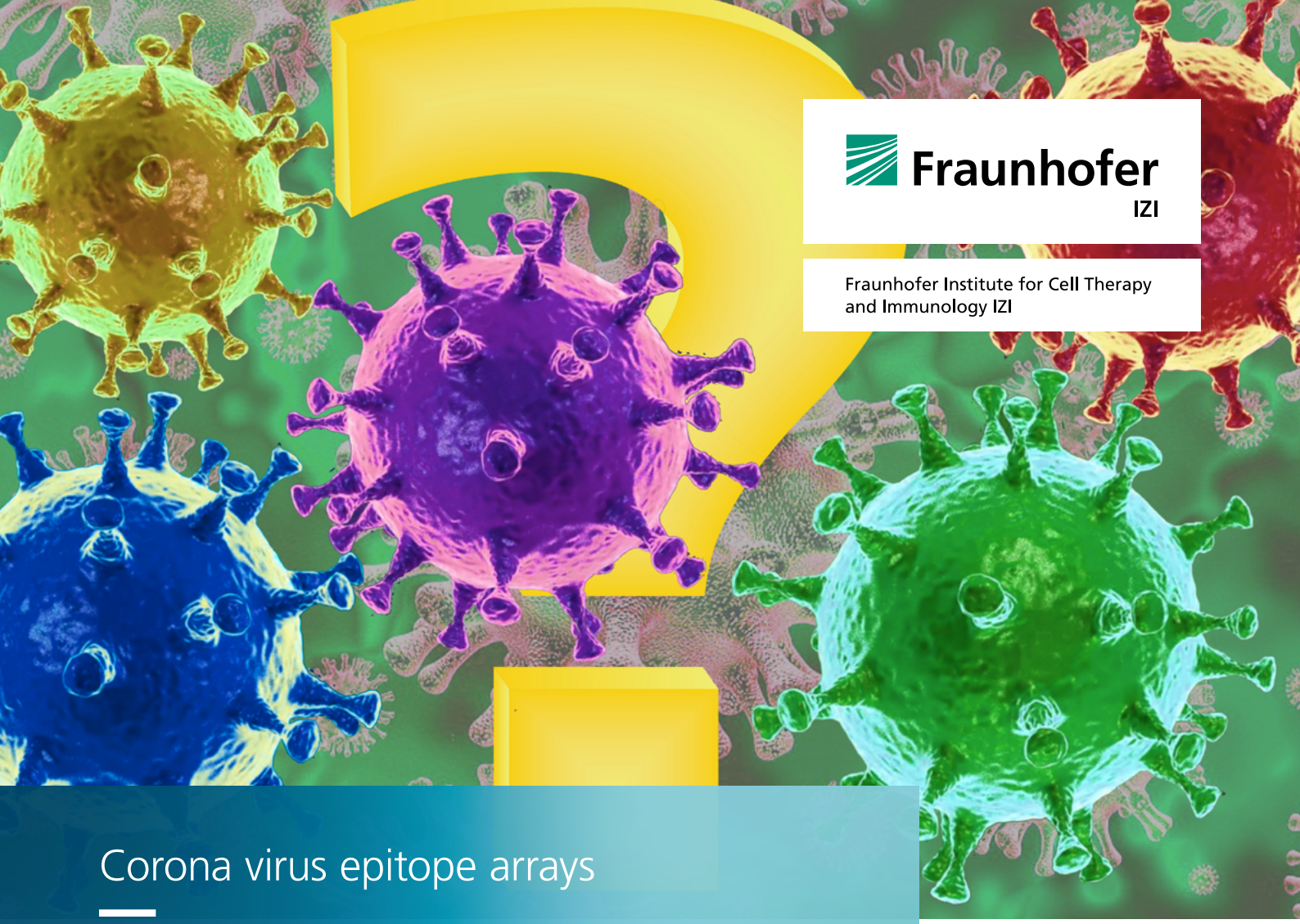


Fraunhofer IZI

Fraunhofer Institute for Cell Therapy and Immunology IZI

Corona virus epitope arrays

Peptide epitopes covering CoV-2 and endemic corona virus

Epitopes for differentiated diagnosis

In the serological diagnosis of COVID-19 it is a particular challenge to make a specific diagnosis against a high background of other coronavirus infections. Genome and thus proteins of the different coronaviruses are up to 50 percent identical. Serological screening for patient antibodies with whole envelope proteins is therefore usually always associated with the risk of a high and variable number of false positive results.

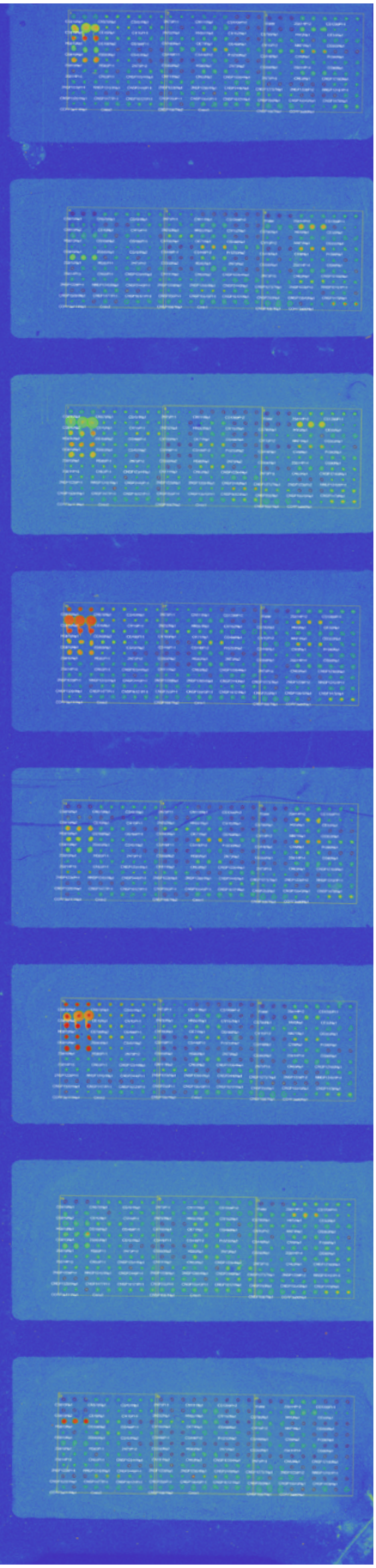
In the EpiCoV2020 project, the Ligand Development Unit and the Cell-Functional Image Analysis Unit of the Fraunhofer Institute for Cell Therapy and Immunology IZI and the Department of Molecular and Cellular Bioanalytics of the Fraunhofer IZI's Bioanalytics and Bioprocesses Division were working on identifying coronavirus-specific epitopes.

Established epitope toolbox

More than 100 epitopes have been identified from all different corona virus strains several different proteins. These are recognized by antibodies from at least a larger fraction of the patients and they are suitable for the specific detection of antibodies against various corona viruses. The identified antibodies give interesting insights into to patients' ongoing and pre-existing immunisation, i.e. the presence of antibodies against other corona viruses as well as the temporal development of the IgM, IgG and IgA response for individual epitopes.

Fraunhofer IZI offers these epitopes and arrays for preclinical and clinical study support and development of diagnostic assays.

There are several endemic corona viruses around. These have generated a general immune reaction also against CoV-2. Peptide epitopes are the only way to differentiate between the individuals and their immunological past. (Picture courtesy from K. Rischka, Fraunhofer IFAM.)



Peptide Arrays

Fingerprinting of potential epitopes in all SARS-CoV-2 virus proteins including all endemic viruses (HKU1, OC43, 229E, NL63) has been carried out. This resulted in the successful identification of more than 100 potential epitopes. The array on the left shows a routine array for the characterization of patient sera with 90 peptides. The usage of epitopes determined at amino acid resolution results in sometimes surprising results for already known immunogenic regions for example of the S-protein.

Special glass slides have been developed that allow the coupling of peptides applying click chemistry. This allows to couple almost any peptide irrespective of its chemical characteristics and makes these arrays a versatile tool to explorer disease specific epitopes rapidly and with a large number of sera. In the ongoing studies more than 1000 different sera have been analyzed for IgG, IgM and IgA binding.

Within the project a special array imaging software has been developed, which allows to identify and quantify the spot intensities on a single array in about 5 min instead of the approximate hour required by commercial semi-automatic software. Only with the help of this software the large number of samples investigated in this study could be handled.

Array slide processed with a specially developed array imaging software: Eight different COVID-19 patients sera on an array with 90 peptides in triplicate. Peptides are taken from all different protein classes and also from the endemic corona viruses.

The entire system allows the measurement of 160 patient samples within 36 hours. Slides are containing individual arrays from 16 x 90 spots (30 peptides) up to 2 x 1000 spots (300 peptides) and can be specifically manufactured for projects including our validated peptides epitopes.

SARS-CoV-2 and Endemic Corona Virus Immune Response

A surprisingly large number of epitopes cross reactive between the different virus strains and an in many cases very early immune response against such epitopes has been found. They role in disease progression and their value as predictive markers for the outcome of the disease is under investigation. Vaccinated patients show an in many aspects similar picture. But we learned from the results with these sera, that there are significant differences to antibodies caused by the COVID-19 disease. These can only be measured, when the exactly fitting epitopes for the different antibody families are used.

Epitope Validation

For the epitope validation larger numbers of COVID-19 and vaccinated sera from the municipal hospital St.Georg in Leipzig (S. Borte), respiratory disease patient sera (InVent, Hennigsdorf) as well as healthy individuals were available. The epitopes in the figure on this page show a good binding of IgG. Variations of the treshhold for a positive measurement have little influence on the number of positive sera. Here the threshold was set either 5-fold or 20-fold background. This is a reliable data basis for the usage of the epitopes in diagnostics and for clinical studies.

Fingerprinting of potential epitopes in all SARS-CoV-2 virus proteins including all endemic viruses (HKU1, OC43, 229E, NL63) has been carried out. This resulted in the successful identification of more than 100 potential epitopes. The array on the left shows a routine array for the characterization of patient sera with 90 peptides. The usage of epitopes determined at amino acid resolution results in sometimes surprising results for already known immunogenic regions for example of the S-protein.

Special glass slides have been developed that allow the coupling of peptides applying click chemistry. This allows to couple almost any peptide irrespective.

			CoV-2 n=53		vaccine n=24		r.d.** n=36		r.d. all n=105	non covid n=16	
peptide code	strain	Protein	T20	T5	T20	T5	T20	T5		T20	T5
C-E-45Na1	SARS-CoV2	E-protein	43%	51%	4%	13%	3%	8%	2%	0%	0%
C-E-55Na1	SARS-CoV2	E-protein	0%	10%	0%	29%	0%	0%	n.d.	0%	13%
O-E-71Na1	OC43	E-protein	49%	63%	71%	96%	6%	58%	7%	6%	6%
H-M-5Na1	HKU1	M-protein	75%	76%	88%	100%	53%	81%	62%	13%	19%
C-M-8Na1	SARS-CoV2	M-protein	59%	73%	25%	83%	8%	28%	n.d.	0%	6%
C-M-15Ph1	SARS-CoV2	M-protein	24%	51%	17%	38%	0%	6%	n.d.	0%	0%
C-M-15Ph2	SARS-CoV2	M-protein	16%	37%	4%	21%	0%	3%	n.d.	0%	0%
C-M-152Na1	SARS-CoV2	M-protein	10%	25%	0%	21%	0%	0%	n.d.	0%	0%
C-N-53Na1*	SARS-CoV2	N-protein	63%	71%	88%	100%	53%	97%	n.d.	6%	19%
C-N-53Na2*	SARS-CoV2	N-protein	61%	73%	96%	100%	8%	64%	n.d.	6%	6%
C-N-53Ph1*	SARS-CoV2	N-protein	4%	16%	0%	17%	0%	3%	n.d.	0%	0%
N-N-71Na1	NL63	N-protein	22%	47%	46%	63%	3%	6%	17%	0%	0%
C-N-111Na1	SARS-CoV2	N-protein	2%	10%	0%	13%	0%	3%	n.d.	0%	0%
H-N-351Na1	HKU1	N-protein	51%	59%	46%	92%	8%	33%	7%	0%	0%
C-N-376Na1	SARS-CoV2	N-protein	55%	63%	25%	33%	3%	8%	n.d.	0%	6%
2-NSP12-28Ph1	229E	NSP-12	18%	59%	13%	54%	0%	3%	8%	0%	6%
C-NSP12-344Na1	SARS-CoV2	NSP-12	29%	51%	21%	50%	0%	22%	n.d.	0%	6%
N-NSP12-431Na1	NL63	NSP-12	10%	25%	4%	21%	6%	11%	9%	0%	6%
N-NSP12-431Ph1	NL63	NSP-12	29%	53%	42%	67%	6%	17%	18%	0%	6%
C-NSP13-543Na1	SARS-CoV2	NSP-13	8%	24%	0%	29%	0%	6%	n.d.	0%	0%
C-NSP13-543Ph1	SARS-CoV2	NSP-13	49%	63%	33%	46%	25%	47%	n.d.	0%	0%
C-NSP14-77Ph1	SARS-CoV2	NSP-14	29%	37%	13%	21%	8%	11%	n.d.	0%	0%
C-NSP14-78Na1	SARS-CoV2	NSP-14	18%	47%	21%	63%	0%	0%	n.d.	0%	6%
C-NSP16-87Na2	SARS-CoV2	NSP-16	6%	29%	29%	29%	0%	3%	n.d.	0%	6%
C-NSP16-121Na1	SARS-CoV2	NSP-16	55%	69%	50%	92%	3%	14%	n.d.	6%	6%
C-NSP16-121Ph1	SARS-CoV2	NSP-16	14%	20%	0%	25%	0%	3%	n.d.	0%	0%
C-NSP2-55Na1	SARS-CoV2	NSP-2	43%	63%	42%	88%	3%	6%	n.d.	0%	6%
C-NSP2-440Na1	SARS-CoV2	NSP-2	41%	63%	58%	92%	3%	25%	n.d.	6%	6%
C-NSP2-440Ph1	SARS-CoV2	NSP-2	33%	47%	17%	33%	11%	19%	n.d.	0%	0%
H-S-85Na1	HKU1	S-protein	20%	37%	17%	21%	6%	8%	8%	0%	0%
H-S-85Ph1	HKU1	S-protein	22%	43%	4%	25%	3%	3%	3%	0%	6%
O-S-168Na2	OC43	S-protein	45%	55%	25%	50%	6%	6%	10%	6%	6%
C-S-431Na2	SARS-CoV2	S-protein	4%	12%	0%	21%	0%	0%	n.d.	0%	0%
C-S-350Na1	SARS-CoV2	S-protein	12%	18%	13%	38%	0%	6%	n.d.	0%	0%
C-S-431Na1	SARS-CoV2	S-protein	37%	67%	25%	79%	8%	14%	n.d.	19%	50%
C-S-448Ph1	SARS-CoV2	S-protein	6%	12%	0%	4%	3%	3%	n.d.	0%	0%
C-S-448Ph2	SARS-CoV2	S-protein	53%	65%	33%	79%	8%	31%	n.d.	0%	6%
O-S-488Ph1	OC43	S-protein	8%	22%	8%	33%	3%	8%	3%	0%	0%
2-S-614Ph4	229E	S-protein	29%	53%	4%	54%	0%	6%	3%	0%	0%
2-S-614Ph2	229E	S-protein	61%	71%	63%	92%	3%	39%	58%	13%	13%
2-S-614Ph3	229E	S-protein	6%	24%	4%	25%	0%	3%	32%	0%	0%
2-S-815Na1	229E	S-protein	84%	88%	79%	96%	33%	58%	22%	6%	6%
C-S-815Na1	SARS-CoV2	S-protein	63%	71%	4%	8%	14%	19%	n.d.	6%	13%
C-S-815Na2	SARS-CoV2	S-protein	76%	78%	46%	54%	17%	25%	10%	0%	0%
H-S-815Na1	HKU1	S-protein	86%	88%	96%	100%	100%	100%	74%	0%	6%
O-S-815Na1	OC43	S-protein	76%	84%	42%	75%	28%	36%	24%	6%	6%
C-S-1209Ph2	SARS-CoV2	S-protein	2%	8%	4%	4%	0%	0%	n.d.	0%	0%
Pr-390-Na1	huProteinase3		14%	25%	4%	8%	3%	3%	n.d.	0%	0%

*List and serological profile of some available corona virus epitopes. Peptide codes: 2/C/H/N/O, first letter of strain 229E. SARS-CoV-2, HKU1. NL63, OC43; S/N/... protein code; Number = first residue position in first discovered sequence; Na naive / Ph phage derived sequences (mimotopes) and consecutive number. The position for identical and homologous regions is kept in the epitope code throughout all strains' sequences. *C-N-53 are cross reactive with an S-protein epitope. Position: First amino acid of underlined motif identified in data; Patient data for 5× and 20× threshold in array measurements: COVID-19 patients; vaccinated (altogether); R.D. respiratory disease with R.D.** potential HKU1 patients only; R.D. all: data from parallel testing with a different set of peptides specific for endemic coronaviruses (see supplement) containing not all of the epitopes; non-COVID-19: healthy control.*

Method Background

Epitope fingerprinting

Epitope fingerprinting is a novel method developed at Fraunhofer IZI in collaboration with epitopic GmbH (Leipzig). Instead of generating novel peptide libraries or arrays specific for each antigen, which is the core of most comparable methods, epitope fingerprinting is applying an advanced combination of gene libraries, NGS and a specially developed software to map not only single antibodies but also the entire spectrum of antibodies found in a single drop of patient serum.

Advanced library design

A patented library design is the basis of a naïve peptide library on phage. This comprises more than 5 billion different 16mer sequences. The library has several unique advantages:

- Statistically well balanced, all sequences occur with the same frequency
- Covers almost all possible combinations of seven amino acids
- Allows constrained (Cys-flanked) as well as linear peptides in one library
- Stable in replication: Unlike similar biological libraries it can be replicated with minimal counterselection of "unfavorable" sequences

State of the art technologies

Peptides are selected by binding to antibody molecules. Because of the library's characteristics only one or wo selection rounds are required. Working with peptide presenting phage only 50-100 need to be enriched, this increases the sensitivity. Only a few hundred antibody molecules are required to identify an epitope, when the selected peptide's genes are identified with NGS.

Novel software

A novel software was developed, which can handle the NGS data and helps to identify individual motifs:

- Reliable recognition of NGS sequencing and removal of suspicious sequences
- Statistical identification of peptide motifs, which is made possible by the reliable amino acid distribution in the patented peptide library
- Rapid comparison with potential antigens and a user-friendly JAVA interface that allows quick analyses of sequences containing potential epitope motifs
- Overall statistical analysis methods allow to identify unusual motifs for unknown targets or sequences enriched by binding to targets like cancer cells.

The final output of the typical analysis yields dozens and more enriched sequences sharing antigen motifs. These allow to assign the probably essential amino acids of an epitope. These are often surrounded by structurally relevant amino acids not present in the linear peptide of the antigen. Since the library peptides are 16 amino acids long and allow cysteines loop structures, unusual folds are often found directly from the analysis of the sequence alignment.

Validation of peptide epitopes in microarrays

Although the statistical predictions are usually conclusive for single antibodies or sera, the diagnostic value of a peptide epitopes needs to be validated by analysis with multiple sera. For this purpose, a method was developed to immobilize any peptide irrespective of its individual chemistry on microarrays. This implies the use of click chemistry and allows to couple each peptide covalently through a linker to a glass slide surface in arrays of up to several hundred peptides. This is a routine method established by Fraunhofer IZI-BB in Potsdam. Minimal quantities of serum are required to measure binding of antibodies from the individual sera. The bound patient antibodies are detected with fluorescently labeled molecules and the arrays can by analyzed in any standard array reader.

A novel software for automatic analysis of array data has been created to overcome the bottleneck of semi manual evaluation. This allows to measure the epitope spectrum of hundreds of sera within a few days.

Contact

Dr. Michael Szardenings
Department of Diagnostics
Ligand Development Unit
Tel. +49 381 494-2805
michael.szardenings@
izi.fraunhofer.de

Fraunhofer Institute
for Cell Therapy and
Immunology
Perlickstraße 1
04103 Leipzig
Germany
www.izi.fraunhofer.de

Dr. Eva Ehrentreich-Förster
Director (acting)
Tel. +49 331 58187-203
eva.ehrentreich@
izi-bb.fraunhofer.de

Fraunhofer Institute
for Cell Therapy and
Immunology, Branch
Bioanalytics and
Bioprocesses IZI-BB
Am Mühlenberg 13
14476 Potsdam
Germany
www.izi-bb.fraunhofer.de