Organic & Biomolecular Chemistry

PAPER

View Article Online
View Journal | View Issue

Open Access Article. Published on 21 April 2015. Downloaded on 12/5/2025 12:49:11 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Cite this: *Org. Biomol. Chem.*, 2015, **13**, 6203

# Hydroxyethylene isosteres introduced in type II collagen fragments substantially alter the structure and dynamics of class II MHC $A^q$/glycopeptide complexes†

Cecilia Lindgren,[a] Ida E. Andersson,[a,b] Lotta Berg,[a] Doreen Dobritzsch,[c] Changrong Ge,[b] Sabrina Haag,[b] Urszula Uciechowska,[a] Rikard Holmdahl,[b] Jan Kihlberg*[c] and Anna Linusson*[a]

Received 26th February 2015,
Accepted 21st April 2015

DOI: 10.1039/c5ob00395d

www.rsc.org/obc

Class II major histocompatibility complex (MHC) proteins are involved in initiation of immune responses to foreign antigens *via* presentation of peptides to receptors of $CD4^+$ T-cells. An analogous presentation of self-peptides may lead to autoimmune diseases, such as rheumatoid arthritis (RA). The glycopeptide fragment CII259–273, derived from type II collagen, is presented by $A^q$ MHCII molecules in the mouse and has a key role in development of collagen induced arthritis (CIA), a validated model for RA. We have introduced hydroxyethylene amide bond isosteres at the $Ala^{261}$–$Gly^{262}$ position of CII259–273. Biological evaluation showed that $A^q$ binding and T cell recognition were dramatically reduced for the modified glycopeptides, although static models predicted similar binding modes as the native type II collagen fragment. Molecular dynamics (MD) simulations demonstrated that introduction of the hydroxyethylene isosteres disturbed the entire hydrogen bond network between the glycopeptides and $A^q$. As a consequence the hydroxyethylene isosteric glycopeptides were prone to dissociation from $A^q$ and unfolding of the $\beta_1$-helix. Thus, the isostere induced adjustment of the hydrogen bond network altered the structure and dynamics of $A^q$/glycopeptide complexes leading to the loss of $A^q$ affinity and subsequent T cell response.

## Introduction

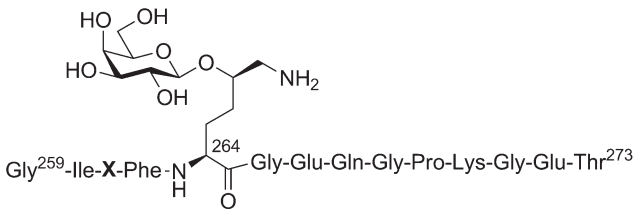
Rheumatoid arthritis (RA) is an autoimmune disease that mainly affects peripheral joints by causing inflammation that destroys cartilage and bone. The cause of the disease is not yet known, but a genetic link has been found to the expression of the class II major histocompatibility complex (MHC) proteins HLA-DR1 and -DR4.[1] In collagen-induced arthritis (CIA), a validated mouse model for studies of RA,[2] development of disease is linked to expression of $A^q$ MHC II molecules.[3] Class II MHC proteins are found in antigen presenting cells (APC) where they are associated with peptides derived from endocytosed protein antigens. The complexes are then transported to the surface of the APC and presented to T-cell receptors (TCR) on circulating $CD4^+$ T-cells, leading to initiation of an immune response towards the original antigen.

CII259–273 **1** (Fig. 1) – a glycopeptide fragment derived from type II collagen – is presented to $CD4^+$ T-cells by $A^q$ MHC molecules in CIA.[4] Glycopeptide **1** has also been proven to work as a vaccine that effectively prevents development of disease in CIA when administered as a complex with $A^q$.[5] The complex also reduces disease progression and severity in a mouse model of chronic relapsing arthritis.[5] In humans, glycopeptide **1** binds to DR4 and the complex is recognized by T-cells isolated from patients suffering from severe RA, suggesting **1** to have a central role in RA just as in CIA.[6] The minimal epitope in **1** required for binding to $A^q$, and stimulation of CII restricted T cell hybridomas obtained in CIA, consists of amino acid residues $Ile^{260}$–$Gln^{267}$.[7] Antigenic peptides bind between the $\alpha_1$- and $\beta_1$-helices of class II MHC proteins

[a]*Department of Chemistry, Umeå University, SE-901 87 Umeå, Sweden. E-mail: anna.linusson@umu.se; Tel: +46-90-7866890*
[b]*Medical Inflammation Research, Department of Medical Biochemistry and Biophysics, Karolinska Institute, SE-171 77 Stockholm, Sweden*
[c]*Department of Chemistry – BMC, Uppsala University, Box 576, SE-751 23 Uppsala, Sweden. E-mail: jan.kihlberg@kemi.uu.se; Tel: +46-18-4713801*

†Electronic supplementary information (ESI) available: $^1$H NMR and $^{13}$C NMR spectra of all new isolated compounds, RMSD for $A^q$ *vs.* simulation time, RMSD for the peptide binding groove *vs.* simulation time, RMSD for glycopeptides **1–4** *vs.* simulation time, RMSD for $Ile^{260}$–$Glu^{266}$ sequences *vs.* simulation time, RMSD for the $\alpha_1$ helix *vs.* simulation time, RMSD for the $\beta_1$ helix *vs.* simulation time, RMSD for the β-sheet of the $\alpha_1\beta_1$ domain *vs.* simulation time, the number of frames occupying each cluster for the MD simulations, superposition of frames from the cluster analysis displaying the glycopeptide/$A^q$ complex, superposition of frames from the cluster analysis displaying the glycopeptides **1–4**, visualization of hydrogen bonds formed between $A^q$ and glycopeptides **1–4**. See DOI: 10.1039/c5ob00395d

Open Access Article. Published on 21 April 2015. Downloaded on 12/5/2025 12:49:11 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.




**1** X = Ala$^{261}$-Gly$^{262}$
**2** X = Ala$^{261}\Psi$[(4*S*)-CH(OH)CH$_2$]Gly$^{262}$
**3** X = Ala$^{261}\Psi$[(4*R*)-CH(OH)CH$_2$]Gly$^{262}$
**4** X = Ala$^{261}\Psi$[COCH$_2$]Gly$^{262}$

**Fig. 1** Structures of the four investigated glycopeptides **1**–**4**. Chemical structure of the T-cell epitope CII259–273, **1**, and the modified glycopeptides **2**–**4**.

through a hydrogen bond network and by anchoring of side-chains in pockets in the binding groove.[8] Synthesis and biological evaluation of a comprehensive panel of synthetic analogues of **1** has revealed Ile$^{260}$ and Phe$^{263}$ to be crucial anchor residues that bind in hydrophobic pockets of A$^q$ (Fig. 2).[9,10] Moreover, the galactosylated hydroxylysine at position 264 protrudes out of the binding site and is recognized by the TCR with exquisite specificity.[11]

Modification of glycopeptides, through the incorporation of amide bond isosteres, has been used to probe molecular interactions in the MHC/glycopeptide/TCR system.[12–17] For instance, aza-amino acids,[15] (*E*)-alkene and ethylene isosteres[13] and oxazole modifications[12] have been employed. The Ile$^{260}$–GalHyl$^{264}$ fragment of glycopeptide **1** has been the target for the majority of these modifications as it contains essential features, such as the two MHC anchor residues[9] and the galactose moiety crucial for TCR recognition.[4,11] Incorporation of isosteres into **1** revealed the importance of having an intact network of hydrogen bonds for the stability of the complex with A$^q$.[13] Furthermore, the dynamics of the complex were found to govern the response from the TCR; complexes that were more rigid than those of A$^q$ and **1** resulted in weaker T cell responses.[10,13]

In this study, we further explore the molecular interactions in the MHC/glycopeptide/TCR system that are crucial for the development of CIA, and most likely also for RA, by synthesis and evaluation of two additional amide bond isosteres incorporated into glycopeptide **1**.

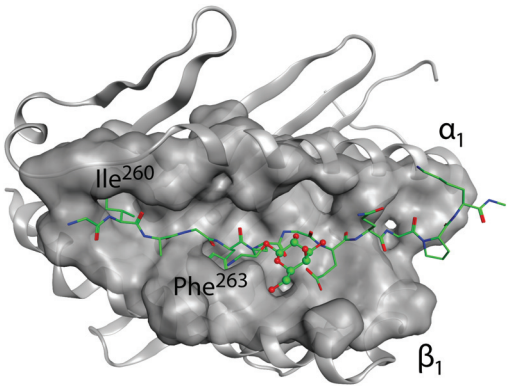



**Fig. 2** CII259–273 **1** in complex with A$^q$. A 3D stick model displaying the general binding mode of the glycopeptide **1** between the $\alpha_1$- and $\beta_1$-helix of A$^q$. The important anchor residues Ile$^{260}$ and Phe$^{263}$ are positioned in their respective binding pockets. The carbohydrate moiety is displayed in ball and stick with oxygen in red and carbon in green. The picture is based on a homology model of A$^q$.[14]

# Results

### Amide bond isosteres

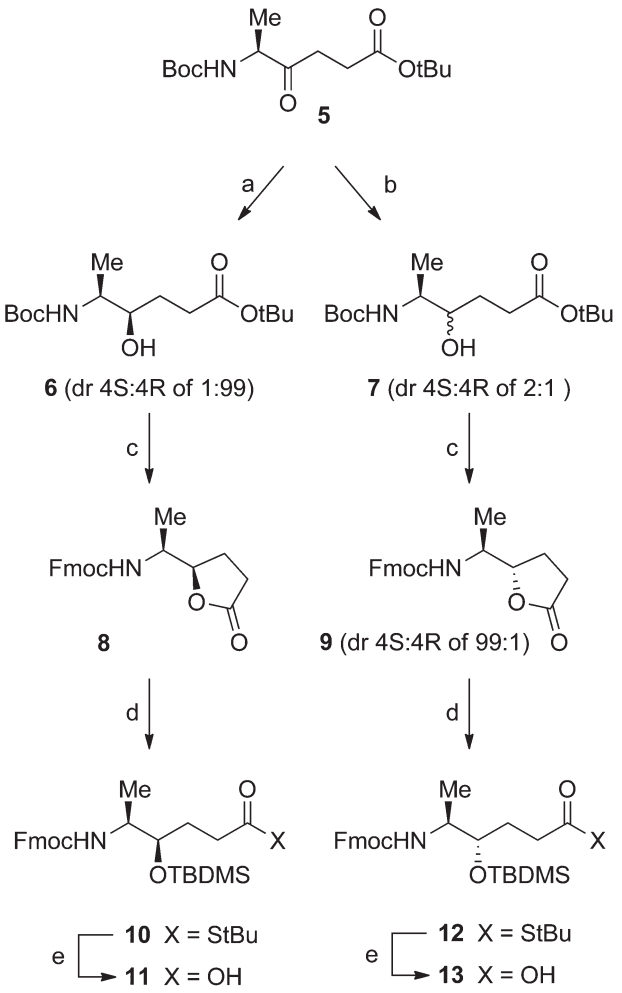
Hydroxyethylene isosteres were selected for incorporation at the Ala$^{261}$–Gly$^{262}$ position into glycopeptide **1** (Fig. 1, compounds **2** and **3**). The choice of isosteres and their location were based on previous investigations, where we explored replacement of this amide bond with ketomethylene, methyleneamine and (*E*)-alkene isosteres.[14] Evaluation of these glycopeptides showed that only ketomethylene-isostere modified glycopeptide **4** displayed comparable affinity for A$^q$ and T-cell responses to **1**, whereas both binding and cellular responses were lost for the others. We thus concluded that a hydrogen bond acceptor was required to maintain binding to A$^q$ and T-cell responses, whereas the hydrogen bond donor of the amide bond could be omitted. The hydroxyethylene isosteres utilized in this study introduce a chiral centre and a hydrogen bond donor capability at the Ala$^{261}$–Gly$^{262}$ position, while retaining the hydrogen bond acceptor capability.

### Synthesis

The diastereomeric Ala$^{261}$–Gly$^{262}$ hydroxyethylene isosteres were synthesized as dipeptide building blocks suitably protected for use in Fmoc-based solid-phase peptide synthesis (SPPS).[18] In short, ketomethylene derivative **5**[14] was first diastereoselectively reduced to the respective alcohol, followed by protective group manipulations and separation of diastereomeric mixtures. Diastereoselective reduction of ketomethylene derivative **5**[14] was performed using conditions guided by the work of Hoffman *et al.*[19] and Våbenø *et al.*[20] (Scheme 1). Treatment of **5** with LiAlH(O-*t*Bu)$_3$ in EtOH[19,20] gave alcohol **6** in 75% yield and with an excellent diastereomeric ratio (dr) 4*S* : 4*R* of 1 : 99 according to $^1$H NMR analysis. When **5** was instead treated with (*S*)-alpine-hydride in THF,[20] alcohol **7** was obtained in 84% yield although with a disappointing 4*S* : 4*R* dr of 2 : 1. The diastereomeric mixture **7** was difficult to separate by column chromatography and therefore additional transformations were performed. Both **6** and **7** were deprotected using TFA to remove the Boc and *t*Bu protecting groups followed by Fmoc protection, which produced the lactones **8** and **9** (4*S* : 4*R* of 2 : 1), respectively. The latter mixture of lactones was at this stage separated by chiral HPLC chromatography to afford **9** with a 4*S* : 4*R* dr of 99 : 1. In the final step, reaction with *t*BuSH-AlMe$_3$ converted the lactones into the corres-

 This journal is © The Royal Society of Chemistry 2015

Open Access Article. Published on 21 April 2015. Downloaded on 12/5/2025 12:49:11 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.




**Scheme 1** Reagents and conditions: (a) $LiAlH(O\text{-}tBu)_3$, EtOH, −78 °C, 75%; (b) (*S*)-alpine-hydride, THF, −78 °C, 84%; (c) (i) TFA, $CH_2Cl_2$, rt; (ii) FmocOSu, $Na_2CO_3$, dioxane, $H_2O$, 0 °C → rt, 86% for **8** from **6**, 48% for **9** from **7** (after separation of the diastereomers by chiral chromatography); (d) (i) *t*BuSH, $AlMe_3$, $CH_2Cl_2$, 0 °C; (ii) TBDMSCl, imidazole, DMF, rt, 55% for **10** from **8** and 85% for **12** from **9**; (e) LiOH, $H_2O_2$, THF, 0 °C → rt, 75% for **11** from **10** and 82% for **13** from **12**.

ponding γ-hydroxy thioesters that were immediately silylated with TBDMSCl to give **10** and **12** in order to prevent relactonization.[21] Finally, base-catalysed hydrolysis of the thioesters provided the target hydroxyethylene isostere building blocks **11** and **13**.

The isostere building blocks **11** and **13** were incorporated into the CII259–273 glycopeptide sequence using Fmoc-based SPPS[18] under standard conditions. The isosteric glycopeptides were cleaved from the solid support, deacetylated, and purified by reversed-phase HPLC to give the trifluoroacetate salts of **2** and **3** (Fig. 1) with overall yields of 18 and 28%, respectively, and in >98% purity. All glycopeptides were homogeneous, according to analytical reversed-phase HPLC and their structures were confirmed by $^1H$ NMR spectroscopy and MALDI-TOF mass spectrometry.

## Binding to the $A^q$ protein and recognition by T-cell hybridomas

The ability of the hydroxyethylene-modified glycopeptides **2** and **3** to bind to the $A^q$ protein compared to **1** was assessed in a competitive ELISA binding assay. Significant optimisation was undertaken to obtain a satisfactory signal to noise ratio in the ELISA, which led to the use of rather high concentrations of the biotinylated class II-associated invariant chain peptide (CLIP) tracer. As the CLIP binds with high affinity to $A^q$ higher concentrations of weaker inhibitors, such as native **1**, had to be employed in the ELISA. No binding could be detected for **2** and **3** even at the highest concentration of 1 mM where **1** completely competed out of the biotinylated CLIP tracer (Fig. 3). However, it is likely that a weak binding of **2** and **3** to $A^q$ would not be detected under the conditions that had to be employed in the ELISA. Replacement of the $Ala^{261}$–$Gly^{262}$ amide bond with the two different hydroxyethylene isosteres thus had severe effects on the glycopeptide abilities to bind to $A^q$. These results underscore our previous results that this particular amide bond is highly sensitive to modifications. At this position the corresponding ketomethylene-isostere **4** is the only modified glycopeptide so far that has been shown to bind equally well to $A^q$ as **1**.[12–14]

The T-cell response was investigated in a cell-based assay, where IL-2 secretion of $A^q$ restricted T-cell hybridoma lines was measured after incubation with antigen-presenting spleen cells expressing $A^q$ and increasing concentrations of the glycopeptides **1**, **2** or **3** (Fig. 4). Four $A^q$-restricted T-cell hybridomas that differed in their TCR specificity for the hydroxyl groups of the galactose moiety were investigated for their ability to recognize the glycopeptides. Despite the fact that no binding of **2** and **3** to $A^q$ was detected by the competition assay, both were

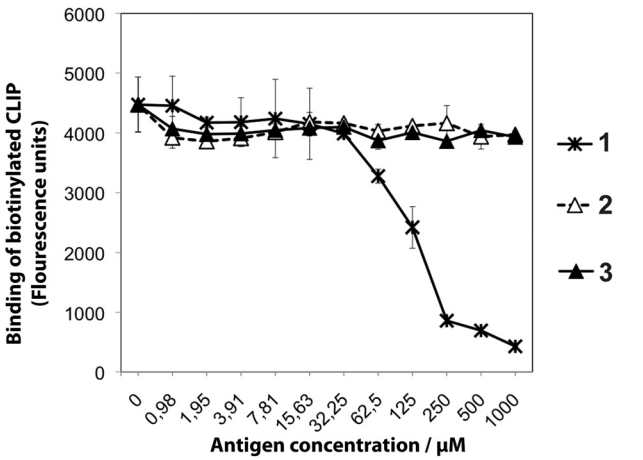



**Fig. 3** Binding of native glycopeptide **1** and hydroxyethylene isosteres **2** and **3** to $A^q$. In the competitive binding assay, increasing concentrations of **1**, **2**, or **3** were incubated with $A^q$ protein and a fixed concentration of a biotinylated CLIP peptide. After the incubation, the $A^q$-bound CLIP peptide was detected in a time-resolved fluoroimmunoassay using europium-labeled streptavidin. The points represent the average of duplicates and error bars are set to ±one standard deviation.

Open Access Article. Published on 21 April 2015. Downloaded on 12/5/2025 12:49:11 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

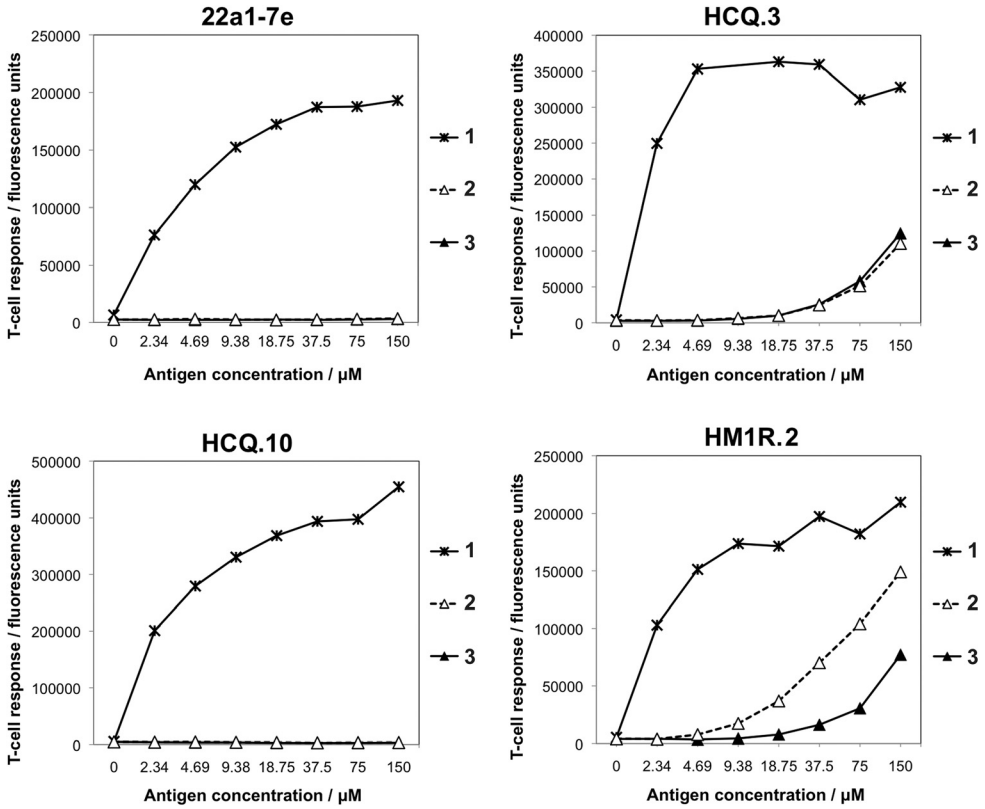



Fig. 4 Recognition of glycopeptides **1–3** by the $A^q$-restricted T-cell hybridomas HCQ.3, HCQ.10, HM1R.2, and 22a1-7e after incubation with antigen-presenting spleen cells expressing $A^q$ and increasing concentrations of **1–3**. Recognition of the peptide/MHC complex by the T-cell hybridomas results in secretion of IL-2, which was quantified by a sandwich ELISA using the DELFIA system.

recognized and stimulated IL-2 secretion by HCQ.3 and HM1R.2, while the other two hybridomas HCQ.10 and 22a1-7e did not give any response (Fig. 4). The HCQ.3 and HM1R.2 responses of **2** and **3** were significantly weaker compared to the responses elicited by both the non-modified **1** and ketomethylene-isosteric glycopeptide **4**; the latter of which elicited intermediate to strong responses from three of the four hybridomas.[14] This suggests that **2** and **3** do in fact bind to $A^q$ and were presented to the T-cell receptor in the functional assay, although too weakly to be detected in the binding assay.

That HCQ.10 and 22a1-7e were more sensitive towards the introduced modifications in the backbone compared to HCQ.3 and HMIR.2 has also been seen in previous studies.[13,14]

### Analysis of glycopeptide/$A^q$-complexes by MD simulations

Glycopeptides **1–4** were modelled into $A^q$ to understand if the structures of the complexes correlated with the binding affinity for $A^q$ and with the subsequent T cell responses. We found that the modified glycopeptides **2–4** displayed a binding mode almost identical to that of the native glycopeptide **1** in their respective complexes. Similar interactions were retained for all four glycopeptides although the position of the hydrogen bond acceptor had been altered for the hydroxylethylene-isosteric glycopeptides **2** and **3** compared to the carbonyl oxygen in **1** and **4**. As the static structures were unable to explain the observed differences in $A^q$ binding and T cell responses the dynamics of the isostere modified glycopeptide/$A^q$-complexes and the native system were studied over time using MD simulations. Repeated 60 ns MD simulations at different initial velocities were conducted for **1** (two times) and the modified glycopeptides **2–4** (three times) in a complex with $A^q$, using the static models as starting structures. The repeated runs will hereafter be referred to as (I), (II) and (III) and can be seen as samples from the complete conformational landscape of the systems. In general, the simulations reached equilibrium after 40 ns according to the root-mean-square deviation (RMSD) values (see the ESI† for RMSD plots). The analyses of the MD simulations were then focused on the binding mode and flexibility of the glycopeptides, and the hydrogen bonding network between the modified glycopeptides and $A^q$ over the last 20 ns of the repeated runs. The system of **1**/$A^q$ served as a reference to which the modified glycopeptides **2–4** have been compared.

### Flexibility of glycopeptide binding to $A^q$

During the simulations, glycopeptide **1** was bound in the groove-shaped binding site of $A^q$ (Fig. 5). The N-terminus remained firmly anchored in the binding groove, mediated by a network of intermolecular hydrogen bonds and by positioning of the anchor residues $Ile^{260}$ and $Phe^{263}$ in the P1 and P4 pockets, respectively. The C-terminus on the other hand

 This journal is © The Royal Society of Chemistry 2015

Open Access Article. Published on 21 April 2015. Downloaded on 12/5/2025 12:49:11 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

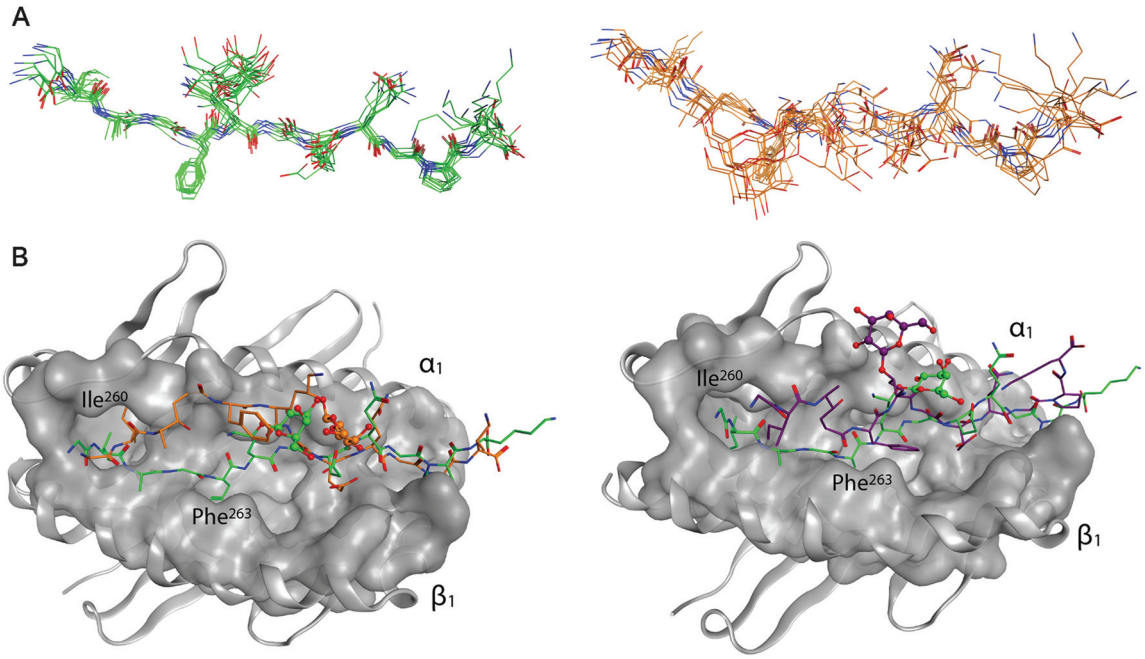



**Fig. 5** Snapshots from the MD-simulations comparing the mobility of the $A^q$/glycopeptide complexes of **1–3**. (A) Superposition of frames extracted over the last 20 ns of the MD simulations, comparing native **1** (I) (in green, left) to **2** (I) (in orange, right). Glycopeptides **2** and **3** (data not shown) both display a higher mobility than the native glycopeptide **1**. In order to extract representative snapshots, frames have been randomly selected from each of the top 10 clusters for the minimal epitope ($Ile^{260}$–$Gln^{267}$) over the last 20 ns of the MD simulations. (B) Comparison of structures of glycopeptides **1–3** bound to $A^q$. The representative snapshots have been randomly selected from the most highly populated cluster for the minimal epitope ($Ile^{260}$–$Gln^{267}$) over the last 20 ns of the MD simulations. In contrast to native **1**, isosteric glycopeptides **2** and **3** both dissociate from the binding groove of $A^q$, here illustrated by comparison of the structure of **1** (I) (green) to **2** (III) (in orange, left) and **3** (II) (in purple, right). A complete list with figures displaying superposed frames of the complexes from the last 20 ns of each MD simulations can be found in the ESI†. See the Experimental section for details on the clustering. The roman numbers I, II and III refer to the repeated MD simulations conducted for the same glycopeptide.

proved to be much more flexible and did not adopt a stable binding mode during the course of the simulations. This is consistent with previous MD simulations of the system,[10,13] as well as the fact that the minimal epitope required for binding to $A^q$ consists of the N-terminal octapeptide $Ile^{260}$–$Gln^{267}$.[7]

The *S*- and *R*-hydroxyethylene-isosteric glycopeptides **2** and **3** displayed a different dynamic pattern compared to **1**, with large variations between the repeated MD simulations (Table 1 and ESI†). The *S*-hydroxyethylene-isostere **2** showed larger movements; the N-terminus of the glycopeptide was not as firmly anchored in the binding groove as for **1** and displayed a high degree of flexibility (Fig. 5 and the ESI†). In fact, the anchor residue $Phe^{263}$ and occasionally also $Ile^{260}$ were leaving their respective binding pockets. In addition, a clear unfolding was observed in the $\beta_1$-helix at the point of the modification of the peptide. The *R*-hydroxyethylene-isostere **3** also displayed larger flexibility; again both anchor residues were leaving their respective binding pockets and the $\beta_1$-helix was unfolded adjacent to the isostere in **3** (Fig. 5). The simulations of the ketomethylene isostere **4** in complex with $A^q$ displayed a similar dynamic behaviour as the native glycopeptide **1**, including firmly anchored side chains of $Phe^{263}$ and $Ile^{260}$ and a stable $\beta_1$-helix.

The results of the repeated MD simulations are different scenarios or time periods. While some of the MD simulations for **2** and **3** displayed a behaviour indicative of a poor binder about to completely dissociate from $A^q$ other simulations indicated a plausible binder instead. It should be noted that in previous MD simulations[10,13] of the native glycopeptide **1**, we have not observed a behaviour similar to **2** and **3**, with the peptide dissociating from the binding groove, anchor residues leaving the P1 and P4 pockets, or unfolding of the $\beta_1$-helix.

### Hydrogen bond network

An extensive hydrogen bond network anchored glycopeptide **1** into $A^q$ during the molecular dynamics simulations. In total, 13 different hydrogen bonds were observed between the backbone of **1** and $A^q$, and 7 of these had over 40% occupancy during the simulation (Tables 1 and 2). The majority of hydrogen bonds were formed between the N-terminal part of **1** ($Gly^{259}$–$Gly^{265}$) and $A^q$, which was consistent with previous findings.[13,14] In this sequence, 8 out of 13 intermolecular hydrogen bonds were present, of which 5 were considered strong bonds. In particular two hydrogen bonds showed a high presence during the MD simulations, $Ala^{261}$(a)–$Asn^{82}$(s) and $Gln^{267}$(a)–$Trp^{61}$(s) (Table 2).

In the case of the modified glycopeptides **2** and **3**, insertion of the hydroxyethylene isostere at the $Ala^{261}$–$Gly^{262}$ position disturbed the overall hydrogen bond pattern during the MD simulations (Tables 1 and 2). Although hydrogen bonds could

Open Access Article. Published on 21 April 2015. Downloaded on 12/5/2025 12:49:11 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

**Table 1** Summary of findings in the MD simulations of $A^q$ in a complex with **1–4**, over the last 20 ns

| | | General comments | | | | |
|---|---|---|---|---|---|---|
| Glycopeptide | MD-run[a] | Increased flexibility of the glycopeptide? | Anchor residues, $Ile^{260}$ and $Phe^{263}$ | Stability of the $\beta_1$-helix | Number of strong hydrogen bonds[b] | Number of weak hydrogen bonds[c] |
| **1** | I | No | Anchored | Stable | 5 | 8 |
| | II | No | Anchored | Stable | 4 | 4 |
| **2** | I | Yes | Anchored | Stable | 0 | 7 |
| | II | No | Anchored | Unfolding | 3 | 5 |
| | III | Yes | Leaving | Unfolding | 2 | 5 |
| **3** | I | No | Anchored | Stable | 6 | 4 |
| | II | Yes | Leaving | Unfolding | 1 | 8 |
| | III | No | Anchored | Unfolding | 3 | 5 |
| **4** | I | No | Anchored | Stable | 4 | 6 |
| | II | No | Anchored | Stable | 5 | 6 |
| | III | No | Anchored | Stable | 4 | 4 |

[a] I, II and III represent the different MD simulations conducted for the same glycopeptide. [b] Strong hydrogen bonds have an occupancy above 40%. [c] Weak hydrogen bonds have an occupancy above 10% and under 40%.

**Table 2** Hydrogen bond occupancy, reported in %, between $A^q$ and **1–4** over the last 20 ns of the MD simulations

| | | **1** | | **2** | | | **3** | | | **4** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Glycopeptide residue (backbone)[a] | $A^q$ residue[b] | I[c] | II | I | II | III | I | II | III | I | II | III |
| $Gly^{259}$ (a) | $\beta His^{81}$ (s) | 22[d] | 46 | —[e] | — | 11 | — | — | — | 34 | — | 19 |
| $Ile^{260}$ (a) | $\alpha Asp^{56}$ (m) | — | — | — | — | — | — | 22 | — | — | — | — |
| $Ile^{260}$ (d) | $\alpha Thr^{53}$ (m) | — | — | — | — | 13 | — | — | — | — | — | — |
| $Ile^{260}$ (d) | $\alpha Leu^{52}$ (m) | — | — | — | — | — | — | — | — | — | — | 14 |
| $Ile^{260}$ (d) | $\alpha Ser^{54}$ (m) | 16 | 16 | — | — | — | — | 17 | 23 | 29 | 23 | — |
| **$Ala^{261}$ (a)**[f] | **$\beta Asn^{82}$ (s)** | **67** | **70** | *[g] | * | * | * | * | * | **66** | **66** | **74** |
| $Ala^{261}$ [OH][h] (a) | $\beta Asn^{82}$ (s) | * | * | — | 33 | — | — | — | — | * | * | * |
| $Ala^{261}$ [OH] (d) | $\beta Thr^{77}$ (m) | * | * | 24 | — | — | — | 10 | — | * | * | * |
| $Ala^{261}$ [OH] (d) | $\beta Val^{78}$ (m) | * | * | — | 57 | — | — | — | — | * | * | * |
| $Ala^{261}$ (d) | $\beta His^{81}$ (s) | — | — | — | — | — | — | — | — | — | 11 | — |
| $Ala^{261}$ (d) | $\beta Asn^{82}$ (s) | 26 | 18 | — | 36 | — | 26 | — | — | 40 | — | 61 |
| $Ala^{261}$ [OH] (d) | $\beta Asn^{82}$ (s) | * | * | — | — | — | 42 | — | — | * | * | * |
| $Phe^{263}$ (a) | $\alpha Tyr^{23}$ (s) | 70 | 19 | — | — | — | 67 | — | 39 | — | — | — |
| $Phe^{263}$ (a) | $\alpha Asn^{63}$ (s) | — | — | 22 | — | 49 | — | 16 | — | — | — | — |
| $Phe^{263}$ (d) | $\alpha Tyr^{9}$ (m) | 42 | 39 | 22 | 40 | — | 18 | — | 51 | 21 | — | 57 |
| $Lys^{264}$ (d) | $\beta Glu^{74}$ (s) | — | — | — | — | — | — | — | — | 36 | — | — |
| $Gly^{265}$ (a) | $\alpha Thr^{66}$ (s) | — | — | 18 | — | — | — | — | — | — | 20 | — |
| $Gly^{265}$ (a) | $\alpha Asn^{63}$ (s) | 25 | — | — | — | — | — | — | — | — | — | — |
| $Gly^{265}$ (a) | $\beta Arg^{70}$ (s) | — | — | — | — | — | — | — | — | — | — | 11 |
| $Gly^{265}$ (d) | $\alpha Asn^{63}$ (s) | 26 | 46 | 18 | 33 | — | 44 | — | 31 | — | 48 | — |
| $Glu^{266}$ (a) | $\alpha Thr^{66}$ (s) | — | — | — | — | 14 | — | — | — | — | — | 57 |
| $Glu^{266}$ (a) | $\alpha Asn^{70}$ (s) | 34 | — | — | 51 | 39 | 70 | 36 | 63 | 65 | 52 | — |
| $Glu^{266}$ (d) | $\beta Tyr^{30}$ (s) | 11 | — | — | 18 | — | 24 | 16 | — | — | 41 | — |
| **$Gln^{267}$ (a)** | **$\beta Trp^{61}$ (s)** | **61** | **60** | **24** | **38** | **52** | **67** | **65** | **57** | **61** | **64** | — |
| $Gly^{268}$ (a) | $\alpha His^{69}$ (s) | — | — | — | — | — | 18 | — | 15 | 21 | 10 | — |
| $Gly^{268}$ (d) | $\alpha Gln^{62}$ (m) | — | — | — | — | — | — | — | — | — | — | 29 |
| $Gly^{268}$ (d) | $\alpha His^{69}$ (s) | | | | | | | | | | | |
| $Gly^{268}$ (d) | $\alpha Asn^{70}$ (s) | 49 | — | — | — | 36 | 47 | 21 | 36 | 38 | 38 | — |
| $Pro^{269}$ (a) | $\alpha His^{69}$ (s) | 23 | — | 13 | — | — | — | 12 | — | — | 11 | — |

[a] a = hydrogen bond acceptor, d = hydrogen bond donor. [b] s = sidechain in $A^q$, m = main chain in $A^q$. [c] I, II and III represent the different MD simulations conducted for the same glycopeptide. [d] The hydrogen bond occupancy has been calculated with the aid of the hydrogen bond extension tool in VMD[23] with a donor–acceptor length of less than 3.3 Å and a donor–H–acceptor angle of minimum 160°. Only hydrogen bonds with an occupancy >10% are reported. [e] — = hydrogen bond occupancy <10%, the occupancy is therefore not reported. [f] The two hydrogen bonds that showed a high presence during the MD simulations; $Ala^{261}$(a)–$Asn^{82}$(s) and $Gln^{267}$(a)–$Trp^{61}$(s), are marked in bold. [g] * = not applicable. [h] [OH] represents the hydroxyethylene-isosteric glycopeptides **2** and **3**.

 This journal is © The Royal Society of Chemistry 2015

Open Access Article. Published on 21 April 2015. Downloaded on 12/5/2025 12:49:11 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

still be formed with the hydroxyl group of the isostere, these were weak, and the remaining hydrogen bond network between the two glycopeptides and $A^q$ was disrupted causing the number of strong hydrogen bonds to show a large variation over time between MD simulations. A weak hydrogen bond was formed with $Asn^{82}$(s) only in one MD-run, and thus no strong hydrogen bond corresponding to the strong $Ala^{261}$(a)–$Asn^{82}$(s) hydrogen bond in the **1**/$A^q$ complex was observed. The ketomethylene isostere introduced in **4**, on the other hand, maintained the $Ala^{261}$(a)–$Asn^{82}$(s) hydrogen bond (Table 2). The pattern of strong hydrogen bonds observed for **4** was more consistent with the pattern seen in the complex of **1** and $A^q$, despite the loss of hydrogen bond donation ability of $Gly^{262}$(d); a finding that agrees well with the observation that no hydrogen bond was formed with this residue in the MD simulations of native **1**.[14]

The disrupted hydrogen bond networks seen in the dynamics analyses of **2** and **3**, as compared with **1** and **4**, resulted in the larger flexibility as seen for those glycopeptides and their dissociation from the binding groove of $A^q$ discussed in the previous section. It is also likely that the altered hydrogen bonding was connected to the unfolding of the $\beta_1$-helix observed in the complexes of **2** and **3**. Thus, the dynamics calculations point to the fact that loss of one important hydrogen bond has a tremendous effect on both the structure and the stability of the entire complex. This loss most likely explains the dramatic difference in binding to $A^q$ between isosteres **2** and **3** and native **1** as seen in the competitive binding assay. Retention of the hydrogen bond between $Gln^{267}$(a)–$Trp^{61}$(s), which has a high occupancy for all four glycopeptides (Table 2), may explain the weak binding of **2** and **3** to $A^q$ as revealed by the stimulation of two of the four T cell hybridomas.

In conclusion, the MD-simulations displayed a far more detailed picture than the static models, and emphasized the importance of studying the dynamics to understand the behaviour and evolution over time of the $A^q$–glycopeptide complexes. While the binding mode for the native glycopeptide **1** and the ketomethylene-isosteric glycopeptide **4** was similar to the static model, the hydroxyethylene-isosteres **2** and **3** displayed a behaviour deviating from what could be seen in static models. Studies of the dynamics of $A^q$–glycopeptide complexes, but not static models, were therefore crucial to explain the results from the *in vitro* studies of class II MHC binding and T cell recognition.

### Alteration of the epitope presented to the T cell

As described above, the T-cell response patterns elicited by the modified glycopeptides **2** and **3** differed from that of **1**, being reduced for two hybridomas and not detectable for the other two. A reduced or eliminated T-cell response may be due to either a weaker binding to $A^q$, and/or that the epitope presented to the TCR has been changed. We conclude that the finding that two of the hybridomas respond weakly to **2** and **3**, while the other two do not respond at all, suggests that both effects are in operation. This conclusion is further supported by the MD simulations that showed increased flexibility of **2** and **3**, relative to **1**, in their complexes with $A^q$ and dissociation from the MHC binding groove indicating a reduced affinity. The increased flexibility, and the predicted unfolding of the $\beta_1$-helix of $A^q$, also suggests that the epitope presented to the T-cell has been altered, including the position of the $GalHyl^{264}$, an important TCR contact point.[4,11] Thus, even a small modification at the $Ala^{261}$–$Gly^{262}$ amide bond, not itself directly in contact with the TCR, had a major impact on the T-cell response pattern.

## Conclusions

Different amide bond isosteres display different properties and are valuable tools to probe interactions in complexes between peptides and proteins. With this in mind, we have further explored the effect of replacing the $Ala^{261}$–$Gly^{262}$ amide bond in the glycopeptide CII259–273 (**1**) with isosteres upon binding to the $A^q$ class II MHC molecule and subsequent T cell recognition of the complex. The results revealed an unexpected fine specificity as hydroxyethylene-isosteric glycopeptides **2** and **3** displayed a substantial loss of $A^q$ binding and T-cell response compared with **1**. In contrast, the previous introduction of a ketomethylene-isostere at the same position, to give glycopeptide **4**, resulted in maintained affinity for $A^q$ and T-cell responses was maintained.[14] Static models of the complexes of **1** and isosteric glycopeptides **2–4** with $A^q$ displayed comparable binding modes and thus did not provide any insight into the observed effects on isostere introduction. MD simulations, however, demonstrated that hydroxyethylene isosteric glycopeptides **2** and **3**, in contrast to **1** and **4**, were more flexible in their complexes with $A^q$ and also prone to dissociation from the binding groove. In addition, glycopeptides **2** and **3** led to unfolding of the $\beta_1$-helix of $A^q$. Analysis of the $A^q$/glycopeptide complexes revealed that replacement of the $Ala^{261}$–$Gly^{262}$ amide bond with hydroxyethylene-isosteres disrupted the hydrogen bond network in the complexes. This, in turn, led to the altered structure and dynamics and explains the dramatic effects on the $A^q$ binding and the T cell recognition.

## Experimental section

### Chemistry

**General.** All reactions were performed under an inert atmosphere with dry solvents under anhydrous conditions, unless otherwise stated. TLC was performed on Silica Gel 60 $F_{254}$ (Merck) with detection by using UV light and staining with alkaline aqueous $KMnO_4$ followed by heating. After workup, organic solutions were dried over $Na_2SO_4$ and filtered before being concentrated under reduced pressure. Flash column chromatography was performed on silica gel (60 Å, 230–400 mesh, Merck grade, 9385). Optical rotations were measured with a Perking-Elmer 343 polarimeter at 20 °C. $^1H$ and $^{13}C$ NMR spectra of the isostere dipeptide derivatives were

Open Access Article. Published on 21 April 2015. Downloaded on 12/5/2025 12:49:11 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

recorded at 298 K on a Bruker DRX-400 spectrometer at 400 MHz and 100 MHz, respectively. Calibration was done using the residual peak of the solvent as an internal standard [$CDCl_3$ ($CHCl_3$ $\delta_H$ 7.26 ppm, $CDCl_3$ $\delta_C$ 77.0 ppm) or $CD_3OD$ ($CD_2HOD$ $\delta_H$ 3.31 ppm, $CD_3OD$ $\delta_C$ 49.0 ppm)]. Spectra of glycopeptides **2** and **3** were recorded at 298 K on a Bruker Avance spectrometer at 500 MHz in $H_2O/D_2O$ (9 : 1) with $H_2O$ ($\delta_H$ 4.76) as an internal standard. HRMS data were recorded with fast atom bombardment ($FAB^+$) ionization. Analytical reversed-phase HPLC was performed on a Beckman System Gold HPLC equipped with a Supelco Discovery® Bio Wide Pore C18 column (250 × 4.6 mm, 5 μm) using a flow-rate of 1.5 mL $min^{-1}$ and detection at 214 nm. Preparative reversed-phase HPLC was performed by using a Supelco Discovery® Bio Wide Pore C18 column (250 × 21.2 mm, 5 μm) using the same eluent as for the analytical HPLC, a flow rate of 11 mL $min^{-1}$, and detection at 214 nm. Compound **5**[14,24] and glycopeptide **1**[4] were synthesized as described in the cited references.

*(4R,5S)-5-tert-Butoxycarbonylamino-4-hydroxy-hexanoic acid tert-butyl ester (**6**).* Ketone **5**[14,24] (160 mg, 0.531 mmol) dissolved in EtOH (8.0 mL) was added dropwise to a solution of $LiAlH(O\text{-}tBu)_3$ (278 mg, 1.06 mmol) in EtOH (5.5 mL) at −78 °C. After stirring for 150 min at −78 °C, a second portion of $LiAlH(O\text{-}tBu)_3$ (28 mg, 0.107 mmol) was added. After stirring for additional 60 min, citric acid (10%, aq.) was added and the mixture was warmed to rt followed by addition of EtOAc and $H_2O$. The aqueous phase was extracted with EtOAc and the combined organic layers were washed with $H_2O$ and brine, dried and concentrated under reduced pressure. The crude was purified by flash chromatography (*n*-heptane/EtOAc 3 : 1) to afford **6** (4*S*:4*R* 1 : 99, 122 mg, 75%) as a white solid. $[\alpha]_D^{20}$ = −19.5 (*c* = 1.0, $CHCl_3$); $^1H$ NMR ($CDCl_3$) $\delta$ 4.77 (br s, 1H, NH), 3.76–3.63 (m, 1H, NCH), 3.63–3.56 (m, 1H, HOC*H*), 2.48–2.32 (m, 2H, $CH_2CO_2$), 1.73–1.57 (m, 2H, $CH_2CH_2CO_2$), 1.43 (s, 9H, $C(CH_3)_3$), 1.42 (s, 9H, $C(CH_3)_3$), 1.10 (d, *J* = 6.8 Hz, 3H, $CH_3$); $^{13}C$ NMR ($CDCl_3$) $\delta$ 173.8 ($CO_2$), 156.0 (NCO), 80.6 ($C(CH_3)_3$), 79.5 ($C(CH_3)_3$), 74.3 (HO*C*H), 51.0 (NCH), 32.4 ($CH_2CO_2$), 28.4 ($C(CH_3)_3$), 28.1 ($CH_2CH_2CO_2$), 28.1 ($C(CH_3)_3$), 14.9 ($CH_3$).

*(4S/R,5S)-5-tert-Butoxycarbonylamino-4-hydroxy-hexanoic acid tert-butyl ester (7).* Ketone **5** (269 mg, 0.892 mmol) was dissolved in dry THF (22 mL) and cooled to −78 °C followed by dropwise addition of (*S*)-alpine-hydride (0.5 M in THF, 1.96 mL, 0.980 mmol). After stirring for 80 min at −78 °C, citric acid (10%, aq.) was added and the mixture was allowed to reach rt followed by the addition of $Et_2O$ and brine. The aqueous phase was extracted with $Et_2O$ and the combined organic layers were dried and concentrated under reduced pressure. The crude was purified by flash chromatography (*n*-heptane/EtOAc 5 : 1 → 2 : 1) to afford **7** (4*S* : 4*R* 2 : 1, 227 mg, 84%) as a colorless oil. **(4*S*)-7** (major isomer): $^1H$ NMR ($CDCl_3$) $\delta$ 4.74 (br s, 1H, NH), 3.66–3.57 (m, 1H, NCH), 3.53–3.47 (m, 1H, HOC*H*), 2.47–2.33 (m, 2H, $CH_2CO_2$), 1.81–1.60 (m, 2H, $CH_2CH_2CO_2$), 1.43 (s, 18H, $C(CH_3)_3$), 1.17 (d, *J* = 6.8 Hz, 3H, $CH_3$); $^{13}C$ NMR ($CDCl_3$) $\delta$ 173.9 ($CO_2$), 156.2 (NCO), 80.7 ($C(CH_3)_3$), 79.4 ($C(CH_3)_3$), 74.6 (HO*C*H), 50.7 (NCH), 32.3 ($CH_2CO_2$), 29.3 ($CH_2CH_2CO_2$), 28.4 ($C(CH_3)_3$), 28.1 ($C(CH_3)_3$), 18.3 ($CH_3$). **(4*R*)-7** (minor isomer): $^1H$ NMR ($CDCl_3$) $\delta$ 4.74 (br s, 1H, NH), 3.75–3.66 (m, 1H, NCH), 3.66–3.57 (m, 1H, HOC*H*), 2.47–2.20 (m, 2H, $CH_2CO_2$), 1.81–1.60 (m, 2H, $CH_2CH_2CO_2$), 1.43 (s, 18H, $C(CH_3)_3$), 1.10 (d, *J* = 6.8 Hz, 3H, $CH_3$); $^{13}C$ NMR ($CDCl_3$) $\delta$ 173.8 ($CO_2$), 156.2 (NCO), 80.7 ($C(CH_3)_3$), 79.5 ($C(CH_3)_3$), 74.4 (HO*C*H), 51.0 (NCH), 32.5 ($CH_2CO_2$), 29.3 ($CH_2CH_2CO_2$), 28.4 ($C(CH_3)_3$), 28.1 ($C(CH_3)_3$), 14.9 ($CH_3$).

*(5S)-5-[(1R)-1-(9H-Fluoren-9-ylmethoxycarbonylamino)-ethyl]-dihydro-furan-2-one (**12**).* Alcohol **6** (137 mg, 0.452 mmol) dissolved in $CH_2Cl_2$ (2.2 mL) was treated with TFA (0.9 mL) at rt for 24 h. The solution was concentrated under reduced pressure, the resulting residue was concentrated from $CHCl_3$ three times and was thereafter dissolved in dioxane (2.2 mL) and $H_2O$ (0.8 mL). FmocOSu (187 mg, 0.542 mmol) was added and the mixture was cooled to 0 °C followed by dropwise addition of $Na_2CO_3$ (58 mg, 0.542 mmol) dissolved in $H_2O$ (0.35 mL). The mixture was allowed to reach rt and was stirred for 24 h before dioxane was removed under reduced pressure. $CHCl_3$ and $H_2O$ were added and the aqueous phase was acidified by the addition of HCl (10%, aq.) at 0 °C to pH 1. The phases were separated and the aqueous phase was extracted with $CHCl_3$. The combined organic layers were washed with brine, dried and concentrated under reduced pressure. The resulting crude was purified by flash chromatography (*n*-heptane/EtOAc/AcOH 2 : 1 : 1% → 1 : 1 : 1%) to afford **8** (136 mg, 86%). $[\alpha]_D^{20}$ = −32.8 (*c* = 1.0, $CHCl_3$); $^1H$ NMR ($CD_3OD$) $\delta$ 7.78 (d, *J* = 7.4 Hz, 2H, Fmoc-arom), 7.62 (d, *J* = 7.4 Hz, 2H, Fmoc-arom), 7.40–7.34 (m, 2H, Fmoc-arom), 7.33–7.26 (m, 2H, Fmoc-arom), 4.47–4.31 (m, 3H, Fmoc-C$H_2$, OCH), 4.18 (t, *J* = 6.4 Hz, 1H, Fmoc-C*H*), 3.85–3.75 (m, 1H, NCH), 2.56–2.46 (m, 2H, $CH_2CO_2$), 2.28–2.14 (m, 1H, OCHC$H_2$), 2.05–1.92 (m, 1H, OCHC$H_2$), 1.17 (d, *J* = 7.0 Hz, 3H, $CH_3$); $^{13}C$ NMR ($CD_3OD$) $\delta$ 179.5 ($CO_2$), 158.2 (NCO), 145.1 (splitted), 142.5, 128.6, 128.0, 126.0 (splitted), 120.8, 84.2 (OCH), 67.4 (Fmoc-$CH_2$), 50.7 (NCH), 48.4 (Fmoc-*C*H), 29.0 ($CH_2CO_2$), 24.8 (OCH$CH_2$), 16.5 ($CH_3$).

*(5S)-5-[(1S/R)-1-(9H-Fluoren-9-ylmethoxycarbonylamino)-ethyl]-dihydro-furan-2-one (**9**).* Alcohol **7** (188 mg, 0.621 mmol) dissolved in $CH_2Cl_2$ (3.0 mL) was treated with TFA (1.3 mL) at rt for 19 h. The solution was concentrated under reduced pressure, the resulting residue was concentrated from $CHCl_3$ three times and was thereafter dissolved in dioxane (3.0 mL) and $H_2O$ (1.2 mL). FmocOSu (257 mg, 0.746 mmol) was added and the mixture was cooled to 0 °C followed by dropwise addition of $Na_2CO_3$ (79 mg, 0.746 mmol) dissolved in $H_2O$ (0.3 mL). The mixture was allowed to reach rt and was stirred for 22 h before dioxane was removed under reduced pressure. $H_2O$ and $CHCl_3$ were added, the phases were separated and the aqueous phase was extracted with $CHCl_3$. The combined organic layers were washed with brine, dried and concentrated under reduced pressure. The resulting crude was purified by flash chromatography (*n*-heptane/EtOAc 2 : 1 → 1 : 1) followed by purification by chiral preparative HPLC (Chiralpak AD column, 250 × 20 mm, 5 μm, with *n*-heptane/EtOH 70 : 30 as the eluent, 18 mL $min^{-1}$ flow rate and UV detection at 300 nm) to afford **9** (4*S* : 4*R* 99 : 1, 108 mg, 48%). $H^1$ NMR ($CD_3OD$)

 This journal is © The Royal Society of Chemistry 2015

Open Access Article. Published on 21 April 2015. Downloaded on 12/5/2025 12:49:11 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

$\delta$ 7.79 (d, $J$ = 7.6 Hz, 2H, Fmoc-arom), 7.65 (d, $J$ = 7.5 Hz, 2H, Fmoc-arom), 7.41–7.36 (m, 2H, Fmoc-arom), 7.34–7.27 (m, 2H, Fmoc-arom), 4.52–4.45 (m, 2H, Fmoc-$CH_2$, OC*H*), 4.39–4.32 (m, 1H, Fmoc-$CH_2$), 4.25–4.18 (m, 1H, Fmoc-C*H*), 3.87–3.79 (m, 1H, NCH), 2.61–2.47 (m, 1H, $CH_2CO_2$), 2.46–2.35 (m, 1H, $CH_2CO_2$), 2.26–2.15 (m, 1H, C$H_2CH_2CO_2$), 1.99–1.85 (m, 1H, C$H_2CH_2CO_2$), 1.19 (d, $J$ = 6.9 Hz, 3H, $CH_3$). A $^{13}$C NMR spectrum could not be obtained due to degradation during NMR analysis.

*(4R,5S)-4-(tert-Butyl-dimethyl-silanyloxy)-5-(9H-fluoren-9-ylmethoxy-carbonylamino)-hexanethioic acid tert-butyl ester (**10**).* $AlMe_3$ (1 M in heptane, 2.50 mL, 2.50 mmol) was added dropwise to a solution of *t*-BuSH (0.28 mL, 2.53 mmol) in $CH_2Cl_2$ (3.6 mL) at 0 °C. The mixture was stirred for 10 min at 0 °C, 10 min at rt followed by 10 min at 0 °C. Lactone **8** (106 mg, 0.301 mmol) dissolved in $CH_2Cl_2$ (1.9 m L) was added dropwise followed by stirring for 3 h and 30 min at 0 °C. The reaction was cooled to −78 °C and $Et_2O$ (11 mL) was added followed by dropwise addition of HCl (1 M, 3 mL). The resulting mixture was then stirred for 10 min before it was allowed to reach 0 °C. The workup was performed under cold conditions to prevent relactonization using ice-cold solvent and aqueous solutions. After addition of $Et_2O$ and HCl (1 M), the layers were separated, and the aqueous layer was extracted with $Et_2O$. The combined organic layers were washed with HCl (1 M, aq.), $NaHCO_3$ (satd, aq.) and brine, dried and concentrated under reduced pressure under cold conditions. The resulting crude product was mixed with TBDMSCl (197 mg 1.27 mmol) and imidazole (173 mg, 2.54 mmol), and dissolved in DMF (0.27 mL). After stirring for 20 h and 30 min at rt, $NaHCO_3$ was added followed by the addition of $H_2O$ and $CH_2Cl_2$. The phases were separated and the aqueous layer was extracted with $CH_2Cl_2$. The combined organic layers were washed with $H_2O$ and brine, dried, and concentrated under reduced pressure. Purification by flash chromatography twice (first using *n*-heptane/EtOAc 8 : 1, then using *n*-heptane/EtOAc 1 : 0 → 8 : 1) gave **10** (93 mg, 55%) as a colorless oil. $[\alpha]_D^{20}$ = −7.2 ($c$ = 1.0, $CHCl_3$); $^1$H NMR ($CDCl_3$) $\delta$ 7.76 (d, $J$ = 7.4 Hz, 2H, Fmoc-arom), 7.62–7.56 (m, 2H, Fmoc-arom), 7.43–7.37 (m, 2H, Fmoc-arom), 7.35–7.28 (m, 2H, Fmoc-arom), 4.85–4.76 (m, 1H, NH), 4.44–4.32 (m, 2H, Fmoc-C$H_2$), 4.21 (t, $J$ = 6.7 Hz, 1H, Fmoc-C*H*), 3.78–3.65 (m, 2H, NCH, OCH), 2.60–2.39 (m, 2H, $CH_2COS$), 1.80–1.70 (m, 2H, C$H_2CH_2COS$), 1.46 (s, 9H, $C(CH_3)_3$), 1.10 (d, $J$ = 5.5 Hz, 3H, $CH_3$), 0.91 (s, 9H, $C(CH_3)_3$), 0.08 (s, 3H, $SiCH_3$), 0.04 (s, 3H, $SiCH_3$); $^{13}$C NMR ($CDCl_3$) $\delta$ 199.7 (COS), 155.6 (NCO), 144.0 (splitted), 141.3, 127.6, 127.0, 125.0, 120.0, 73.5 (SiO*C*H), 66.5 (Fmoc-*C*$H_2$), 50.1 (NCH), 48.0 (*C*$(CH_3)_3$), 47.3 (Fmoc-*C*H), 40.5 (*C*$H_2$COS), 29.8 (C(*C*$H_3)_3$), 29.4 (*C*$H_2CH_2$COS), 25.9 (C(*C*$H_3)_3$), 18.1 (*C*$(CH_3)_3$), 14.2 ($CH_3$), −4.2 ($SiCH_3$), −4.6 ($SiCH_3$).

*(4S,5S)-4-(tert-Butyl-dimethyl-silanyloxy)-5-(9H-fluoren-9-ylmethoxy-carbonylamino)-hexanethioic acid tert-butyl ester (**12**).* $AlMe_3$ (2 M in heptane, 1.15 mL, 2.30 mmol) was added dropwise to a solution of *t*-BuSH (0.26 mL, 2.31 mmol) in $CH_2Cl_2$ (2.5 mL) at 0 °C. The mixture was stirred for 10 min at 0 °C, 10 min at rt followed by 10 min at 0 °C. Lactone **9** (96 mg, 0.273 mmol) dissolved in $CH_2Cl_2$ (3.5 mL) was added dropwise followed by stirring for 3 h at 0 °C. The solution was cooled to −78 °C and $Et_2O$ (10 mL) was added followed by dropwise addition of HCl (1 M, 2.7 mL). The resulting mixture was then stirred for 10 min before it was allowed to reach 0 °C. The workup was performed using ice-cold solvent and aqueous solutions to prevent relactonization. After addition of $Et_2O$ and HCl (1 M), the phases were separated, and the aqueous layer was extracted with $Et_2O$. The combined organic layers were washed with HCl (1 M, aq.), $NaHCO_3$ (satd, aq.) and brine, dried and concentrated under reduced pressure and cold conditions. The resulting crude product was immediately mixed with TBDMSCl (182 mg 1.20 mmol) and imidazole (157 mg, 2.30 mmol) and dissolved in DMF (1.0 mL). After stirring for 16 h at rt, $NaHCO_3$ (aq., satd.) was added followed by the addition of $H_2O$ and $CH_2Cl_2$. The phases were separated and the aqueous layer was extracted with $CH_2Cl_2$. The combined organic layers were washed with $H_2O$ and brine, dried, and concentrated under reduced pressure. Purification by flash chromatography (*n*-heptane/EtOAc 1 : 0 → 8 : 1) gave **12** (102 mg, 85%) as a colorless oil. $[\alpha]_D^{20}$ = +8.9 ($c$ = 1.0, $CHCl_3$); $^1$H NMR ($CDCl_3$) $\delta$ 7.77 (d, $J$ = 7.5 Hz, 2H, Fmoc-arom), 7.60 (d, $J$ = 7.5 Hz, 2H, Fmoc-arom), 7.43–7.37 (m, 2H, Fmoc-arom), 7.34–7.29 (m, 2H, Fmoc-arom), 4.87 (d, $J$ = 8.9 Hz, 1H, NH), 4.46–4.39 (m, 1H, Fmoc-C$H_2$), 4.39–4.31 (m, 1H, Fmoc-C$H_2$), 4.25 (t, $J$ = 7.0 Hz, 1H, Fmoc-C*H*), 3.83–3.73 (m, 1H, NCH), 3.68–3.61 (m, 1H, OC*H*), 2.50 (t, $J$ = 7.5 Hz, 2H, $CH_2CO$), 1.82–1.67 (m, 2H, C$H_2CH_2O$), 1.45 (s, 9H, $SC(CH_3)_3$), 1.15 (d, $J$ = 6.7 Hz, 3H, $CH_3$), 0.93 (s, 9H, $SiC(CH_3)_3$), 0.11 (s, 3H, $SiCH_3$), 0.10 (s, 3H, $SiCH_3$); $^{13}$C NMR ($CDCl_3$) $\delta$ 199.9 (COS), 155.9 (NCO), 144.0, 141.3, 127.6, 127.0, 125.1 (splitted), 119.9, 73.4 (CHO), 66.6 (Fmoc-*C*$H_2$), 49.1 (NCH), 47.9 (S*C*$(CH_3)_3$), 47.3 (Fmoc-*C*H), 40.5 (*C*$H_2CO$), 29.8 (SC(*C*$H_3)_3$), 29.2 (C$H_2CH_2CO$), 25.9 (SiC(*C*$H_3)_3$), 18.1 (Si*C*$(CH_3)_3$), 17.9 ($CH_3$), −4.3 ($SiCH_3$), −4.6 ($SiCH_3$).

*(4R,5S)-4-(tert-Butyl-dimethyl-silanyloxy)-5-(9H-fluoren-9-ylmethoxy-carbonylamino)-hexanoic acid (**11**).* A premixed solution of $H_2O_2$ (30% aq. solution, 2.70 mL) and LiOH (0.2 M, 1.38 mL, 0.276 mmol) was added dropwise to **10** (81 mg, 0.145 mmol) dissolved in THF (10 mL) at 0 °C. The mixture was allowed to reach rt followed by stirring for 3.5 h and then it was cooled to 0 °C. $Na_2SO_3$ (2 M, 5 mL) was added dropwise followed by stirring for 40 min at 0 °C. The mixture was acidified by the addition of HCl (10%, aq.), $Et_2O$ was added, the phases were separated and the aqueous layer was extracted with $Et_2O$. The combined organic layers were washed with $NaHCO_3$ (satd, aq.), dried and concentrated under reduced pressure. Purification by flash chromatography (*n*-heptane/EtOAc/AcOH 6 : 1 : 1% → 3 : 1 : 1%) gave **11** (53 mg, 75%) as a white solid. HRMS ($FAB^+$) calcd for $C_{27}H_{36}NNa_2O_5Si^+$ (M − H + 2Na)$^+$ 528.2153, found 528.2178. $[\alpha]_D^{20}$ = −7.7 ($c$ = 0.5, $CHCl_3$); H$^1$ NMR ($CD_3OD$) $\delta$ 7.78 (d, $J$ = 7.4 Hz, 2H, Fmoc-arom), 7.65 (d, $J$ = 7.4 Hz, 2H, Fmoc-arom), 7.41–7.35 (m, 2H, Fmoc-arom), 7.33–7.27 (m, 2H, Fmoc-arom), 6.73 (d, $J$ = 8.0 Hz, 1H, NH), 4.43–4.28 (m, 2H, Fmoc-C$H_2$), 4.19 (t, $J$ = 6.6 Hz, 1H, Fmoc-C*H*), 3.79–3.70 (m, 1H, OC*H*), 3.66–3.55 (m, 1H, NCH), 2.43–2.30 (m, 2H, $CH_2CO_2$), 1.80–1.72 (m, 2H, C$H_2CH_2CO_2$), 1.10 (d, $J$ = 6.6 Hz, 3H, $CH_3$), 0.90 (s, 9H, $C(CH_3)_3$), 0.07 (s, 3H, $SiCH_3$), 0.05 (s,

Open Access Article. Published on 21 April 2015. Downloaded on 12/5/2025 12:49:11 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

3H, $SiCH_3$); $^{13}C$ NMR (MeOD) $\delta$ 177.3 ($CO_2H$), 158.2 (splitted, NCO), 145.4, 142.6, 128.7, 128.1, 126.2, 120.9, 74.8 (O*C*H), 67.5 (Fmoc-*C*$H_2$), 51.5 (splitted, NCH), 48.5 (Fmoc-*C*H), 30.3 (*C*$H_2CO_2$), 29.7 (*C*$H_2CH_2CO_2$), 26.4 (C(*C*$H_3)_3$), 18.9 (*C*$(CH_3)_3$), 16.1 ($CH_3$), −4.2 ($SiCH_3$), −4.3 ($SiCH_3$).

*(4S,5S)-4-(tert-Butyl-dimethyl-silanyloxy)-5-(9H-fluoren-9-ylmethoxy-carbonylamino)-hexanoic acid (**13**).* A premixed solution of $H_2O_2$ (30% aq., 3.0 mL) and LiOH (0.2 M, 1.5 mL, 0.300 mmol) was added dropwise to **12** (91 mg, 0.165 mmol) in THF (11 mL) at 0 °C. The mixture was allowed to reach rt followed by stirring for 3.5 h and then it was cooled to 0 °C. $Na_2SO_3$ (2 M, 5.6 mL) was added dropwise followed by stirring for 40 min at 0 °C. The mixture was acidified by addition of HCl (10% aq.), $Et_2O$ was added, the phases were separated and the aqueous phase was extracted with $Et_2O$. The combined organic layers were washed with $NaHCO_3$ (satd, aq.), dried and concentrated under reduced pressure. Purification by flash chromatography (*n*-heptane/EtOAc/AcOH 6 : 1 : 1%) gave **13** (65 mg, 82%) as a white solid. HRMS ($FAB^+$) calcd for $C_{27}H_{36}NNa_2O_5Si^+$ (M − H + 2Na)$^+$ 528.2153, found 528.2175. $[\alpha]_D^{20}$ = +5.5 (*c* = 1.0, $CHCl_3$); $H^1$ NMR ($CD_3OD$) $\delta$ 7.79 (d, *J* = 7.5 Hz, 2H, Fmoc-arom), 7.68–7.63 (m, 2H, Fmoc-arom), 7.41–7.35 (m, 2H, Fmoc-arom), 7.33–7.27 (m, 2H, Fmoc-arom), 6.73 (d, *J* = 8.1 Hz, 1H, NH), 4.42 (dd, *J* = 10.5, 6.8 Hz, 1H, Fmoc-C*H*$_2$), 4.32 (dd, *J* = 10.4, 6.8 Hz, 1H, Fmoc-C*H*$_2$), 4.23–4.18 (m, 1H, Fmoc-C*H*), 3.76–3.67 (m, 2H, NCH and NCHC*H*), 2.42–2.26 (m, 2H, $CH_2CO_2$), 1.80–1.69 (m, 1H, C*H*$_2CH_2CO_2$), 1.69–1.58 (m, 1H, C*H*$_2CH_2CO_2$), 1.09 (d, *J* = 6.5 Hz, 3H, $CH_3$), 0.92 (s, 9H, C$(CH_3)_3$), 0.11 (s, 3H, $SiCH_3$), 0.10 (s, 3H, $SiCH_3$); $^{13}C$ NMR (MeOD) $\delta$ 177.4 ($CO_2H$), 158.2 (NCO, splitted), 145.3 (splitted), 142.6, 128.7, 128.1 (splitted), 126.2 (splitted), 120.9, 74.5 (NCH*C*H), 67.5 (Fmoc-*C*$H_2$), 51.4 (NCH, splitted), 48.6 (Fmoc-*C*H), 31.3 (*C*$H_2CO_2$), 28.4 (*C*$H_2CH_2CO_2$), 26.4 (C(*C*$H_3)_3$), 18.9 (*C*$(CH_3)_3$), 15.4 ($CH_3$), −4.3 ($SiCH_3$), −4.4 ($SiCH_3$).

**General procedure for solid-phase peptide synthesis.** Glycopeptides **2** and **3** were synthesized in mechanically agitated reactors on a TentaGel-S-PHB-Thr(*t*Bu)-Fmoc resin (40 μmol) using standard solid-phase methodology essentially as described elsewhere.[25] All couplings were performed in DMF. $N^{\alpha}$-Fmoc amino acids with standard side-chain protecting groups (4 equiv.) were activated with 1-hydroxybenzotriazole (HOBt, 6 equiv.) and 1,3-diisopropylcarbodiimide (DIC, 3.9 equiv.) and the couplings were monitored using bromophenol blue as the indicator. (5*R*)-$N^{\alpha}$-(Fluoren-9-ylmethoxycarbonyl)-$N^{\varepsilon}$-benzyloxycarbonyl-5-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-5-hydroxy-L-lysine[14,26] (1.5 equiv.) and the isostere building blocks **11** and **13** (1.5 equiv.) were activated with 7-aza-1-hydroxybenzotriazole (HOAt, 3 equiv.) and DIC (1.5 equiv.) and coupled for 24 h. Fmoc deprotection after each coupling cycle was accomplished by treatment with 20% piperidine in DMF for 10 min. The glycopeptides were cleaved from the resin with trifluoroacetic acid/$H_2O$/thioanisole/ethanedithiol (35 : 2 : 2 : 1, 40 mL) for 3 h at 40 °C with workup performed essentially as described elsewhere.[25] Purification by reversed-phase HPLC was followed by deacetylation with NaOMe in MeOH (20 mM, 1 mL per mg peptide) for 2–3 h at rt (monitored by analytical reversed-phase HPLC). The mixture was neutralized by addition of AcOH and concentrated under reduced pressure, and the residue was finally purified using reversed-phase HPLC followed by lyophilization.

*Glycyl-L-isoleucyl-L-alanylψ[(4S)-CHOHCH$_2$]glycyl-L-phenylalanyl-(5R)-5-O-(β-D-galactopyranosyl)-5-hydroxy-L-lysylglycyl-Lglutam-1-yl-L-glutaminylglycyl-L-prolyl-L-lysylglycyl-L-glutam-1-yl-L-threonine (**2**).* Synthesis was performed with building block **13** according to the general procedure described above. This afforded the trifluoroacetate salt of **2** (12.4 mg, 18% yield based on the amount of resin used) as a white amorphous solid after lyophilization. MS (MALDI-TOF) calcd 1654.82 for $[M + H]^+$, found 1654.82. $^1H$ NMR data are given in Table 3.

*Glycyl-L-isoleucyl-L-alanylψ[(4R)-CHOHCH$_2$]glycyl-L-phenylalanyl-(5R)-5-O-(β-D-galactopyranosyl)-5-hydroxy-L-lysylglycyl-L-glutam-1-yl-L-glutaminylglycyl-L-prolyl-L-lysylglycyl-L-glutam-1-yl-L-threonine (**3**).*

**Table 3** $^1H$ NMR chemical shifts for CII259–273 with Ala$^{261}$ψ[(4*S*)-CHOHCH$_2$]Gly$^{262}$ **2**[a]

| Residue | NH | Hα | Hβ | Hγ | Others |
|---|---|---|---|---|---|
| Gly$^{259}$ | | 3.86[b] | | | |
| Ile$^{260}$ | 8.39 | 4.17 | 1.81 | 1.41, 1.14, 0.88 ($CH_3$) | 0.84 (Hδ) |
| Ala$^{261}$ψ[CHOHCH$_2$] | 7.92 | 3.86 | 1.10 | | 3.41 ψ[C*H*OHCH$_2$], 1.55[b] ψ[CHOHC*H*$_2$] |
| Gly$^{262}$ | | 2.31[b] | | | |
| Phe$^{263}$ | 8.20 | 4.62 | 3.10, 2.98 | | 7.25 (Hδ), 7.35 (Hε), 7.31 (Hζ) |
| Hyl$^{264}$ | 8.40 | 4.29 | 2.00, 1.75 | 1.61[b] | 4.02 (Hδ), 3.18 and 2.98 (Hε), 7.62 (εNH$_2$)[c] |
| Gly$^{265}$ | 8.00 | 3.92[b] | | | |
| Glu$^{266}$ | 8.23 | 4.37 | 2.10, 1.94 | 2.43[b] | |
| Gln$^{267}$ | 8.49 | 4.37 | 2.10, 1.96 | 2.35[b] | 7.50 and 6.83 (δNH$_2$) |
| Gly$^{268}$ | 8.27 | 4.11, 3.98 | | | |
| Pro$^{269}$ | | 4.40 | 2.26, 1.90 | 2.00[b] | 3.58[b] (Hδ) |
| Lys$^{270}$ | 8.45 | 4.29 | 1.84, 1.75 | 1.43[b] | 1.65[b] (Hδ), 3.00[b] (Hε), 7.50 (εNH$_2$) |
| Gly$^{271}$ | 8.35 | 3.94[b] | | | |
| Glu$^{272}$ | 8.21 | 4.47 | 2.14, 1.96 | 2.45[b] | |
| Thr$^{273}$ | 8.09 | 4.33 | 4.33 | 1.16 | |

[a] Measured at 500 MHz and 298 K in water containing 10% $D_2O$ with $H_2O$ ($\delta_H$ 4.76 ppm) as an internal standard. [b] Degeneracy has been assumed. [c] Chemical shifts for the galactose moiety: δ 4.43 (H1), 3.91 (H4), 3.76 (H6), 3.68 (H5), 3.64 (H3), and 3.52 (H2).

 This journal is © The Royal Society of Chemistry 2015

Open Access Article. Published on 21 April 2015. Downloaded on 12/5/2025 12:49:11 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Table 4 $^1$H NMR chemical shifts for CII259–273 with Ala$^{261}\psi$[(4*R*)-CHOHCH$_2$]Gly$^{262}$ **3**[a]

| Residue | NH | H$\alpha$ | H$\beta$ | H$\gamma$ | Others |
|---|---|---|---|---|---|
| Gly$^{259}$ | | 3.84[b] | | | |
| Ile$^{260}$ | 8.41 | 4.13 | 1.81 | 1.43, 1.16, 0.88 (CH$_3$) | 0.84 (H$\delta$) |
| Ala$^{261}\psi$[CHOHCH$_2$] | 8.11 | 3.82 | 1.08 | | 3.49 $\psi$[C*H*OHCH$_2$], 1.67 and 1.47 $\psi$[CHOHC*H*$_2$] |
| Gly$^{262}$ | | 2.33, 2.24 | | | |
| Phe$^{263}$ | 8.19 | 4.60 | 3.06, 2.98 | | 7.25 (H$\delta$), 7.34 (H$\varepsilon$), 7.30 (H$\zeta$) |
| Hyl$^{264}$ | 8.39 | 4.29 | 2.00, 1.75 | 1.59[b] | 4.02 (H$\delta$), 3.16 and 2.98 (H$\varepsilon$), 7.62 ($\varepsilon$NH$_2$)[c] |
| Gly$^{265}$ | 7.94 | 3.88[b] | | | |
| Glu$^{266}$ | 8.22 | 4.37 | 2.10, 1.94 | 2.43[b] | |
| Gln$^{267}$ | 8.49 | 4.37 | 2.10, 1.96 | 2.35[b] | 7.50 and 6.83 ($\delta$NH$_2$) |
| Gly$^{268}$ | 8.27 | 4.13, 4.00 | | | |
| Pro$^{269}$ | | 4.40 | 2.26, 1.90 | 1.98[b] | 3.58[b] (H$\delta$) |
| Lys$^{270}$ | 8.45 | 4.29 | 1.84, 1.75 | 1.43[b] | 1.65[b] (H$\delta$), 2.98[b] (H$\varepsilon$), 7.50 ($\varepsilon$NH$_2$) |
| Gly$^{271}$ | 8.35 | 3.94[b] | | | |
| Glu$^{272}$ | 8.20 | 4.47 | 2.14, 1.96 | 2.45[b] | |
| Thr$^{273}$ | 8.10 | 4.33 | 4.33 | 1.16 | |

[a] Measured at 500 MHz and 298 K in water containing 10% $D_2O$ with $H_2O$ ($\delta_H$ 4.76 ppm) as an internal standard. [b] Degeneracy has been assumed. [c] Chemical shifts for the galactose moiety: $\delta$ 4.43 (H1), 3.92 (H4), 3.76 (H6), 3.68 (H5), 3.64 (H3), and 3.52 (H2).

Synthesis was performed with building block **11** according to the general procedure described above. This afforded the trifluoroacetate salt of **3** (22.5 mg, 28% yield based on the amount of resin used) as a white amorphous solid after lyophilization. MS (MALDI-TOF) calcd 1654.82 for $[M + H]^+$, found 1654.81. $^1$H NMR data are given in Table 4.

### A$^q$ binding assay

The abilities of the backbone-modified glycopeptides **2** and **3** and non-modified CII259–273 **1**[4] to bind to A$^q$ were determined relative to a biotinylated CLIP peptide with a competitive inhibition assay performed using 96-well microtiter assay plates essentially as described elsewhere.[9,27] Briefly, purified soluble recombinant A$^q$ proteins (1 μM) expressed in *Drosophila melanogaster* SL2 cells were incubated with a fixed concentration of the CLIP-bio (3 μM, sequence KPVSKMRMATPLLMQALPM) and increasing concentrations of test peptides **1**, **2** and **3** (0, 0.98, 1.95, 3.91, 7.81, 15.63, 32.25, 62.5, 125, 250, 500 and 1000 μM) in PBS for 48 h at room temperature. This mixture also contained a cocktail of protease inhibitors (Complete™, Boehringer, Mannheim). During this incubation, new 96-well microtiter assay plates were precoated with Y3P monoclonal antibodies (mAb) (10 μg mL$^{-1}$) by incubation overnight at 4 °C followed by blocking with PBS containing 2% low fat milk and washing with PBS containing 0.1% Tween 20. The incubated mixtures containing A$^q$ proteins, CLIP-bio, and test peptides (90 μL) were transferred to the plates precoated with mAb followed by incubation overnight at 4 °C. After washing with PBS (0.1% Tween 20) the amount of CLIP-bio bound to the A$^q$ proteins captured by the mAb in the wells was quantified using the dissociation-enhanced lanthanide fluoroimmunoassay (DELFIA®) system based on the time-resolved fluoroimmunoassay technique with europium labeled streptavidin (Wallac, Turku), according to the manufacturer's instructions. The experiments were performed in duplicates.

### T-cell activation assays

Responses of the A$^q$ restricted T-cell hybridoma lines (*i.e.*, the amount of IL-2 secreted following incubation of the hybridoma with antigen-presenting spleen cells expressing A$^q$ and increasing concentrations of the glycopeptides) were determined essentially as described elsewhere,[28] with slight modifications. Glycopeptides **1–3** were evaluated with the A$^q$-restricted T-cell hybridomas[16,29] HCQ.3, HCQ.10, HM1R.2 and 22a1-7E. In brief, T-cell hybridomas ($5 \times 10^4$) were co-cultured with syngeneic spleen cells ($5 \times 10^5$) and increasing concentrations of the test glycopeptides (0, 0.01, 0.048, 0.24, 1.2, 6.0, 30, and 150 μM) in a volume of 200 μL in 96-well U-bottom microtiter plates. After 24 h, 140 μL supernatant was removed, transferred to a V-bottom microtiter plate, and spun down to avoid transfer of T-cell hybridomas to the ELISA plate. The contents of IL-2 in the culture supernatant (100 μL) were measured by sandwich ELISA (capturing mAb: purified rat anti-mouse IL-2, JES6-IA 12; detecting mAb, biotinylated anti-Mouse Interleukin-2 mAb 5H4, Mabtech AB) using the DELFIA® system (Wallac, Turku, Finland), according to the manufacturer's instructions. The culture supernatant from ConA activated splenocytes served as a positive control.

### Molecular modelling

**Preparation of the complexes.** The initial coordinates for A$^q$ in a complex with **1** were obtained from an X-ray crystal structure[22] and prepared using the Protein Preparation Wizard[30] in Maestro.[31] Bond orders were assigned, hydrogens added, disulfide bonds created, termini were capped, and no water molecules were included. The hydrogen bond assignment was optimized and the hydrogen atoms were energy minimized with the OPLS 2005 force field[32] to converge to RMSD 0.30 Å. Due to low electron density, Glu$^{108}$–Leu$^{110}$ in one loop in the $\beta_2$ domain and the carbohydrate moiety GalHyl$^{264}$ were not modelled in the crystal structure. Glu$^{108}$–Leu$^{110}$ were added using the building module in Maestro. Arg$^{106}$–His$^{112}$ were

Open Access Article. Published on 21 April 2015. Downloaded on 12/5/2025 12:49:11 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

thereafter energy minimized using the OPLS 2005 force field and a dielectric constant of 80 as implemented in the Macro model[33] within Maestro, while the rest of the complex was frozen. Manual hydroxylation and glycosylation of $Lys^{264}$ were also performed with the building module in Maestro. A conformational search was performed for the carbohydrate moiety (while the rest of the complex was frozen). The generated conformations were thereafter energy minimized using the OPLS 2005 force field and a dielectric constant of 80 as implemented in the Macro model[33] within Maestro while the rest of the proteins were constrained with a force constant of 100 kJ $mol^{-1}$ $Å^2$. The conformation with the lowest energy was thereafter selected. The native glycopeptide **1** was manually mutated into the amide bond isostere modified glycopeptides **2–4** and thereafter energy minimized in the binding grove with the Macro model and a flat bottom constraint on the entire complex, with a force constant of 100 kJ $mol^{-1}$ $Å^2$ and a width of 0.2 Å.

**Parameter preparation GalHyl[264].** Partial charges for the non-standard residue $GalHyl^{264}$ were computed using the R.E.D. Server/R.E.D IV[34–37] Seven conformations of the ACE–$GalHyl^{264}$–NME capped amino acid were used, extracted from a conformational search using the OPLS 2005 force field and a dielectric constant of 80 as implemented in the Macro model[33] within Maestro. The partial charges were thereafter fitted to reproduce molecular electrostatic potential (MEP). The intramolecular charge constraints for ACE and NME were set to zero during the charge fitting step and are thereafter removed from the molecule to give a force field library for the central fragment only. Using Ante_R.E.D. 2.0, P2N files were generated, thereafter manually modified to correct the atom order and connectivity, and finally joined together into one single file. The R.E.D IV server with the RESP-A1 (HF/6-31G*) charge model was thereafter applied.

The GLYCAM force field atom types were assigned for the carbohydrate part and AMBER force field atom types for the amino acid part.

**Amide bond isosteres.** The coordinates for the three amide bond isosteres were taken from the prepared complexes as described above. The partial charges for the amide bond isosteres were thereafter derived following a similar procedure as described above for $GalHyl^{264}$ using the R.E.D server and the RESP-A1 (HF/6-31G*) charge model. The central fragment procedure was applied using the capped amide bond isostere fragment with an acetylated N-terminus (ACE) and a methylated C-terminus (NME) using one conformation for each amide bond isostere.

**Molecular dynamic simulations.** The AMBER12[38] simulation package was used for the MD simulations with the ff99SB force field together with GLYCAM_06h for the carbohydrate moiety. Hydrogen atoms were deleted from the complex before preparation with the xleap module. Hydrogen atoms were added, disulphide bridges were specified, and $Na^+$ counter ions were added to neutralize the system. Thereafter the system was solvated with the TIP3P water model with a 12 Å octahedral box, using periodic boundary conditions. Langevin dynamics was used to control the temperature during the MD simulations, the SHAKE algorithm constrain bonds involving hydrogens, and a time step of 2 fs were used. The Particle Mesh Ewald method was applied to treat long range electrostatic interactions and an 8 Å cut off was used for the short range non-bonded interactions. A first minimization was performed where the protein was constrained with a force constant of 500 kJ $mol^{-1}$ $Å^2$, followed by another minimization without constraints. Thereafter the system was heated from 0 to 300 K with a 20 ps simulation, with constraints added to the complex with a force constant of 10 kJ $mol^{-1}$ $Å^2$. This was followed by a 100 ps simulation to equilibrate the system at 300 K. Constraints with a force constant of 100 kJ $mol^{-1}$ $Å^2$ were added to the backbone of the end terminus residues of the $\alpha_2\beta_2$ domains, since these residues are membrane bound. Constant pressure periodic boundaries with an average pressure of 1 atm were used with isotropic position scaling to maintain the pressure. The production runs, for in total 60 ns, were performed at 298.15 K with constraints on the end terminal residues. All the MD simulations were executed on the High Performance Computing Centre North[39] (HPC2N).

Analysis was performed with the aid of the Visual Molecular Dynamics (VMD)[23] software. Frames were extracted and visualized with a time step of 2 ps. RMSD values were calculated with the RMSD trajectory tool to confirm stability of the complexes. The hydrogen bond occupancy was calculated with the hydrogen bond extension tool, with a donor–acceptor length of less than 3.3 Å and a donor–H–acceptor angle of minimum 160°. Only hydrogen bonds with an occupancy >10% are reported, and hydrogen bonds with an occupancy >40% are considered strong. Cluster analysis was performed with the clustering plugin[40] over the last 20 ns. Clustering was based on the minimal epitope ($Ile^{260}$–$Gln^{267}$) after superposing the complex against the binding groove, defined as $\alpha Ala^3$–$Asn^{85}$, $\beta Glu^4$–$Gln^{97}$. An RMSD cut-off of 0.8 Å was used and a maximum number of 20 clusters were extracted. Representative snapshots were thereafter randomly selected from each cluster. Figures have been prepared with MOE.[41]

## Acknowledgements

This study was supported by the Swedish Research Council, Swedish Foundation for Strategic Research, Knut and Alice Wallenberg Foundation, and the EU IMI Project BeTheCure. The access to synchrotron radiation and data collection support by the beamline staff at the European Synchrotron Radiation Facility (Grenoble, France) are gratefully acknowledged. We also thank HPC2N (High Performance Computing Center North) for HPC resources and technical support.

## Notes and references

1 P. K. Gregersen, J. Silver and R. J. Winchester, *Arthritis Rheum.*, 1987, **30**, 1205–1213.

 This journal is © The Royal Society of Chemistry 2015

Open Access Article. Published on 21 April 2015. Downloaded on 12/5/2025 12:49:11 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

2 M. Hegen, J. C. Keith, M. Collins and C. L. Nickerson-Nutter, *Ann. Rheum. Dis.*, 2008, **67**, 1505–1515.
3 U. Brunsberg, K. Gustafsson, L. Jansson, E. Michaelsson, L. Ahrlund-Richter, S. Pettersson, R. Mattsson and R. Holmdahl, *Eur. J. Immunol.*, 1994, **24**, 1698–1702.
4 J. Broddefalk, J. Bäcklund, F. Almqvist, M. Johansson, R. Holmdahl and J. Kihlberg, *J. Am. Chem. Soc.*, 1998, **120**, 7676–7683.
5 B. Dzhambazov, K. S. Nandakumar, J. Kihlberg, L. Fugger, R. Holmdahl and M. Vestberg, *J. Immunol.*, 2006, **176**, 1525–1533.
6 J. Backlund, S. Carlsen, T. Hoger, B. Holm, L. Fugger, J. Kihlberg, H. Burkhardt and R. Holmdahl, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 9960–9965.
7 L. Holm, P. Kjellen, R. Holmdahl and J. Kihlberg, *Bioorg. Med. Chem.*, 2005, **13**, 473–482.
8 E. F. Rosloniec, K. B. Whittington, D. D. Brand, L. K. Myers and J. M. Stuart, *Cell. Immunol.*, 1996, **172**, 21–28.
9 P. Kjellen, U. Brunsberg, J. Broddefalk, B. Hansen, M. Vestberg, I. Ivarsson, A. Engstrom, A. Svejgaard, J. Kihlberg, L. Fugger and R. Holmdahl, *Eur. J. Immunol.*, 1998, **28**, 755–767.
10 I. E. Andersson, C. D. Andersson, T. Batsalova, B. Dzhambazov, R. Holmdahl, J. Kihlberg and A. Linusson, *PLoS One*, 2011, 6.
11 B. Holm, J. Backlund, M. A. F. Recio, R. Holmdahl and J. Kihlberg, *ChemBioChem*, 2002, **3**, 1209–1222.
12 I. E. Andersson, T. Batsalova, B. Dzhambazov, L. Edvinsson, R. Holmdahl, J. Kihlberg and A. Linusson, *Org. Biomol. Chem.*, 2010, **8**, 2931–2940.
13 I. E. Andersson, T. Batsalova, S. Haag, B. Dzhambazov, R. Holmdahl, J. Kihlberg and A. Linusson, *J. Am. Chem. Soc.*, 2011, **133**, 14368–14378.
14 I. E. Andersson, B. Dzhambazov, R. Holmdahl, A. Linusson and J. Kihlberg, *J. Med. Chem.*, 2007, **50**, 5627–5643.
15 M. Hart and C. Beeson, *J. Med. Chem.*, 2001, **44**, 3700–3709.
16 L. Holm, R. Bockermann, E. Wellner, J. Bäcklund, R. Holmdahl and J. Kihlberg, *Bioorg. Med. Chem.*, 2006, **14**, 5921–5932.
17 A. M. H. Boots, H. Hubers, M. Kouwijzer, L. D. Zandbrink, B. M. Westrek-Esselink, C. van Doorn, R. Stenger, E. S. Bos, M. J. C. van Lierop, G. F. Verheijden, C. M. Timmers and C. J. van Staveren, *Arthritis Res. Ther.*, 2007, **9**, R71.
18 E. Atherton and R. C. Sheppard, *Solid phase peptide synthesis*, IRL Press at Oxford University Press, Oxford, England, 1989.
19 R. V. Hoffman, N. Maslouh and F. Cervantes-Lee, *J. Org. Chem.*, 2002, **67**, 1045–1056.
20 J. Våbenø, M. Brisander, T. Lejon and K. Luthman, *J. Org. Chem.*, 2002, **67**, 9186–9191.
21 T. M. Gierasch, Z. J. Shi and G. L. Verdine, *Org. Lett.*, 2003, **5**, 621–624.
22 D. Dobritzsch, H. Uysal, C. Ge and R. Holmdahl, unpublished results.
23 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33–38.
24 R. Déziel, R. Plante, V. Caron, L. Grenier, M. Llinas-Brunet, J. S. Duceppe, E. Malenfant and N. Moss, *J. Org. Chem.*, 1996, **61**, 2901–2903.
25 J. Broddefalk, M. Forsgren, I. Sethson and J. Kihlberg, *J. Org. Chem.*, 1999, **64**, 8948–8953.
26 B. M. Syed, T. Gustafsson and J. Kihlberg, *Tetrahedron*, 2004, **60**, 5571–5575.
27 C. M. Hill, A. Liu, K. W. Marshall, J. Mayer, B. Jorgensen, B. Yuan, R. M. Cubbon, E. A. Nichols, L. S. Wicker and J. B. Rothbard, *J. Immunol.*, 1994, **152**, 2890–2898.
28 E. Michaëlsson, M. Andersson, A. Engström and R. Holmdahl, *Eur. J. Immunol.*, 1992, **22**, 1819–1825.
29 A. Corthay, J. Bäcklund, J. Broddefalk, E. Michaëlsson, T. J. Goldschmidt, J. Kihlberg and R. Holmdahl, *Eur. J. Immunol.*, 1998, **28**, 2580–2590.
30 Schrödinger Suite 2009 Protein Preparation Wizard; Epik version 2.0, Schrödinger, LLC, New York, NY, 2009; Impact version 5.5, Schrödinger, LLC, New York, NY, 2009; Prime version 2.1, Schrödinger, LLC, New York, NY, 2009.
31 Maestro, *version 9.1*, Schrödinger, LLC, New York, NY, 2010.
32 J. L. Banks, H. S. Beard, Y. X. Cao, A. E. Cho, W. Damm, R. Farid, A. K. Felts, T. A. Halgren, D. T. Mainz, J. R. Maple, R. Murphy, D. M. Philipp, M. P. Repasky, L. Y. Zhang, B. J. Berne, R. A. Friesner, E. Gallicchio and R. M. Levy, *J. Comput. Chem.*, 2005, **26**, 1752–1780.
33 MacroModel, *version 9.8*, Schrödinger, LLC, New York, NY, 2010.
34 E. Vanquelef, S. Simon, G. Marquant, E. Garcia, G. Klimerak, J. C. Delepine, P. Cieplak and F.-Y. Dupradeau, *Nucleic Acids Res.*, 2011, **39**, W511–W517.
35 F.-Y. Dupradeau, A. Pigache, T. Zaffran, C. Savineau, R. Lelong, N. Grivel, D. Lelong, W. Rosanski and P. Cieplak, *Phys. Chem. Chem. Phys.*, 2010, **12**, 7821–7839.
36 C. I. Bayly, P. Cieplak, W. D. Cornell and P. A. Kollman, *J. Phys. Chem.*, 1993, **97**, 10269–10280.
37 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09*, Gaussian, Inc., Wallingford, CT, USA, 2009.

38 D. A. Case, T. A. Darden, T. E. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A. W. Goetz, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, R. M. Wolf, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko and P. A. Kollman, *AMBER 12*, University of California, San Fransisco, 2012.
39 High Performance Computer Center North (HPC2N); HPC2N: Umeå, Sweden; http://www.hpc2n.umu.se/.
40 L. Gracia, Weill Cornell Medical College, 1300 York Ave, New York, 10065 NY.
41 *Molecular Operating Environment (MOE)*, 2012.10; Chemical Computing Group Inc., 1010 Sherbrooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2012.

Open Access Article. Published on 21 April 2015. Downloaded on 12/5/2025 12:49:11 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.


This journal is © The Royal Society of Chemistry 2015