
YghJ: Preclinical Vaccination and Human Disease Studies at the Epitope Level

Statistical Analyses based on NGS and Novel Library Technologies



SVEET

Sugar-modified Vaccine Epitope
Exploration and Translation



Klinikum | **St.GEORG**



SVEET

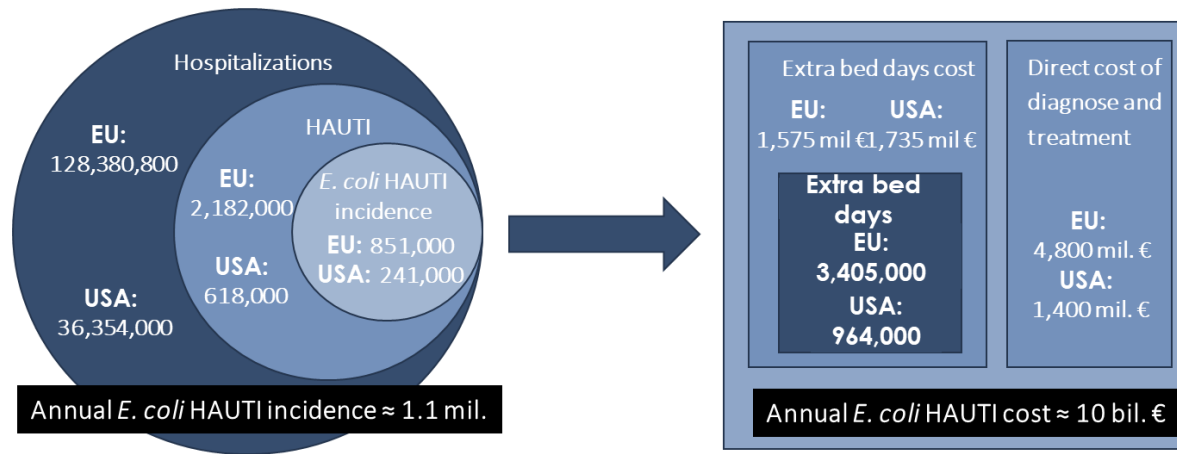
Sugar-modified Vaccine Epitope
Exploration and Translation



ABOUT SVEET

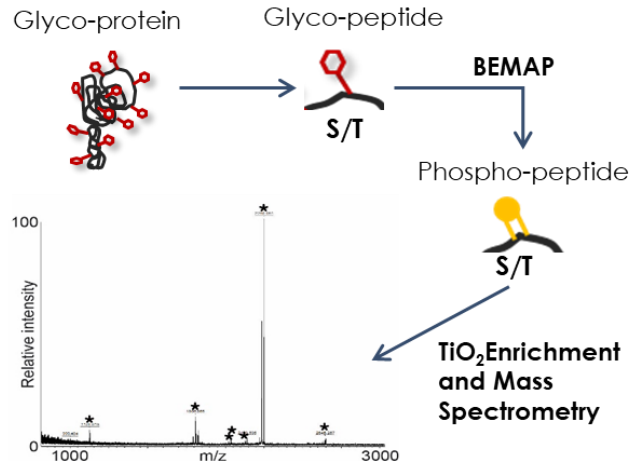
Background

- Urinary tract infections caused by *E. coli* (UPEC)
 - Simple UTI affect 40–50% of all women in their lifetime
 - 25% of these experience recurrent infections
- Hospital-acquired urinary tract infections (HAUTI)
 - Risk factors: age (>65 years), insertion of urine catheters and the length of hospitalization
 - Increasing challenge due to development of resistance to treatment (AMR)
 - Currently no effective way of preventing HAUTI



BEMAP – GLYPROVAC PROPRIETARY TECHNOLOGY

BEMAP is a mass spectrometry-based method, which allows one to **pinpoint** the exact location of any type of O-linked protein glycosylation.



BEMAP analysis has revealed a staggering number of glycoproteins particularly in disease causing bacteria.

Bacterial strain	Relevant organism Disease/ Indication	Number of identified glycoproteins	
		Known prior to BEMAP	Identified with BEMAP
<i>E. coli</i> K12	None	0	34
ETEC H10407	Diarrhoea	2	>300
NMEC IHE3034	Neonatal meningitis	0	>150
UPEC UT189	Urinary tract infections	0	>250
<i>Shigella sonnei/flexneri</i>	Bloody Diarrhoea	0	>250
<i>Pseudomonas aeruginosa</i> PA14/LESB58	Skin, Lung and urinary tract infections	<5	>250

The table mainly contains unpublished data

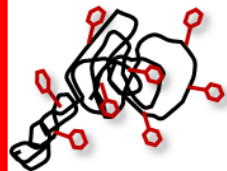
BEMAP – GLYPROVAC PROPRIETARY TECHNOLOGY

BEMAP is a mass spectrometry-based method, which allows one to **pinpoint** the exact location of any type of O-linked protein glycosylation.

BEMAP analysis has revealed a staggering number of glycoproteins particularly in disease causing bacteria.

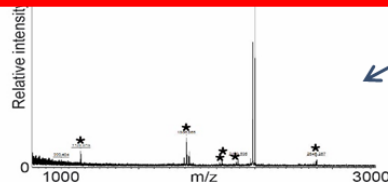
Relevant organism

Number of identified glycoproteins



The GlyProVac hypothesis:

Inclusion of glycosylations in vaccine design ensures potent vaccine candidates



TiO₂ Enrichment and Mass Spectrometry

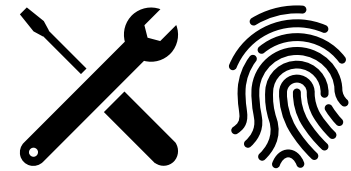
<i>Shigella sonnei/flexneri</i>	Bloody Diarrhoea	0	>250
<i>Pseudomonas aeruginosa</i> PA14/LESB58	Skin, Lung and urinary tract infections	<5	>250

The table mainly contains unpublished data

SVEET – Epitope tasks

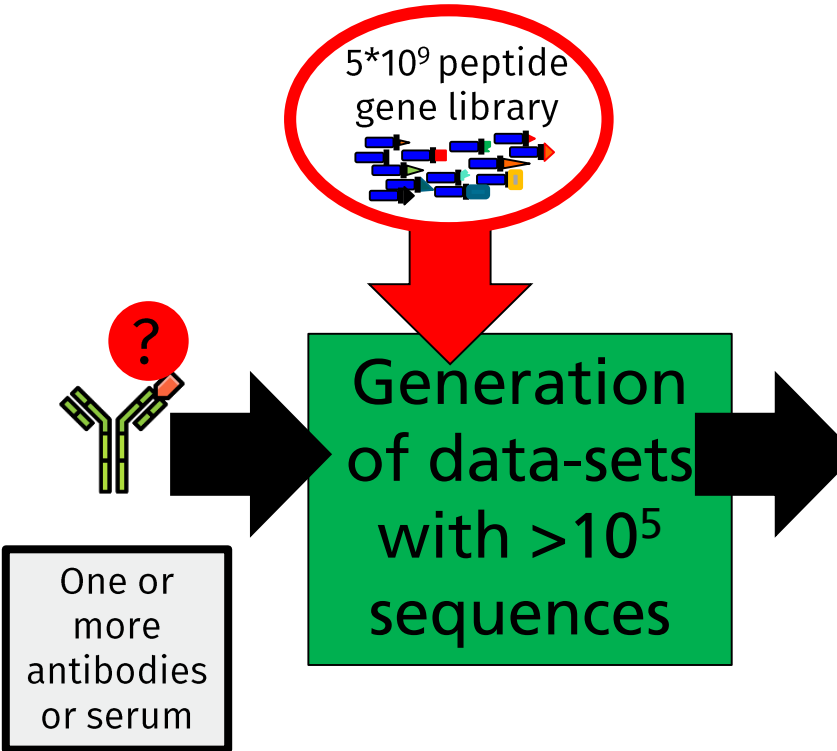
- Identify epitopes from preclinical studies and from patients
- Identify differences between glycosylated and non-glycosylated vaccines
- Provide arrays for routine (pre)clinical studies

- Identify peptide mimotopes for glycosylation sites
- Evaluate „manual“ mimotope identification methods and turn them into algorithms



ABOUT EPITOPE FINGERPRINTING

How Epitope-Fingerprinting Works



Result: Hundreds of sequences weakly enriched **sharing epitope motifs** are identified entirely in silico for any chosen antigen!

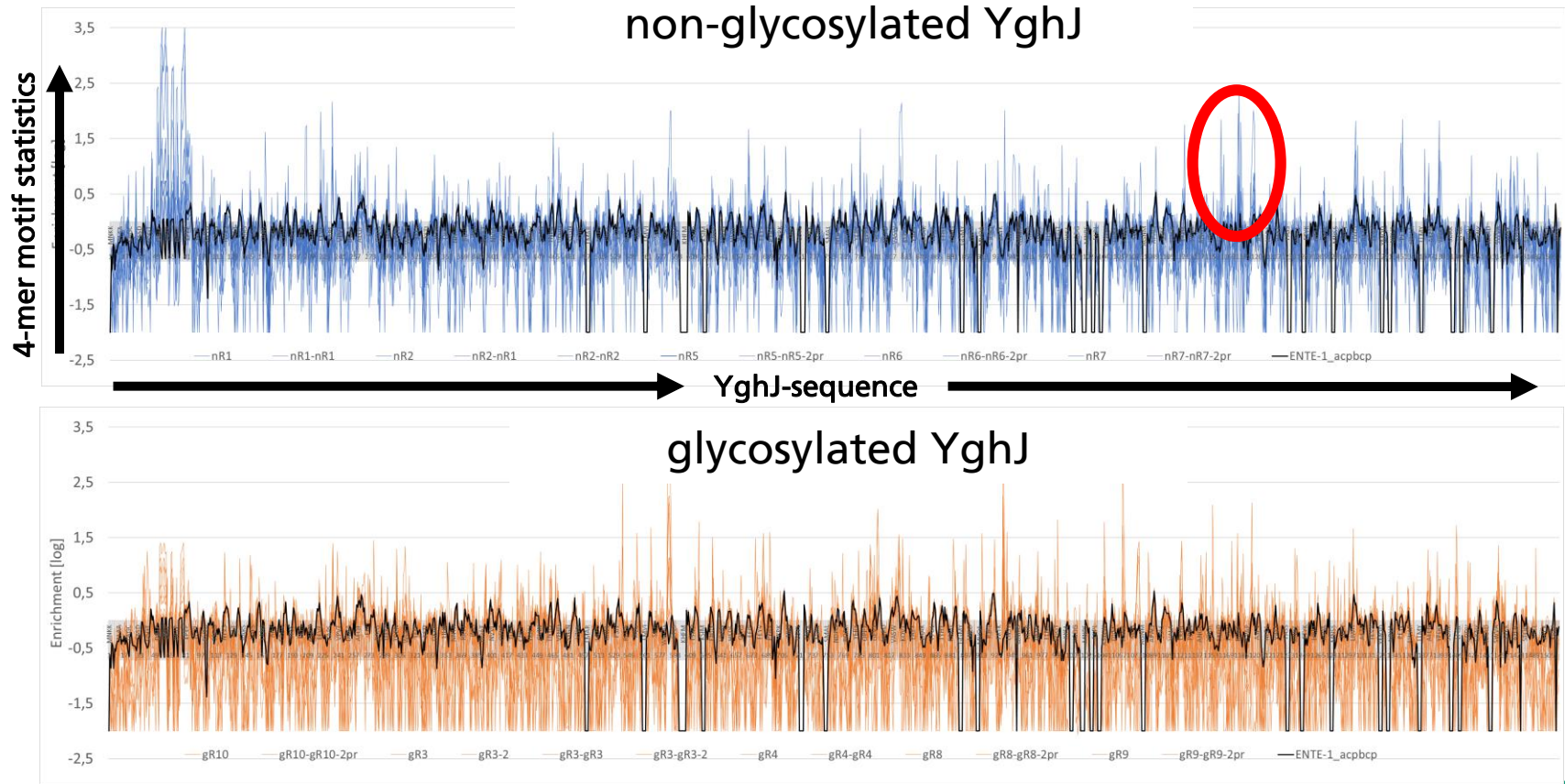
Consensus	----- I f S c L k d r h d f -----
huINF_alpha_2	M R R I S L F S C L K D R H D F G F P Q
seq5 S2-12616729-1pr count=5	----- N F S C L S D T R Q C T N K L
seq1 S2-12616729-2pr count=9	----- D F S C L E F R A C E P V E V
seq14 S2-12616729-1pr count=4	N R W C S L F S E Q Y D V Q D I -----
seq5 S2-12616729-2pr count=4	- S H N S L F S P L F Q P H Q H Y - - -
seq12 S2-12616729-2pr count=5	----- S L F S S P K H N Q Q Q Y Q C C
seq16 S2-12616729-2pr count=4	----- S L F S G L S L V F Q V H H N F
seq28 S2-12616729-2pr count=4	----- S L F S S L L P A V E K S V V C
seq1 S2-12616729-2pr count=5	----- F P A C K D R H R L Q V V Y
seq9 S2-12616729-1pr count=4	----- F K S C L K T I H N G F A A Y
seq3 S2-12616729-2pr count=4	H N F C P D Y S S N K D R H D F - - - -
seq7 S2-12616729-2pr count=8	W Q S C F D H Y C I K P R H D F - - - -
seq7 S2-12616729-2pr count=5	----- K Y V S K W S H D F G H Q F

Starting from a very large set of random sequences binding and statistical analyses result in epitope motifs and hundreds or more sequences of binding peptides.

How Epitope-Fingerprinting Works

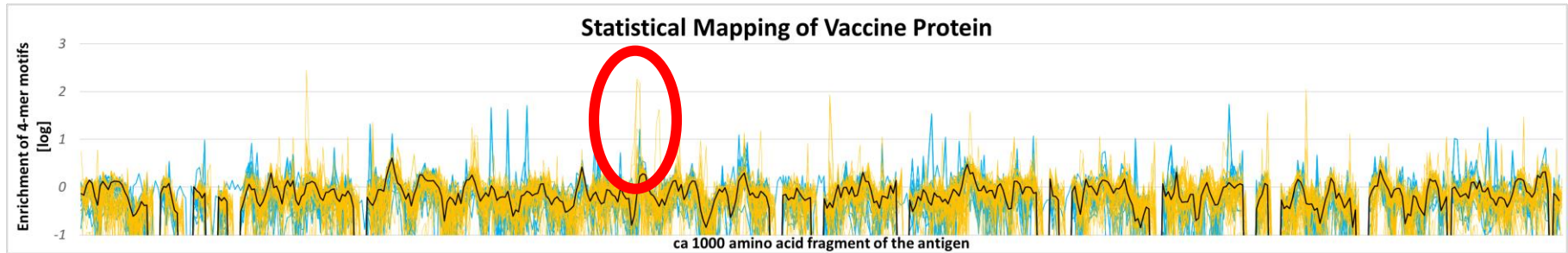
- ...applying in silico analyses to retrieve „binding“ peptide information from NGS data sets of peptide phage display
 - One or two selection rounds on a few thousand antibody molecules
 - Requires 100 µl serum
 - **Analyses motif statistics** instead of individual full length sequences


Comparing g/ng-YghJ in rabbits – enrichment of 4-mers



Patients vs Rabbits

The blue statistical curve for patient sera and another group of rabbits (yellow) receiving different vaccines.



„Traue keiner Statistik, die du nicht selber gefälscht hast“ 
(Don't trust statistics, which you have not falsified yourself)

IDENTIFICATION OF EPITOPE SEQUENCES

Epitope fingerprints by statistics

Motif 1178_PKFYKDGGEWK

Identified more than 2500 sequences containing one of the 4-mer motifs occur more than twice in six rabbit data-sets.

Displayed are only sequences occurring at least 5x and sharing at least 5 amino acids identity.

Ruler 1	1	10	20	30	40
Consensus	- - - - - v n h s f n k d g e x v q f - - - - -				
YGHJ_1161-1209	S K D D V S A N F T F T G V V K A P F Y K D G E W K N D L D S P A P L G E L E S				
seq87 K12-2prfcount=6	- - - - - Y T P C V I N D G E W Y I L C - - - - -				
seq94 K12-2prfcount=25	- - - - - A P A C W V D E G E W K V D I D - - - - -				
seq179 K13-2prfcount=26	- - - - - S W H S F N K D G E Q V H G I Y - - - - -				
seq180 K13-2prfcount=8	- - - - - F N H S D N K D G E W V Q F A Y - - - - -				
seq184 K13-2prfcount=13	- - - - - Y Y N S P F N K D G E F A G T Y - - - - -				
seq202 K13-2prfcount=5	- - - - - H Y S S V F N K D G E R I G L F - - - - -				
seq211 K13-2prfcount=8	- - - - - F N H S D N K D G E W V Q F A Y - - - - -				
seq322 K14-2prfcount=17	- - - - - I H H S A K V S I N K D G E N I - - - - -				
seq323 K14-2prfcount=12	- - - - - E H W C Q Y D V F N K D G E C Y - - - - -				
seq325 K14-2prfcount=7	- - - - - D H M S Q V F N K D G E Q Q L I - - - - -				
seq331 K14-2prfcount=21	- - - - - S W H S F N K D G E Q V H G I Y - - - - -				
seq335 K14-2prfcount=5	- - - - - W K H C N F N K D G E C I F I A - - - - -				
seq363 K14-2prfcount=6	- - - - - F N H S D N K D G E W V Q F A Y - - - - -				
seq600 K16-1prfcount=17	- - - - - F N H S D N K D G E W V Q F A Y - - - - -				
seq616 K16-1prfcount=5	- - - - - S W H S F N K D G E Q V H G I Y - - - - -				
seq652 K16-1prfcount=17	- - - - - F N H S D N K D G E W V Q F A Y - - - - -				
seq661 K16-2prfcount=36	- - - - - E V L C L W F Y K D Q E C L N Y - - - - -				
seq665 K16-2prfcount=12	- - - - - P E G C F Y K D E E L A F S D - - - - -				
seq676 K16-2prfcount=11	- - - - - V H P C I Q W Y V Y K D G E F C - - - - -				
seq679 K16-2prfcount=21	- - - - - V E - L C Y K D G E R P C F L R D - - - - -				
seq685 K16-2prfcount=7	- - - - - P I A C Q W V Y K D G E C P A A - - - - -				
seq691 K16-2prfcount=733	- - - - - F N H S D N K D G E W V Q F A Y - - - - -				
seq692 K16-2prfcount=428	- - - - - S W H S F N K D G E Q V H G I Y - - - - -				
seq693 K16-2prfcount=26	- - - - - P K F C N F N K D G E R A V S I Y - - - - -				
seq697 K16-2prfcount=88	- - - - - Y Y N S P F N K D G E F A G T Y - - - - -				
seq701 K16-2prfcount=20	- - - - - N R F S I F N K D G E Q H F Q A - - - - -				
seq702 K16-2prfcount=86	- - - - - E H W C Q Y D V F N K D G E C Y - - - - -				
seq703 K16-2prfcount=8	- - - - - G G V C S Y N K D G E H Y N E C - - - - -				
seq708 K16-2prfcount=21	- - - - - V E - L C Y K D G E R P C F L R D - - - - -				
seq709 K16-2prfcount=112	- - - - - D H M S Q V F N K D G E Q Q L I - - - - -				
seq712 K16-2prfcount=32	- - - - - G N P C S Q N K D G E Q T N V D - - - - -				
seq713 K16-2prfcount=40	- - - - - E H V C V N K D G E I T R D C C - - - - -				
seq715 K16-2prfcount=49	- - - - - W K H C N F N K D G E V Q V Q F - - - - -				
seq720 K16-2prfcount=29	- - - - - P W R S S T N K D G E E L Q D F - - - - -				
seq723 K16-2prfcount=24	- - - - - D Y Q C P H N K D G E V S Q P Y - - - - -				
seq724 K16-2prfcount=23	- - - - - W K H C N F N K D G E E N Y R D - - - - -				
seq725 K16-2prfcount=11	- - - - - V H P C I Q W Y V Y K D G E F C - - - - -				
seq727 K16-2prfcount=29	- - - - - P F V C R N V N K D G E C H Y I - - - - -				
seq731 K16-2prfcount=7	- - - - - G D V S E Y F P Y N K D G E N F - - - - -				
seq736 K16-2prfcount=37	- - - - - I H H S A K V S I N K D G E N I - - - - -				
seq739 K16-2prfcount=65	- - - - - I N V C D W R N K D G E T T H C - - - - -				
seq742 K16-2prfcount=8	- - - - - F Y T C G E Q K D G E Y V C S D - - - - -				
seq747 K16-2prfcount=13	- - - - - W K H C N F N K D G E C I F I A - - - - -				
seq748 K16-2prfcount=7	- - - - - H D V C D S K N K D G E N I Y C - - - - -				
seq749 K16-2prfcount=7	- - - - - P I A C Q W V Y K D G E C P A A - - - - -				
seq756 K16-2prfcount=7	- - - - - H V S S V F N K D G E A E Y E N - - - - -				
seq758 K16-2prfcount=8	- - - - - Q E G C P Q D N K D G E E I Y D - - - - -				
seq763 K16-2prfcount=8	- - - - - D Y N S P I N K D G E D A F S A - - - - -				
seq769 K16-2prfcount=6	- - - - - Y Y N S P F N K D G E T C N A C - - - - -				
seq796 K16-2prfcount=12	- - - - - H Y S S V F N K D G E E V H F C - - - - -				
seq951 K17-2prfcount=5	- - - - - H P F C N T F D K D G E Q I R C - - - - -				

Epitope fingerprints by statistics

Motif 1178_PKFYKDGEWK

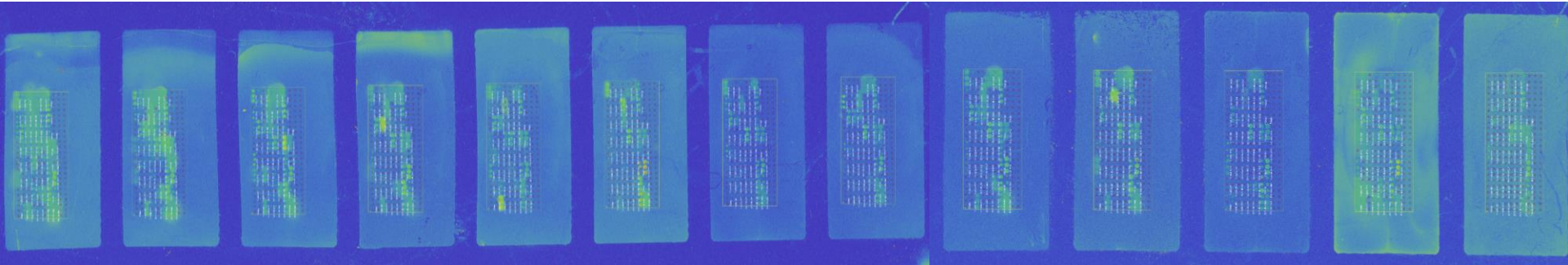
The same epitope is recognized by patient sera as well, but selecting sequences from the data sets with the same stringency yields less sequences, due to a lower antibody concentration.

Ruler 1	
Consensus	----- x s w p F y k D G E W k e x -----
YGHJ_1161-1209	S K D D V S A N F T F T G V V K A P F Y K D G E W K N D L D S P A P L G E L E S A S F V Y T
PS10-2-2pr seq286 count=4	----- Y S W C F G T T G E W K I G E I -----
PS11-1-2pr seq448 count=4	----- S P F S H W D Y P H Y K D G V F -----
PS12-1-2pr seq720 count=7	----- D E N S E P F Y K D R E E V C F -----
PS12-1-2pr seq723 count=5	----- H Q M S R Q E V P F W P F Y K D -----
PS12-1-2pr seq724 count=4	----- Q V P S W P F Y K D E D E I Y A -----
PS12-1-2pr seq729 count=24	----- V A N S F D D D G E W F E A L I -----
PS12-2-2pr seq914 count=6	----- D E N S E P F Y K D R P C I E C -----
PS12-2-2pr seq929 count=13	----- V A N S F D D D G E W F E A L I -----
PS13-1-2pr seq1005 count=4	----- D H M C H K Q H G E W K C D S Y -----
PS14-1-2pr seq1169 count=4	----- Y S H S G H K Y - G E W K Q Q F I -----
PS15-2-2pr seq1444 count=5	-- V W V S P G F E F ----- F W K D G E C -----
PS15-2-2pr seq1453 count=5	----- V Y N C - G E W K T T Q Q Q S G D -----
PS15-2-2pr seq1460 count=6	----- N W W S A E D N G E W K V D Q F -----
PS15-2-2pr seq1463 count=4	----- P Y M C T G E W K S Q S Q F L N -----

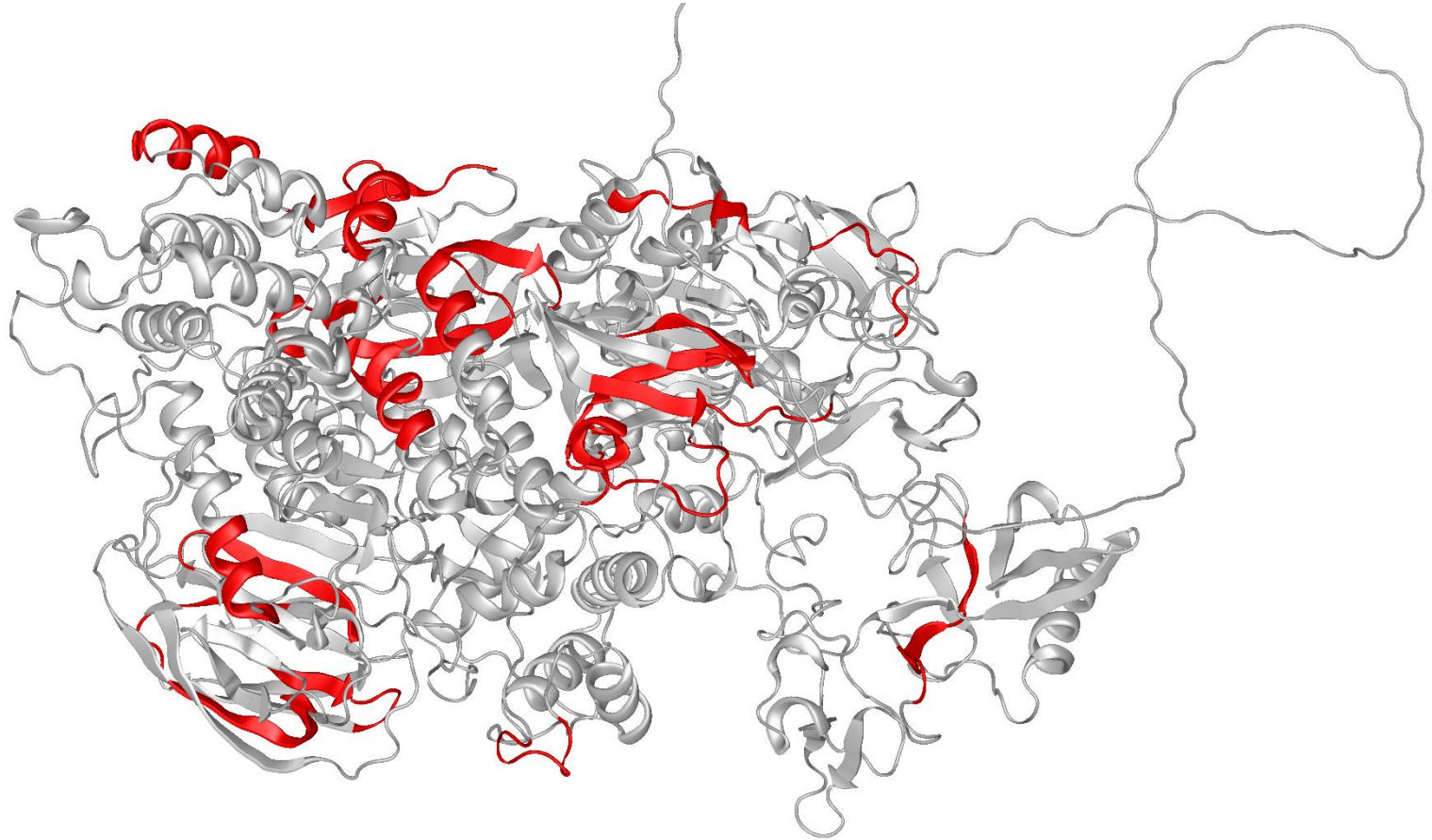
Validaton of Epitopes

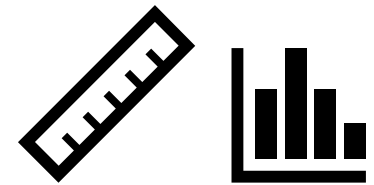
More than 40 epitopes have been identified from patient and rabbit sera. But to finally validate them, they need to be tested in peptide arrays!

Peptides have been synthesized and immobilized with click-chemistry on peptide arrays. Detection with goat-anti-human/rabbit/pig pAb and Cy5-labelled mouse-anti-goat mAb.



YghJ Epitopes

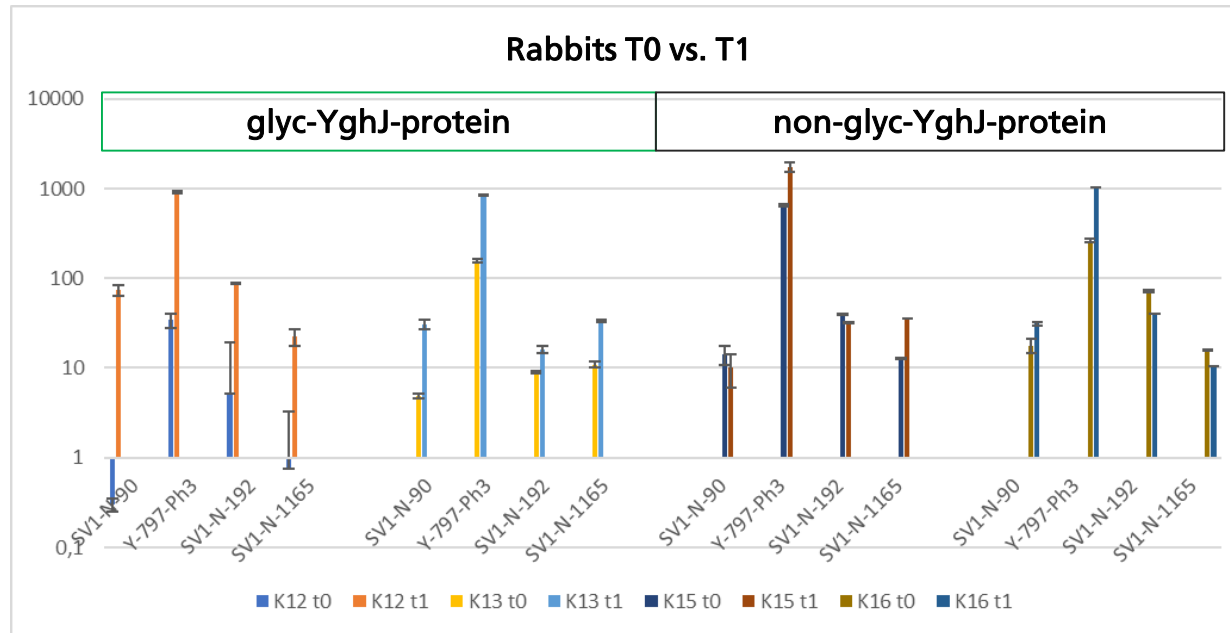




RESULTS FROM PEPTIDE EPITOPE MEASUREMENTS & CONCLUSIONS

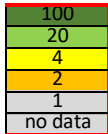
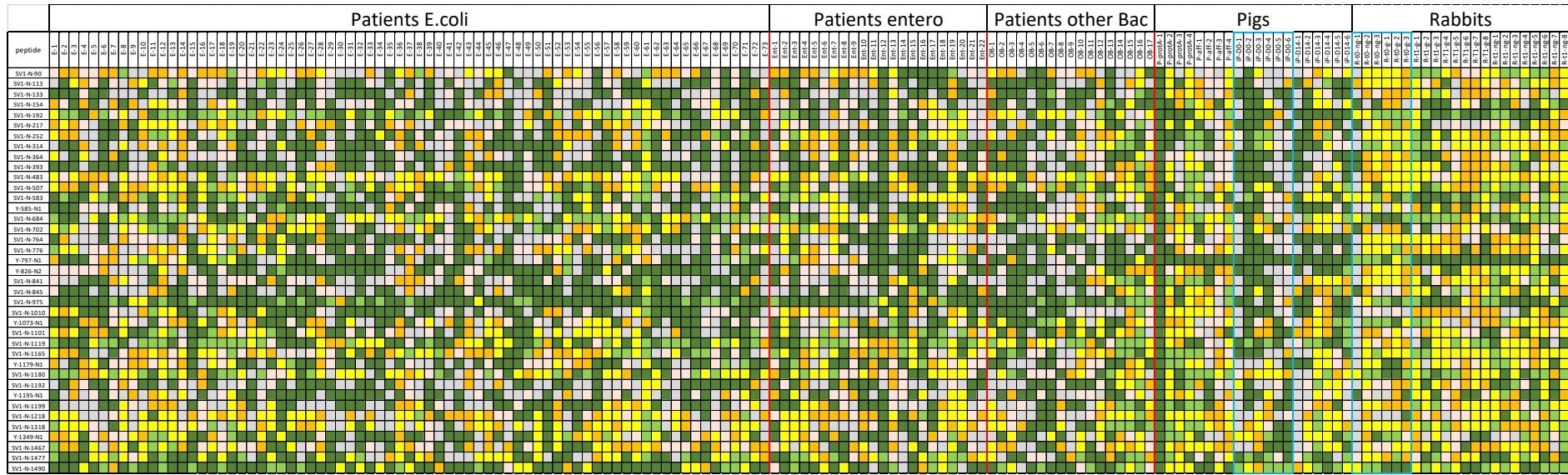
Are there effects of glycosylation?

Comparing the array data for four selected epitope peptide in four rabbit sera taken before and after immunisation. (*intensities n-fold background in log scale*)



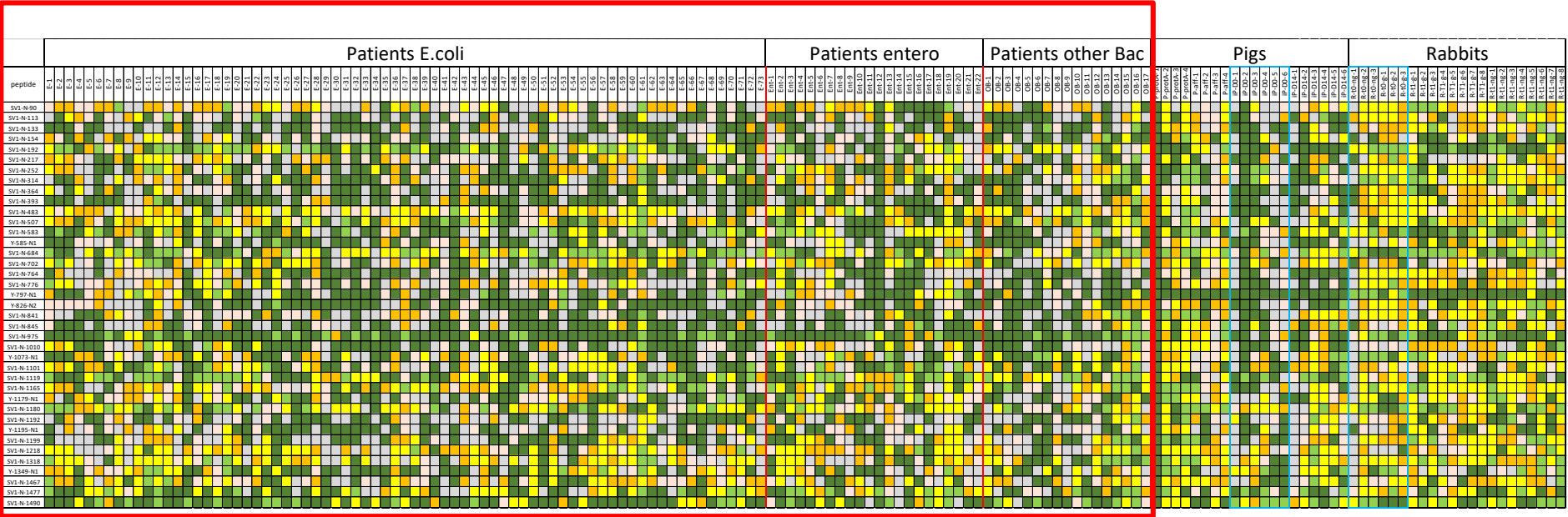
Yes!

Comparing Patients, Pigs and Rabbits



Sera from UTI patients with confirmed infections, pigs vaccinated and from infection models as well as immunized rabbits have been tested in a peptide epitope array. The signal cutoffs are signal over background (empty spots).

Comparing Patients, Pigs and Rabbits



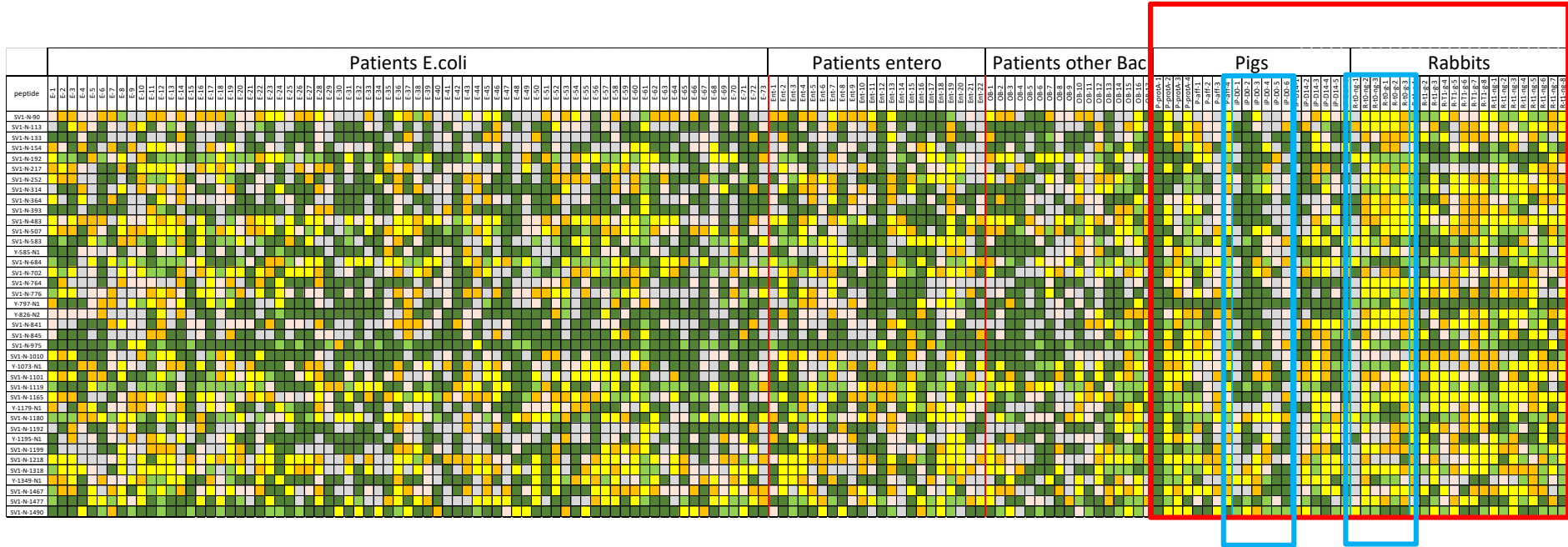
For the patients the epitope patterns are not correlating to infections with *E.coli* or related bacteria

YghJ related proteins are highly conserved

Natural YghJ-like proteins are well conserved in particular in the epitopes' regions. Cross reactivity of antibodies is therefore likely.



Comparing Patients, Pigs and Rabbits



Pigs have antibodies already prior to infection/immunisation, rabbits sometimes, probably due to exposure in foods or just the microbiome. The situation is probably the same for human patients, who are likely to have already established antibodies.

Thanks to

Fraunhofer IZI, Leipzig: Nicolas Delaroque and Andreas Kölsch, Karolin Kern

Fraunhofer IZI, Potsdam: Eva Ehrentreich-Förster



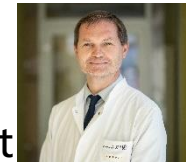
GlyProVac: Ann Zahle Andersen, Anders Boysen



University Hospital Odense, Denmark: Karin Andersen, Thomas Emil, Kristian Staerk, Lars Lund



epitopic GmbH: Marcus Puder , Karolin Kern



Municipal Hospital St.Georg, Leipzig: Ralph Wendt

Contact:

Michael Szardenings

Ligand Development Unit

Fraunhofer Institute for Cell Therapy and Immunology

Perlickstr. 1, 04103 Leipzig, Germany

Tel: 0049-341-35536 2805

Cell: 0049 159 044 268 49

Fax: 0049-341-35536 82805

michael.szardenings@izi.fraunhofer.de

www.izi.fraunhofer.de



SPONSORED BY THE



Federal Ministry
of Education
and Research

FKZ 01QE2109B



This project has received funding from the Eurostars-2 joint programme with co-funding from the European Union Horizon 2020 research and innovation programme

Thank you for your attention