from *S. cerevisiae* as exemplary model antigen to test the hypothesis. TOP1 from this fungus was amongst the predicted hits with the strongest structural similarity with human TOP1 and contains some homologous stretches of amino acids (Table S1, Suppl. Fig. S1). In addition, *S. cerevisiae* is abundantly present at mucosal sites, including the lungs[35-37]. Immunological exposure to *S. cerevisiae* can induce anti-*S.*

cerevisiae antibodies (ASCAs) which are associated with Crohn's disease but have also been detected in healthy donors (HDs) and SSc patients[38-40]. Therefore, we recombinantly produced and subsequently purified human and *S. cerevisiae* TOP1 (Suppl. Fig. S2). As expected, human TOP1 was recognized specifically by ATA$^+$ SSc patients but not by anti-centromere autoantibody (ACA)$^+$ SSc patients and HDs (cohort 1, Fig. 2A, Suppl. Fig. S3A, Table 1). Interestingly, a subset of ATA$^+$ SSc patients (45%) showed also reactivity to *S. cerevisiae* TOP1 (Fig. 2B, Suppl. Fig. S3B). This reactivity was not observed in plasma from ACA$^+$ SSc patients and HDs (Fig. 2B, Suppl. Fig. S3B). Recognition of human and *S. cerevisiae* TOP1 weakly correlated (r_s=0.35, p=0.12), with strong variation in the recognition of *S. cerevisiae* TOP1 among ATA$^+$ SSc patients among ATA$^+$ SSc patients with comparable anti-human TOP1-IgG levels (Fig. 2C, Suppl. Fig. S3C).

To independently replicate these findings, we measured *S. cerevisiae* TOP1 reactivity in

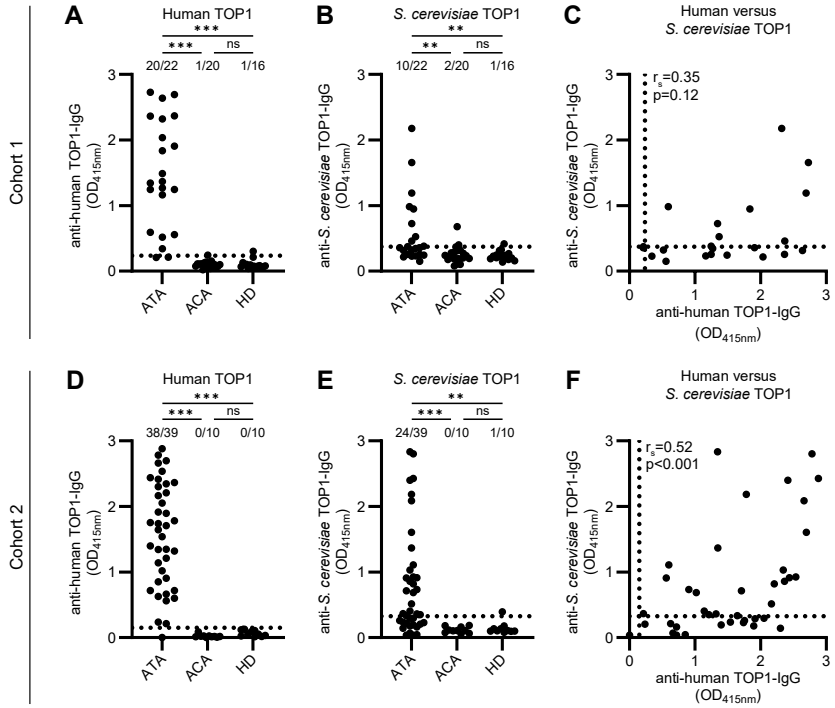


Figure 2. Recognition of *S. cerevisiae* TOP1 by circulating IgG of ATA+ SSc patients. (A-B) Reactivity of IgG in plasma of ATA$^+$ SSc patients (n=22), ACA$^+$ SSc patients (n=20) and healthy donors (HD)(n=16) included in cohort 1 towards human **(A)** and *S. cerevisiae* **(B)** TOP1. **(C)** Correlation between the recognition of human and *S. cerevisiae* TOP1 by ATA$^+$ SSc patients in cohort 1. **(D-E)** Binding of human **(D)** and *S. cerevisiae* **(E)** TOP1 by IgG in serum of ATA$^+$ SSc patients (n=39), ACA$^+$ SSc patients (n=10) and healthy donors (HD)(n=10) included in cohort 2. **(F)** Relation between the reactivity towards human and *S. cerevisiae* TOP1 by ATA$^+$ SSc patients in cohort 2. **(A-F)** Optical density was measured at 415 nanometers (OD_{415nm}) and blank was subtracted. Saturated samples (OD>2.5) were diluted further to determine antibody levels in arbitrary units per milliliter (aU/ml), see Fig. S3. Dotted lines represent cut-off: mean + 2x standard deviation of HDs. **(A-B, D-E)** Kruskal-Wallis test combined with Dunn's multiple comparison test was used to test for statistical significant differences. ns = not significant, * = p<0.05, ** = p<0.005, *** = p<0.001. **(C, F)** Correlations were described using the Spearman's rank correlation coefficient (r_s).

serum of ATA$^+$ SSc patients in a second cohort (cohort 2, table 2). As in cohort 1, human TOP1 was recognized by IgG of ATA$^+$ SSc patients, whereas age- and sex-matched ACA$^+$ SSc patients and HDs did not recognize this antigen (Fig. 2D, Suppl. Fig. S3D). *S. cerevisiae* TOP1 was also recognized by a subset of ATA$^+$ SSc patients in this second cohort (62%), but not by ACA$^+$ SSc patients and HDs (Fig. 2E, Suppl. Fig. S3E). A moderate correlation (r_s=0.52, p<0.001) was found between the recognition of human and *S. cerevisiae* TOP1 (Fig. 2F, Suppl. Fig. S3F). Similar to the observations in cohort 1, strong heterogeneity was observed in the recognition of *S. cerevisiae* TOP1 between patients with comparable reactivity towards human TOP1. Together, these data indicate that a subgroup of ATA$^+$ SSc patients, but not healthy or disease control individuals harbor reactivity towards TOP1 of microbial origin, supporting the concept of potential cross-reactivity.

Cross-reactivity of patient-derived ATA monoclonal antibodies towards S. cerevisiae TOP1

To more directly investigate cross-reactivity of human ATA to *S. cerevisiae* TOP1, we next used seven patient-derived ATA monoclonal antibodies (ATA mAbs) which were generated based on the BCR sequence of TOP1-reactive B cells from one ATA$^+$ SSc patient (Suppl. Table S4)[26]. All mAbs showed the expected composition and molecular weight upon purification (Suppl. Fig. S4A). In line, all ATA mAbs, but not a tetanus toxoid (TT)-reactive control mAb, reacted to human TOP1 (Fig. 3A). Also, none of the mAbs bound the control protein centromere protein B, except for 8D7 which generated a weak signal at high concentrations (Suppl. Fig. S4B). Reactivity to TOP1 from *S. cerevisiae* was observed for four of the seven ATA mAbs (2F8, 6E3, 8D7, 9D11) in a concentration-dependent manner (Fig. 3B). Binding of ATA-IgG 9D11 to *S. cerevisiae* TOP1 was strong and detectable at lower concentrations (<0.01µg/ml), whereas ATA-IgG 2F8, 6E3 and 8D7 showed detectable

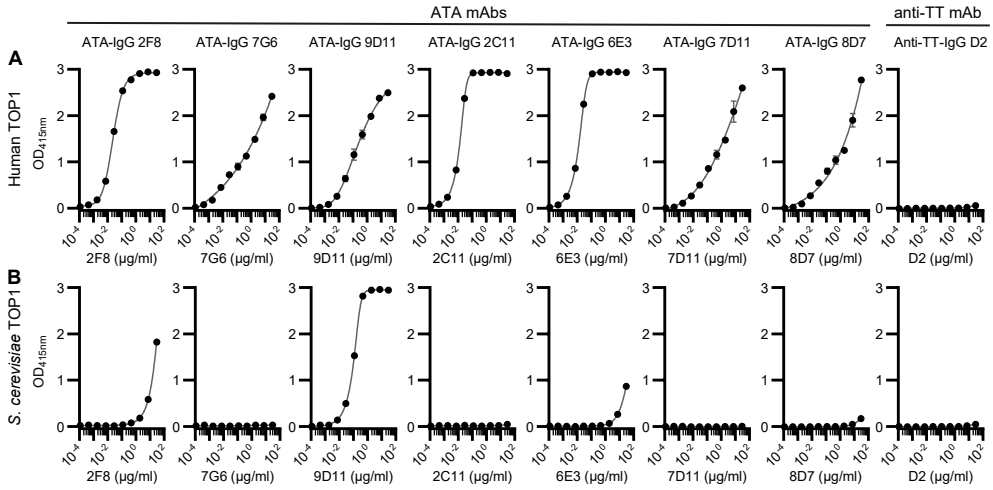


Figure 3. Reactivity of patient-derived ATA mAbs towards human and *S. cerevisiae* TOP1. (A-B) Reactivity of patient-derived ATA mAbs towards human **(A)** and *S. cerevisiae* TOP1 **(B)**. The mAbs were applied in a 4-fold serial dilution starting at 30µg/ml. A tetanus toxoid (TT)-reactive mAb (anti-TT-IgG D2) was used as negative control. Optical density was measured at 415 nanometers (OD_{415nm}) and blank was subtracted. Error bars represent the standard deviation of technical duplicates. Data are representative of three independent experiments. Individual datapoints were connected using a five-parameter logistic regression.

binding only at higher concentrations (>1µg/ml). In contrast, the TT-reactive control mAb D2 and three other ATA mAbs (7G6, 2C11 and 7D11) did not recognize *S. cerevisiae* TOP1, even at higher concentrations. Together, these data provide direct evidence that a subset of ATA mAbs reacts to both human and *S. cerevisiae* TOP1.

***S. cerevisiae* TOP1 can activate an ATA-expressing B cell line**

We next investigated the potential functional relevance of the observations described above by assessing the ability of *S. cerevisiae* TOP1 to stimulate B cells directed against human TOP1. To this end, we generated anti-human TOP1 BCR-expressing Ramos B cell lines and studied the ability of *S. cerevisiae* TOP1 to activate these cells in an antigen-specific manner (Fig. 4A). Ramos B cells expressing anti-TOP1-IgG as functional BCR were generated by retroviral transduction of constructs encoding the BCR sequence of B cell clones 2F8, 7G6 and 9D11. In the receiving Ramos B cell line, genes encoding for IgD, IgM and activation-induced cytidine deaminase (AID) were knocked out (MDL-AID KO Ramos cells)[28,29]. Upon transduction, Ramos B cells expressed GFP and IgG as BCR on their surface (Suppl. Fig. S5A-B). All TOP1-reactive Ramos B cell lines were activated upon stimulation with anti-IgG (positive control) and human TOP1 (Fig. 4B, Suppl. Fig. S5C), as evidenced by the phosphorylation of Syk. Importantly, the Ramos B cell line expressing anti-human TOP1 BCR 9D11 was also activated by *S. cerevisiae* TOP1, in contrast to the cell lines expressing 7G6 and 2F8 (Fig. 4A-B, Suppl. Fig. S5C), corroborating the findings obtained by ELISA. A control Ramos B cell line expressing a citrullinated protein-reactive BCR (ACPA-IgG 3F3) was activated by anti-human IgG but not by human and *S. cerevisiae* TOP1. These data show that TOP1 of microbial origin cannot only be recognized by human ATA, but that it can also activate autoreactive B cells expressing cross-reactive BCRs derived from patients.

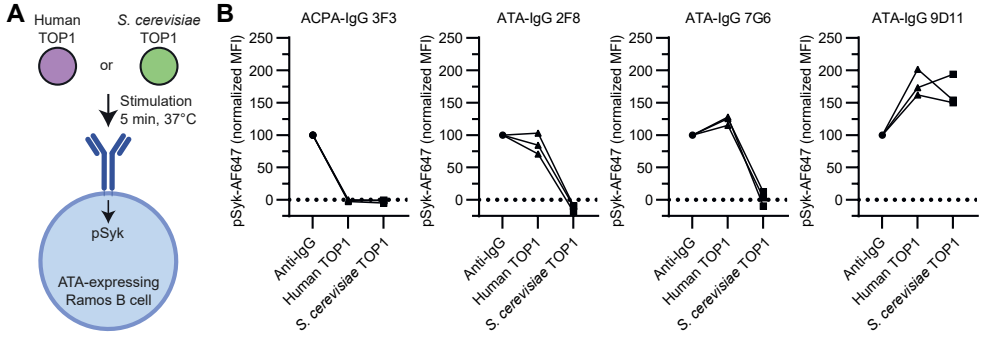


Figure 4. Stimulation of ATA-expressing Ramos B cell lines with human and *S. cerevisiae* TOP1. (A) Experimental procedure to measure the activation of Ramos B cell lines upon stimulation. ATA-expressing Ramos B cell lines and an ACPA-expressing Ramos B cell line (negative control) were stimulated with 5µg/ml anti-IgG (positive control), human TOP1 or *S. cerevisiae* TOP1 for 5 minutes at 37°C. Subsequently, the expression of pSyk by stimulated RAMOS B cells was measured by flow cytometry. **(B)** Expression of pSyk by Ramos B cell lines upon stimulation with anti-IgG, human TOP1 or *S. cerevisiae* TOP1. MFIs were normalized to the MFI generated in the corresponding anti-IgG condition upon subtracting the MFI obtained in the unstimulated condition. Data is derived from three independent experiments, independent experiment are connected by lines.

Recognition of S. cerevisiae TOP1 is enhanced in patients with severe disease

We previously demonstrated a possible link between the activity of the TOP1-reactive B cell response and severe disease[9,10]. Given the ability of *S. cerevisiae* TOP1 to activate human ATA-expressing B cells, we next questioned whether the recognition of *S. cerevisiae* TOP1 by a subset of ATA$^+$ SSc patients would associate with severity of the clinical phenotype. Therefore, we measured anti-human TOP1-IgG and anti-*S. cerevisiae* TOP1-IgG levels in a cohort (cohort 2) consisting of ATA$^+$ SSc patients at both ends of the SSc disease severity spectrum: 18 patients with severe SSc and 21 with mild disease (Table 2). In this cohort, levels of anti-human TOP1-IgG were higher in patients with severe disease compared to those with mild SSc (Fig. 5A). Moreover, sera of patients with severe SSc more often harbored antibodies recognizing *S. cerevisiae* TOP1 in comparison to patients with mild SSc (83% versus 43%) (Fig. 5B). Interestingly, the recognition of *S. cerevisiae* TOP1 was particularly observed in patients with clinically relevant ILD (Fig. 5C-D). Male patients and patients with diffuse cutaneous SSc also tended to have higher levels of anti-*S. cerevisiae* TOP1-IgG (Suppl. Fig. S6A-J). In conclusion, based on the evaluation of patients at the extreme ends of the severity spectrum, we observed that recognition of microbial TOP1 was more prevalent in severe SSc, in particular in those with clinically relevant ILD.

To replicate the association and evaluate the possible relevance of *S. cerevisiae* TOP1 recognition for clinically relevant ILD in ATA$^+$ SSc in a cohort without a selection bias for disease severity, we measured anti-human and *S. cerevisiae* TOP1-IgG levels in serum samples from all available ATA$^+$ SSc patients (cohort 3, n=163)[21]. In this cohort, human TOP1 was recognized by the majority of ATA$^+$ SSc patients and not by age- and sex-matched ACA$^+$ SSc patients and HDs (Fig. 5E, Suppl. Fig. S7A). A subset of ATA$^+$ SSc patients (77%) recognized *S. cerevisiae* TOP1, while most ACA$^+$ SSc patients and HDs were negative (Fig. 5F, Suppl. Fig. S7B). Again, anti-human TOP1- and *S. cerevisiae* TOP1-IgG levels correlated (r_s=0.61, p<0.001) (Fig. 5G, Suppl. Fig. S7C). Interestingly, patients with clinically relevant ILD were more commonly anti-*S. cerevisiae* TOP1-IgG$^+$ (90% seropositivity in patients with clinically relevant ILD versus 72% in patients without clinically relevant ILD) (Fig. 5H-I). This corresponded to an adjusted odds ratio of 3.4 (95% confidence interval: 1.1 - 10.3, p=0.034), when corrected for sex, disease subset and elevated C-reactive protein levels. Moreover, anti-*S. cerevisiae* TOP1-IgG levels, but not anti-human TOP1-IgG levels, were significantly increased in patients with clinically relevant ILD (Fig. 5H-I). This finding remained statistically significant when patients in cohort 2, who were also included in cohort 3, were excluded (Suppl. Fig. S7D-E). Also in cohort 3, patients with diffuse cutaneous SSc and male patients tended to have higher levels of *S. cerevisiae* TOP1-IgG (Suppl. Fig. S7F-O). In conclusion, antibodies recognizing *S. cerevisiae* TOP1 associate with clinically relevant ILD in ATA$^+$ SSc.

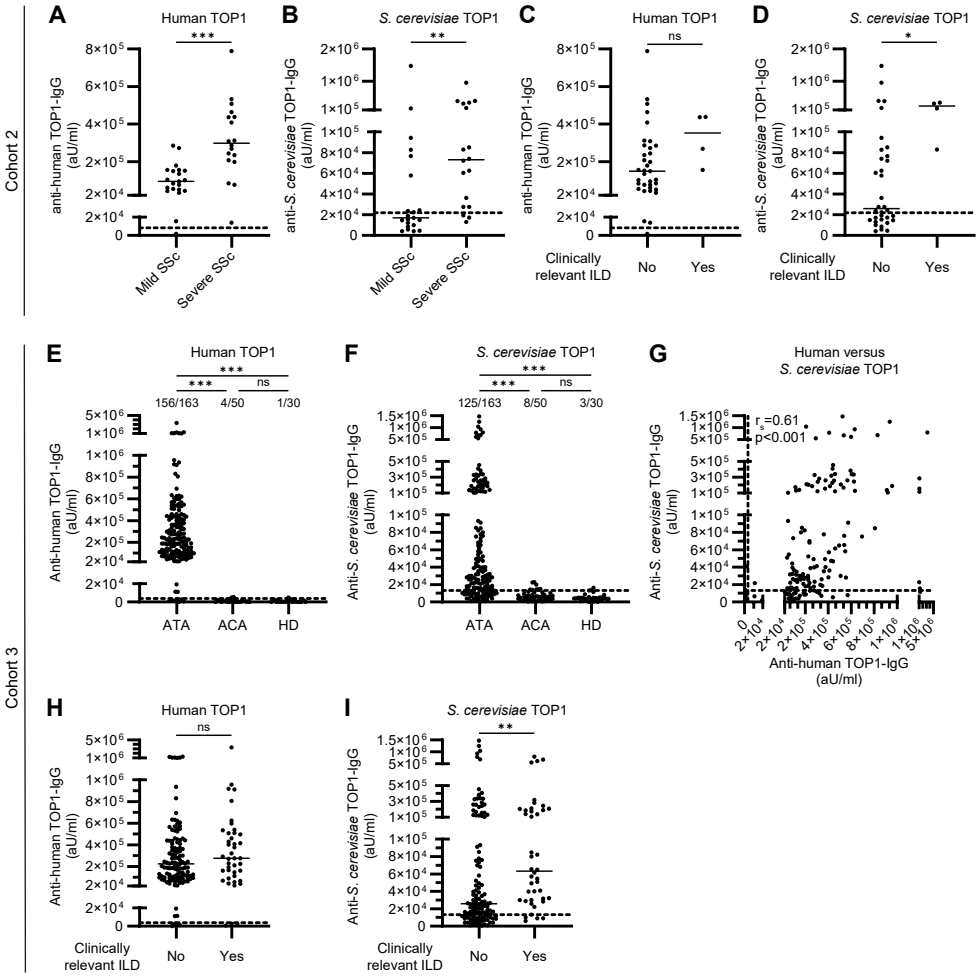


Figure 5. Association between the recognition of *S. cerevisiae* TOP1 and clinical disease in ATA⁺ SSc. (A-B) Levels of anti-human TOP1-IgG **(A)** and anti-*S. cerevisiae* TOP1-IgG **(B)** in serum of ATA⁺ SSc patients included in cohort 2 with mild (n=21) versus severe disease (n=18). **(C-D)** Comparison of anti-human TOP1-IgG **(C)** and anti-*S. cerevisiae* TOP1-IgG **(B)** levels in serum between ATA⁺ SSc patients included in cohort 2 with and without clinically relevant ILD. **(E-F)** Reactivity of IgG in plasma of ATA⁺ SSc patients (n=163), ACA⁺ SSc patients (n=50) and healthy donors (HD)(n=30) included in cohort 3 towards human **(E)** and *S. cerevisiae* **(F)** TOP1. Kruskal-Wallis test combined with Dunn's multiple comparison test was used to test for statistical significant differences. **(G)** Correlation between the recognition of human and *S. cerevisiae* TOP1 by ATA⁺ SSc patients in cohort 3. Correlations were described using the Spearman's rank correlation coefficient (r_s). **(H-I)** Comparison of anti-human TOP1-IgG **(C)** and anti-*S. cerevisiae* TOP1-IgG **(B)** levels in serum between ATA⁺ SSc patients included in cohort 3 with and without clinically relevant ILD. **(A-I)** Levels are expressed in arbitrary units per milliliter (aU/ml). Striped lines represent cut-off: mean + 2x standard devation of HD samples. **(A-D, H-I)** Mann-Whitney U test was used to test for significant differences. **(A-F, H-I)** ns = not significant, * = p<0.05, ** = p<0.005, *** = p<0.001.

Discussion

The clinical management of autoimmune diseases such as SSc is highly complex due to heterogeneity within patient populations and lack of curative treatments. To understand such heterogeneity in the development of disease complications, it is important to understand the mechanisms driving and maintaining autoimmunity. SSc is hallmarked by a break in B cell tolerance towards nuclear proteins. At present, it is unclear how these B cell responses are initiated and activated. Here, we provide evidence that a subset of patient-derived ATAs can cross-react to microbial TOP1 derived from an environmental yeast. More specifically, we found that sera from a subset of ATA$^+$ SSc patients and multiple ATA mAbs recognized TOP1 from *S. cerevisiae*, a microbial TOP1 which shares homology with human TOP1. *In vitro*, *S. cerevisiae* TOP1 induced in the activation of TOP1-reactive B cells. Furthermore, the recognition of *S. cerevisiae* TOP1 was enhanced in patients with severe ATA$^+$ SSc and in patients with clinically relevant ILD.

These findings are intriguing as they point to the possibility that microbial antigens initiate and activate TOP1-reactive B cells, a notion supported by a recent large epidemiological study implicating exposure to certain yeast strains to high rates of SSc diagnoses[41]. Microbes have frequently been implicated in the break of B cell tolerance related to autoimmune disease[13,42]. One important mechanism relies on cross-reactivity of autoreactive BCRs to microbial antigens, a feature which has been described for various types of autoantibodies[11,12,43]. An important consideration in this context is that microbial antigen recognition by autoreactive B cells not only activates the BCR (known as signal 1) and, potentially, also innate receptors (signal 2)[14,44,45]. In fact, such B cells may recruit help from microbial-directed T cells (signal 3) to promote their activation, thereby bypassing the need to break T cell tolerance to human TOP1. In this study, we found that microbial TOP1 can activate B cells expressing an anti-human TOP1 BCR based on cross-reactivity. In this context, it is important to note that we used TOP1 from *S. cerevisiae* as a prototypic yeast-derived TOP1 molecule, but that multiple microbes share TOP1 molecules with structural similarities. This means that not *S. cerevisiae*-derived TOP1 specifically may have a causative role in SSc pathogenesis, but that also TOP1 molecules from other microbial sources could be involved. Nonetheless, irrespective of the origin of microbial TOP1, an intriguing yet unexplored possibility is that the initial B cell response primarily targets the pathogen, and that cross-reactivity to human TOP1 is enhanced or even generated de novo by somatic hypermutation. B cells recruit help from T cells by presenting microbial peptides on MHC class II[46,47]. Help from T cells is essential for the maturation of B cell responses by inducing isotype switching, somatic hypermutation and epitope spreading[48,49]. Thus, based on our observations, we consider this mechanistic framework relevant for the activation of autoreactive B cells in SSc, a process which likely facilitates the initiation and subsequent maturation of the ATA B cell response in the context of SSc pathogenesis.

Besides this intriguing role that microbial TOP1 might play in the initial break of B cell tolerance towards human TOP1, BCR cross-reactivity could also determine the subsequent dynamics of the ATA B cell response during chronic disease. Exposure to one or multiple microbes, whether in the context of acute infection or chronic exposure (e.g. commensals), could act as continuous source of microbial triggering of TOP1-reactive B

cells. As it is possible that ATAs can recognize TOP1 from other microbes as well, variations might exist in avidity and in between patients. Interestingly, TOP1 of fungi showed the strongest structural similarities with human TOP1. Fungi represent a substantial fraction of the human microbiome which, in the context of autoimmunity, has so far received little attention[50,51]. Fungi can induce systemic immune reactions and were recently linked to the pathogenesis of various other immune-related diseases, like Crohn's disease and severe COVID-19[38,50-52]. In the context of SSc, a higher abundance of Rhodotorula glutinis RNA sequence reads was detected in the skin of SSc patients, in comparison to HDs[19]. In addition, reactivity towards the fungal counterparts of human fibrillarin and α-enolase derived from *S. cerevisiae* was detected in sera of SSc patients[53,54]. Hence, our findings warrant further exploration of the mycobiome of SSc patients and extended analyses of BCR cross-reactivity to delineate whether fungi indeed serve as a continuous source of antigens triggering autoreactive B cells in this disease.

The association with severe disease makes anti-*S. cerevisiae* TOP1-IgG a promising clinical biomarker. This is particularly important as within patients with reactivity to human TOP1, or ATA positivity, still large heterogeneity exists in clinical disease course for which biomarkers are missing that can stratify between subsets of patients[22]. The increased recognition of *S. cerevisiae* TOP1, but not human TOP1, by patients with clinically relevant ILD indicates that measuring *S. cerevisiae* TOP1 reactivity might be complementary to measuring human TOP1 reactivity. This additive value is based on the remarkable variability between patients in the recognition of *S. cerevisiae* TOP1, even in patients with comparable anti-human TOP1-IgG levels. Such a biomarker might, ideally in combination with other standard diagnostics, aid in the stratification of ATA^+ patients to tailor monitoring, management and treatment of the disease. Additional replication of these findings and longitudinal analyses are warranted to explore the potential of anti-*S. cerevisiae* TOP1-IgG as clinical biomarker. In future steps, it will also be interesting to determine the recognition of microbial TOP1 by individuals with very early SSc, to assess germline-reverted ATA mAbs, and the recognition of other, not yet studied microbial proteins.

In conclusion, the recognition of microbial TOP1 by human ATA and the activation of TOP1-reactive B cells by microbial TOP1 support a role for microbes in the activation of the ATA B cell response in SSc. The increased recognition in patients with severe SSc and in patients with clinically relevant ILD indicates that microbial TOP1 could be a clinically relevant trigger of ATA B cells which relates to disease heterogeneity. Together, these results enhance our understanding of mechanisms underlying the breach of tolerance towards human autoantigens. Also, they hold promise to stratify clinically heterogeneous patient populations based on recognition of a microbial antigen to which the human autoreactive B cell response cross-reacts, and highlight the human mycobiome as an interesting area for further research.

Acknowledgements
R.Q. Kim and A. Moutsiopoulou (Protein Facility of the Leiden University Medical Center) are thanked for the production of human TOP1. We thank J.F.X. Diffley (The Francis Crick Institute) for providing the yAE42 strain used for production and purification of *S. cerevisiae* TOP1. Flow cytometry was performed at the Flow Cytometry Core Facility of the Leiden University Medical Center.

Conflict of interest
SN, JKdVB, REMT and HUS are listed as inventors of a filed patent on patients stratification in SSc based on autoreactive B cell and autoantibody cross-reactivity to microbial antigens (NL4000202, "Systemic Sclerosis Patient Stratification"). The authors declare no other conflicts of interest.

Funding
This work was funded by the Dutch Arthritis Foundation (grants 17-1-402, 15-2-402, 18-1-205, LLP5), the IMI funded project RTCure (777357), Health~Holland Life Sciences & Health sector programs Target to B! (LSHM18055-5GF) and Immune HealthSeed (LSHM22042-SFG), a European Research Council Advanced grant (AdG2019-884796, to REMT) and a NWO-ZonMW VIDI grant (09150172010067, to HUS).

Author contributions
Conceptualization: SN, SIEL, JKdVB, REMT, HUS. Methodology: SN, SIEL, TvL NEWL, NHD, JKdVB, REMT, HUS. Investigation: SN. Analysis of the data: SN, SIEL, EMH. Interpretation of the data: SN, SIEL, TvL, EMH, CMF, NHD, JKdVB, REMT, HUS. Supervision: JKdVB, REMT, HUS. Writing – original draft: SN, HUS. Writing – review & editing: SN, SIEL, TvL, EMH, CMW, NEWL, CMF, NHD, JKdVB, REMT, HUS.

Data and materials availability
Data will be made available upon reasonable request.

References

1 Liem, S. I. E. *et al*. Autoreactive B cell responses targeting nuclear antigens in systemic sclerosis: Implications for disease pathogenesis. *Semin Arthritis Rheum* **58**, 152136 (2023).
2 Wang, X. *et al*. Allogeneic CD19-targeted CAR-T therapy in patients with severe myositis and systemic sclerosis. *Cell* **187**, 4890-4904 e4899 (2024).
3 Auth, J. *et al*. CD19-targeting CAR T-cell therapy in patients with diffuse systemic sclerosis: a case series. *Lancet Rheumatol* **7**, e83-e93 (2025).
4 Goswami, R. P. *et al*. Rituximab in the treatment of systemic sclerosis-related interstitial lung disease: a systematic review and meta-analysis. *Rheumatology (Oxford)* **60**, 557-567 (2021).
5 Reveille, J. D., Solomon, D. H. & American College of Rheumatology Ad Hoc Committee of Immunologic Testing, G. Evidence-based guidelines for the use of immunologic tests: anticentromere, Scl-**70**, and nucleolar antibodies. *Arthritis Rheum* **49**, 399-412 (2003).
6 Walker, U. A. *et al*. Clinical risk assessment of organ manifestations in systemic sclerosis: a report from the EULAR Scleroderma Trials And Research group database. *Ann Rheum Dis* **66**, 754-763 (2007).
7 Nihtyanova, S. I. *et al*. Using autoantibodies and cutaneous subset to develop outcome-based disease classification in systemic sclerosis. *Arthritis Rheumatol* **72**, 465-476 (2020).
8 Hesselstrand, R., Scheja, A., Shen, G. Q., Wiik, A. & Akesson, A. The association of antinuclear antibodies with organ involvement and survival in systemic sclerosis. *Rheumatology (Oxford)* **42**, 534-540 (2003).
9 Boonstra, M. *et al*. Association of anti-topoisomerase I antibodies of the IgM isotype with disease progression in anti-topoisomerase I-positive systemic sclerosis. *Arthritis Rheumatol* **72**, 1897-1904 (2020).
10 Wortel, C. M. *et al*. Anti-topoisomerase, but not anti-centromere B cell responses in systemic sclerosis display active, Ig-secreting cells associated with lung fibrosis. *RMD Open* **9**, e003148 (2023).
11 Laman, J. D., Huizinga, R., Boons, G. J. & Jacobs, B. C. Guillain-Barre syndrome: expanding the concept of molecular mimicry. *Trends Immunol* **43**, 296-308 (2022).
12 Lanz, T. V. *et al*. Clonally expanded B cells in multiple sclerosis bind EBV EBNA1 and GlialCAM. *Nature* **603**, 321-327 (2022).
13 Rojas, M. *et al*. Molecular mimicry and autoimmunity. *J Autoimmun* **95**, 100-123 (2018).
14 Greiling, T. M. *et al*. Commensal orthologs of the human autoantigen Ro60 as triggers of autoimmunity in lupus. *Sci Transl Med* **10**, eaan2306 (2018).
15 Volkmann, E. R., Andreasson, K. & Smith, V. Systemic sclerosis. *Lancet* **401**, 304-318 (2023).
16 Hu, P. Q., Fertig, N., Medsger, T. A., Jr. & Wright, T. M. Correlation of serum anti-DNA topoisomerase I antibody levels with disease severity and activity in systemic sclerosis. *Arthritis Rheum* **48**, 1363-1373 (2003).
17 Volkmann, E. R. *et al*. Association of systemic sclerosis with a unique colonic microbial consortium. *Arthritis Rheumatol* **68**, 1483-1492 (2016).
18 Johnson, M. E. *et al*. Microbiome dysbiosis is associated with disease duration and increased inflammatory gene expression in systemic sclerosis skin. *Arthritis Res Ther* **21**, 49 (2019).
19 Arron, S. T. *et al*. High Rhodotorula sequences in skin transcriptome of patients with diffuse systemic sclerosis. *J Invest Dermatol* **134**, 2138-2145 (2014).
20 van den Hoogen, F. *et al*. 2013 classification criteria for systemic sclerosis: an American college of rheumatology/European league against rheumatism collaborative initiative. *Ann Rheum Dis* **72**, 1747-1755 (2013).
21 Meijs, J. *et al*. Therapeutic and diagnostic outcomes of a standardised, comprehensive care pathway for patients with systemic sclerosis. *RMD Open* **2**, e000159 (2016).
22 Groseanu, L. *et al*. Do we have good activity indices in systemic sclerosis? *Curr Rheumatol Rev* **18**, 39-47 (2022).
23 Valentini, G., D'Angelo, S., Della Rossa, A., Bencivelli, W. & Bombardieri, S. European Scleroderma

Study Group to define disease activity criteria for systemic sclerosis. *IV*. Assessment of skin thickening by modified Rodnan skin score. *Ann Rheum Dis* **62**, 904-905 (2003).
24 van Kempen, M. *et al*. Fast and accurate protein structure search with Foldseek. *Nat Biotechnol* **42**, 243-246 (2024).
25 Redinbo, M. R., Champoux, J. J. & Hol, W. G. Novel insights into catalytic mechanism from a crystal structure of human topoisomerase I in complex with DNA. *Biochemistry* **39**, 6832-6840 (2000).
26 Neppelenbroek, S. *et al*. Enhanced recognition of topoisomerase 1 upon DNA binding by a subset of anti-topoisomerase 1 autoantibodies in systemic sclerosis. *bioRxiv,* 2025.05.26.656088 (2025).
27 Yeeles, J. T. P., Janska, A., Early, A. & Diffley, J. F. X. How the eukaryotic replisome achieves rapid and efficient DNA replication. *Mol Cell* **65**, 105-116 (2017).
28 He, X. *et al*. Continuous signaling of CD79b and CD19 is required for the fitness of Burkitt lymphoma B cells. *EMBO J* **37**, e97980 (2018).
29 Kissel, T. *et al*. Antibodies and B cells recognising citrullinated proteins display a broad cross-reactivity towards other post-translational modifications. *Ann Rheum Dis* **79**, 472-480 (2020).
30 Liakouli, V., Ciancio, A., Del Galdo, F., Giacomelli, R. & Ciccia, F. Systemic sclerosis interstitial lung disease: unmet needs and potential solutions. *Nat Rev Rheumatol* **20**, 21-32 (2024).
31 Volkmann, E. R. & Fischer, A. Update on morbidity and mortality in systemic sclerosis-related interstitial lung disease. *J Scleroderma Relat Disord* **6**, 11-20 (2021).
32 Hoffmann-Vold, A. M. *et al*. Progressive interstitial lung disease in patients with systemic sclerosis-associated interstitial lung disease in the EUSTAR database. *Ann Rheum Dis* **80**, 219-227 (2021).
33 Guler, S. *et al*. Phenotyping by persistent inflammation in systemic sclerosis associated interstitial lung disease: a EUSTAR database analysis. *Thorax* **78**, 1188-1196 (2023).
34 Muangchan, C. *et al*. Association of C-reactive protein with high disease activity in systemic sclerosis: results from the Canadian Scleroderma Research Group. *Arthritis Care Res (Hoboken)* **64**, 1405-1414 (2012).
35 Zhang, F., Aschenbrenner, D., Yoo, J. Y. & Zuo, T. The gut mycobiome in health, disease, and clinical applications in association with the gut bacterial microbiome assembly. *Lancet Microbe* **3**, e969-e983 (2022).
36 Nash, A. K. *et al*. The gut mycobiome of the Human Microbiome Project healthy cohort. *Microbiome* **5**, 153 (2017).
37 Cui, L. *et al*. Topographic diversity of the respiratory tract mycobiome and alteration in HIV and lung disease. *Am J Respir Crit Care Med* **191**, 932-942 (2015).
38 Kusakabe, T. *et al*. Fungal microbiota sustains lasting immune activation of neutrophils and their progenitors in severe COVID-19. *Nat Immunol* **24**, 1879-1889 (2023).
39 Gaifem, J. *et al*. A unique serum IgG glycosylation signature predicts development of Crohn's disease and is associated with pathogenic antibodies to mannose glycan. *Nat Immunol* **25**, 1692-1703 (2024).
40 Fedrigo, A., Skare, T. L., Bortoluzzi, A. L. & Nisihara, R. ASCA (anti-*Saccharomyces cerevisiae* antibody) in patients with scleroderma. *J Clin Rheumatol* **25**, 24-27 (2019).
41 Kosarek, N. N. *et al*. Geographic clustering of systemic sclerosis in areas of environmental pollution. *Arthritis Care Res (Hoboken)* **77**, 855-866 (2025).
42 Pisetsky, D. S. Pathogenesis of autoimmune disease. *Nat Rev Nephrol* **19**, 509-524 (2023).
43 Scherer, H. U., van der Woude, D. & Toes, R. E. M. From risk to chronicity: evolution of autoreactive B cell and antibody responses in rheumatoid arthritis. *Nat Rev Rheumatol* **18**, 371-383 (2022).
44 Ruff, W. E., Greiling, T. M. & Kriegel, M. A. Host-microbiota interactions in immune-mediated diseases. *Nat Rev Microbiol* **18**, 521-538 (2020).
45 Ruff, W. E. *et al*. Pathogenic autoreactive T and B cells cross-react with mimotopes expressed by a common human gut commensal to trigger autoimmunity. *Cell Host Microbe* **26**, 100-113

e108 (2019).
46 Sanderson, N. S. *et al*. Cocapture of cognate and bystander antigens can activate autoreactive B cells. *PNAS* **114**, 734-739 (2017).
47 Fahlquist-Hagert, C. *et al*. Antigen presentation by B cells enables epitope spreading across an MHC barrier. *Nat Commun* **14**, 6941 (2023).
48 Akkaya, M., Kwak, K. & Pierce, S. K. B cell memory: building two walls of protection against pathogens. *Nat Rev Immunol* **20**, 229-238 (2020).
49 Cornaby, C. *et al*. B cell epitope spreading: mechanisms and contribution to autoimmune diseases. *Immunol Lett* **163**, 56-68 (2015).
50 Limon, J. J., Skalski, J. H. & Underhill, D. M. Commensal fungi in health and disease. *Cell Host Microbe* **22**, 156-165 (2017).
51 Martini, G. R. *et al*. Selection of cross-reactive T cells by commensal and food-derived yeasts drives cytotoxic T(H)1 cell responses in Crohn's disease. *Nat Med* **29**, 2602-2614 (2023).
52 Doron, I. *et al*. Human gut mycobiota tune immunity via CARD9-dependent induction of anti-fungal IgG antibodies. *Cell* **184**, 1017-1031.e1014 (2021).
53 Aris, J. P. & Blobel, G. cDNA cloning and sequencing of human fibrillarin, a conserved nucleolar protein recognized by autoimmune antisera. *PNAS* **88**, 931-935 (1991).
54 Terrier, B. *et al*. Antifibroblast antibodies from systemic sclerosis patients bind to alpha-enolase and are associated with interstitial lung disease. *Ann Rheum Dis* **69**, 428-433 (2010).

Supplementary Tables

Supplementary tables S1-3 are available online at: https://www.biorxiv.org/content/10.64898/2026.03.17.712484v1.supplementary-material or scan the QR-code

Supplementary Table S4. B cell receptor sequence details of mAbs.

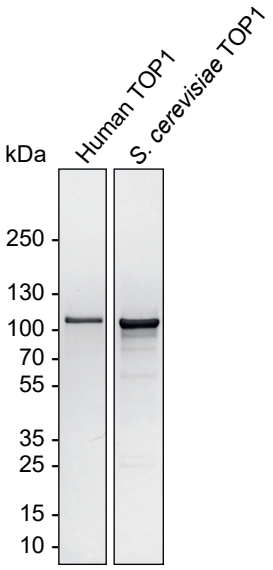
Clone	Antigen	HC	IGHV	IGHV identity (%)	IGHD	IGHJ	IGH-CDR3	LC	IGK/LV	IGK/LV identity (%)	IGK/LJ	IGK/L-CDR3
2F8	H. TOP1	IGHA1	V1-18	95.14	D6-6	J4	CARGRDFSSSPYDYW	λ	V3-21	97.49	J2/J3	CQVWDRSSDDVIF
7G6	H. TOP1	IGHG1	V3-48	90-97	D6-6	J5	CATTPRTGRPW	κ	V4-1	95.96	J4	CQQYLTPLQIF
9D11	H. TOP1	IGHG3	V1-3	93-75	D2-15	J4	CARDGGSSVAVISATLDSW	λ	V2-23	97.22	J2/J3	CCSYAGSRGVVF
2C11	H. TOP1	IGHG1	V3-30	96.88	D5-18	J4	CAKDGSDVRRGYIYGSFDYW	κ	V1-33	97.77	J4	CQQYANLPLTF
6E3	H. TOP1	IGHG1	V3-23	89.58	D2-2	J4	CAKQRCEATGCPSDFW	λ	V3-21	95.34	J2/J3	CQVWDRYSDTIFF
7D11	H. TOP1	IGHG1	V1-18	83.68	D3-16	J4	CARGGDSSSAASDYW	λ	V3-21	94.27	J2/J3	CQVWDNSDDSVLF
8D7	H. TOP1	IGHG1	V3-48	89.24	D1-14	J5	CATTPKAGRPW	κ	V4-1	96.63	J4	CQQYYTTFQTF
D2	TT	IGHG1	V4-39	92.44	D3-16	J4	CARHVLITFGGLIEQTYYFDSW	κ	V2-28	96.60	J5	CMGGLQTPTF
3F3	CCP2	IGH1G2/4	V1-2	80.56	D2-8	J3	CARGTYLPVDESAAFDVW	κ	V4-1	87.88	J2	CQQYTYEAPYTF

Antigen = antigen used for isolation of the B cell, H. TOP1 = human TOP1, TT = tetanus toxoid, CCP2 = cyclic citrullinated peptide 2, HC = heavy chain, IGHV = heavy chain V gene, IGHV identity (%) = similarity of IGHV compared to closest germline IGHV expressed in a percentage, IGHD = heavy chain D gene, IGHJ = heavy chain J gene, IGH-CDR3 = amino acid sequence of heavy chain complementarity-determining region 3, LC = light chain, κ = kappa, λ = lambda, IGK/LV = light chain V gene, IGK/LV identity (%) = similarity of IGK/LV compared to germline IGK/LV expressed in a percentage, IGK/LJ = light chain J gene, IGK/L-CDR3 = amino acid sequence of light chain complementarity-determining region 3.

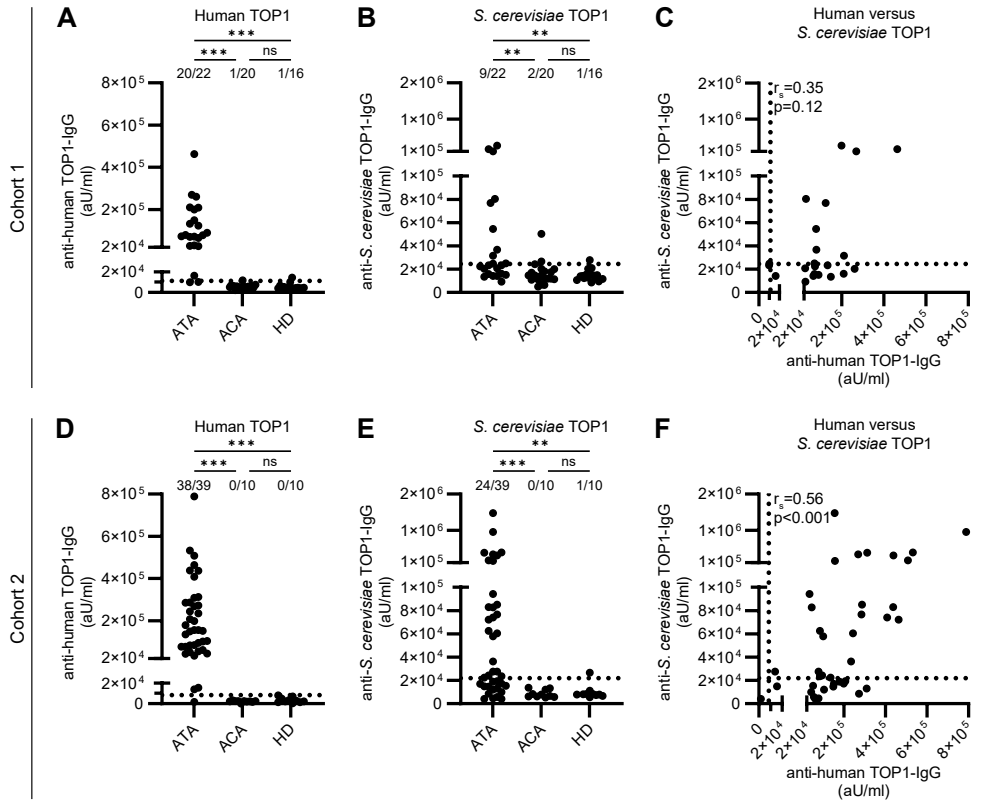
Supplementary Figures

```
Human TOP1              1       WKWWEEERYPEGIKWKFLEHKGPVFAPPYEPLPENVKFYYDGKVMKLSPKAEEVATFFAK      62
                                +KWWE+E   + IKW  L+H G +F PPY+PLP ++K+YYDGK + L P+AEEVA FFA
S. cerevisiae TOP1      131     YKWWEKENEDDTIKWVTLKHNGVIFPPPYQPLPSHIKLYYDGKPVDLPPQAEEVAGFFAA      190

Human TOP1              61      MLDHEYTTKEIFRKNFFKDWRKEMTNEEKNI------ITNLSKCDFTQMSQYFKAQTEAR      114
                                +L+ ++    +F+KNFF D+ +    +E         I  +S+CDFT+M  YF+ Q E +
S. cerevisiae TOP1      191     LLESDHAKNPVFQKNFFNDFLQV—LKESGGPLNGIEIKEFSRCDFTKMFDYFQLQKEQK      248

Human TOP1              115     KQMSKEEKLKIKEENEKLLKEYGFCIMDNHKERIANFKIEPPGLFRGRGNHPKMGMLKRR      174
                                KQ++ +EK +I+ E EK+ ++Y FC +D  +E+++NFK+EPP LFRGRG HPK G LKRR
S. cerevisiae TOP1      249     KQLTSQEKKQIRLEREKFEEDYKFCELDGRREQVGNFKVEPPDLFRGRGAHPKTGKLKRR      308

Human TOP1              175     IMPEDIIINCSKDAKVPSPPPGHKWKEVRHDNKVTWLVSWTENIQGSIKYIMLNPSSRIK      234
                                + PEDI++N SKDA VP  P GHKW E+RHDN V WL  W ENI  S KY+ L  +S +K
S. cerevisiae TOP1      309     VNPEDIVLNLSKDAPVPPAPEGHKWGEIRHDNTVQWLAMWRENIFNSFKYVRLAANSSLK      368

Human TOP1              235     GEKDWQKYETARRLKKCVDKIRNQYREDWKSKEMKVRQRAVALYFIDKLALRAGNEKEEG      294
                                G+ D++K+E AR+LK  +D IR  Y  + KSK M  RQ+AVA+Y+ID +ALRAG EK E
S. cerevisiae TOP1      369     GQSDYKKFEKARQLKSYIDAIRRDYTRNLKSKVMLERQKAVAIYLIDVFALRAGGEKSED      427

Human TOP1              295     ETADTVGCCSLRVEHINLHPELDGQEYVVEFDFLGKDSIRYYNKVPVEKRVFKNLQLF-M      353
                                  ADTVGCCSLR EH+ L P        V FDFLGKDSIR+Y++V V+K+VFKNL +F
S. cerevisiae TOP1      428     -EADTVGCCSLRYEHVTLKPPN-----TVIFDFLGKDSIRFYQEVEVDKQVFKNLTIFKR      482

Human TOP1              354     ENKQPEDDLFDRLNTGILNKHLQDLMEGLTAKVFRTYNASITLQQQLKELTAPDENIPAK      413
                                  KQP   LFDRL+  ILNK+LQ+ M GLTAKVFRTYNAS T+Q QL +L     ++  K
S. cerevisiae TOP1      483     PPKQPGHQLFDRLDPSILNKYLQNYMPGLTAKVFRTYNASKTMQDQL-DLIPNKGSVAEK      541

Human TOP1              414     ILSYNRANRAVAILCNH----------------------------------QIA------      432
                                IL YN ANR VAILCNH                                    Q+
S. cerevisiae TOP1      542     ILKYNAANRTVAILCNHQRTVTKGHAQTVEKANNRIQELEWQKIRCKRAILQLDKDLLKK      601

Human TOP1              433     ------------------------------------------------------------      433

S. cerevisiae TOP1      602     EPKYFEEIDDLTKEDEATIHKRIIDREIEKYQRKFVRENDKRKFEKEELLPESQLKEWLE      661

Human TOP1              433     ------------------------------------------------------------LG      435
                                                                                             LG
S. cerevisiae TOP1      662     KVDEKKQEFEKELKTGEVELKSSWNSVEKIKAQVEKLEQRIQTSSIQLKDKEENSQVSLG      721

Human TOP1              436     TSKLNFLDPRITVAWCKKWGVPIEKIYNKTQREKFAWAIDMADEDYEF                  483
                                TSK+N++DPR++V +CKK+ VPIEKI+ KT REKF WAI+  DE++ F
S. cerevisiae TOP1      722     TSKINYIDPRLSVVFCKKYDVPIEKIFTKTLREKFKWAIESVDENWRF                  769
```

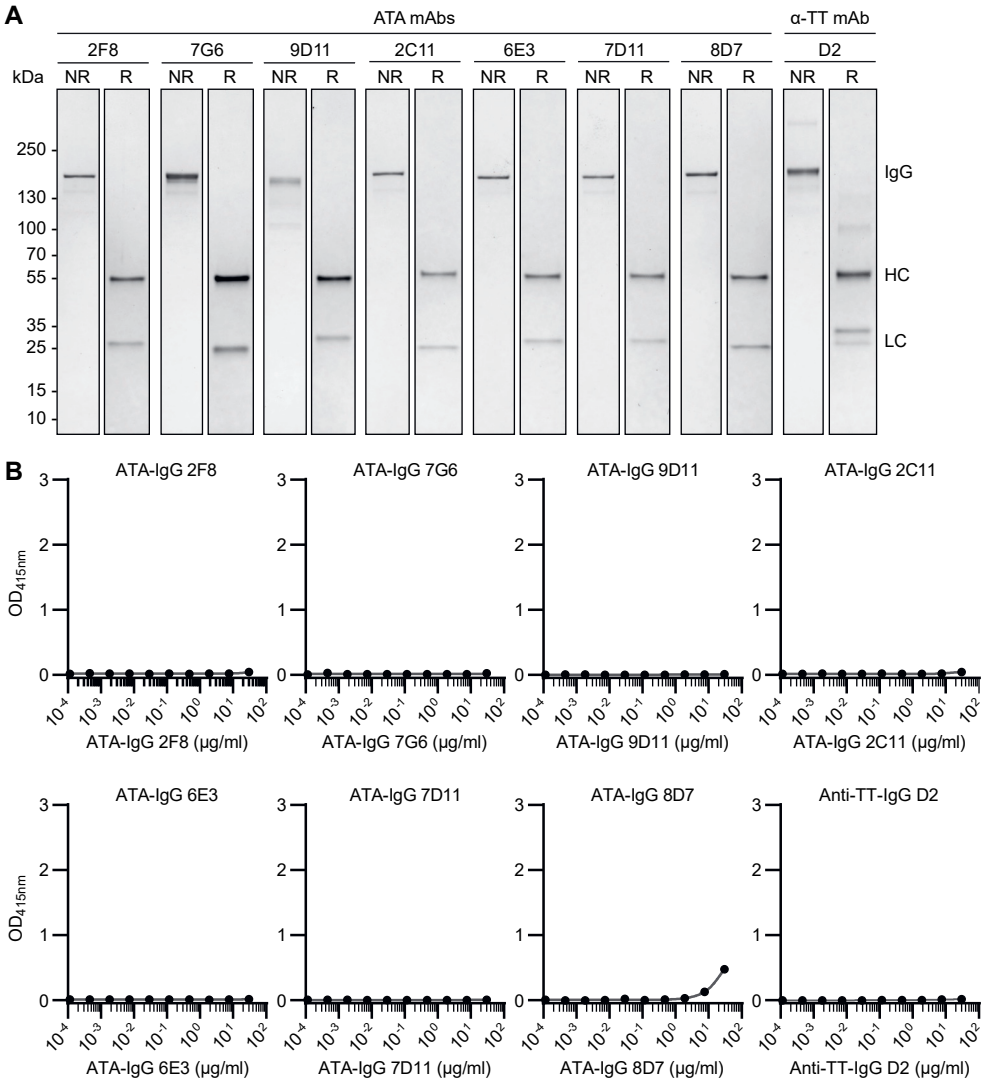
Supplementary Figure S1. Sequence homology between human and *S. cerevisiae* TOP1. Alignment of the amino acid sequences of human and *S. cerevisiae* TOP1 as determined by a Foldseek search using a crystal structure of human TOP1 (1EJ9, Protein Data Bank). The sequence of *S. cerevisiae* is derived from the AlphaFold/Proteome (v4) database.



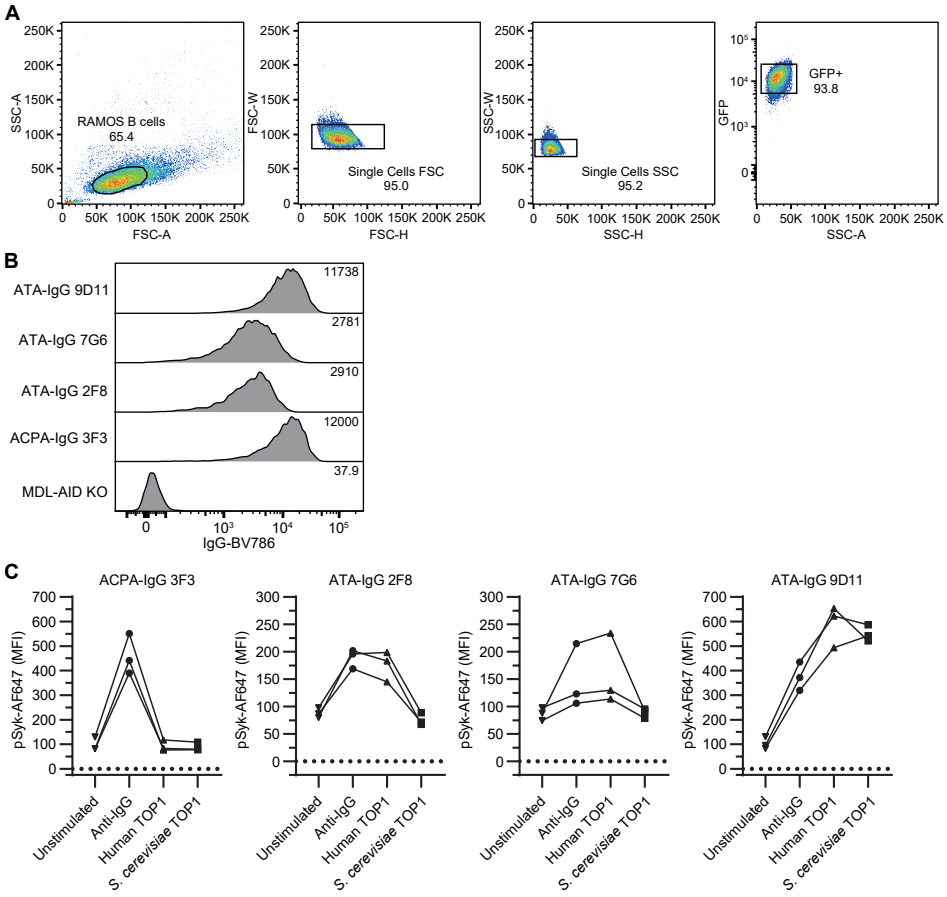
Supplementary Figure S2. Gel electrophoresis of human and *S. cerevisiae* TOP1. Gel electrophoresis of human and *S. cerevisiae* TOP1 upon purification under non-reducing conditions on a polyacrylamide gel with a 4-15% gradient. Protein was visualized with a Coomassie protein stain and molecular weight in kilodaltons (kDa) was estimated using a protein ladder.



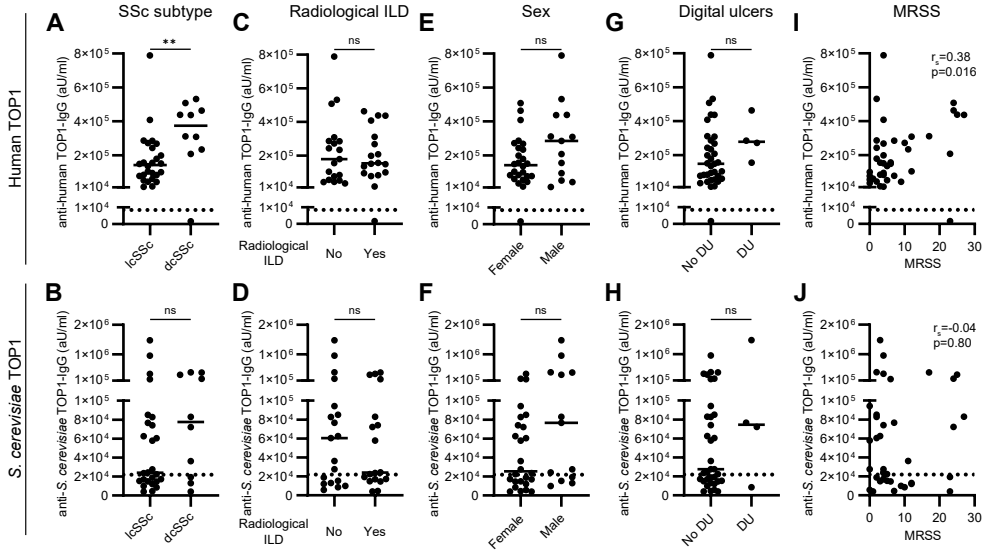
Supplementary Figure S3. Recognition of *S. cerevisiae* TOP1 by circulating IgG of ATA⁺ SSc patients. (A-B) Levels of anti-human TOP1-IgG **(A)** and anti-*S. cerevisiae* TOP1-IgG **(B)** in plasma of ATA⁺ SSc patients (n=22), ACA⁺ SSc patients (n=20) and healthy donors (HD)(n=16) included in cohort 1. **(C)** Correlation between the anti-human TOP1-IgG and anti-*S. cerevisiae* TOP1-IgG levels in the ATA⁺ SSc patients of cohort 1. **(D-E)** Anti-human TOP1-IgG **(A)** and anti-*S. cerevisiae* TOP1-IgG **(B)** levels in serum of ATA⁺ SSc patients (n=39), ACA⁺ SSc patients (n=10) and healthy donors (HD)(n=10) included in cohort 2. **(F)** Association between the levels of anti-human TOP1-IgG and anti-*S. cerevisiae* TOP1-IgG in ATA⁺ SSc patients of cohort 2. **(A-F)** Levels are expressed in arbitrary units per milliliter (aU/ml). **(A-B, D-E)** Kruskal-Wallis test combined with Dunn's multiple comparison test was used to test for statistical significant differences. ns = not significant, * = $p<0.05$, ** = $p<0.005$, *** = $p<0.001$. **(C, F)** Correlations were described using the Spearman's rank correlation coefficient (r_s).



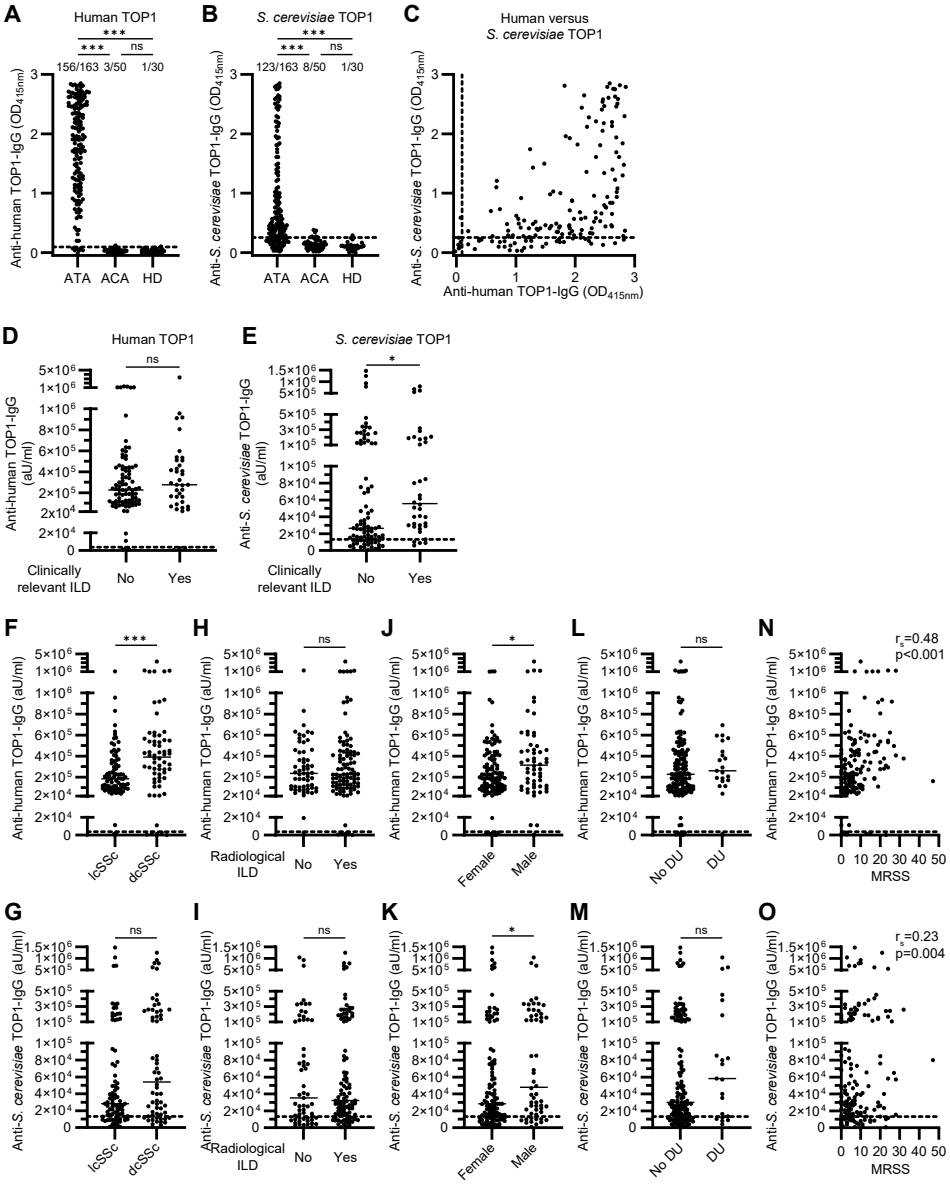
Supplementary Figure S4. Generation of patient-derived ATA mAbs. (A) Gel electrophoresis of patient-derived ATA mAbs and an anti-TT IgG mAb upon Protein G purification under non-reducing (NR) and reducing (R) conditions on a polyacrylamide gel with a 4-15% gradient. Protein was visualized with a Coomassie protein stain and molecular weight in kilodaltons (kDa) was estimated using a protein ladder. **(B)** Reactivity of recombinantly produced ATA-IgG mAbs and anti-TT-IgG D2 towards centromeric protein B (a control protein) in an anti-centromeric protein B antibody-IgG (ACA-IgG) ELISA. mAbs were applied in a 4-fold serial dilution starting at 30µg/ml. Optical density was measured at 415 nm (OD_{415nm}) and blank was subtracted. Error bars represent standard deviation of technical duplicates. Data is representative of two independent experiments.



Supplementary Figure S5. Generation of ATA-expressing Ramos B cell lines. (A) Gating strategy used to identify Ramos B cells by flow cytometry based on forward and side scatter area (FSC-A and SSC-A, respectively), on the exclusion of doublets based on FSC-width (FSC-W) versus FSC-height (FSC-H) and SSC-width (SSC-W) versus SSC-height (FSC-H), and based on the expression of green fluorescent protein (GFP). **(B)** Expression of IgG on the surface of the ATA-expressing Ramos B cell lines (ATA-IgG 2F8, 7G6 and 9D11) and an ACPA-IgG 3F3-expressing Ramos B cell line. Numbers indicate the median fluorescence intensity (MFI). MDL-AID KO cells were used as negative control. **(C)** Expression of pSyk by ATA-expressing Ramos B cell lines upon stimulation for 5 minutes with 5µg/ml anti-IgG, human TOP1 and *S. cerevisiae* TOP1. ACPA-IgG 3F3-expressing Ramos B cell line was used as negative control for the stimulation with human and *S. cerevisiae* TOP1. Expression of pSyk is expressed as MFI. Data is derived from three independent experiments, independent experiment are connected by lines.



Supplementary Figure S6. Association between the recognition of *S. cerevisiae* TOP1 and clinical disease parameters in ATA⁺ SSc patients included in cohort 2. (A-H) Comparison of the levels of anti-human TOP1-IgG and anti-*S. cerevisiae* TOP1-IgG between ATA⁺ SSc patients included in cohort 2 (n=39) with limited cutaneous SSc (lcSSc) versus diffuse cutaneous SSc (dcSSc) **(A-B)**, with versus without interstitial lung disease (ILD) on high resolution computed tomography (HRCT) **(C-D)**, between male and female subjects **(E-F)**, and patients with or without digital ulcers **(G-H)**. Mann-Whitney U test was used to test for significant differences. ns = not significant, * = $p<0.05$, ** = $p<0.005$, *** = $p<0.001$. **(I-J)** Association between the recognition of human and *S. cerevisiae* TOP1 by circulating IgG and MRSS of ATA⁺ SSc patients included in cohort 2 (n=39). Correlations were described using the Spearman's rank correlation coefficient (r_s). **(A-J)** Levels are expressed in arbitrary units per milliliter (aU/ml).


A
Human TOP1
***
***
ns
156/163 3/50 1/30
Anti-human TOP1-IgG (OD415nm)
ATA ACA HD
B
S. cerevisiae TOP1
***
***
ns
123/163 8/50 1/30
Anti-S. cerevisiae TOP1-IgG (OD415nm)
ATA ACA HD
C
Human versus
S. cerevisiae TOP1
Anti-S. cerevisiae TOP1-IgG (OD415nm)
Anti-human TOP1-IgG (OD415nm)
D
Human TOP1
ns
Anti-human TOP1-IgG (aU/ml)
Clinically relevant ILD
No Yes
E
S. cerevisiae TOP1
*
Anti-S. cerevisiae TOP1-IgG (aU/ml)
Clinically relevant ILD
No Yes
Human TOP1
F
***
lcSSc dcSSc
H
ns
Radiological ILD
No Yes
J
*
Female Male
L
ns
No DU DU
N
rs=0.48
p<0.001
MRSS
S. cerevisiae TOP1
G
ns
lcSSc dcSSc
I
ns
Radiological ILD
No Yes
K
*
Female Male
M
ns
No DU DU
O
rs=0.23
p=0.004
MRSS

Supplementary Figure S7. Association between the recognition of *S. cerevisiae* TOP1 and clinical disease parameters in ATA$^+$ SSc patients included in cohort 3. (A-B) Binding of human **(A)** and *S. cerevisiae* **(B)** TOP1 by IgG in serum of ATA$^+$ SSc patients (n=163), ACA$^+$ SSc patients (n=50) and healthy donors (HD)(n=30) included in cohort 3. Optical density was measured at 415 nanometers (OD_{415nm}) and blank was subtracted. Saturated samples (OD>2.5) were diluted further to determine antibody levels in arbitrary units per milliliter (aU/ml), see Figure 5. Kruskal-Wallis test combined with Dunn's multiple comparison test was used to test for statistical significant differences. **(C)** Relation between the reactivity towards human and *S. cerevisiae* TOP1 by ATA$^+$ SSc patients in cohort 3. **(D-E)** Comparison of anti-human TOP1-IgG **(D)** and anti-*S. cerevisiae* TOP1-IgG **(E)** levels in serum between ATA$^+$ SSc patients included in cohort 3, with exclusion of patients which are also included in cohort 2, with and without clinically relevant ILD. **(F-M)** Comparison of the levels of anti-human TOP1-IgG and anti-*S. cerevisiae* TOP1-IgG between ATA$^+$ SSc patients included in cohort 3 with limited cutaneous SSc (lcSSc) versus diffuse cutaneous SSc (dcSSc) **(F-G)**, with versus without interstitial lung disease (ILD) on high resolution computed tomography (HRCT) **(H-I)**, between male and female subjects **(J-K)**, and patients with or without digital ulcers **(L-M)**. **(N-O)** Correlation between the recognition of human and *S. cerevisiae* TOP1 and MRSS by circulating IgG of ATA$^+$ SSc patients included in cohort 3. **(A-O)** Striped lines represent cut-off: mean + 2x standard devation of HDs. **(D-O)** Levels are expressed in arbitrary units per milliliter (aU/ml). **(C, N-O)** Correlations were described using the Spearman's rank correlation coefficient (r_s). **(D-M)** Mann-Whitney U test was used to test for significant differences. **(A-B, D-M)** ns = not significant, * = p<0.05, ** = p<0.005, *** = p<0.001.