Journal of Medicinal Chemistry

Article

Dynamics determine signaling in a multicomponent system associated with rheumatoid arthritis

Cecilia Lindgren, Mohit Tyagi, Johan Viljanen, Johannes Toms, Changrong Ge, Naru Zhang, Rikard Holmdahl, Jan Kihlberg, and Anna Linusson

J. Med. Chem., **Just Accepted Manuscript** • DOI: 10.1021/acs.jmedchem.7b01880 • Publication Date (Web): 04 May 2018

ACS Publications

is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Dynamics determine signaling in a multicomponent system associated with rheumatoid arthritis

Cecilia Lindgren,[‡,†] Mohit Tyagi,[‡,§] Johan Viljanen,[§] Johannes Toms,[§] Changrong Ge,[∥] Naru Zhang,[∥,▱] Rikard Holmdahl,[∥,▱] Jan Kihlberg,[§,] Anna Linusson.[†,*]*

[1]Department of Chemistry, Umeå University, SE-901 87 Umeå, Sweden. [§]Department of Chemistry – BMC, Uppsala University, Box 576, SE-751 23 Uppsala, Sweden. [∥]Medical Inflammation Research, Department of Medical Biochemistry and Biophysics, Karolinska Institute, SE-171 77 Stockholm, Sweden. [▱]School of Pharmaceutical Science, Southern Medical University, Guangzhou, China. [‡]These authors contributed equally.

KEYWORDS. Rheumatoid arthritis, glycopeptide, DR4, T-cell, multicomponent system, MHCII, molecular dynamics, hydrogen bonding

ABSTRACT: Strategies that target multiple components are usually required for treatment of diseases originating from complex biological systems. The multicomponent system consisting of the DR4 major histocompatibility complex type II molecule, the glycopeptide CII259–273 from type II collagen, and a T-cell receptor is associated with development of rheumatoid arthritis (RA). We introduced non-native amino acids and amide bond isosteres into CII259–273 and investigated the effect on binding to DR4 and the subsequent T-cell response. Molecular dynamics simulations revealed that complexes between DR4 and derivatives of CII259–273

were highly dynamic. Signaling in the overall multicomponent system was found to depend on formation of an appropriate number of dynamic intramolecular hydrogen bonds between DR4 and CII259–273, together with the positioning of the galactose moiety of CII259–273 in the DR4 binding groove. Interestingly, the system tolerated modifications at several positions in CII259–273, indicating opportunities to use analogues to increase our understanding of how rheumatoid arthritis develops and for evaluation as vaccines to treat RA.

INTRODUCTION

Modern medicinal chemistry has mainly focused on the discovery of drugs that act at one specific target with high potency and specificity. However, this can result in a minimal impact on the biological system and failure of the drug candidate when the system is regulated through multiple components and pathways. Such, increasingly complex biological systems are now often being investigated in drug discovery and pose significant challenges for medicinal chemists. For instance, there is an increasing interest in the design of drugs that will modulate multiple targets of a biological system, *i.e.* drugs that display polypharmacology.[1,2] Alternatively, drugs targeting multicomponent systems are sought for, where the successful drug must be able to interact with, or influence, two or more macromolecular species simultaneously to obtain the desired effect..

One example of a multicomponent system is the complex formed between a T-cell receptor (TCR), a peptide and a major histocompatibility complex class II (MHCII) protein present on an antigen-presenting cell. This multicomponent system is essential for the function of the immune system, and enables it to distinguish between the body's own constituents and foreign antigens so that an immune response is initiated towards the latter. The detailed mechanisms behind the

initiation of an immune response are still not fully understood, but are known to require additional interactions between CD3- and CD4 co-receptors on the two interacting immune cells.[3] However, as for many other biological systems,[4] dynamics[5] have been suggested to play an important role for signaling in the TCR/peptide/MHCII complex and have been studied by molecular dynamics (MD) simulations in a few cases.[6,7]

Development of autoimmune diseases are assumed to originate from a failure of the immune system to differentiate between self and non-self, resulting in that an immune response is mounted towards the body's own constituents. Rheumatoid arthritis (RA) is such an autoimmune disorder that mainly affects the peripheral joints of hands and feet, leading to a chronic inflammation that destroys bone and cartilage, involving the degradation of type II collagen (CII).[8] The cause of the disease is not yet known, but it has been associated to the expression of the MHCII allelic DRA/DRB1*0101 (DR1) and DRA/DRB1*0401 (DR4) proteins.[9] Interestingly, the glycopeptide CII259–273 (**1**, Figure 1A**)**, which is derived from CII, is an MHCII-restricted self-antigen that is targeted in the immune response associated with RA.[10] When presented by MHCII DR1 or DR4 proteins on antigen presenting cells, glycopeptide **1** is recognized by T-cells isolated from patients that suffer from RA, and by T-cell hybridomas from transgenic mice that express human DR4.[10] Moreover, studies have shown that vaccination with glycopeptide **1** alone, or with MHCII complexes of **1** can reduce the progression of arthritis in collagen induced arthritis (CIA), an established mouse model for RA.[11] Vaccination with MHCII complexes of **1** was particularly effective and reduced the incidence of arthritis in CIA from >90% to <20%.[12] Most recently it was found that T cells specific for **1** escape tolerance induction in the thymus,[13] which explains why **1** is a target for the autoimmune response in CIA and also in patients with RA.[10]

Structural models suggest that glycopeptide **1** binds to the human MHCII proteins between the α_1- and β_1-helices,[7c, 14] consistent with how **1** is bound by the murine MHCII allelic protein A^q.[15] It is kept in the binding groove by a hydrogen bond network that is formed to the backbone of the peptide, but also through anchoring of Phe^{263} in the deep hydrophobic P1 pocket and Glu^{266} in the shallow P4 pocket (Figure 1B).[16] The galactosylated hydroxylysine ($GalHyl^{264}$) then protrudes from the MHC binding groove, forming a crucial contact point for the TCR,[10, 17] which will approach from above (Figure 1C). Thus, the MHCII/peptide-complex will interact with the variable domain of the TCR, forming a multicomponent system that is important for development of RA.

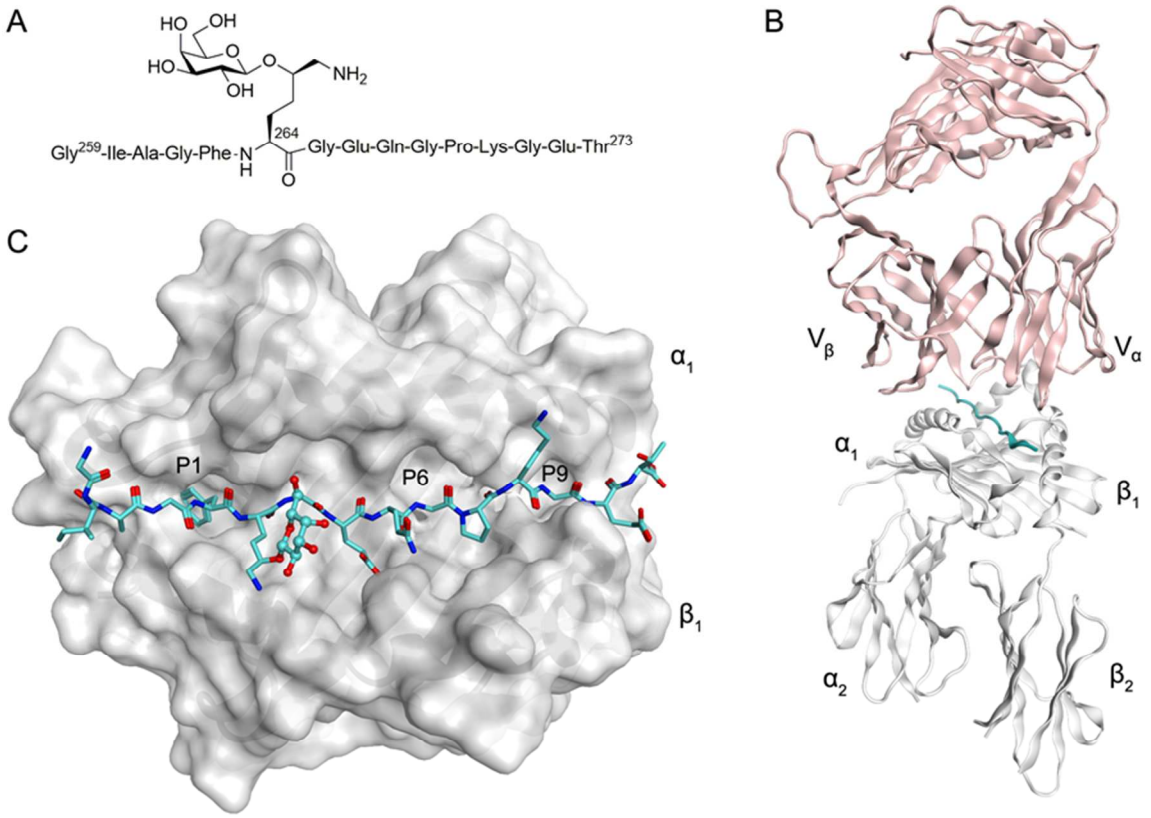


Figure 1. Structure of glycopeptide **1**, the DR4/**1** complex, and the multicomponent system consisting of the DR4/peptide/TCR complex. A) Chemical structure of glycopeptide CII259–

273, **1**. B) A three dimensional structure of the multicomponent system consisting of the DR4/peptide/TCR complex (based on PDB:1J8H[14a], where an influenza virus peptide is positioned in the binding groove). The peptide is located between the α_1- and β_1-helices of DR4, being presented to the TCR that is approaching from above. DR4 is displayed in grey, the peptide in cyan, and the TCR in pink. C) A structural model of the glycopeptide **1** bound to DR4. Phe^{263} is positioned in the P1 pocket, Gly^{268} and Gly^{271} are positioned next to the to P6 and P9 pockets, respectively. The carbohydrate moiety of $GalHyl^{264}$ is protruding out from the binding pocket (displayed in ball and stick). A surface has been added to the DR4 protein to display the binding pockets. Carbon, oxygen, and nitrogen atoms in **1** are displayed in cyan, red, and blue, respectively.

In previous studies, we have focused on the multicomponent system consisting of the murine MHCII protein A^q, glycopeptide **1** and TCRs as this system regulates development of disease in collagen induced arthritis (CIA), a validated model for RA.[7, 11a, 17] Glycopeptide **1** was modified through introduction of non-native amino acids or amide bond isosteres in different positions in order to probe molecular interactions and to increase resistance to degradation. We found that such modifications could affect the binding to A^q as well as recognition by TCRs, *i.e.* signaling in the multicomponent system. Here, we explore how chemical variations of **1** influence the binding to DR4 and the subsequent T-cell response of the multicomponent system linked to development of RA in humans. This information is important in future investigations of how tolerance may be induced. Amino acids in **1** important for binding to DR4 were varied, and amide bond isosteres of the backbone were introduced. The dynamics of the resulting complexes were studied by MD simulations, and correlated to experimental data for binding of the different glycopeptide variants of **1** to DR4 and to subsequent T-cell responses.

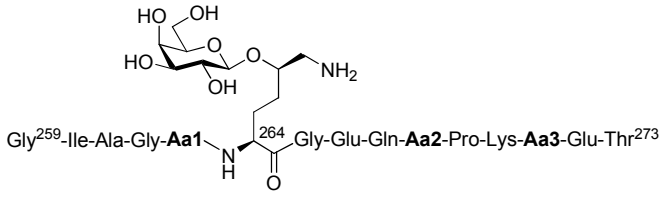
RESULTS AND DISCUSSION

Design of Modified Glycopeptides and Peptide Isosteres

Three amino acids in glycopeptide **1** (Phe^{263}, Gly^{268}, and Gly^{271}) were selected for substitution with non-native amino acids. The purpose was to investigate how the variations influence the binding affinity of glycopeptides to DR4, and investigate how that would affect the subsequent T-cell response. Phe^{263} is an important anchor residue positioned in the P1 pocket of DR4,[7c] while Gly^{268} and Gly^{271} are positioned next to the P6 and P9 pockets, respectively (Figure 1B). Visual inspection of structures of DR4 in complex with other peptides[14] showed that the latter two pockets could accommodate larger lipophilic or polar side chains, which we hypothesized would provide stronger binding to DR4.

In a previous study,[7c] it has been shown that Phe at position 263 can be substituted for non-native amino acids with maintained DR4 binding and subsequent T-cell response. Here, we decided to explore *m*-fluoro-phenylalanine (*m*FPhe), cyclohexylalanine (Cha), and Trp together with the native Phe in this position. Positions 268 and 271 were investigated by incorporation of Val or Thr and Leu or Gln, respectively, in addition to the native Gly. A D-optimal design was applied to select a smaller subset of glycopeptides for synthesis (G-efficiency of 64.6, condition number of 2.4), resulting in a final selection of 12 modified glycopeptides (**2**–**13**, Figure 2). Positions Gln^{267}–Gly^{268} and Lys^{270}–Gly^{271} were further investigated through the introduction of alkane and alkene amide bond isosteres (**14**–**16**, Figure 2). The alkane isostere was chosen as it cannot form backbone hydrogen bonds and has an increased flexibility as compared to an amide bond. Because of these differences it is suitable to probe molecular interactions in the DR4/glycopeptide/TCR system. A Z-alkene isostere became available for incorporation at

position Gln^{267}–Gly^{268} as a result of the selected synthetic route, and glycopeptide **15** was investigated as a "negative control".



Modified Glycopeptides			
ID	**Substitutions**		
	Aa1 (p263)	**Aa2 (p268)**	**Aa3 (p271)**
1	Phe	Gly	Gly
2	*m*FPhe[a]	Val	Leu
3	Cha[b]	Gly	Gln
4	Trp	Gly	Gly
5	Phe	Thr	Gly
6	Phe	Val	Gly
7	Phe	Gly	Leu
8	*m*FPhe	Gly	Gly
9	Cha	Val	Gly
10	*m*FPhe[a]	Thr	Gln
11	Cha	Thr	Leu
12	Trp	Val	Gln
13	Trp	Thr	Leu
Amide bond isosteres			
14	$Gln^{267}[CH_2CH_2]Gly^{268}$		
15	$Gln^{267}[(Z)\text{-}CH_2{=}CH_2]Gly^{268}$		
16	$Lys^{270}[CH_2CH_2]Gly^{271}$		

Figure 2. Structure of native CII259–273 (**1**), modified glycopeptides **2**–**13** with substituted amino acids at three positions, and isosteric glycopeptides **14**–**16**. Apart from the amide bond isostere, **14**–**16** had identical sequences as native **1**. [a]*m*FPhe = *m-f*luoro-phenylalanine. [b]Cha = cyclohexylalanine.

Synthesis

The designed glycopeptides **2**–**13** (Figure 2) were synthesized by Fmoc-based solid-phase peptide synthesis[18] using commercially available amino acids and protected β-D-galactosylated hydroxylysine, prepared as reported previously.[19] After cleavage from the solid support and purification by reversed-phase HPLC the galactose moiety was deacetylated by treatment with methanolic sodium methoxide. Final purification by reversed-phase HPLC then afforded **2**–**13** in 6–52% overall yield based on the capacity of the resin. All glycopeptides had a purity of >95% according to analytical reversed-phase HPLC, and their structures were confirmed by MALDI-TOF mass spectrometry (Table 3).

Synthesis of glycopeptides **14** and **16** required access to alkane isosteres **17** as Fmoc protected dipeptide building blocks, ready to be incorporated at the Gln^{267}–Gly^{268} and Lys^{270}–Gly^{271} positions of **1**. A Wittig reaction between Gln and Lys derived aldehydes **19** and 1-(2-triphenylphosphoniummethyl)-4-methyl-2,6,7-trioxa-bicyclo-[2,2,2]-oxetane bromide (**20**)[20] constitutes the key step in synthesis of these isosteres (Figure 3).

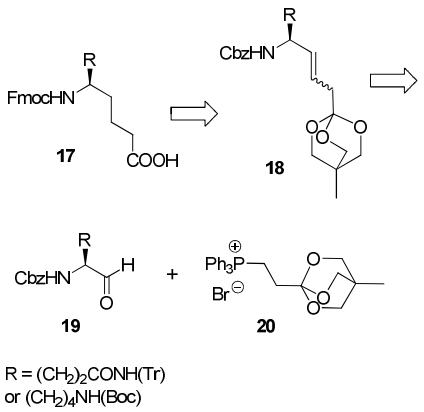
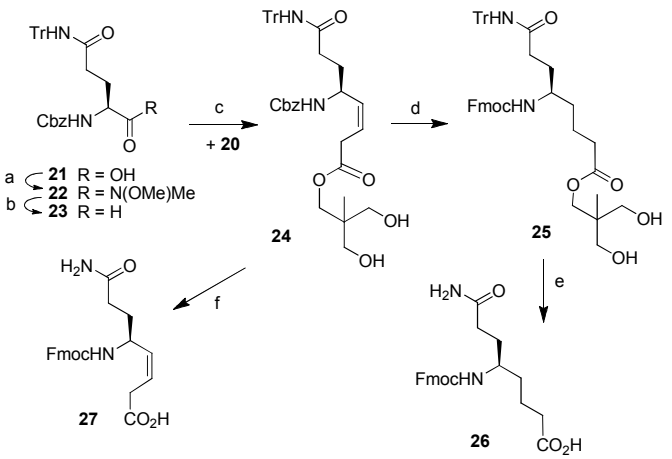


Figure 3. Retrosynthetic analysis of Gln^{267}–Gly^{268} and Lys^{270}–Gly^{271} dipeptide isosteres **17**.

Preparation of the Gln^{267}–Gly^{268} isostere began by reaction of protected glutamine **21** with *N*,*O*-dimethylhydroxylamine, using HATU as promoter and TEA as a base, to give Weinreb amide **22** in 92% yield (Scheme 1).[21] Compound **22** was then reduced to aldehyde **23** using DIBAL-H at -78 °C in dry THF (90% yield).[22] Triphenylphosphonium bromide (**20**) was reacted with KHMDS in dry THF to generate the corresponding ylide, which was used in a Wittig reaction with aldehyde **23**.[23] As aldehyde **23** and the alkene obtained in the Wittig reaction could not be separated by flash column chromatography on silica gel, reversed-phase HPLC was used for purification. Interestingly, the orthoester functionality underwent hydrolysis under the acidic conditions of the purification resulting in isolation of alkene **24** as an ester in 62% yield and with 100% *Z*-alkene selectivity. Changing the base to NaHMDS, or LiHMDS did not change the selectivity of the Wittig reaction and no traces of *E*-isomer could be detected in the crude products by NMR spectroscopy. As (*Z*)-isomer **24** can be used for synthesis of the desired alkane isostere it was subjected to hydrogenation over Pd/C in order to obtain simultaneous reduction of the alkene and deprotection of the Cbz group. Subsequent Fmoc protection using *N*-(9-fluorenylmethoxycarbonyl)oxysuccinimide (Fmoc-OSu) then gave **25** in 68% yield over two steps.[7a] Finally, the trityl and orthoester moieties were removed by reacting **25** with acetyl chloride in a mixture of *t*-BuOH and water to give the desired Gln^{267}–Gly^{268} alkane isostere **26** in 83% yield. Since Fmoc protected (*Z*)-isostere **27** can be made in two steps by acid catalyzed hydrolysis of **24**, followed by Fmoc protection, it was also prepared for incorporation into glycopeptide **1**.

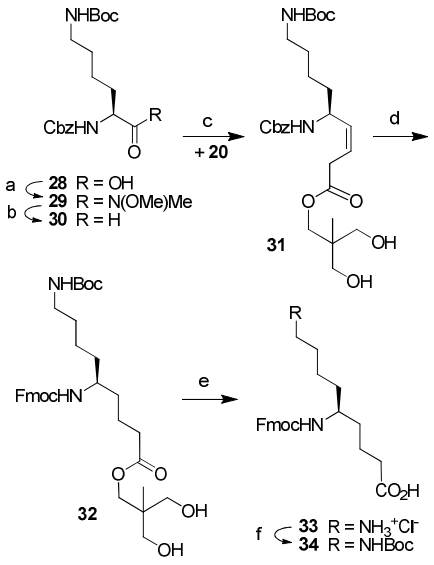
Scheme 1. Synthesis of Gln^{267}–Gly^{268} dipeptide isosteres **26** and **27**[a]



[a]Reagents and conditions: (a) TEA, HATU, $Cl^-NH_3^+OMe$, 3 h, 92%; (b) DIBAL-H, THF, -78 °C, 4 h, 90%; (c) **20**, KHMDS, THF -30 °C to rt to -30 °C again (1.5 h), then **23** was added (-30 °C for 1 h, then 1 h at rt.), 62%; (d) (i) Pd/C, H_2, MeOH, (ii) Fmoc-OSu, 10% Na_2CO_3, H_2O:acetonitrile (1:1), 68% over two steps; (e) *t*-BuOH, H_2O, acetyl chloride, 0 °C to rt. 24 h, 83%; (f) (i) *t*-BuOH, H_2O, acetyl chloride, 0 °C to rt. 24 h, (ii) Fmoc-OSu, 10% Na_2CO_3 (H_2O:acetonitrile 1:1), 82% over two steps.

Lys^{270}–Gly^{271} alkane isostere **34** was synthesized using a similar sequence as used for the two Gln^{267}–Gly^{268} isosteres (Scheme 2). Protected lysine **28** was first converted into aldehyde **30**, *via* the corresponding Weinreb amide in 59% yield over two steps. A Wittig reaction of **30** with the ylid obtained from **20**, followed by purification by reversed-phase HPLC and concomitant acid-catalyzed rearrangement of the orthoester, gave alkene **31** in 84% yield and 100% (*Z*)-selectivity. Reduction of the double bond and cleavage of the Cbz group of **31** by hydrogenation over Pd/C and subsequent Fmoc protection of the α-amino group gave **32** in 98% yield. Acid catalyzed hydrolysis of the ester of **32** also resulted in cleavage of the Boc group at the α-amino group (**33**, 79%), which was therefore reprotected to give Lys^{270}–Gly^{271} alkane isostere **34** in 55% yield.[24]

Scheme 2. Synthesis of Lys^{270}–Gly^{271} dipeptide isostere **34**[a]



[a]Reagents and conditions: (a) TEA, HATU, $Cl^-NH_3^+OMe$, 3 h, 81%; (b) DIBAL-H, THF, -78 °C, 4 h, 73%; (c) **20**, KHMDS, THF -30 °C to rt (1.5 h), then **30** was added (-30 °C for 1 h then 2 h at rt.), 84%; (d) (i) Pd/C, H_2, MeOH, 76% (ii) Fmoc-OSu, 10% Na_2CO_3, H_2O:acetonitrile (1:1), 98%; (e) *t*-BuOH, H_2O, acetyl chloride, 0 °C to rt, 24 h, 79%; (f) $(Boc)_2O$, $NaHCO_3$, H_2O:dioxane (3:5), 55%.

The isosteric amino acids **26**, **27** and **34** were utilized in Fmoc-based solid-phase glycopeptide synthesis using HATU as coupling agent and DIPEA as base. After completion of the solid-phase synthesis, cleavage from the resin, purification by reversed-phase HPLC, deacetylation and purification again by HPLC, glycopeptides **14**, **15** and **16** were obtained in 11–23% yield based on the capacity of the resin (Table 3). The three glycopeptides had a purity of >95% according to analytical reversed-phase HPLC and their structures were confirmed by MALDI-TOF mass spectrometry (Table 3).

Binding to the DR4 Protein and T-cell Receptor Recognition

Glycopeptides **1**–**13** and glycopeptide isosteres **14**–**16** were evaluated for their binding to DR4 and subsequent T-cell response using a competitive ELISA and a cell-based assay, respectively. All of the side-chain modified glycopeptides **2**–**13** bound to DR4, while none of the glycopeptide isosteres **14**–**16** displayed any binding (Figure 4). Among **2**–**13**, six glycopeptides (**5**, **6** and **10**–**13**) bound somewhat better to DR4, while **4** and **7** bound somewhat weaker.

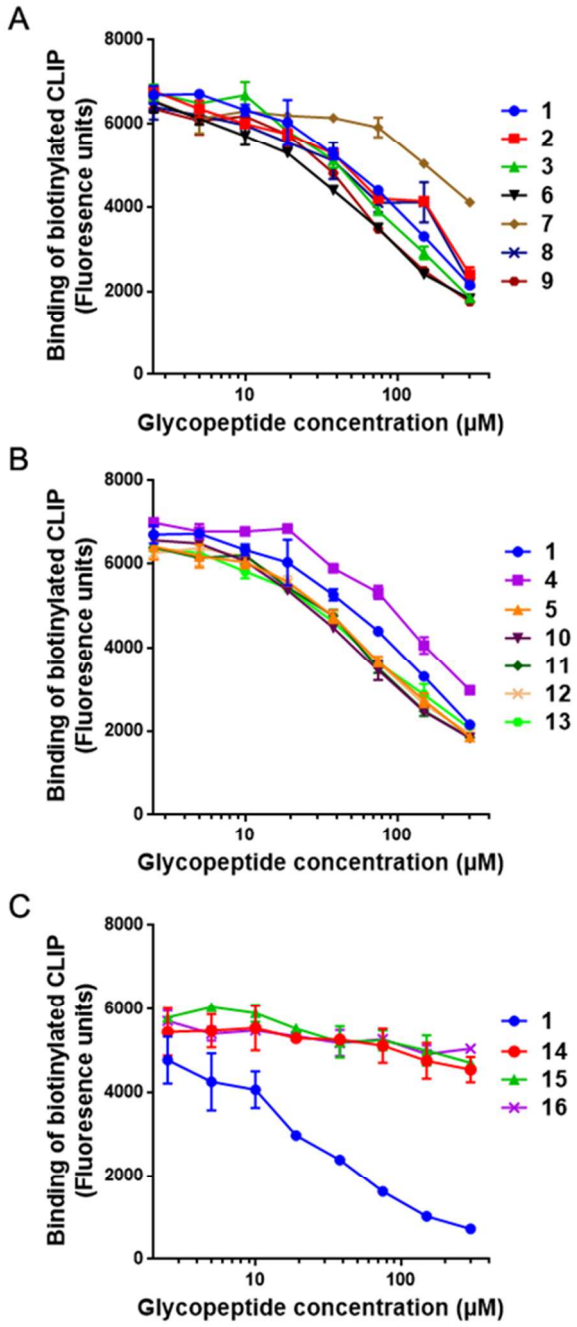


Figure 4. Binding of glycopeptide **1**, modified glycopeptides **2**–**13**, and glycopeptide isosteres **14**–**16**, to DR4. Increasing concentrations of **1**–**16** and a fixed concentration of biotinylated CLIP tracer were incubated with DR4. The amount of CLIP tracer bound to DR4 was thereafter

measured in a time-resolved fluoroimmunoassay. Each point represents the average of duplicates and error bars show ± one standard deviation. A) Glycopeptide **1** compared to **2**–**3** and **6**–**9**, which are the glycopeptides that elicit T-cell responses when bound to DR4. B) Glycopeptide **1** compared to **4**–**5** and **10**–**13**, which did not elicit T-cell responses in complexes with DR4. C) Glycopeptide **1** compared to isosteres **14**–**16**.

Stimulation of DR4 restricted T-cell hybridoma mDR1.1 was determined through measurements of IL-2 secretion. Half of the modified glycopeptides, i.e. **2**–**3** and **6**–**9**, elicited a T-cell response (Figure 4A and Figure 5A). For glycopeptides **3** and **8**, the response was similar to the native glycopeptide (**1**), while **6**, **7**, and **9** elicited a slightly lower T-cell response, and glycopeptide **2** had a significantly lower response than **1**. The other glycopeptides (**4**–**5**, **10**–**13**, Figure 4B) did not stimulate any IL-2 secretion at the concentrations investigated (Figure 5A). Interestingly, the binding of the glycopeptides to DR4 did not correlate well with the T-cell response (*cf.* Figure 4A for glycopeptides that did elicit T-cell responses, and Figure 4B for those that did not). For instance, glycopeptides that bound weaker (**7**), somewhat better (**6**) or equally well (**3**, **8** and **9**, Figure 4A) to DR4 as **1** all elicited a similar or slightly lower T-cell response as **1** (Figure 5A). In contrast, several glycopeptides that bound somewhat better than **1** to DR4, *i.e.* **5** and **10**–**13** (Figure 4B), where among those that did not elicit any T-cell response at all (Figure 5A). As expected, no IL-2 secretion was detected at the tested concentrations for glycopeptide isostere **14**–**16** that showed no binding to DR4 (Figure 5B).

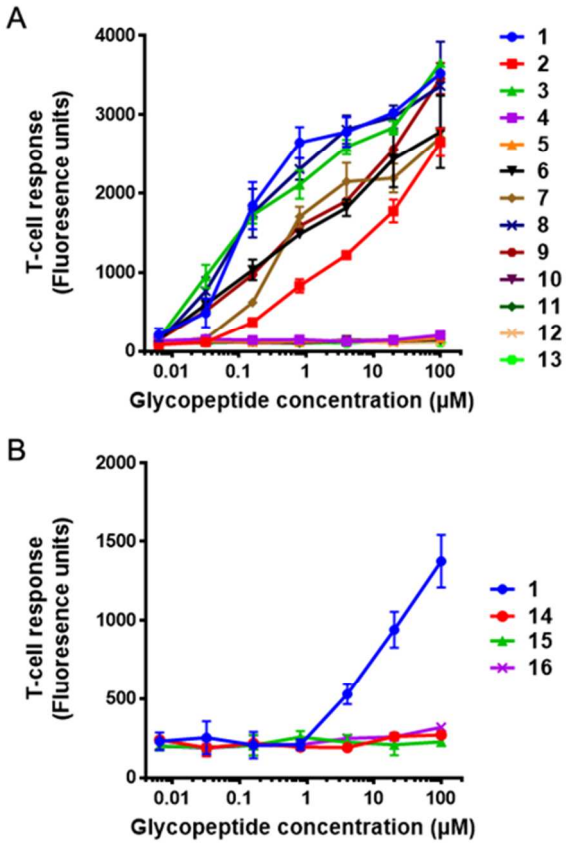


Figure 5. Recognition of glycopeptide **1**, modified glycopeptides **2**–**13** (A), and glycopeptide isosteres **14**–**16** (B), by the DR4 restricted T-cell hybridoma mDR1.1, after incubation with antigen presenting spleenocytes and increasing concentration of the glycopeptides. If the DR4-glycopeptide complex is recognized by the T-cell hybridoma, secretion of IL-2 into the supernatant will occur, which is quantified by an ELISA assay. Each point represents the average of duplicates and error bars show ± one standard deviation.

Structure-Activity Relationships for DR4 Binding and TCR Recognition

In general, the P1 pocket of DR4 was able to accommodate a variation of amino acids at position 263 of the glycopeptides. Both the somewhat larger *m*FPhe and Trp, and the non-aromatic Cha were tolerated by DR4. In a previous study,[7c] we found that *m*-methylphenylalanine, *p*-

fluorophenylalanine, and tyrosine were tolerated in this position as well. The exchange of Phe for *m*FPhe led to similar binding to DR4 compared to the native **1** (*cf.* **8** and **1**), while replacement of Phe for Trp (*cf.* **4** and **1**) or Cha (*cf.* **9** and **6**) resulted in a somewhat weaker binding. The Gly at positions 268 and 271 positioned next to the pockets P6 and P9 of DR4, respectively, were successfully exchanged for amino acids with larger side-chains with similar or increased binding to DR4. The change of Gly^{268} to Thr^{268} or Val^{268} somewhat increased the binding to DR4 (**5** and **6**), while introducing Gln at position 271 appeared to result in similar binding affinity. The exchange of Gly^{271} for Leu resulted in a decreased binding to DR4 (**7**).

The structure-activity relationship (SAR) for the T-cell response differed dramatically from the SAR of DR4 binding, as mentioned above. Two of the modifications abolished the T-cell response, the exchange of Phe^{263} for Trp^{263} and Gly^{268} for Thr^{268}; glycopeptides with these modifications (**4**, **5** and **10–13**) did not result in any IL-2 secretion despite that their binding to DR4 were similar to other glycopeptides that were recognized by the TCR (Figures 4B and 5A). The other seven glycopeptides with non-native amino acids in positions 263, 268 and/or 271 did result in T-cell responses (**2**, **3** and **6–9**, Figure 5A), although no clear SAR could be distinguished among these.

In comparison, the glycopeptide isosteres **14–16** displayed neither binding to DR4, nor a T-cell response. The fact that they did not bind to DR4 probably originates from the increased flexibility of **14** and **16**, and/or loss of hydrogen bonding abilities. Glycopeptide **15** has a substantially different conformational landscape compared to **1**, which explains its lack of binding to DR4.

In summary, glycopeptides **2–13** that had one, two or three non-native amino acids displayed only minor variations in their binding to DR4 and there was no clear correlation between DR4

binding and their T-cell responses. This enigma was further illustrated by contrasting the glycopeptides that had Trp at position 263 to those that had Thr at position 268. The former all bound weaker to DR4 than **1**, while the latter bound somewhat stronger. However, both categories of glycopeptides failed to elicit a T-cell response.

Dynamics of Glycopeptides in Complex with DR4

As the SAR analysis failed to explain what governs the response of the multicomponent system consisting of DR4-expressing antigen-presenting cells, glycopeptide **1** and T-cell hybridoma mDR1.1, we decided to investigate the dynamics of the complex between DR4 and some of the glycopeptides. Based on their T-cell responses, four of the modified glycopeptides were selected for MD simulations, just as the native **1**. Glycopeptides **2** and **3** were included as they both bound to DR4 *and* elicited a T-cell response. Glycopeptide **2** displayed the lowest T-cell response of the peptides that were recognized, *i.e.* a response approximately two orders of magnitude weaker than native **1**, while **3** displayed a response comparable to **1**. Neither of the modified glycopeptides with the non-native amino acids Trp^{263} or Thr^{268} displayed any T-cell response, thus **4** and **5** were included as examples of glycopeptides that bound to DR4 but *did not* give rise to IL-2 secretion. It should be noted that **2** and **3** have non-native amino acids at all three or two of the modified positions, respectively, while **4** and **5** only had single replacements at positions 263 and 268.

Modified glycopeptides **2–5** were modeled *in silico* into DR4, based on four previously published X-ray crystal structures of DR4 in complex with other peptides (see experimental section for further details).[14a, 14c, 14d, 25] The resulting static models displayed very similar binding modes as for the native **1** and could therefore not explain the differences observed in the

biological assays. The dynamics of the complexes were instead studied over time using MD simulations where the conformational space was sampled by six parallel 60 ns MD simulations for each complex. In general, according to the root-mean-square deviation (RMSD, Figure S62), all MD simulations were converged after 30 ns. The fast stabilizations of the complexes are consistent with other MD studies of MHCII/peptide systems.[26] We focused our analyses of the MD simulations on the binding mode and dynamics of the glycopeptide in the binding groove, and on the hydrogen bond network formed between the peptide backbone and DR4.

For the six MD simulations of the glycopeptide **1** in complex with DR4, the DR4 protein was stable throughout the simulations (Figures S62–S66). Phe^{263} in the glycopeptide is positioned in the P1 pocket (Figure S70), firmly anchoring the residues of the glycopeptide adjacent to galactosylated hydroxylysine^{264} (positions 262–265) between the α_1- and β_1-helices, while the C-terminus (positions 267–273) was more mobile (Figure 6A and S67–69).

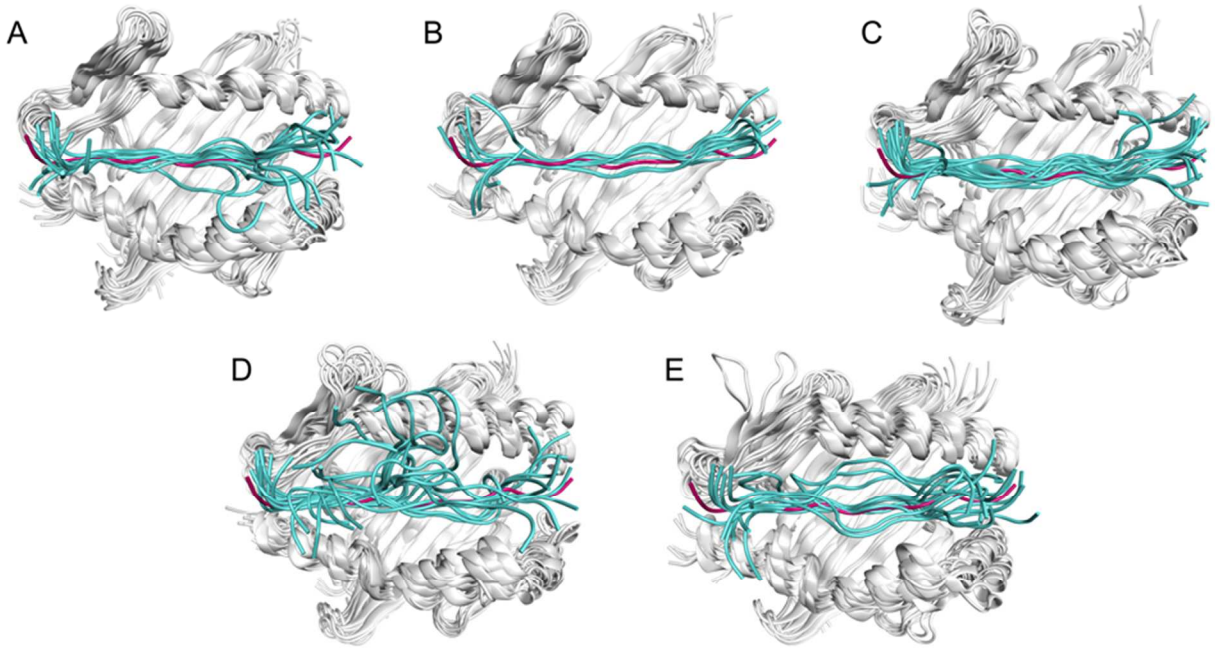


Figure 6. Snapshots from the MD simulations displaying the dynamics of the DR4 systems. A) Glycopeptide **1**, B) glycopeptide **2**, C) glycopeptide **3**, D) glycopeptide **4**, E) glycopeptide **5**, all in complex with DR4. For each simulation trajectory a cluster analysis of the glycopeptide was performed based on snapshots over the last 30 ns. One representative glycopeptide conformation for each cluster with at least 40 members is displayed. The complexes are displayed with the proteins in gray and the glycopeptides in green. The conformation of the glycopeptide **1** from the homology model is displayed in magenta, as a reference.

The MD simulations of DR4 in complex with modified glycopeptides **2** and **3** displayed that **2** was more firmly anchored in the binding groove than native **1**, while the behavior of **3** was very similar to that of **1** (Figure 6A–C and S67–69). For **2**, this is most likely due to the larger side chains of the non-native amino acids, positioned in the P1, P6 and P9 pockets of DR4, which led to decreased mobility of the peptide in accordance with the design. Interestingly, the reduced mobility of glycopeptide **2** in the DR4 binding groove also resulted in the lowest T-cell response of all glycopeptides that were recognized, while **3** that displayed a similar mobility as **1** gave an almost identical T-cell response. This suggests that rigidification of the peptide epitope prevents it from adapting to the TCR and eliciting a response upon formation of the multicomponent system, while modified peptide epitopes having similar flexibility as **1** elicit identical T-cell responses.

The complexes of DR4 and glycopeptides **4** and **5**, respectively, had a different dynamic pattern than **1**–**3** (Figure 6D–E). Glycopeptide **4**, with Trp^{263} instead of Phe^{263}, displayed a much higher degree of mobility in the binding groove than **1**, and large variations were observed for the multiple simulation trajectories (Figure S67–69). The high mobility did not only include the C-terminus of the peptide, but now also the N-terminus, which may explain the somewhat lower

binding affinity of **4**. In addition, the DR4 protein also displayed an increased mobility, apparent when analyzing the α_1- and β_1 helices (Figures S64 and S65). Glycopeptide **5** showed a mobility that was higher than **1** (Figures S67–S69) primarily in the central and C-terminal part of the epitope (*cf.* Figure 6A and Figure 6E). Interestingly, the introduction of Thr^{268} affected the dynamics of the $\alpha_1\beta_1$ β-sheet forming the bottom of the DR4 binding groove (Figure S66). Overall, the binding strength of **5** to DR4 was not affected by the changed dynamic pattern, but it appears to have led to a more flexible epitope being presented to the TCR in the center of the glycopeptide, which may explain why **5** did not give a T-cell response.

Hydrogen Bond Network between the Glycopeptides and DR4

Throughout the MD simulations, an extended hydrogen bond network was formed between the residues of DR4 and the backbone of the native glycopeptide **1**, keeping it anchored in the binding groove. During the last 30 ns of the MD simulations there were on average 13.5 hydrogen bonds formed between the backbone of **1** and the residues of DR4. Out of these hydrogen bonds, on average 7 had occupancy of at least 40% (Table 1). Notably, four of the highest occupancy hydrogen bonds (>60%) were formed between DR4 and the backbone of residues 262–264 of **1,** which includes Phe^{263} that anchors the glycopeptide in the P1 pocket of DR4 and Gal–Hyl^{264} known to be essential for contacting the TCR (Table 2).[10, 17] This finding suggests that the exact positioning of these two residues is of high importance for DR4 binding and T-cell response, *i.e.* for signaling in the multicomponent system.

Table 1. The number of hydrogen bonds formed between residues of DR4 and the backbone of glycopeptides **1–5**, during the last 30 ns of the MD simulations

Glyco-peptide	Occupancy[a] above 40%		Occupancy[a] above 10%	
	Mean	SD	Mean	SD
1	7.0	0.8	13.5	2.2
2	9.3	1.8	13.5	3.1
3	7.7	1.9	13.3	3.0
4	4.8	3.5	9.3	3.9
5	6.0	2.9	11.2	3.2

[a]The hydrogen bond occupancy has been calculated using the hydrogen bond extension tool in VMD[27], with a donor-H-acceptor angle of minimum 140° and a donor-acceptor length of maximum 3.4 Å.

Table 2. Occupancy of the four most populated hydrogen bonds formed between the backbone of glycopeptides **1**–**5** and residues of DR4 over the last 30 ns of the MD simulations

Glycopeptide	Gly^{262}(a)[a]–βHis^{81}(s)[b]		Phe^{263}(d)[c]–αSer^{53}(m)[d]		GalHyl^{264}(a)–βAsn^{82}(s)		GalHyl^{264}(d)–βAsn^{82}(s)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1	63.0	31.6	79.5	18.6	88.7	7.1	89.5	10.4
2	74.3	7.1	70.0	32.1	95.3	3.7	96.3	2.4
3	65.7	24.2	77.8	20.6	64.8	35.8	67.7	35.2
4	37.2	36.0	50.8	41.2	61.8	44.2	61.2	43.4
5	34.7	37.5	40.8	35.1	58.2	41.9	62.5	44.3

The hydrogen bond occupancy has been calculated using the hydrogen bond extension tool in VMD, with a donor-H-acceptor angle of minimum 140° and a donor-acceptor length of maximum 3.4 Å. [a]a = hydrogen bond acceptor. [b]s = sidechain in DR4. [c]d = hydrogen bond donor. [d]m = main chain in DR4.

The glycopeptides that were recognized by the T-cell receptor (**1**–**3**) had a larger number of average hydrogen bonds present over time compared to **4** and **5** that did not elicit IL-2 secretion (Tables 1 and 2). The least mobile glycopeptide in the DR4 binding groove (**2**) had the largest average number of hydrogen bonds with an occupancy >40%, while **3** was more similar to **1**. In contrast, the glycopeptides that were not recognized by the TCR (**4** and **5**) had lower hydrogen bond occupancies, in particular of the four important hydrogen bonds (Table 2). However, **5** had higher hydrogen bond occupancies than **4** (Table 1), in line with the lower flexibility displayed by **5** in the MD simulations.

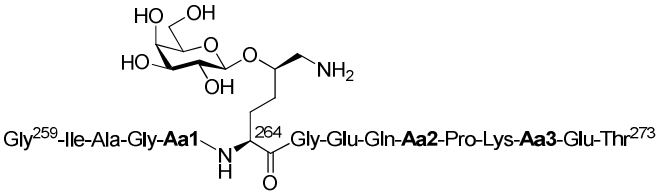
For glycopeptide **4**, the exchange of Phe^{263} to Trp^{263} in the P1 pocket of DR4 led to a slight repositioning of the backbone in the binding groove. This weakened the four important hydrogen bonds (Table 2), which increased the dynamics of the complex and allowed Trp^{263} to leave the P1 pocket occasionally (Figure S64 and S68). The modified glycopeptide **5** also had lower occupancy of the four key hydrogen bonds to DR4, despite having the native Phe^{263} in the P1 pocket. Consequently, the reduced occupation of the key hydrogen bonds at positions 262–264 could be linked to the exchange of Gly^{268} for Thr^{268}, positioned in the P6 pocket of DR4. New hydrogen bonds were formed between the hydroxyl group of Thr^{268} of **5** and residues in P6 pocket of DR4 (αAsn^{62} and/or αAsp^{66}; Table S3), which altered the conformational landscape of the complex and led to a changed hydrogen bonding network. The resulting altered dynamic pattern of the complex of **5** with DR4 as compared to **1**, apparently had a tremendous effect on the T-cell response. It is also possible that steric hindrance between Thr^{268} in **5** and the TCR contribute to the lack of response to **5**.

Challenging the Multicomponent System with Additional Glycopeptides

Based on the MD simulations, we conclude that a retained biological response to **1** is highly dependent on the dynamic pattern of the DR4/glycopeptide complexes. Appropriate dynamics could be traced to a large number of hydrogen bonds of varying occupancy between the glycopeptide and the binding groove of DR4, including four of high occupancy that involved residue Phe^{263} and adjacent residues of the glycopeptides. The results indicated that a reduced mobility of the glycopeptides, originating from an increased number of moderate (>40%) occupancy hydrogen bonds, might lead to reduced T-cell responses, just as an increased mobility. In order to develop our model, and to further investigate the system, we designed three additional glycopeptides (**35**, **36**, and **37**) that were modified at the same positions as **2**–**13** but had no overlap with the previously selected non-native amino acids. All three positions were modified in all three glycopeptides and the selection of amino acids was on purpose made to challenge the limits of our understanding of this multicomponent system.

From the results obtained with **1**–**13**, it was clear that Phe^{263} in the glycopeptides could be replaced by larger amino acids with maintained binding to DR4. However, the indole of Trp^{263} disrupted the dynamic pattern of the complex, which led to loss of the T-cell response. Phe^{263} was now replaced by *m*-bromo-phenylalanine or benzothiophenealanine (Figure 7) to explore larger lipophilic amino acids, with the latter being equal in size to Trp. At position 268, a replacement of Gly for Val was tolerated with maintained DR4 binding and T-cell response; in the new set Gly^{268} was replaced by cyclopropylalanine in order to increase both size and lipophilicity. The third position, Gly^{271}, tolerated both larger lipophilic (Leu^{271}) and larger polar (Gln^{271}) amino acids. To investigate the limits of this position further we incorporated pyridine or His, both of which provide increased non-polar and polar surfaces for interactions with DR4

and the TCR. This resulted in glycopeptides **35**–**37** that contain significantly different side chains at all three positions compared to native **1** and the modified glycopeptides **2**–**13**.



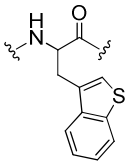
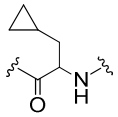
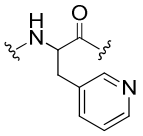
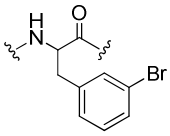
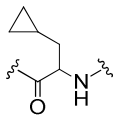
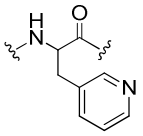
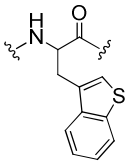
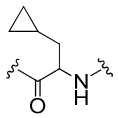
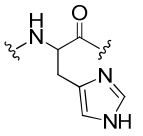
ID	Substitutions		
	Aa1 (p263)	Aa2 (p268)	Aa3 (p271)
35			
36			
37			

Figure 7. Structure of native CII259–273 (**1**) and glycopeptides **35**–**37** substituted with non-native amino acids at three positions.

The dynamics of the three complexes of DR4 and **35**–**37**, respectively, were investigated using MD simulations in an attempt to predict their DR4 binding and subsequent T-cell response before synthesis and biological evaluation. During the MD simulations of the DR4/**35** complex, both benzothiophenealanine^{263} and cyklopropylalanine^{268} remained in their respective DR4 pockets throughout the MD simulations, while pyridine^{271} occasionally dissociated from P9. Thus, the glycopeptide displayed a low mobility in the binding groove, similar to **2** in the DR4/**2** complex (Figures S67–S69). The replacement of benzothiophenealanine^{263} with *m*-bromo-phenylalanine^{263} to give glycopeptide **36** increased the dynamics of the glycopeptide in the

complex with DR4. The positioning of *m*-bromo-phenylalanine^{263} in P1 resulted in a decrease in the hydrogen bond occupancy between GalHyl^{264}(a) of the glycopeptide and βAsn^{82}(s) of DR4, and that the cyclopropylalanine^{268} of **36** dissociated occasionally from the P6 pocket (Figure 8 and Table S2). Finally, during the simulations of the DR4/**37** complex, the benzothiophenealanine^{263} was firmly anchored in the P1 pocket, while both cyklopropylalanine^{268} and His^{271} left their respective pockets during the simulations, resulting in a highly mobile C-terminus and a somewhat larger overall mobility than displayed by **1** (Figure 8, S67 and S68). In addition, the β_1-helix of DR4 occasionally unfolded around the C-terminus of **37** during the simulations (Figure 8 and S65).

The hydrogen bonding networks formed between DR4 and the backbone of glycopeptides **35**–**37**, respectively, differed somewhat from the DR4/**1** complex. Although the average number of hydrogen bonds between DR4 and **35**–**37**, respectively, having occupancy higher than 40%, were similar to DR4/**1**, a higher variability was observed between the different simulated trajectories compared to native system. Furthermore, the total number of hydrogen bonds above 10% occupancy was somewhat lower compared to the DR4/**1** complex (Table S1). This reflects a slight overall alteration in the dynamic patterns of the complexes formed by DR4 and **35**–**37**, respectively, compared to the native system. Still, the four essential hydrogen bonds were formed in high occupancy (>60%) for all new complexes, except for **36** as mentioned above (Table S2). Based on the MD simulations, we concluded that glycopeptide **35**–**37** should all bind to DR4, and that **36** and **37** are likely to elicit a T-cell response. The greater rigidity of the DR4/**35** complex, which resembles that of the DR4/**2** complex, should result in a low T-cell response.

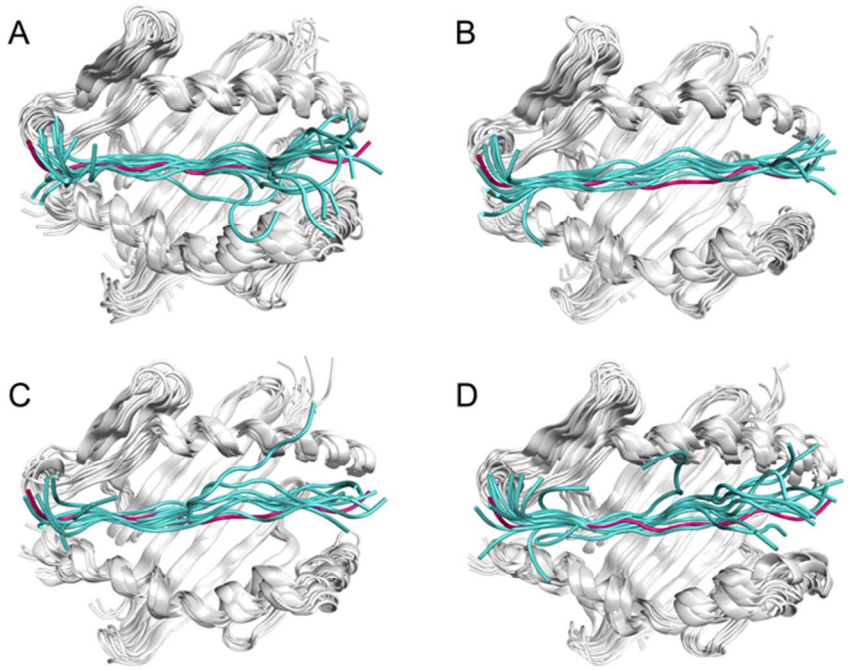


Figure 8. Snapshots from the MD simulations displaying the dynamics of the DR4 systems. A) Glycopeptide **1**, B) glycopeptide **35**, C) glycopeptide **36** and D) glycopeptide **37**. For each MD simulation a cluster analysis of the glycopeptides has been performed based on snapshots over the last 30 ns. One representative glycopeptide conformation for each cluster with at least 40 members is displayed. The tertiary structure is displayed with the DR4 protein in gray and the glycopeptide in yellow. The conformation of the glycopeptide **1** from the prepared homology model is displayed in magenta as a reference.

Determination of DR4 binding and TCR recognition revealed that **35**–**37** all bound to DR4 with similar strength as **1** (Figure 9A), and that both **36** and **37** gave a T-cell response, although decreased compared to **1**, while the rigid DR4/**35** complex did not elicit a significant response at the tested concentrations (Figure 9). These data agree well with our predictions based on the MD simulations of complexes between DR4 and **35**–**37**. The elicited T-cell response suggests that

substantial chemical modifications can be made to the glycopeptides, while still being recognized by the TCRs. However, the lower T-cell response towards DR4 in complex with **36** and **37**, respectively, compared to DR4/**1**, indicates that the outer boarder of chemical alterations has been reached.

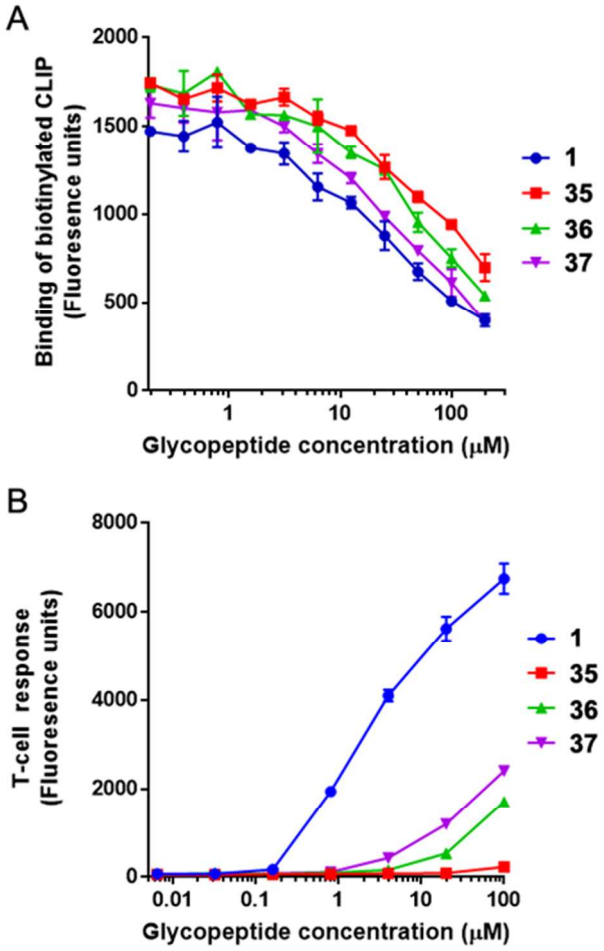


Figure 9. Binding to DR4 and TCR recognition for glycopeptide **1** compared to **35**–**37**. A) Binding of glycopeptide **1** to DR4 compared to modified glycopeptides **35**–**37**. Increasing concentrations of **1** and **35**–**37** and a fixed concentration of biotinylated CLIP tracer were incubated with DR4. The amount of CLIP tracer bound to DR4 was thereafter measured in a time-resolved fluoroimmunoassay. Each point represents the average of duplicates and error bars

show ± one standard deviation. B) TCR recognition of glycopeptide **1** and modified glycopeptides **35**–**37** by the DR4 restricted T-cell hybridoma mDR1.1, after incubation with antigen presenting spleenocytes and increasing concentration of the glycopeptides. If the DR4-glycopeptide complex is recognized by the T-cell hybridoma, secretion of IL-2 into the supernatant will occur, which is quantified by a sandwich ELISA. Each point represents the average of duplicates and error bars show ± one standard deviation.

Similarities and Differences between the Human and Murine the MHC/glycopeptide/TCR Multicomponent Systems

In CIA, the glycopeptide CII259–273 (**1**) is presented to TCRs by the murine MHCII molecule A^q.[24] Just as for presentation of **1** to T-cells by DR4, signaling in the A^q/glycopeptide/TCR multicomponent system has been linked to development of disease. Although both MHCII molecules bind the same glycopeptide, the epitope is shifted with Ile^{260} binding in the P1 pocket of A^q,[15] while Phe^{263} fills this pocket in DR4.[28] As a consequence $GalHyl^{264}$, which has been found to be a crucial residue for contacting the TCR,[10, 17] is shifted from the center of the binding grove in A^q towards one of the ends of the DR4 groove. The results presented in this manuscript now, for the first time allow us to compare signaling in the two related multicomponent systems to each other.

The two analogous multicomponent systems displayed both similarities and important differences. Just as for the complexes of glycopeptides with DR4, reported herein, static models were unable to explain experimentally observed differences in glycopeptide binding to A^q and subsequent T-cell responses. Instead, insight into the dynamics of the systems was required to understand their signaling.[7a, 7b] Binding of **1** to both MHCII molecules relied on a network of

hydrogen bonds, that were in dynamic equilibrium so that just under 50% of the total number was present at any single point in time. In addition, hydrogen bond networks in both systems appear to be crucial also for eliciting a T-cell response; *i.e.* the removal or weakening of even just one important hydrogen bond can alter the response from the entire multicomponent system.[7a, 7b]

Interestingly, the most populated hydrogen bonds in both MHCII/glycopeptide complexes are formed between the N-terminal residues of **1** when bound in the MCHII groove. However, the two complexes differ regarding the number and position of residues involved in formation of this core of strong hydrogen bonds. For DR4 only three residues (262–264) in **1** participate in the core, as compared to the A^q complex that involves an additional three residues (259–264) of **1**.[7a] Both cores contain $GalHyl^{264}$, the residue crucial for contacting the TCR. This strengthens our hypothesis that exact positioning of the galactose moiety is crucial for eliciting a T-cell response. Another difference is that the opportunities for modification of **1** with maintained MHCII binding affinity is larger for the DR4 complexes than for A^q. Thus, out of 20 glycopeptides with modifications of Ile^{260} and Phe^{263} in **1**, which bind in the P1 and P4 pockets of A^q, only two had a similar affinity as native **1**.[7c] When varying the amino acids of **1** that bind in the P1, P6 and P9 pockets of DR4 all of the synthesized glycopeptides bound to DR4, and most bound with similar affinity as **1**. In addition, several of the analogues of **1** were also able to elicit a significant T-cell response that ranged from weak to equally strong as that generated by **1**. This can be expected to provide larger opportunities to make novel glycopeptides with maintained and modified T-cell response for further investigations of signaling in the DR4 multicomponent system as compared to the A^q system.

CONCLUSIONS

We have investigated the effect on the MHCII/glycopeptide/TCR multicomponent system upon modification of the glycopeptide CII259–273 (**1**). Signaling in the multicomponent system relies on that glycopeptide **1**, or analogues, binds to DR4 and that the resulting complex is recognized by the TCR. Unexpectedly, the functional multicomponent system was compatible with significant structural variation at Phe^{263} of **1**, the key anchor residue in DR4, including replacement by a benzothiophenealanine moiety. It is equally surprising that major variations at the P6 and P9 positions of **1** were also allowed, e.g. substitutions of Gly^{268} with Val and cyclopropylalanine, and of Gly^{271} with Leu or Gln. It is possible that the rather large variation allowed at these positions depends on that the T-cell hybridoma investigated herein is highly specific for the $GalHyl^{264}$ moiety in the P2 position of **1**.

Static structural models of the multicomponent system were unable to explain why some analogues of **1** elicited a T-cell response and others failed to do so, in spite of displaying almost identical affinities for DR4. Instead, MD simulations revealed that analogues of **1**, such as **3**, that displayed similar dynamics of their complexes with DR4 as **1** were able to elicit an equal T-cell response. DR4/glycopeptide complexes that were more rigid or substantially more dynamic than the DR4/**1** complex, gave either a weak or no T-cell response. Importantly, complexes with similar dynamics as the DR4/**1** complex, such as the DR4/**3** complex, were characterized by having a similar number of hydrogen bonds of high, medium and low occupancy between the bound glycopeptide and DR4 as the DR4/**1** complex.

Comparison of the DR4/glycopeptide/TCR multicomponent system with the corresponding murine A^q-based system revealed both similarities and differences. Both MHCII molecules bound **1** by a dynamic network of cooperative hydrogen bonds, in which loss of a

single hydrogen bond might have a major effect on the response of the system. Both MHCII complexes with **1** also utilized a set of high occupancy hydrogen bonds to position the crucial TCR contact, GalHyl[264], for interactions with the TCR. However, the number and distribution of these hydrogen bonds differed between the two systems with the DR4/**1** complex having them located to a smaller number of residues at the N-terminus of **1**. Another, and perhaps more striking, difference between the two systems is that the one based on DR4 allows much larger variations in the structure of **1** while still maintaining signaling in the multicomponent system.

In conclusion, the DR4/glycopeptide/TCR system tolerated modifications at several positions in CII259–273, indicating opportunities to design and synthesize analogues to be used as chemical probes in order to increase our understanding the underlying causes of rheumatoid arthritis and for evaluation as vaccines for treatment of RA. Glycopeptide analogues of **1** that display altered dynamics when bound to DR4, while still eliciting a significant T cell response, are of particular interest for such investigations. Specific examples for *in vivo* studies include glycopeptides such as **3** that binds more rigidly to DR4 than **1** while eliciting a somewhat weaker T cell response, and **37** that binds more flexibly and elicits a significantly weaker response.

EXPERIMENTAL SECTION

Chemistry

General Synthetic Procedures. All reactions were carried out under argon atmosphere with dry solvents under anhydrous conditions, unless otherwise stated. THF was distilled from sodium and benzophenone and collected over freshly dried 4 Å molecular sieves, while DCM, methanol and acetonitrile (99.9%) were purchased from Aldrich and used directly. Reactions were monitored by TLC and LC-MS. TLC was performed on silica gel 60 F_{254} (Merck) with detection

by UV light and staining with phosphomolybdic acid in ethanol or ethanolic H_2SO_4 (5% v/v). LC-MS was performed on an Agilent 1100 series HPLC equipped with a C18 Atlantis T3 column (3.0 x 50 mm, 5 μm) and a Waters micromass Z_Q (model code: MM1) mass spectrometer using an acetonitrile–water gradient with a flow rate of 0.75 mL/min over 6 min. The molecular ion peak was detected in electrospray ionization mode. After workup, organic phases were dried over Na_2SO_4/$MgSO_4$ before being concentrated under reduced pressure. Flash column chromatography was performed on silica gel (Matrex, 60 Å, 35–70 μm, Grace Amicon). Optical rotations were measured with a Perking-Elmer model 343 polarimeter at 20 °C. ^{1}H and ^{13}C NMR spectra for the $Gln^{267}[CH_2CH_2]Gly^{268}$ (**26**), $Gln^{267}[(Z)\text{-}CH_2{=}CH_2]Gly^{268}$ (**27**), $Lys^{270}[CH_2CH_2]Gly^{271}$ (**34**) isostere building blocks and their intermediates were recorded at 298 K on an Agilent Technologies 400 MR spectrometer at 400 MHz and 100 MHz, respectively, and assigned considering the residual peak of solvent as internal standard [$CDCl_3$ ($CHCl_3$ δH 7.26 ppm, $CDCl_3$ δC 77.0 ppm) or CD_3SOCD_3 (CD_2HSOCD_3 δH 2.50 ppm, CD_3SOCD_3 δC 39.52 ppm)]. Proton resonances were assigned from COSY experiments. Analytical reversed-phase HPLC was performed on a Gilson HPLC equipped with either a Kromasil C8 column (250 × 4.6 mm, 5 μm) or a Varian 940 HPLC using a GraceVydac C18 column (150 × 4.6 mm, 5 μm) using acetonitrile–water gradients as eluents with a flow-rate of 2.0 mL/min and detection at 220 nm. Preparative reversed-phase HPLC was performed on either a Kromasil C8 column (250 × 21.2 mm, 5 μm) on a Gilson HPLC or using a Varian 940 HPLC equipped with a GraceVydac C8 column (150 × 21.2 mm, 10 μm) using acetonitrile-water gradients as eluents with flow-rate of 15 mL/min and detection at 214 or 220 nm. HRMS for the $Gln^{267}[CH_2CH_2]Gly^{268}$ (**26**), $Gln^{267}[(Z)\text{-}CH_2{=}CH_2]Gly^{268}$ (**27**) and $Lys^{270}[CH_2CH_2]Gly^{271}$ (**34**) isostere building blocks and their intermediates were recorded with electro spray (ESI) ionization. The identities of the

synthesised glycopeptides were confirmed by detecting the molecular ion peak in MALDI-TOF MS in positive ion mode using a α-cyano-4-hydroxycinnamic acid matrix. (5*R*)-N^{α}-(Fluoren-9-ylmethoxycarbonyl)-N^{ε}-benzyloxycarbonyl-5-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-5-hydroxy-L-lysine was prepared according to a published procedure.[19a] The unnatural amino acids Fmoc-3-fluoro-L-phenylalanine and Fmoc-3-bromo-L-phenylalanine were purchased from ChemPep Inc, Fmoc-3-(3-benzothienyl)-L-alanine and Fmoc-3-(3-pyridyl)-L-alanine from Alfa Aesar, Fmoc-L-cyclohexylalanine from Iris Biotech GmbH, and Fmoc-β-cyclopropyl-L-alanine from NeoMPS.

Benzyl (S)-(1-(methoxy(methyl)amino)-1,5-dioxo-5-(tritylamino)pentan-2-yl)carbamate (22):[21b, 22, 29] Cbz-Gln(Trt)-OH **21** (2.09 g, 4.00 mmol) was dissolved in DCM (40 mL) by addition of triethyl amine (560 μL, 4.0 mmol) followed by addition of 2-(7-aza-1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU, 1.52 g, 4.00 mmol). The reaction mixture was stirred for 15 min at rt before adding *N*,*O*-dimethylhydroxylamine hydrochloride (390 mg, 4.00 mmol) and triethyl amine (560 µL, 4.0 mmol). TLC (EtOAc:*n*-hexane 1: 1) showed completion of reaction in 3 h, DCM (100 mL) was then added to the reaction mixture, which was washed with aqueous HCl (1 M, 2x100 mL), saturated aqueous $NaHCO_3$ solution (100 mL), brine (2x50 mL), dried, filtered and concentrated under reduced pressure. The residue was purified by silica gel column chromatography (50% EtOAc in *n*-hexane) to give **22** (2.08 g, 92%) as a white solid.

Benzyl (S)-(1,5-dioxo-5-(tritylamino)pentan-2-yl)carbamate (23):[21b, 22] Compound **22** (1.90 g, 3.35 mmol) was dissolved in THF (30 mL) and DIBAL-H (8.37 mL of a 1 M solution in hexane, 8.37 mmol) was added to it dropwise at -78 °C under stirring. TLC (EtOAc:*n*-hexane 2:3) showed completion of reaction in 5 h, then MeOH (2 mL) was added to the solution

followed by aqueous HCl (1 M, 4 mL) after 5 min. Diethyl ether (150 mL) was then added to the reaction mixture which was washed with aqueous HCl (0.1 M, 2x100 mL), half-saturated aqueous $NaHCO_3$ (1x100 mL), saturated aqueous sodium potassium tartrate (1x100 mL), dried and concentrated under reduced pressure to give **23** (1.52 g, 90%) as white solid that was used in the next step without further purification.

3-Hydroxy-2-(hydroxymethyl)-2-methylpropyl (S,Z)-5-(((benzyloxy)carbonyl)amino)-8-oxo-8-(tritylamino)oct-3-enoate (24): Compound **20** (1.92 g, 3.85 mmol) and KHMDS (768 mg, 3.85 mmol) were weighed in a glove box under argon atmosphere and dry THF (24 mL) was added at -30 °C. The cooling bath was removed after 15 min and the reaction mixture was stirred at rt for 1 h and again at -30 °C for 15 min. Aldehyde **23** (1.50 g, 2.96 mmol) dissolved in THF (12 mL) was added dropwise to the reaction mixture at -30 °C and stirring was continued for 1 h at -30 °C and then for 1 h at rt. Aldehyde **23** and compound **24** had identical R_f on silica gel TLC plates (50% EtOAc in *n*-hexane), therefore reaction was monitored by LC-MS. It was quenched by adding water (5 mL) and concentrated under reduced pressure. The residue was dissolved in DCM (150 mL) and washed with water (2x100 mL) and brine (2x100 mL), dried, filtered, concentrated under reduced pressure and purified by reversed-phase HPLC (50 -> 90% acetonitrile in water over 25 min) to give **24 (**1.04 g, 62%) as a white solid. $[\alpha]_D$ +2.8 (*c* 1.0, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.36–7.17 (m, 21H, Cbz-*Ar*, Trt-*Ar*, Cbz-N*H*), 6.88 (s, 1H, Trt-N*H*), 5.64 (q, *J* = 8.6 Hz, 1H, -CH=C*H*CH_2-), 5.37 (ddt, *J* = 10.8, 9.4, 1.5 Hz, 1H, -CHC*H*=CH-), 5.08–5.01 (m, 2H, CbzCH_2-), 4.32 (d, *J* = 8.8 Hz, 1H, Cbz-NHC*H*-), 4.12 (s, 2H, -COOCH_2C(CH_3)-), 3.47 (m, 4H, -C(CH_3)(CH_2OH)$_2$), 3.22 (d, *J* = 7.7 Hz, 2H, -CH=CHCH_2-), 2.45 (s, 2H, -C(CH_3)(CH_2O*H*)$_2$), 2.34 (q, *J* = 6.9, 5.8 Hz, 2H, Trt-NHCOCH_2-), 1.81 (q, *J* = 7.3 Hz, 2H, Trt-NHCOCH_2CH_2-), 0.77 (s, 3H, -C(CH_3)(CH_2OH)$_2$). ^{13}C NMR (101 MHz, $CDCl_3$) δ

172.0, 171.4, 156.1, 144.5, 136.2, 133.4, 128.6, 128.5, 128.4, 128.2, 128.1, 128.0, 127.9, 127.0, 123.4, 77.3, 77.2, 77.0, 76.7, 70.6, 67.5, 67.50, 66.8, 66.7, 48.4, 40.5, 33.5, 33.3, 30.3, 16.9. HRMS (ESI) m/z calcd for $C_{40}H_{44}N_2O_7$ $[M + Na]^+$ 687.3041, found 687.3035.

3-Hydroxy-2-(hydroxymethyl)-2-methylpropyl (R)-5-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-8-oxo-8-(tritylamino)octanoate (25): A solution of **24** (388 mg, 0.60 mmol) in MeOH (8 mL) was added to Pd/C (10%, 25 mg) under argon. After applying vacuum and adding H_2 six times hydrogenation was continued for 24 h. The mixture was filtered through a celite pad and the solution was concentrated under reduced pressure. The product was dissolved in aqueous Na_2CO_3 solution (10%, 8 mL) after which acetonitrile (8 mL) and *N*-(9-fluorenylmethoxycarbonyloxy)succinimide (243 mg, 0.72 mmol) were added at rt. After 8 h the solution was concentrated under reduced pressure, dichloromethane (100 mL) was added and the mixture was washed with water (2x50 mL) and brine (2x50 mL). The organic layer was dried over sodium sulfate, filtered and concentrated under reduced pressure. The residue was purified by silica gel column chromatography (pre-treated with 5% triethyl amine; 50% EtOAc in *n*-hexane) to give 25 (300 mg, 68%) as a white solid. $[\alpha]_D$ +0.5 (*c* 4.0, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.74 (dd, *J* = 7.6, 0.9 Hz, 2H, Fmoc-*Ar*), 7.57 (d, *J* = 7.5 Hz, 2H, Fmoc-*Ar*), 7.42–7.17 (m, 20H, Fmoc-*Ar,* Trt-*Ar*, Fmoc-N*H*), 6.95 (d, *J* = 5.7 Hz, 1H, Trt-N*H*), 4.76 (m, 1H, Fmoc-NHC*H*-), 4.47–4.39 (m, 2H, FmocCH_2-), 4.18 (t, *J* = 6.4 Hz, 1H, FmocC*H*-), 4.13–4.11 (m, 2H, -COOCH_2C(CH_3)(CH_2OH)$_2$), 3.54–3.42 (m, 5H, -COOCH_2C(CH_3)(CH_2OH)$_2$), Fmoc-NHC*H*CH_2CH_2-), 2.68–2.50 (m, 2H, -COOCH_2C(CH_3)(CH_2O*H*)$_2$), 2.34 (dt, *J* = 11.6, 5.7 Hz, 2H, -CH_2COOCH_2C(CH_3)(CH_2OH)$_2$), 2.26 (t, *J* = 7.0 Hz, 2H, -CH_2CONHTrt), 1.70–1.52 (m, 4H, Fmoc-NHCHCH_2CH_2-), 1.50–1.36 (m, 2H, Fmoc-NHCHCH_2-), 0.79 (s, 3H, -COOCH_2C(CH_3)(CH_2OH)$_2$). ^{13}C NMR (101 MHz, $CDCl_3$) δ 174.0, 171.8, 156.6, 144.5, 143.9,

143.7, 141.3, 141.3, 128.6, 127.9, 127.6, 127.6, 127.0, 126.9, 124.9, 119.9, 77.3, 77.2, 77.0, 76.6, 70.6, 67.6, 66.4, 66.1, 50.6, 47.4, 40.4, 34.8, 33.8, 33.7, 31.0, 21.3, 16.9, 1.8. HRMS (ESI) m/z calcd for $C_{47}H_{50}N_2O_7$ $[M + H]^+$ 755.3691, found 755.3696.

(R)-5-((((9H-Fluoren-9-yl)methoxy)carbonyl)amino)-8-amino-8-oxooctanoic acid (26): Compound **25** (110 mg, 0.27 mmol) was taken-up in *t*-BuOH: water (1: 1, 3 mL) and acetyl chloride (3.5 mL) was added dropwise at 0 °C. The ice bath was removed and the reaction was stirred at rt and monitored with LC-MS for 24 h. After completion of the reaction the mixture was concentrated under reduced pressure and lyophilized after the residue had been dissolved in 0.1 M aqueous acetic acid. The crude product was washed with cold diethyl ether and dissolved in 30% acetic acid in acetonitrile and purified by reverse phase preparative HPLC (50 -> 90% acetonitrile in water over 25 minutes) to give **26** (50 mg, 83%) as a white solid. ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.96 (s, 1H), 7.87 (dt, J = 7.6, 0.9 Hz, 2H, Fmoc-*Ar*), 7.68 (dd, J = 7.5, 1.0 Hz, 2H, Fmoc-*Ar*), 7.41–7.36 (m, 2H, Fmoc-*Ar*), 7.31 (td, J = 7.4, 1.2 Hz, 2H, Fmoc-*Ar*), 7.19 (s, 1H, -CONH_2), 7.09 (d, J = 8.9 Hz, 1H, Fmoc-N*H*), 6.66 (s, 1H, -CONH_2), 4.34–4.15 (m, 3H, FmocC*H*-, FmocCH_2-), 3.34 (m, 1H, Fmoc-NHC*H*-), 2.16 (m, 2H, -CH_2CONH$_2$), 2.01 (t, J = 7.8 Hz, 2H, -CH_2COOH), 1.67–1.29 (m, 6H, FmocNHCHCH_2-, FmocNHCHCH_2CH_2-). ^{13}C NMR (101 MHz, DMSO-d_6) δ 174.8, 174.4, 156.3, 144.4, 144.3, 141.1, 128.0, 127.4, 125.6, 125.6, 120.5, 65.5, 50.6, 47.2, 40.6, 40.4, 40.3, 40.2, 40.1, 39.9, 39.7, 39.5, 39.3, 34.4, 33.9, 32.2, 30.8, 21.5. HRMS (ESI) m/z calcd for $C_{23}H_{26}N_2O_5$ $[M + H]^+$ 411.1914, found 411.1929.

(S,Z)-5-((((9H-Fluoren-9-yl)methoxy)carbonyl)amino)-8-amino-8-oxooct-3-enoic acid (27): Compound **25** (300 mg, 0.46 mmol) was taken up in *t*-BuOH: water (1: 4, 5 mL) followed by dropwise addition of acetyl chloride (8 mL) at 0 °C. The ice bath was removed and the reaction was stirred at rt and monitored with LC-MS for 24 h. After completion of the reaction

the mixture was concentrated under reduced pressure and lyophilized after the residue had been dissolved in 0.1 M aqueous acetic acid. The solid crude product was dissolved in a mixture of 10% aqueous Na_2CO_3 and acetonitrile (1:1, 20 mL), *N*-(9-fluorenylmethoxycarbonyloxy)succinimide (188 mg, 0.56 mmol) was added and the mixture was stirred at rt. After 8 h 3 M HCl was added dropwise to neutralize the reaction mixture which was then concentrated under reduced pressure, lyophilized and purified by reversed-phase preparative HPLC (50 -> 90% acetonitrile in water over 25 min) to give **27** (122 mg, 82%) as a white solid. $[\alpha]_D$ +3.0 (*c* 0.5, $CHCl_3$: EtOH 1: 1); ^{1}H NMR (400 MHz, DMSO-d_6) δ 7.85 (d, *J* = 8.0 Hz, 2H, Fmoc-*Ar*), 7.66 (d, *J* = 8.0 Hz, 2H, Fmoc-*Ar*), 7.41 - 7.35 (m, 3H, Fmoc-*Ar*, Fmoc-N*H*), 7.31 (td, *J* = 7.4, 1.2 Hz, 2H, Fmoc-*Ar*), 7.20 (s, 1H, -CONH_2), 6.69 (s, 1H, -CONH_2), 5.53 (dt, *J* = 10.9, 6.9 Hz, 1H, HOOCCH$_2$C*H*=CH-), 5.32 (t, *J* = 10.1 Hz, 1H, HOOCCH$_2$CH=C*H*-), 4.29–4.08 (m, 4H, FmocC*H*-, FmocCH_2-, Fmoc-NHC*H*-), 3.09 (d, *J* = 6.9 Hz, 2H, HOOCCH_2CH=CH-), 2.01 (t, *J* = 7.7 Hz, 2H, -CH_2CONH$_2$), 1.72–1.50 (m, 2H, -CH_2CH$_2$CONH$_2$). ^{13}C NMR (101 MHz, DMSO-d_6) δ 174.2, 172.9, 155.9, 144.3, 141.1, 133.3, 128.0, 127.5, 125.6, 123.5, 120.5, 65.7, 48.3, 47.1, 40.6, 40.5, 40.4, 40.3, 40.1, 39.9, 39.7, 39.5, 39.3, 33.0, 31.8, 30.8. HRMS (ESI) m/z calcd for $C_{23}H_{24}N_2O_5$ $[M + H]^+$ 409.1758, found 409.1762.

Benzyl tert-butyl (6-(methoxy(methyl)amino)-6-oxohexane-1,5-diyl)(S)-dicarbamate (29): Cbz-Lys(Boc)-OH **28** (3.00 g, 7.89 mmol) was dissolved in dry DCM (50 mL) by adding triethylamine (733 µL, 7.89 mmol) and subsequently HATU (3.00 g, 7.89 mmol). The solution was stirred for 15 min before adding *N*,*O*-dimethylhydroxylamine hydrochloride (770 mg, 7.89 mmol) and another equivalent of triethylamine (733 µL, 7.89 mmol). TLC (EtOAc:*n*-hexane 1:1) showed complete consumption of **28** in 3 h. DCM (150 mL) was then added and the reaction mixture was washed with water (1x100 mL), saturated aqueous sodium bicarbonate solution

(1x100 mL) and brine (1x100 mL), dried, filtered and concentrated under reduced pressure. The residue was purified on silica gel using 50% EtOAc in *n*-hexane as mobile phase to give **29** (1.792 g, 81%) as a colourless syrup. $[\alpha]_D$ -4.2 (*c* 5.0, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.37–7.28 (m, 5H, Cbz-*Ar*), 5.49 (d, *J* = 9.0 Hz, 1H, CbzN*H*-), 5.13–5.05 (dd, *J* = 16.0 Hz, *J* = 4.0 Hz, 2H, CbzCH_2-), 4.72 (s, 1H, Cbz-NHC*H*-), 4.57 (s, 1H, BocN*H*-), 3.77 (s, 3H, -OCH_3), 3.20 (s, 3H, -NCH_3), 3.08 (bs, 2H, BocNHCH_2-), 1.78–1.69 (m, 2H, BocNHCH$_2$CH_2-), 1.62–1.46 (m, 13H, *Boc*NHCH$_2$CH$_2$(CH_2)$_2$-). ^{13}C NMR (101 MHz, $CDCl_3$) δ 172.5, 156.1, 155.9, 136.3, 128.4, 128.0, 128.0, 79.0, 77.3, 77.2, 77.0, 76.6, 66.8, 61.5, 50.7, 40.1, 32.4, 32.0, 29.4, 28.4, 22.4. HRMS (ESI) m/z calcd for $C_{21}H_{33}N_3O_6$ $[M + H]^+$ 424.2442, found 424.2433.

Benzyl tert-butyl (6-oxohexane-1,5-diyl)(S)-dicarbamate (30): Compound **29** (1.26 g, 2.98 mmol) was dissolved in dry THF (55 mL) and DIBAL-H (7.44 mL of a 1M solution in hexane, 7.44 mmol) was added dropwise at -78 °C under argon. TLC revealed completion of the reaction in 4 h after which it was terminated by addition of methanol (2 mL) followed by aqueous HCl (1 M, 5.5 mL) after 5 min. The reaction mixture was allowed to reach room temperature, diethyl ether (100 mL) was added and the organic phase was washed with aqueous HCl (0.1 M, 3×100 mL), half-saturated aqueous $NaHCO_3$ (2x100 mL), brine (1×100 mL), dried, filtered and concentrated under reduced pressure. The residue was purified by silica gel column chromatography using 50% EtOAc in *n*-hexane as mobile phase to give **30 (**690 mg, 73%) as a colourless syrup. $[\alpha]_D$ +3.5 (*c* 2.0, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.57 (s, 1H, -C*H*O), 7.36–7.29 (m, 5H, Cbz-*Ar*), 5.52 (bs, 1H, CbzN*H*-), 5.14–5.08 (m, 2H, CbzCH_2-), 4.58 (s, 1H, BocN*H*-), 4.31–4.23 (m, 1H, Cbz-NHC*H*-), 3.10 (m, 2H, BocNHCH_2-), 1.71–1.59 (m, 2H, BocNHCH$_2$CH_2-), 1.55–1.32 (m, 13H, *Boc*NHCH$_2$CH$_2$(CH_2)$_2$-). ^{13}C NMR (101 MHz, $CDCl_3$) δ 199.3, 156.1, 136.0, 128.5, 128.4, 128.2, 128.1, 128.0, 128.0, 79.2, 77.3, 77.2, 77.0, 76.6, 67.1,

66.6, 60.0, 39.6, 29.7, 29.6, 28.4, 28.4, 28.3, 22.0. HRMS (ESI) m/z calcd for $C_{19}H_{28}N_2O_5$ $[M + H]^+$ 365.2071, found 365.2074.

3-Hydroxy-2-(hydroxymethyl)-2-methylpropyl (S,Z)-5-(((benzyloxy)carbonyl)amino)-9-((tert-butoxycarbonyl)amino)non-3-enoate (31): Potassium bis(trimethylsilyl)amide (436 mg, 2.19 mmol) and triphenyl phosphonium salt **20** (1.09 g, 2.19 mmol) were weighed under argon in a glove box, THF (10 mL) was added at -30°C and the solution was stirred for 1.5 h. Then a solution of aldehyde **30** (611 mg, 1.68 mmol) in dry THF (6 mL) was added dropwise over the period of 10 min. The reaction mixture was stirred for 1 h at -30 °C and another 2 h at rt. TLC (33% EtOAc in *n*-hexane) revealed formation of product at the same R_f as of **30**, therefore the reaction was monitored by LC-MS. After completion the reaction was quenched with water (1 mL) and concentrated under reduced pressure. DCM (150 mL) was added and the solution was washed with water (2x100 mL), dried, filtered and concentrated under reduced pressure to give a yellow oil. The residue was purified by reversed-phase HPLC (50% -> 90% acetonitrile in water over 25 minutes) to give **31** (715 mg, 84%) as a colourless syrup. $[\alpha]_D$ +2.0 (*c* 2.0, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.37–7.27 (m, 5H, Cbz-*Ar*), 5.68 (q, *J* = 8.0 Hz, 1H, -CH=C*H*CH_2-), 5.40–5.35 (t, *J* = 8.0 Hz, 1H, -C*H*=$CHCH_2$-), 5.04 (dd, *J* = 2 and 1.2 Hz, 2H, CbzCH_2-), 4.88 (s, 1H, CbzN*H*-), 4.32 (s, 1H, Cbz-NHC*H*-), 4.22–4.11 (m, 2H, -COOCH_2C(CH_3)- ($CH_2OH)_2$), 3.58–3.50 (m, 4H, -$COOCH_2C(CH_3)(CH_2OH)_2$), 3.30–3.21 (m, 2H, BocNHCH_2-), 3.09–3.05 (m, 2H, $BocNHCH_2CH_2$-), 2.86 (s, 2H, -$COOCH_2C(CH_3)(CH_2OH)_2$), 1.63 - 1.26 (m, 13H, *Boc*$NHCH_2CH_2(CH_2)_2$-), 0.83 (s, 3H, -$COOCH_2C(CH_3)(CH_2OH)_2$). ^{13}C NMR (101 MHz, $CDCl_3$) δ 172.1, 156.2, 155.8, 136.2, 133.6, 128.5, 128.1, 128.0, 123.2, 79.3, 77.3, 77.2, 77.0, 76.6, 67.5, 67.4, 66.8, 66.7, 48.3, 40.5, 40.1, 34.7, 33.1, 29.7, 29.6, 28.3, 22.6, 16.9. HRMS (ESI) m/z calcd for $C_{27}H_{42}N_2O_8$ $[M + H]^+$ 523.3014, found 523.3061.

3-Hydroxy-2-(hydroxymethyl)-2-methylpropyl (S)-5-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-9-((tert-butoxycarbonyl)amino)nonanoate (32): A solution of compound **31** (693 mg, 1.38 mmol) in MeOH (18.5 mL) was added to Pd/C (10%, 125 mg) under argon. After applying vacuum and adding H_2 six times hydrogenation was continued for 19 h. The mixture was filtered through a celite pad and concentrated under reduced pressure to give a colorless oil (525 mg) that was used without further purification. It was dissolved in aqueous Na_2CO_3 (10%, 18.5 mL) and acetonitrile (18.5 mL) and *N*-(9-fluorenylmethoxycarbonyloxy)succinimide (386 mg, 1.14 mmol) was added while stirring at rt. After 5.5 h the mixture was concentrated under reduced pressure, the residue was dissolved in DCM (150 mL) and washed with water (2x100 mL) and brine (2x100 mL). The organic phase was dried over sodium sulfate, filtered and concentrated under reduced pressure. The residue was purified by silica gel column chromatography (pre-treated with 5% triethyl amine; 50% EtOAc in *n*-hexane) to give **32** (701 mg, 98%) as a colorless syrup. $[\alpha]_D$ –1.8 (*c* 1.0, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.75 (d, *J* = 8.0, 2H, Fmoc-*Ar*), 7.59 (d, *J* = 8.0 Hz, 2H, Fmoc-*Ar*), 7.41–7.37 (t, *J* = 8.0 Hz, 2H, Fmoc-*Ar*), 7.31 (td, *J* = 8.0 and 1.2 Hz, 2H, Fmoc-*Ar*), 4.70 (d, *J* = 9.1 Hz, 1H, FmocN*H*-), 4.45–4.38 (m, 2H, FmocC*H*, CH_2-), 4.21–4.11 (m, 3H, FmocC*H*, CH_2-/-COOCH_2C(CH_3)(CH_2OH)$_2$), 3.58–3.49 (m, 4H, -COOCH_2C(CH_3)(CH_2OH)$_2$), Fmoc-NHC*H*-), 3.08 (t, *J* = 6.8 Hz, 2H, BocNHCH_2-), 2.43–2.29 (m, 2H, -COOCH_2C(CH_3)(CH_2O*H*)$_2$, -CH_2COOCH_2C(CH_3)(CH_2OH)$_2$), 1.72–1.25 (m, 19H, *Boc*NHCH_2(CH_2)$_3$-, -(CH_2)$_2$CH_2COOCH_2C(CH_3)(CH_2OH)$_2$), 0.82 (s, 3H, -COOCH_2C(CH_3)(CH_2OH)$_2$). ^{13}C NMR (101 MHz, $CDCl_3$) δ 174.2, 156.3, 156.1, 143.9, 143.9, 141.3, 127.6, 127.0, 124.9, 124.6, 119.9, 79.2, 77.3, 77.2, 77.0, 76.6, 67.6, 67.5, 66.5, 66.3, 51.3, 50.7, 47.3, 40.5, 40.2, 34.7, 34.6, 33.8,

29.7, 28.4, 22.8, 21.2, 16.9. HRMS (ESI) m/z calcd for $C_{34}H_{48}N_2O_8$ $[M + H]^+$ 613.3483, found 613.3461.

(S)-5-((((9H-Fluoren-9-yl)methoxy)carbonyl)amino)-8-carboxyoctan-1-aminium chloride (33): Compound **32** (500 mg, 0.84 mmol) was taken-up in *t*-BuOH: water (1:1, 6 mL) and acetyl chloride (10 mL) was added dropwise at 0 °C. The ice bath was removed, the reaction was stirred at room temperature for 24 h and monitored by LC-MS. The solution was then concentrated under reduced pressure, the residue was dissolved in 0.1 M acetic acid and the solution was lyophilized. The residue was washed with cold diethyl ether and purified by reversed-phase HPLC (50 -> 90% acetonitrile in water over 25 min) to give **33** (271 mg, 79%) as a white solid. $[\alpha]_D$ +29.8 (*c* 1.0, $CHCl_3$); ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.98 (s, 1H, COO*H*), 7.87 (d, *J* = 8.0, 2H, Fmoc-*Ar*), 7.80–7.58 (m, 5H, Fmoc-*Ar*, NH_3^+), 7.42–7.36 (m, 2H, Fmoc-*Ar*), 7.33–7.28 (m, 2H, Fmoc-*Ar*), 7.06 (d, *J* = 8.9 Hz, 1H, FmocN*H*), 4.35–4.17 (m, 3H, FmocC*H*, CH_2), 3.35 (m, 1H, Fmoc-NHC*H*-), 2.72 (t, *J* = 7.6 Hz, 2H, -CH_2NH$_3^+$), 2.16 (td, *J* = 7.1 and 4.3 Hz, 2H, -CH_2COOH), 1.53–1.19 (m, 10H, -(CH_2)$_3$CH$_2$NH$_3^+$, -(CH_2)$_2$CH$_2$COOH). ^{13}C NMR (101 MHz, DMSO-d_6) δ 174.8, 158.4, 158.1, 156.3, 144.3, 141.1, 129.9, 128.0, 127.4, 127.4, 125.6, 125.5, 124.3, 121.6, 120.5, 65.3, 50.4, 47.3, 40.6, 40.5, 40.4, 40.3, 40.2, 40.2, 40.1, 39.9, 39.7, 39.5, 39.3, 39.1, 34.5, 34.5, 33.9, 27.2, 22.9, 21.5. HRMS (ESI) m/z calcd for $C_{24}H_{31}N_2O_4^+$ $[M + H]^+$ 411.2278, found 411.2280.

(S)-5-((((9H-Fluoren-9-yl)methoxy)carbonyl)amino)-9-((tert-butoxycarbonyl)amino)nonanoic acid (34): Compound **33** (200 mg, 0.49 mmol) was dissolved in water:dioxane (3:5, 8 mL) and $NaHCO_3$ (84 mg, 1.00 mmol) was added followed by *tert*-butyl dicarbonate (144 mg, 0.66 mmol) over 20 min and stirred overnight at rt. The mixture was then acidified to pH 3 using 1 M aqueous $KHSO_4$ solution, dioxane was removed under reduced

pressure, DCM (100 mL) was added and the organic phase was washed with water (1×50 mL), dried, filtered and concentrated under reduced pressure. The residue was purified by reversed-phase HPLC (50 -> 90% acetonitrile in water over 25 min) to give **34** (137 mg, 55%) as a white solid. $[\alpha]_D$ –1.2 (*c* 1.0, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.75 (d, *J* = 7.5 Hz, 2H, Fmoc-*Ar*), 7.58 (d, *J* = 7.6 Hz, 2H, Fmoc-*Ar*), 7.38 (t, *J* = 8.0 Hz, 2H, Fmoc-*Ar*), 7.30 (td, *J* = 7.5 and 1.2 Hz, 2H, Fmoc-*Ar*), 4.61–4.56 (bs, 2H, FmocC*H*, Fmoc-N*H*-), 4.46–4.38 (m, 2H, FmocCH_2, BocN*H*-), 4.20 (m, 1H, FmocCH_2), 3.63–3.57 (bs, 1H, Fmoc-NHC*H*-), 3.09–2.98 (m, 2H, -CH_2NHBoc), 2.39–2.30 (m, 2H, -CH_2COOH), 1.67–1.20 (m, 20H, -(CH_2)$_3$CH$_2$NH*Boc*, -(CH_2)$_2$CH$_2$COOH). ^{13}C NMR (101 MHz, $CDCl_3$) δ 177.6, 156.2, 143.9, 141.3, 127.6, 127.0, 125.0, 124.9, 124.6, 119.9, 79.1, 77.3, 77.2, 77.0, 76.6, 66.2, 51.5, 50.8, 47.3, 40.2, 34.7, 33.5, 29.7, 28.4, 22.8, 20.9. HRMS (ESI) m/z calcd for $C_{29}H_{38}N_2O_6$ $[M + H]^+$ 511.2803, found 511.2803.

General procedure for solid-phase glycopeptide synthesis. Glycopeptides **2**–**16**, were synthesized on a 50 μmol scale using a Fmoc-Thr(tBu) TentaGel S PHB resin (0.22 mmol/g) following standard solid-phase methodology as described elsewhere with minor modifications.[30] N^{α}-Fmoc amino acids with standard side-chain protecting groups (5 equiv) were activated with HCTU (4.5 equiv) and DIPEA (10 equiv) and couplings were performed in NMP for 2 x 20 min using a Prelude (Protein Technologies) peptide synthesizer. (5*R*)-N^{α}-(Fluoren-9-ylmethoxycarbonyl)-N^{α}-benzyloxycarbonyl-5-*O*-(2,3,4,6-*tetra*-*O*-acetyl-β-D-galactopyranosyl)-5-hydroxy-L-lysine[19] (1.5 equiv) and the isostere building blocks **26**, **27**and **34** (1.5 equiv) were activated with HATU (1.4 equiv) and DIPEA (4 equiv) and coupled manually for at least 3h. Fmoc deprotection after each coupling cycle was accomplished by treatment with 20% piperidine in NMP for 2 x 5 min. The glycopeptides were cleaved from the resin with

trifluoroacetic acid/H_2O/thioanisole/ethanedithiol/TIS (94:2:2:1:1) for 3 h at 40 °C with workup performed essentially as described elsewhere.[30] Purification by reversed-phase HPLC and lyophilization was followed by deacetylation with NaOMe in MeOH (20 mM, 1 mL/mg peptide) for 2-3 h at room temperature (monitored by analytical reversed-phase HPLC). Neutralization was achieved by addition of AcOH and the solution was then concentrated under reduced pressure, after which the residue was purified using reversed-phase HPLC followed by lyophilization. The purity of the glycopeptides was determined to be >95% according to analytical reversed-phase HPLC (SI) and their identities were confirmed by MALDI-TOF mass spectrometry (Table 3).

Table 3. Yield and molecular weights determined by MALDI-TOF for glycopeptides **2**–**16**

Glycopeptide	**Yield (%)**	**MW calcd $[M+H]^+$**	**MW found $[M+H]^+$**
2	11.9	1769.8	1769.6
3	45.6	1730.9	1730.7
4	46.0	1692.8	1692.7
5	49.1	1697.8	1697.7
6	22.0	1695.8	1695.7
7	46.8	1709.9	1709.7
8	6.0	1671.7	1671.7
9	44.7	1701.9	1701.8
10	52.2	1786.9	1786.7
11	25.9	1760.0	1759.7

12	20.1	1806.0	1805.8
13	16.6	1793.0	1792.6
14	10.7	1638.8	1638.7
15	16.7	1636.8	1636.7
16	22.9	1638.8	1638.7
35	55.0	1855.6	1856.1
36	52.8	1878.3	1877.1
37	48.4	1844.5	1844.0

DR4 Binding Assay

The ability of the glycopeptides **1**–**13**, **35**–**37**, and glycopeptide isosteres **14**–**16** to bind to DR4 was determined in a competitive inhibition binding assay, essentially described elsewhere.[15, 31] Briefly, a mixture of 0.1 μM recombinant DR4 protein, 1.5 μM biotinylated CLIP (CLIP-bio: KPVSKMRMATPLLMQALPM), and increasing concentrations of the investigated glycopeptides 1–16 (2.5, 5.0, 10.0, 19.0, 38.0, 75.0, 150.0, 300.0 μM) were incubated in PBS containing protease inhibitors (Complete™ Protease Inhibitor Cocktail (Roche)) at rt for 48 h. Then the mixture (100 μL) was transferred to L243-precoated microtiter assay plates to capture the recombinant DR4 molecules. The plates were washed with PBS containing 0.05% Tween-20 to remove excess peptides and the amount of bound CLIP-bio was quantified by the dissociation-enhanced lanthanide fluoroimmunoassay (DELFIA®) system with europium-labeled streptavidin (Perkinelmer), according to the manufacturer's instructions. All experiments were performed in duplicates.

T-Cell Hybridoma Activation Assay

The response of the DR4 restricted T-cell hybridoma line mDR1.1 (the amount of secreted IL-2 following incubation of the hybridoma with antigen-presenting spleenocytes expressing DR4, with increasing concentrations of the glycopeptides **1**–**13**, **35**–**37**, and glycopeptide isosteres **14**–**16**) was determined as essentially described elsewhere.[32] Briefly, T-cell hybridomas (5×10^4), syngeneic spleen cells (5×10^5) from DRA/DRB1*0401 mice, and increasing concentrations of the glycopeptides and glycopeptide isosteres (0.0, 0.0064, 0.032, 0.16, 0.8, 4.0, 20.0, 100.0 μM) were co-cultured in DMEM medium (supplemented with 5% FCS, 10 U/ml penicillin and 100 μg/ml streptomycin) in 96-well flat bottom microtiter plates. The plates were incubated at 37 °C for 24h. Thereafter 50μl supernatant from each well was transferred into IL-2 capturing antibody-coated ELISA plates and incubated at rt for 2h. After washing, biotin-labeled IL-2 detection antibody was added into each well and the plates were incubated at rt for 1h. After washing, 1000 fold diluted Eu^{3+}-conjugated streptavidin (DELFIA; PerkinElmer) was added into each well and the plates were incubated at rt for 1h. Finally, enhancement solution (DELFIA; PerkinElmer) was added into each well and the fluorescence was detected on a Synergy 2 multi-mode microplate reader. All values are represented as relative fluorescence units. The experiments were performed in duplicates.

Molecular Modeling

Sequence and structure alignments

Four published X-ray crystal structures (1J8H,[14a] 2SEB,[25] 3TOE[14c] and 4IS6[14d]) were used to align the CII259–273 sequence to the sequences of peptides that bind to DR4 (Table S5). Superposition of the four crystal structures based on heavy atoms in the binding groove (the α_1- and β_1-helices and the $\alpha_1\beta_1$ β-sheet) gave the sequence alignment of the DR4 peptides (Table S5). It also revealed aromatic or aliphatic peptide residues placed in P1 pocket of DR4, and acidic or

amide-functionalized peptide residues in P4 pocket of DR4 (Table S5 and Figure S71). Furthermore, it showed the substantial H-bonding network between the backbone of the peptides and residues of DR4, as reported before also for other MCHII proteins in complex with peptides.[33] Altogether, the resulting alignment of the CII259–273 sequence to the DR4 peptides (Table S5) suggested that Phe^{263} and Glu^{266} to be placed in the P1 and P4 pockets of DR4, respectively, and that Gly^{268} and Gly^{271} are positioned next to the P6 and P9 pockets of DR4. This information was used further in the design of modified glycopeptides and in the homology modelling of DR4-glycopeptide complexes (see below).

Design of Modified Glycopeptides

The amino acids chosen to replace Phe^{263} (mFPhe, Trp, and Cha) Gly^{268} (Thr and Val) and Gly^{271} (Leu and Gln) in CII259–273 **1** were selected in order to vary both the size and polarity of the side chain, guided by our previous results[7c] and the investigated X-ray crystal structures of peptides in complex with DR4 (Table S5 and Figure S71). [14b14c14d14a]In order to select a small diverse subset of glycopeptides for synthesis a linear D-optimal design was applied to a matrix with the substituted amino acids as qualitative variables. The MODDE[34] software was used and 12 modified glycopeptides were selected, with a G-efficiency of 64.6 and a condition number of 2.4.

Homology Modeling of CII259–273 in complex with DR4

The superposition of the four published crystal structures of DR4 in complex with peptides (1J8H,[14a] 2SEB,[25] 3TOE[14c] and 4IS6[14d]) showed a high degree of conformational similarity, where the pairwise RMSD values of main chain heavy atoms were between 0.78-1.03Å for the whole complexes and between 0.65-1.61Å for the peptides (Figure S73 and S75). The model of DR4 in complex with CII259–273, **1**, was based on the DR4 X-ray crystal structure 1J8H (a

ternary complex between DR4, an influenza peptide, and a TCR with a resolution of 2.4Å), while the modelling of CII259–273 in the complex was based on all four complexes. The resulting RMSD values between the main chain heavy atoms of 1J8H and DR4- CII259–273 were 0.47Å and 0.77Å for the whole complex and peptides, respectively. The complex 1J8H was prepared using the protein preparation wizard[35] in Maestro.[36] Bond orders were assigned, hydrogen atoms were added, disulfide bonds were created, termini were capped, and crystal waters were deleted. The hydrogen bond network was optimized followed by minimization of hydrogen atoms. The two missing loops (α Lys^{105}–His^{112} in DR4 and Vα Asp^{130}–Lys^{132} in the TCR) were manually added in Maestro followed by an energy minimization using MacroModel[37] with the OPLS_2005 force field. The added amino acids were allowed to move freely while the protein was kept rigid. LOOP within MacroModel was thereafter used to generate multiple loop conformations that were subjected to an energy minimization where the modeled loop atoms including two extra amino acids at each end were minimized using the same procedure as described above.

For the homology modelling of CII259–273 **1** in complex with DR4 the TCR was deleted. The peptide sequence was manually aligned with that of the influenza peptide in the template to make sure that Phe^{263} would overlap with Tyr^{308} positioned in the P1 pocket in DR4 according to the sequence alignment (Figure S5). 100 different models were created for the side chains using the homology model tool implemented in MOE[38] without performing energy minimization. The generated conformations of the side chains were compared to the four X-ray crystal structures of DR4-peptides. The selection of conformations were guided by similarity to the X-ray crystal structures and the ability for the residues to participate in interactions with DR4. The conformations of Lys^{264} that were solvent exposed were favored since a carbohydrate moiety

was to be attached. Thereafter the complex was energy minimized using MacroModel. The peptide backbone and Lys^{264} were constrained using a flat bottom constraint with a force constant of 100 kJ mol^{-1} $Å^2$ and a width of 0.2 Å, with the protein backbone kept rigid. Thereafter Lys^{264} was manually hydroxylated and galactosylated. A conformational search where the carbohydrate moiety was allowed to move freely, while the rest of the complex was kept rigid, was performed. The carbohydrate conformations were thereafter allowed to move freely during an energy minimization with the rest of the complex constrained with the same type of flat bottom constraint on the protein backbone as described above. The resulting models were visually inspected, and the model where the carbohydrate moiety was mostly solvent exposed was selected. The resulting pairwise RMSD values between the main chain heavy atoms of the four X-ray crystal structures and DR4- CII259–273 were between 0.47-1.11Å and 0.-1.45Å for the whole complexes and peptides, respectively.

Preparation of the DR4 Complexes with Modified Glycopeptides

The DR4-glycopeptide complexes with substituted amino acids (**2**–**5**, **35**–**37**,) were based on the DR4-CII259–273 model, where the four X-ray crystal structures were used in the selection of final models. Note that none of the substituted amino acids in the modified glycopeptides was larger than amino acids already present in peptides in the X-ray crystal structures. The substituted residues in glycopeptide **1** were modified *in silico* in Maestro into modified glycopeptide **2**–**5** or **35**–**37**. A conformational search was performed where the substituted residues were allowed to move freely while the rest of the complex was kept rigid. The resulting conformations were subjected to an energy minimization where the substituted amino acid was allowed to move freely while the rest of the glycopeptide and the DR4 protein was restrained with a flat bottom constraint as described above, with the backbone of DR4 kept rigid. No

clashes were observed in the generated conformations. The final models were selected based on comparison of the resulting glycopeptide conformations with the peptide conformations of the four X-ray crystal structures (1J8H,[14a] 2SEB,[25] 3TOE[14c] and 4IS6[14d]). The resulting pairwise RMSD values between the main chain heavy atoms of the four X-ray crystal structures and DR4-CII259–273 were between 0.14–1.19Å and 0.14–1.46Å for the whole complexes and peptides, respectively.

Parameter Preparation for Non-Native Amino Acids

Parameters were prepared for the unnatural amino acids *m*FPhe, Cha, benzothiophenealanine, *m*-bromo phenylalanine, cyclopropylalanine, and pyridine. Each amino acid was built and capped (the N-terminus was acetylated and the C-terminus was methylated) using the Maestro software. A conformational search using the OPLS_2005 force field was performed and the parameter preparation was based on in total six conformations. Partial charges were derived using the R.E.D Server[39] and fitted to reproduce the molecular electrostatic potential using an intramolecular charge constraint set to zero for the termini. The RESP-A1 (HF/6-31G*) charge model was used with the central fragment procedure. Parameters for the carbohydrate moiety were prepared previously, as described elsewhere.[7b]

MD Simulations of Glycopeptide/DR4 Complexes

The MD simulations were run using the Amber12[40] simulation package. Hydrogen atoms were deleted from the complexes before preparation with the xleap module. The ff99SB force field was used together with GLYCAM_06h for the carbohydrate moiety. Hydrogen atoms were added, termini were capped, and disulfide bridges specified. Sodium ions were added to neutralize the system and the complex was solvated using a 12 Å octahedral TIP3P water box with periodic boundary conditions. Bonds involving hydrogen atoms were constrained using the SHAKE algorithm, and a time step of 2 fs were used. The temperature was controlled using the

Langevin thermostat. Long range electrostatic interactions were treated by the Particle Mesh Ewald method and an 8 Å cutoff was used for short range non bonded interactions. The complex was subjected to an energy minimization where the glycopeptide was allowed to move freely while the protein was constrained with a force constant of 500 kJ mol^{-1}, followed by a minimization without constraints. Thereafter the complex was heated from 0 to 300 K during a 20 ps simulation with constraints with a force constant of 10 kJ mol^{-1} on the entire complex. This was followed by a 100 ps equilibration run at 300 K where weak constraints of 1 kJ mol^{-1} were added to the backbone of the terminal residues of the $\alpha_2\beta_2$ domain since they are membrane bound. The same type of weak constraint of 1 kJ mol^{-1} was added to the backbone of 23 amino acids in the α_1-helix (αAla^{56}–Tyr^{79}, the total number of amino acids is 368) in order to mimic the environment in the presence of a TCR, while still allowing the α_1-helix to move. (40 ns MD simulations of the ternary template reveals a stable α_1-helix, while unfolding of the α_1-helix occurs immediately when running MD simulations of the binary template. Apart from the constraints on the α-helix the same settings as for the binary complex with DR4 and glycopeptides have been used for the ternary complex, where constraints with a force constant of 100 kJ mol^{-1} were used for the backbone of the terminal residues of the $\alpha_2\beta_2$ domain.) Constant pressure periodic boundaries were used with an average pressure of 1 atm together with isotropic position scaling to maintain the pressure. The production runs were performed at 300 K with the same constraints as for the equilibration and in total 60 ns. MD simulations on each system were repeated six times with different initial velocities, based on the RMSD values the simulations converged after 30 ns. MD simulations were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) at the High Performance Computing Centre North (HPC2N).[41]

Analysis was performed with the software VMD.[27] Frames were extracted and visualized using a time step of 50 or 4 ps. RMSD values were calculated using the RMSD trajectory tool. The hydrogen bond extension tool was used to calculate hydrogen bond occupancy, using a donor-acceptor length of less than 3.4 Å and a donor-H-acceptor angle of at least 40°. The clustering plugin[42] was used to perform cluster analysis for 30–60 ns, based on the backbone of Ala^{261}–Gly^{271}, an RMSD cutoff of 1.5 Å and a maximum of 10 clusters. Figures were prepared using MOE.[38]

ASSOCIATED CONTENT

Supporting Information.

The Supporting Information is available free of charge.

NMR spectra for compounds **24**–**27** and **29**–**34**. HPLC chromatograms for compound **34**, glycopeptides **2**–**16**, and **35**–**37**. RMSD plots, a summary of the hydrogen bonds formed to the glycopeptides, the number of clusters extracted from the MD simulations of DR4 in complex with glycopeptides **2**–**5** and **35**–**37**, sequence alignment of peptides, overlay of DR4-CII259-273 and published X-ray crystal structures, RMSD values for DR4-peptide models (PDF). Molecular formula strings (CSV).

AUTHOR INFORMATION

Corresponding Author

*E-mail: anna.linusson@umu.se, phone: +46-90-7866890. E-mail: jan.kihlberg@kemi.uu.se, phone +46-18-4713801.

ACKNOWLEDGMENTS

This work was funded by the Swedish Research Council (Dnr: 2014-4675), the KA Wallenberg foundation, the Swedish Foundation for Strategic Research (SSF), Pharmacy School Southern Medical University, Guangzhou, China and a grant from the Guangdong province (201001Y04675344). The MD simulations were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) at the High Performance Computing Center North (HPC2N).

ABBREVIATIONS

GalHyl, β-D-galactopyronasyl hydroxylysine; *m*FPhe, *m*-fluoro phenylalanine; MHC, major histocompatibility complex; MD, molecular dynamics; RA, rheumatoid arthritis; TCR, T-cell receptor.

REFERENCES

1. Medina-Franco, J. L.; Giulianotti, M. A.; Welmaker, G. S.; Houghten, R. A. Shifting from the single to the multitarget paradigm in drug discovery. *Drug Discovery Today* **2013,** *18* (9-10), 495-501.

2. Anighoro, A.; Bajorath, J.; Rastelli, G. Polypharmacology: Challenges and opportunities in drug discovery. *J.Med. Chem.* **2014,** *57* (19), 7874-7887.

3. Rudolph, M. G.; Stanfield, R. L.; Wilson, I. A. How TCRs bind MHCs, peptides, and coreceptors. *Annu. Rev. Immunol.*, 2006; Vol. 24, pp 419-466.

4. (a) Eisenmesser, E. Z.; Millet, O.; Labeikovsky, W.; Korzhnev, D. M.; Wolf-Watz, M.; Bosco, D. A.; Skalicky, J. J.; Kay, L. E.; Kern, D. Intrinsic dynamics of an enzyme underlies catalysis. *Nature* **2005,** *438* (7064), 117-121; (b) Tai, K.; Shen, T. Y.; Borjesson, U.; Philippopoulos, M.; McCammon, J. A. Analysis of a 10-ns molecular dynamics simulation of

mouse acetylcholinesterase. *Biophys. J.* **2001,** *81* (2), 715-724; (c) Dobson, C. M.; Sali, A.; Karplus, M. Protein folding: a perspective from theory and experiment. *Angew. Chem., Int. Ed.* **1998,** *37* (7), 868-893; (d) Manglik, A.; Kobilka, B. The role of protein dynamics in GPCR function: insights from the β_2AR and rhodopsin. *Curr. Opin. Cell Biol.* **2014,** *27*, 136-143; (e) Langosch, D.; Scharnagl, C.; Steiner, H.; Lemberg, M. K. Understanding intramembrane proteolysis: from protein dynamics to reaction kinetics. *Trends Biochem. Sci.* **2015,** *40* (6), 318-327; (f) Gunasekaran, K.; Ma, B. Y.; Nussinov, R. Is allostery an intrinsic property of all dynamic proteins? *Proteins: Struct., Funct., Bioinf.* **2004,** *57* (3), 433-443.

5. Henzler-Wildman, K.; Kern, D. Dynamic personalities of proteins. *Nature* **2007,** *450* (7172), 964-972.

6. (a) van der Merwe, P. A.; Dushek, O. Mechanisms for T cell receptor triggering. *Nat. Rev. Immunol.* **2011,** *11* (1), 47-55; (b) Knapp, B.; Demharter, S.; Esmaielbeiki, R.; Deane, C. M. Current status and future challenges in T-cell receptor/peptide/MHC molecular dynamics simulations. *Briefings Bioinf.* **2015,** *16* (6), 1035-1044; (c) Stavrakoudis, A. Conformational flexibility in designing peptides for immunology: the molecular dynamics approach. *Curr. Comput.-Aided Drug Des.* **2010,** *6* (3), 207-222.

7. (a) Andersson, I. E.; Batsalova, T.; Haag, S.; Dzhambazov, B.; Holmdahl, R.; Kihberg, J.; Linusson, A. (*E*)-Alkene and ethylene isosteres substantially alter the hydrogen-bonding network in class II MHC A^q/glycopeptide complexes and affect T-cell recognition. *J. Am. Chem. Soc.* **2011,** *133* (36), 14368-14378; (b) Lindgren, C.; Andersson, I. E.; Berg, L.; Dobritzsch, D.; Ge, C. R.; Haag, S.; Uciechowska, U.; Holmdahl, R.; Kihlberg, J.; Linusson, A. Hydroxyethylene isosteres introduced in type II collagen fragments substantially alter the structure and dynamics of class II MHC A^q/glycopeptide complexes. *Org. Biomol. Chem.* **2015,** *13* (22), 6203-6216; (c)

Andersson, I. E.; Andersson, C. D.; Batsalova, T.; Dzhambazov, B.; Holmdahl, R.; Kihlberg, J.; Linusson, A. Design of glycopeptides used to investigate Class II MHC binding and T-Cell responses associated with autoimmune arthritis. *Plos One* **2011,** *6* (3).

8. McInnes, I. B.; Schett, G. The pathogenesis of rheumatoid arthritis. *N. Engl. J. Med.* **2011,** *365* (23), 2205-2219.

9. Gregersen, P. K.; Silver, J.; Winchester, R. J. The shared epitope hypothesis - an approach to understanding the molecular-genetics of susceptibility to rheumatoid-arthritis. *Arthritis Rheum.* **1987,** *30* (11), 1205-1213.

10. Backlund, J.; Carlsen, S.; Hoger, T.; Holm, B.; Fugger, L.; Kihlberg, J.; Burkhardt, H.; Holmdahl, R. Predominant selection of T cells specific for the glycosylated collagen type II epitope (263-270) in humanized transgenic mice and in rheumatoid arthritis. *Proc. Natl. Acad. Sci. U. S. A.* **2002,** *99* (15), 9960-9965.

11. (a) Backlund, J.; Treschow, A.; Bockermann, R.; Holm, B.; Holm, L.; Issazadeh-Navikas, S.; Kihlberg, J.; Holmdahl, R. Glycosylation of type II collagen is of major importance for T cell tolerance and pathology in collagen-induced arthritis. *Eur. J. Immunol.* **2002,** *32* (12), 3776-3784; (b) Hegen, M.; Keith, J. C.; Collins, M.; Nickerson-Nutter, C. L. Utility of animal models for identification of potential therapeutics for rheumatoid arthritis. *Ann. Rheum. Dis.* **2008**, *67* (11), 1505-1515.

12. Dzhambazov, B.; Nandakumar, K. S.; Kihlberg, J.; Fugger, L.; Holmdahl, R.; Vestberg, M. Therapeutic vaccination of active arthritis with a glycosylated collagen Type II peptide in complex with MHC class II molecules. *J. Immunol.* **2006,** *176* (3), 1525-1533.

13. Raposo, B.; Merky, P.; Lundqvist, C.; Yamada, H.; Urbonaviciute, V.; Niaudet, C.; Viljanen, J.; Kihlberg, J.; Kyewski, B.; Ekwall, O.; Holmdahl, R.; Backlund, J. T cells specific

for post-translational modifications escape intrathymic tolerance induction. *Nat. Commun.* **2018,** *9*.

14. (a) Hennecke, J.; Wiley, D. C. Structure of a complex of the human alpha/beta T cell receptor (TCR) HA1.7, influenza hemagglutinin peptide, and major histocompatibility complex class II molecule, HLA-DR4 (DRA*0101 and DRB1*0401): Insight into TCR cross-restriction and alloreactivity. *J. Exp. Med.* **2002,** *195* (5), 571-581; (b) Bolin, D. R.; Swain, A. L.; Sarabu, R.; Berthel, S. J.; Gillespie, P.; Huby, N. J. S.; Makofske, R.; Orzechowski, L.; Perrotta, A.; Toth, K.; Cooper, J. P.; Jiang, N.; Falcioni, F.; Campbell, R.; Cox, D.; Gaizband, D.; Belunis, C. J.; Vidovic, D.; Ito, K.; Crowther, R.; Kammlott, U.; Zhang, X. L.; Palermo, R.; Weber, D.; Guenot, J.; Nagy, Z.; Olson, G. L. Peptide and peptide mimetic inhibitors of antigen presentation by HLA-DR class II MHC molecules. Design, structure-activity relationships, and X-ray crystal structures. *J. Med. Chem.* **2000,** *43* (11), 2135-2148; (c) Yin, Y. Y.; Wang, X. X.; Mariuzza, R. A. Crystal structure of a complete ternary complex of T-cell receptor, peptide-MHC, and CD4. *Proc. Natl. Acad. Sc. U. S. A.* **2012,** *109* (14), 5405-5410; (d) Chen, S. M.; Li, Y. L.; Depontieu, F. R.; McMiller, T. L.; English, A. M.; Shabanowitz, J.; Kos, F.; Sidney, J.; Sette, A.; Rosenberg, S. A.; Hunt, D. F.; Mariuzza, R. A.; Topalian, S. L. Structure-based design of altered MHC class II-restricted peptide ligands with heterogeneous immunogenicity. *J. of Immunol.* **2013,** *191* (10), 5097-5106.

15. Kjellen, P.; Brunsberg, U.; Broddefalk, J.; Hansen, B.; Vestberg, M.; Ivarsson, I.; Engstrom, A.; Svejgaard, A.; Kihlberg, J.; Fugger, L.; Holmdahl, R. The structural basis of MHC control of collagen-induced arthritis; binding of the immunodominant type II collagen 256-270 glycopeptide to H-2A^{q} and H-2A^{p} molecules. *Eur. J. Immunol.* **1998,** *28* (2), 755-767.

16. Rosloniec, E. F.; Whittington, K. B.; Brand, D. D.; Myers, L. K.; Stuart, J. M. Identification of MHC class II and TCR binding residues in the type II collagen immunodominant determinant mediating collagen-induced arthritis. *Cell. Immunol.* **1996,** *172* (1), 21-28.

17. Holm, B.; Backlund, J.; Recio, M. A. F.; Holmdahl, R.; Kihlberg, J. Glycopeptide specificity of helper T cells obtained in mouse models for rheumatoid arthritis. *Chembiochem* **2002,** *3* (12), 1209-1222.

18. Atherton, E.; Sheppard, R. C. *Solid phase peptide synthesis: A practical approach*. Oxford University Press: Oxford, 1989.

19. (a) Syed, B. M.; Gustafsson, T.; Kihlberg, J. 9-BBN as a convenient protecting group in functionalisation of hydroxylysine. *Tetrahedron* **2004,** *60* (26), 5571-5575; (b) Ankner, T.; Norberg, T.; Kihlberg, J. Mild oxidative cleavage of 9-BBN-protected amino acid derivatives. *Eur. J. Org. Chem.* **2015,** (17), 3767-3770.

20. (a) Corey, E. J.; Shimoji, K. Total synthesis of the major human urinary metabolite of prostaglanding-D2, a key diagnostic indicator. *J. Am. Chem. Soc.* **1983,** *105* (6), 1662-1664; (b) Keinan, E.; Sinha, S. C.; Singh, S. P. Thermostable enzymes in organic-synthesis. 5. Total Synthesis of S-(+)-Z-dodec-3-en-11-olide (ferrulactone-II) using a TBADH-generated bifunctional chiron. *Tetrahedron* **1991,** *47* (26), 4631-4638.

21. (a) Nahm, S.; Weinreb, S. M. N-methoxy-N-methylamides as effective acylating agents. *Tetrahedron Lett.* **1981,** *22* (39), 3815-3818; (b) Hussey, R. J.; Coates, L.; Gill, R. S.; Wright, J. N.; Sarwar, M.; Coker, S.; Erskine, P. T.; Cooper, J. B.; Wood, S.; Clarke, I. N.; Lambden, P. R.; Broadbridge, R.; Shoolingin-Jordan, P. M. Crystallization and preliminary X-ray diffraction

analysis of the protease from Southampton norovirus complexed with a Michael acceptor inhibitor. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2010,** *66*, 1544-1548.

22. Pelay, G. M.; Garcia-Ramos, Y.; Tulla, P. J.; Albericio, P. F.; Martin, L. M. J. Synthetic process for the manufacture of pipecolidepsin compounds, WO 2014/108526 A l. 2014.

23. Lee, C. S.; Choy, N.; Son, Y. C.; Park, C.; Choi, H.; Moon, K. Y.; Jung, W. H.; Kim, S. C.; Yoon, H. A key intermediate for irreversible HIV protease inhibitors: Synthesis of optically pure N-Cbz-5S-amino-6-phenyl-hex-3Z-enoic acid. *Bull. Korean Chem. Soc.* **1996,** *17* (3), 294-295.

24. Broddefalk, J.; Backlund, J.; Almqvist, F.; Johansson, M.; Holmdahl, R.; Kihlberg, J. T cells recognize a glycopeptide derived from type II collagen in a model for rheumatoid arthritis. *J. Am. Chem. Soc.* **1998,** *120* (31), 7676-7683.

25. Dessen, A.; Lawrence, C. M.; Cupo, S.; Zaller, D. M.; Wiley, D. C. X-ray crystal structure of HLA-DR4 (DRA*0101, DRB1*0401) complexed with a peptide from human collagen II. *Immunity* **1997,** *7* (4), 473-481.

26. (a) Ferrante, A.; Templeton, M.; Hoffman, M.; Castellini, M. J. The Thermodynamic mechanism of peptide-MHC class II complex formation is a determinant of susceptibility to HLA-DM. *J. Immunol.* **2015,** *195* (3), 1251-1261; (b) Bello, M.; Correa-Basurto, J. Molecular dynamics simulations to provide insights into epitopes coupled to the soluble and membrane-bound MHC-II complexes. *Plos One* **2013,** *8* (8); (c) Nguyen, T. B.; Jayaraman, P.; Bergseng, E.; Madhusudhan, M. S.; Kim, C. Y.; Sollid, L. M. Unraveling the structural basis for the unusually rich association of human leukocyte antigen DQ2.5 with class-II-associated invariant chain peptides. *J. Biol. Chem.* **2017,** *292* (22), 9218-9228.

27. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graphics Modell.* **1996,** *14* (1), 33-38.

28. Andersson, E. C.; Hansen, B. E.; Jacobsen, H.; Madsen, L. S.; Andersen, C. B.; Engberg, J.; Rothbard, J. B.; McDevitt, G. S.; Malmstrom, V.; Holmdahl, R.; Svejgaard, A.; Fugger, L. Definition of MHC and T cell receptor contacts in the HLA-DR4-restricted immunodominant epitope in type II collagen and characterization of collagen-induced arthritis in HLA-DR4 and human CD4 transgenic mice. *Proc. Natl. Acad. Sci. U. S. A.* **1998,** *95* (13), 7574-7579.

29. Oost, T.; Sukonpan, C.; Brewer, M.; Goodnough, M.; Tepp, W.; Johnson, E. A.; Rich, D. H. Design and synthesis of substrate-based inhibitors of botulinum neurotoxin type B metalloprotease. *Biopolymers* **2003,** *71* (6), 602-619.

30. Broddefalk, J.; Forsgren, M.; Sethson, I.; Kihlberg, J. Preparation of a diglycosylated hydroxylysine building block used in solid-phase synthesis of a glycopeptide from type II collagen. *J. Org. Chem.* **1999,** *64* (24), 8948-8953.

31. Hill, C. M.; Liu, A.; Marshall, K. W.; Mayer, J.; Jorgensen, B.; Yuan, B.; Cubbon, R. M.; Nichols, E. A.; Wicker, L. S.; Rothbard, J. B. Exploration of requirements for peptide binding to HLA DRB1*0101 and DRB1*0401. *J. Immunol.* **1994,** *152* (6), 2890-2898.

32. Michaelsson, E.; Andersson, M.; Engstrom, A.; Holmdahl, R. Identification of an immunodominant type-II collagen peptide recognized by T-cells in H-2Q mice - self tolerance at the level of determinant selection. *Eur. J. Immunol.* **1992,** *22* (7), 1819-1825.

33. (a) Stern, L. J.; Brown, J. H.; Jardetzky, T. S.; Gorga, J. C.; Urban, R. G.; Strominger, J. L.; Wiley, D. C. Crystal structure of the human class II MHC protien HLA-DR1 complexed with an influenza virus peptide. *Nature* **1994,** *368* (6468), 215-221; (b) Stern, L. J.; Wiley, D. C. Antigenic peptide binding by class I and class II histocompatibility proteins. *Structure* **1994,** *2*

(4), 245-251; (c) Jardetzky, T. S.; Brown, J. H.; Gorga, J. C.; Stern, L. J.; Urban, R. G.; Strominger, J. L.; Wiley, D. C. Crystallographic analysis of endogenous peptides associated with HLA-DR1 suggests a common, polyproline II-like conformation for bound peptides. *Proc. Natl. Acad. Sci. U. S. A.* **1996,** *93* (2), 734-738; (d) Ghosh, P.; Amaya, M.; Mellins, E.; Wiley, D. C. The structure of an intermediate in class II MHC maturation: CLIP bound to HLA-DR3. *Nature* **1995,** *378* (6556), 457-462; (e) Fremont, D. H.; Hendrickson, W. A.; Marrack, P.; Kappler, J. Structures of an MHC class II molecule with covalently bound single peptides. *Science* **1996,** *272* (5264), 1001-1004; (f) McFarland, B. J.; Katz, J. F.; Sant, A. J.; Beeson, C. Energetics and cooperativity of the hydrogen bonding and anchor interactions that bind peptides to MHC class II protein. *J. Mol. Biol.* **2005,** *350* (1), 170-183; (g) McFarland, B. J.; Katz, J. F.; Beeson, C.; Sant, A. J. Energetic asymmetry among hydrogen bonds in MHC class II-peptide complexes. *Proc. Natl. Acad. Sci. U. S. A.* **2001,** *98* (16), 9231-9236; (h) McFarland, B. J.; Beeson, C.; Sant, A. J. Cutting edge: A single, essential hydrogen bond controls the stability of peptide-MHC class II complexes. *J. Immunol.* **1999,** *163* (7), 3567-3571.

34. *MODDE*, 11; Umetrics AB: Umeå, Sweden, 2015.

35. (a) *Protein Preparation Wizard 2014-1; Epik version 2.4, Schrödinger, LLC, New York, NY, 2014; Impact version 5.9, Schrödinger, LLC, New York, NY, 2014; Prime version 3.2, Schrödinger, LLC, New York, NY, 2014*; (b) Olsson, M. H. M.; Sondergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions. *J. Chem. Theory Comput.* **2011,** *7* (2), 525-537; (c) Sastry, G. M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput.-Aided Mol. Des.* **2013,** *27* (3), 221-234.

36. *Schrödinger Release 2014-1: Maestro, version 9.7, Schrödinger, LLC, New York, NY, 2014.*

37. *MacroModel, version 10.3, Schrödinger, LLC, New York, NY, 2014*.

38. *Molecular Operating Environment (MOE)*, Chemical Computing Croup Inc., 1010 Sherbrooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2014, 2014.0901.

39. (a) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges - the RESP model. *J. Phys. Chem.* **1993,** *97* (40), 10269-10280; (b) Dupradeau, F.-Y.; Pigache, A.; Zaffran, T.; Savineau, C.; Lelong, R.; Grivel, N.; Lelong, D.; Rosanski, W.; Cieplak, P. The R.E.D. tools: advances in RESP and ESP charge derivation and force field library building. *Phys. Chem. Chem. Phys.* **2010,** *12* (28), 7821-7839; (c) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Jr., J. A. M.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Gaussian, Inc., Wallingford CT, 2009; (d) Wang, F.; Becker, J.-P.; Cieplak, P.; Dupradeau, F.-Y., R.E.D. Python: Object oriented programming

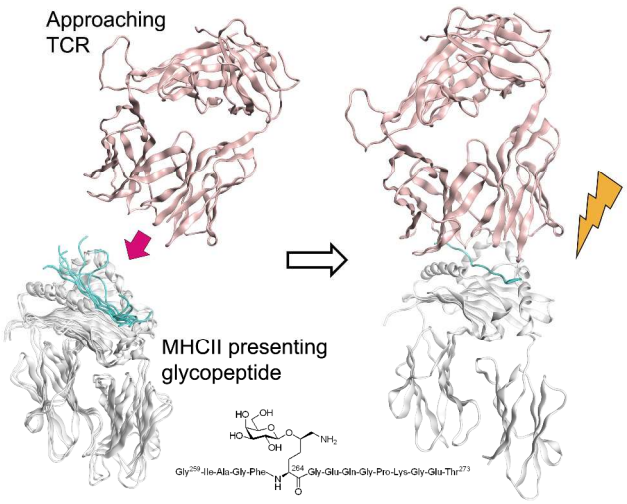
for Amber force fields. *Université de Picardie - Jules Verne, Sanford Burnham Prebys Medical Discovery Institute* **2013**; (e) Vanquelef, E.; Simon, S.; Marquant, G.; Garcia, E.; Klimerak, G.; Delepine, J. C.; Cieplak, P.; Dupradeau, F.-Y. R.E.D. Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments. *Nucleic Acids Res.* **2011,** *39*, W511-W517.

40. Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; B. Roberts; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; A. W. Goetz; Kolossváry, I.; Wong, K. F.; Paesani, F.; J. Vanicek; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; T. Steinbrecher; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER 12*, University of California, San Fransisco, 2012.

41. High Performance Computer Center North (HPC2N), http://www.hpc2n.umu.se/. Umeå, Sweden.

42. Gracia, L. Weill Cornell Medical College, 1300 York Ave, New York, 10065 NY.

Table of Contents graphic.


Approaching
TCR
MHCII presenting
glycopeptide
Gly259-Ile-Ala-Gly-Phe-
264
Gly-Glu-Gln-Gly-Pro-Lys-Gly-Glu-Thr273

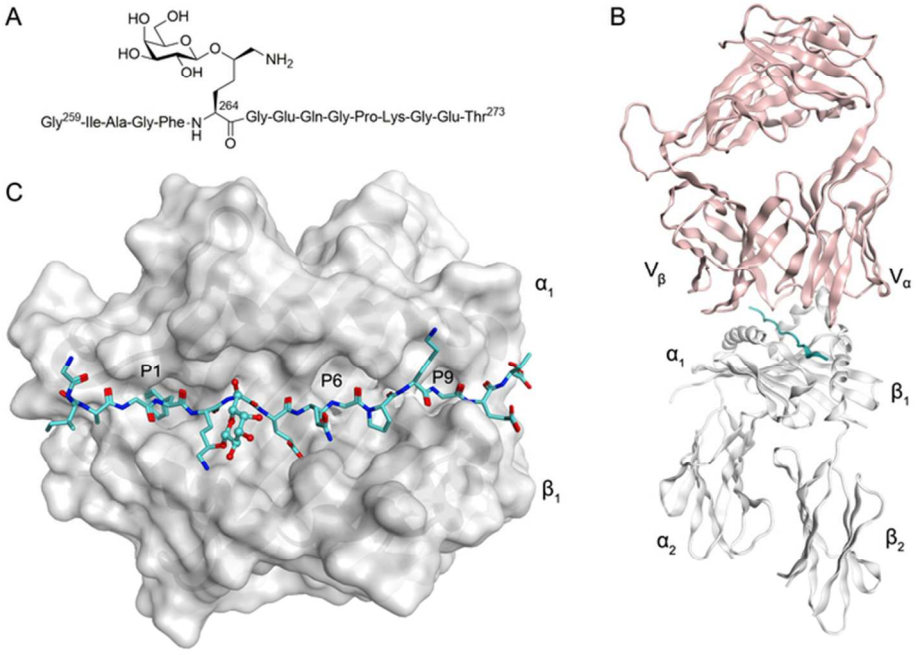


Figure 1. Structure of glycopeptide 1, the DR4/1 complex, and the multicomponent system consisting of the DR4/peptide/TCR complex. A) Chemical structure of glycopeptide CII259–273, 1. B) A three dimensional structure of the multicomponent system consisting of the DR4/peptide/TCR complex (based on PDB:1J8H12a, where an influenza virus peptide is positioned in the binding groove). The peptide is located between the α1- and β1-helices of DR4, being presented to the TCR that is approaching from above. DR4 is displayed in grey, the peptide in cyan, and the TCR in pink. C) A structural model of the glycopeptide 1 bound to DR4. Phe263 is positioned in the P1 pocket, Gly268 and Gly271 are positioned next to the to P6 and P9 pockets, respectively. The carbohydrate moiety of GalHyl264 is protruding out from the binding pocket (displayed in ball and stick). A surface has been added to the DR4 protein to display the binding pockets. Carbon, oxygen, and nitrogen atoms in 1 are displayed in cyan, red, and blue, respectively.

67x45mm (300 x 300 DPI)

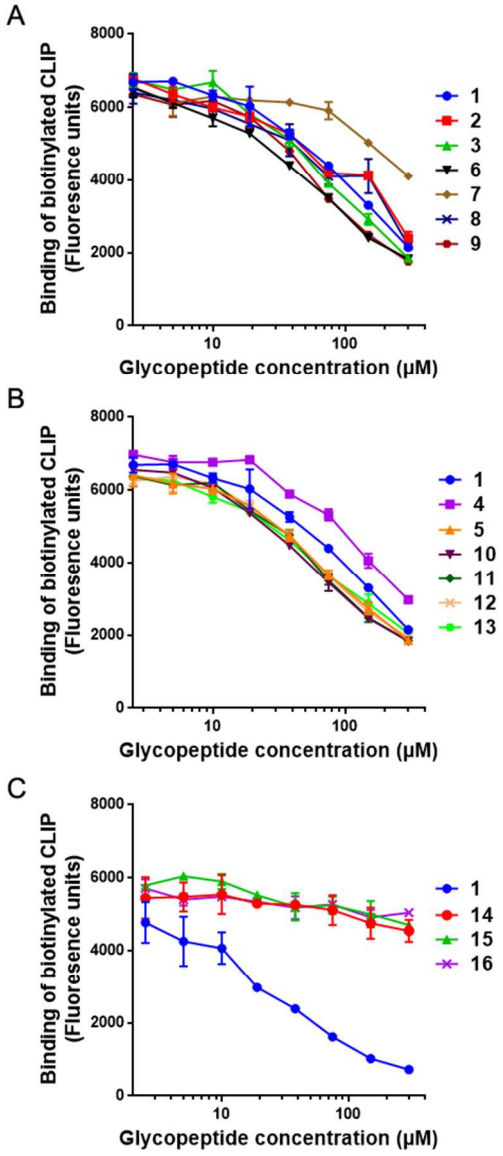


Figure 4. Binding of glycopeptide 1, modified glycopeptides 2–13, and glycopeptide isosteres 14–16, to DR4. Increasing concentrations of 1–16 and a fixed concentration of biotinylated CLIP tracer were incubated with DR4. The amount of CLIP tracer bound to DR4 was thereafter measured in a time-resolved fluoroimmunoassay. Each point represents the average of duplicates and error bars show ± one standard deviation. A) Glycopeptide 1 compared to 2–3 and 6–9, which are the glycopeptides that elicit T-cell responses when bound to DR4. B) Glycopeptide 1 compared to 4–5 and 10–13, which did not elicit T-cell responses in complexes with DR4. C) Glycopeptide 1 compared to isosteres 14–16.

177x400mm (72 x 72 DPI)

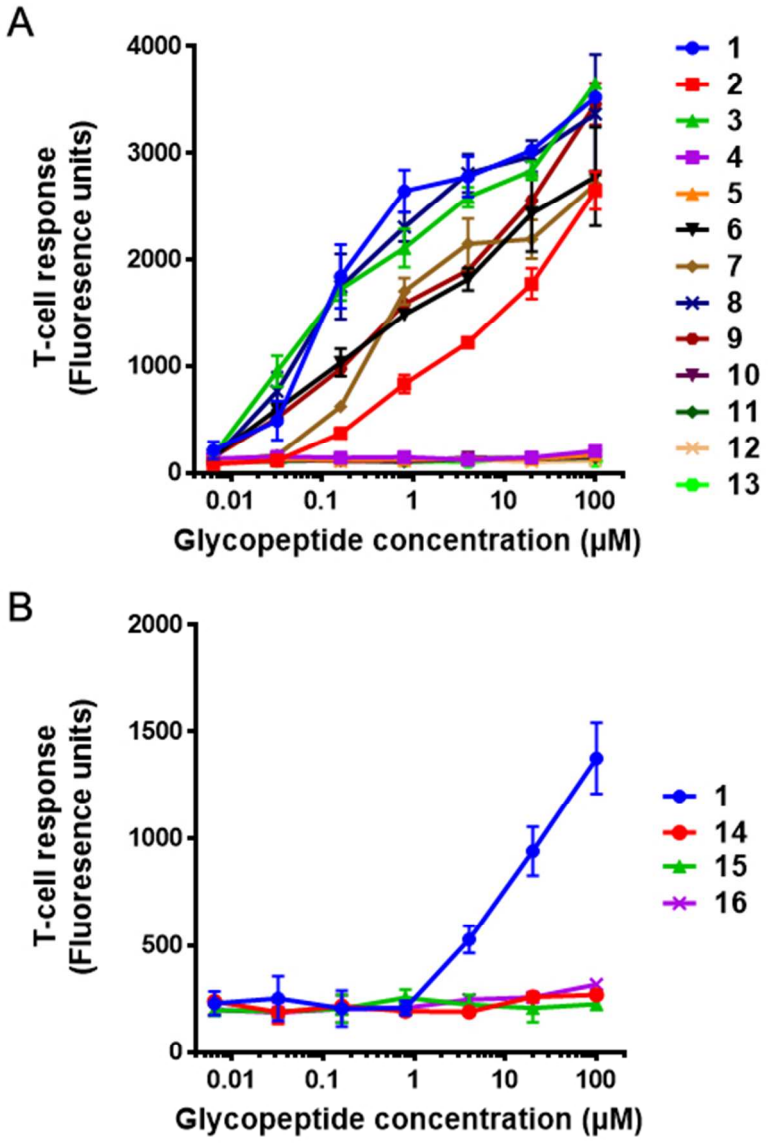


Figure 5. Recognition of glycopeptide 1, modified glycopeptides 2–13 (A), and glycopeptide isosteres 14–16 (B), by the DR4 restricted T-cell hybridoma mDR1.1, after incubation with antigen presenting spleenocytes and increasing concentration of the glycopeptides. If the DR4-glycopeptide complex is recognized by the T-cell hybridoma, secretion of IL-2 into the supernatant will occur, which is quantified by an ELISA assay. Each point represents the average of duplicates and error bars show ± one standard deviation.

183x272mm (72 x 72 DPI)

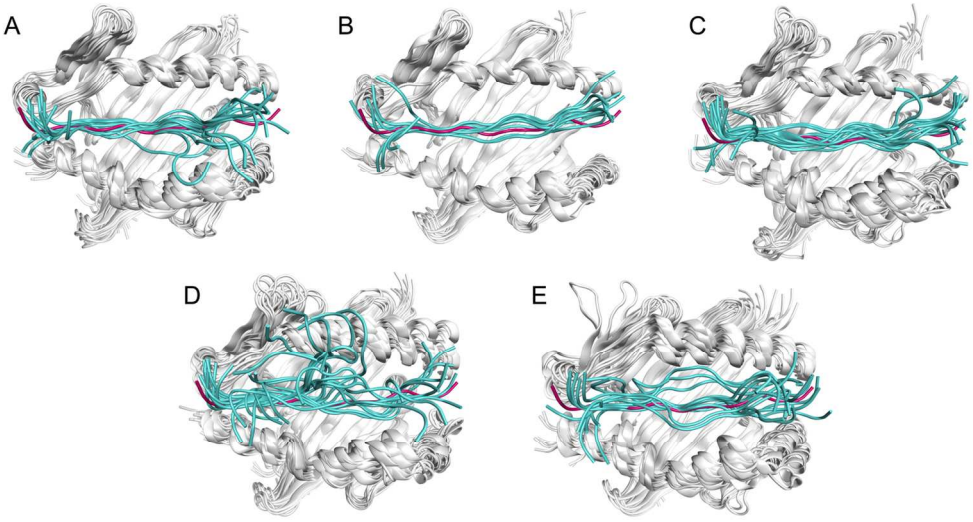


Figure 6. Snapshots from the MD simulations displaying the dynamics of the DR4 systems. A) Glycopeptide 1, B) glycopeptide 2, C) glycopeptide 3, D) glycopeptide 4, E) glycopeptide 5, all in complex with DR4. For each simulation trajectory a cluster analysis of the glycopeptide was performed based on snapshots over the last 30 ns. One representative glycopeptide conformation for each cluster with at least 40 members is displayed. The complexes are displayed with the proteins in gray and the glycopeptides in green. The conformation of the glycopeptide 1 from the homology model is displayed in magenta, as a reference.

139x75mm (300 x 300 DPI)

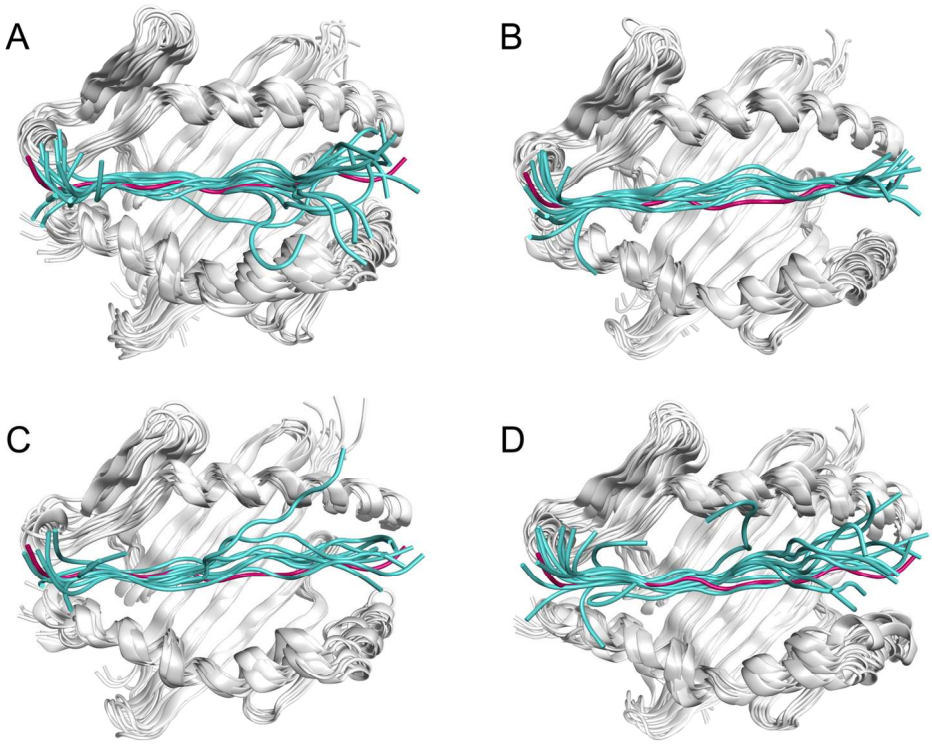


Figure 8. Snapshots from the MD simulations displaying the dynamics of the DR4 systems. A) Glycopeptide 1, B) glycopeptide 35, C) glycopeptide 36 and D) glycopeptide 37. For each MD simulation a cluster analysis of the glycopeptides has been performed based on snapshots over the last 30 ns. One representative glycopeptide conformation for each cluster with at least 40 members is displayed. The tertiary structure is displayed with the DR4 protein in gray and the glycopeptide in yellow. The conformation of the glycopeptide 1 from the prepared homology model is displayed in magenta as a reference.

139x108mm (300 x 300 DPI)

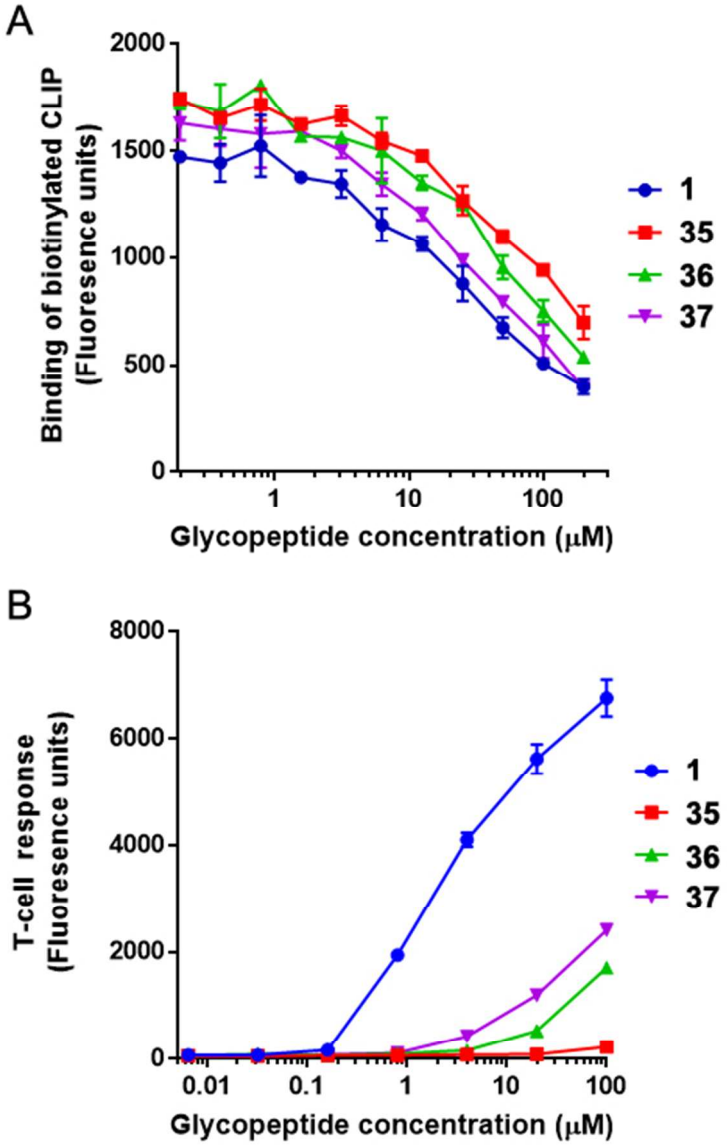


Figure 9. Binding to DR4 and TCR recognition for glycopeptide 1 compared to 35–37. A) Binding of glycopeptide 1 to DR4 compared to modified glycopeptides 35–37. Increasing concentrations of 1 and 35–37 and a fixed concentration of biotinylated CLIP tracer were incubated with DR4. The amount of CLIP tracer bound to DR4 was thereafter measured in a time-resolved fluoroimmunoassay. Each point represents the average of duplicates and error bars show ± one standard deviation. B) TCR recognition of glycopeptide 1 and modified glycopeptides 35–37 by the DR4 restricted T-cell hybridoma mDR1.1, after incubation with antigen presenting spleenocytes and increasing concentration of the glycopeptides. If the DR4-glycopeptide complex is recognized by the T-cell hybridoma, secretion of IL-2 into the supernatant will occur, which is quantified by a sandwich ELISA. Each point represents the average of duplicates and error bars show ± one standard deviation.

172x272mm (72 x 72 DPI)

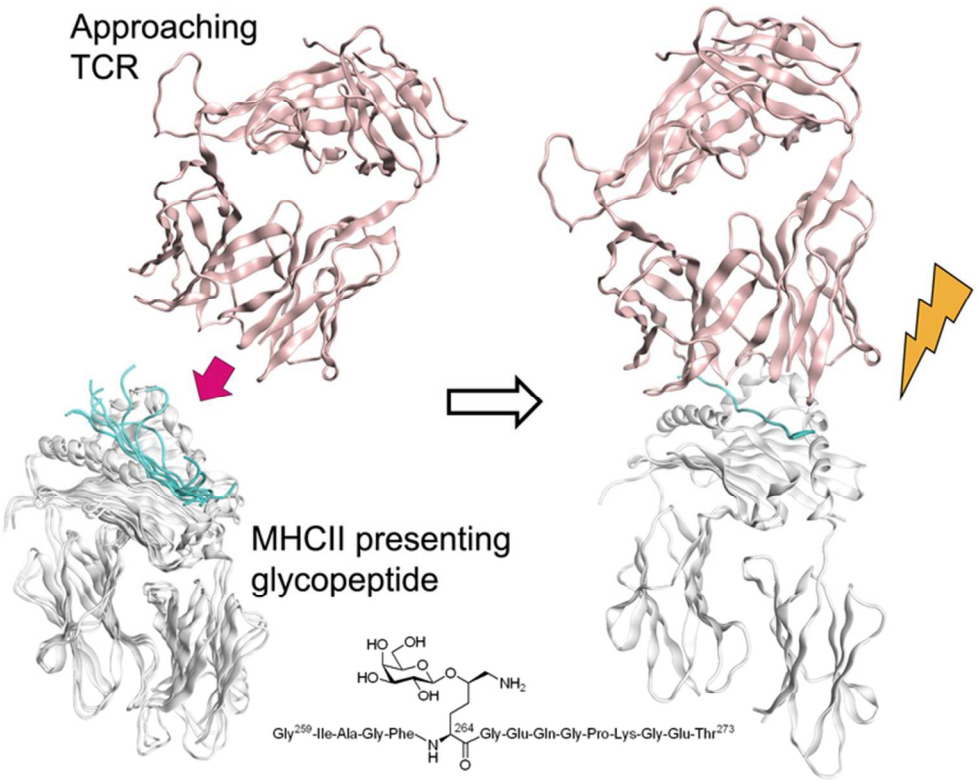


Table of contents graphic

66x52mm (300 x 300 DPI)