Structural Basis of Multivalent Binding to Wheat Germ Agglutinin

David Schwefel,† Caroline Maierhofer,‡ Johannes G. Beck,‡ Sonja Seeberger,‡ Kay Diederichs,† Heiko M. Möller,*,‡ Wolfram Welte,*,† and Valentin Wittmann*,‡

‡ Department of Chemistry, † Department of Biology, Universität Konstanz, Universitätsstr. 10, 78457 Konstanz, Germany

Supporting Information

Contents

Inhibition curves (Figures S1–S3) .......................................................... S2
Electron density plots (Figure S4) ......................................................... S4
Tables S1–S4 ........................................................................................ S5
Figures S5–S7 ...................................................................................... S8
NMR spectra of compounds 5, 15–24, and 28 ................................. S10
Figure S1. Dose-response curves for inhibition of binding of horseradish peroxidase-labeled WGA to β-GlcNAc-modified microtiter plates by GlcNAc, 1, and 29.

Figure S2. Dose-response curves for inhibition of binding of horseradish peroxidase-labeled WGA to β-GlcNAc-modified microtiter plates by GlcNAc, 2, 20, and 24.
Figure S3. Dose-response curves for inhibition of binding of horseradish peroxidase-labeled WGA to β-GlcNAc-modified microtiter plates by GlcNAc, 21, 22, 23, and 28.
Figure S4. Unbiased $F_o-F_c$ (green, contoured at 2.5$\sigma$) and refined 2$F_o-F_c$ (blue, contoured at 1$\sigma$) electron densities for (A) propyl carbamate 20 at binding site B1C2, (B) divalent ligand 23 at the binding site pair B1C2 – C2B1, and (C) visible substructure of cyclic peptide 2 at the binding site pair B1C2 – C2B1 The refined atomic models of the respective compounds are shown in stick representation with carbon atoms colored white, nitrogens blue and oxygens red.
Table S1. $^1$H and $^{13}$C NMR chemical shifts (ppm) of neoglycopeptide 2 in H$_2$O/D$_2$O (95:5, pH 5) at 300 K. $^1$H and $^{13}$C Chemical shifts were referenced to 3-(trimethylsilyl)-2,2,3,3-tetadeuteropropionic acid (TSP). n.d., not determined

<table>
<thead>
<tr>
<th></th>
<th>$^1$H</th>
<th>$^13$C</th>
<th>$^1$H</th>
<th>$^13$C</th>
<th>$^1$H</th>
<th>$^13$C</th>
<th>others</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$^1$H$^a$</td>
<td>$^1$H$^b$</td>
<td>$^1$H$^c$</td>
<td>$^1$H$^d$</td>
<td>$^1$H$^e$</td>
<td>$^1$H$^f$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D-Dab-2</td>
<td>8.499</td>
<td>4.320</td>
<td>52.24</td>
<td>1.938 / 2.093</td>
<td>29.93</td>
<td>3.132 / 3.265</td>
<td>37.15-37.30</td>
<td>$^1$H$^f$: 7.274</td>
</tr>
<tr>
<td>D-Val-3</td>
<td>7.812</td>
<td>4.123</td>
<td>59.57</td>
<td>2.085</td>
<td>0.922</td>
<td>17.76 / 18.52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D-Dab-4</td>
<td>8.322</td>
<td>4.398</td>
<td>51.84</td>
<td>1.93 / 2.054</td>
<td>30.70</td>
<td>3.216</td>
<td>37.15-37.30</td>
<td>$^1$H$^f$: 7.255</td>
</tr>
<tr>
<td>D-Dab-5</td>
<td>8.283</td>
<td>4.499</td>
<td>50.91</td>
<td>1.874 / 2.073</td>
<td>30.88</td>
<td>3.164 / 3.236</td>
<td>37.15-37.30</td>
<td>$^1$H$^f$: 7.222</td>
</tr>
<tr>
<td>D-Val-6</td>
<td>8.131</td>
<td>4.010</td>
<td>60.96</td>
<td>2.127</td>
<td>27.23</td>
<td>0.975</td>
<td>18.50</td>
<td></td>
</tr>
<tr>
<td>D-Dab-7</td>
<td>8.594</td>
<td>4.240</td>
<td>52.18</td>
<td>1.988 / 2.111</td>
<td>3.234</td>
<td>37.15-37.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu-8</td>
<td>7.911</td>
<td>4.177</td>
<td>53.98</td>
<td>2.066 / 2.093</td>
<td>2.287 / 2.356</td>
<td>32.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>β-Ala-9</td>
<td>8.051</td>
<td>3.404 / 3.503</td>
<td>34.55</td>
<td>2.490</td>
<td>35.86</td>
<td></td>
<td>CONH$_2$: 6.863 / 7.534</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$^1$H</th>
<th>$^13$C</th>
<th>$^1$H</th>
<th>$^13$C</th>
<th>$^1$H</th>
<th>$^13$C</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H-1 / C-1</td>
<td>H-2 / C-2</td>
<td>H-3 / C-3</td>
<td>H-4 / C-4</td>
<td>H-5 / C-5</td>
<td>H-6 / C-6</td>
<td>$^1$H$^N$</td>
<td>H$^{Me}$/ C$^{Me}$</td>
</tr>
<tr>
<td>GlcNAc</td>
<td>5.91-5.94 / 91.6</td>
<td>4.03-4.06 / 52.6</td>
<td>3.77 / 70.7</td>
<td>3.57-3.60 / 69.4</td>
<td>3.72 / 73.7</td>
<td>3.80 / 60.3</td>
<td>8.26-8.29</td>
<td>2.020-2.032 / 21.8</td>
</tr>
</tbody>
</table>

|        | $^1$H$^\text{Bu}$ | $^13$C$^\text{Me}$ | $^13$C$^\text{quart}$ |        |
|--------|--------------------|---------------------|----------------------|        |
| Boc    | 1.427              | 27.47               | n.d.                 |        |
Table S2. Structure determination of neoglycopeptide 2 in solution. Internuclear distances as determined from NOESY cross peak integrals were used as upper distance bounds during molecular dynamics simulated annealing within AMBER 8.

<table>
<thead>
<tr>
<th>Atom 1</th>
<th>Atom 2</th>
<th>Distance [Å]</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-Dab-4/5/7-H^{1/2}</td>
<td>D-Val-3</td>
<td>H^{1/2}</td>
</tr>
<tr>
<td>D-Dab-4/5/7-H^{1/2}</td>
<td>D-Val-6</td>
<td>H^{1/2}</td>
</tr>
<tr>
<td>Lys-1 H^{α}</td>
<td>D-Val-3</td>
<td>H^{α}</td>
</tr>
<tr>
<td>Lys-1 H^{N}</td>
<td>D-Val-3</td>
<td>H^{N}</td>
</tr>
<tr>
<td>Lys-1 H^{β1/2}</td>
<td>D-Val-3</td>
<td>H^{β1/2}</td>
</tr>
<tr>
<td>Lys-1 H^{β}</td>
<td>D-Val-3</td>
<td>H^{β1/2}</td>
</tr>
<tr>
<td>D-Dab-2 H^{N}</td>
<td>Lys-1 H^{α}</td>
<td></td>
</tr>
<tr>
<td>D-Dab-2 H^{N}</td>
<td>Lys-1 H^{β1/2}</td>
<td></td>
</tr>
<tr>
<td>D-Dab-2 H^{N}</td>
<td>D-Val-3 H^{N}</td>
<td></td>
</tr>
<tr>
<td>D-Dab-2 H^{α}</td>
<td>D-Val-3 H^{N}</td>
<td></td>
</tr>
<tr>
<td>D-Dab-2 H^{N}</td>
<td>D-Val-3/4-H^{β1/2}</td>
<td></td>
</tr>
<tr>
<td>D-Dab-4 H^{N}</td>
<td>D-Val-3 H^{N}</td>
<td></td>
</tr>
<tr>
<td>D-Dab-4 H^{N}</td>
<td>D-Val-3 H^{β1/2}</td>
<td></td>
</tr>
<tr>
<td>D-Dab-4 H^{α}</td>
<td>D-Val-5 H^{N}</td>
<td></td>
</tr>
<tr>
<td>D-Dab-5 H^{α}</td>
<td>D-Val-6 H^{N}</td>
<td></td>
</tr>
<tr>
<td>D-Dab-5 H^{β1/2}</td>
<td>D-Val-6 H^{N}</td>
<td></td>
</tr>
<tr>
<td>D-Val-6 H^{N}</td>
<td>D-Val-4/5/7-H^{β1/2}</td>
<td></td>
</tr>
<tr>
<td>D-Val-6 H^{α}</td>
<td>D-Val-7 H^{N}</td>
<td></td>
</tr>
<tr>
<td>D-Val-7 H^{N}</td>
<td>D-Val-6 H^{β1/2}</td>
<td></td>
</tr>
<tr>
<td>D-Val-7 H^{N}</td>
<td>D-Val-6 H^{N}</td>
<td></td>
</tr>
<tr>
<td>D-Dab-7 H^{N}</td>
<td>Glu-8 H^{N}</td>
<td></td>
</tr>
<tr>
<td>D-Dab-7 H^{α}</td>
<td>Glu-8 H^{N}</td>
<td></td>
</tr>
<tr>
<td>Glu-8 H^{β1/2}</td>
<td>Lys-1 H^{β1/2}</td>
<td></td>
</tr>
<tr>
<td>Glu-8 H^{α}</td>
<td>Lys-1 H^{α}</td>
<td></td>
</tr>
<tr>
<td>Glu-8 H^{N}</td>
<td>D-Val-6 H^{N}</td>
<td></td>
</tr>
<tr>
<td>Glu-8 H^{N}</td>
<td>D-Val-6 H^{β1/2}</td>
<td></td>
</tr>
<tr>
<td>Glu-8 H^{α}</td>
<td>β-Ala-9 H^{N}</td>
<td></td>
</tr>
<tr>
<td>β-Ala-9 H^{N}</td>
<td>D-Val-6 H^{β1/2}</td>
<td></td>
</tr>
<tr>
<td>β-Ala-9 H</td>
<td>Glu-8 H^{β1/2}</td>
<td></td>
</tr>
</tbody>
</table>
Table S3. Structure determination of neoglycopeptide 2 in solution. \( ^3J_{\text{HNH}\alpha} \) coupling constants as determined from DQF-COSY spectra were employed during molecular dynamics simulated annealing within AMBER 8 as direct \( J \)-coupling restraints with the following Karplus parameters:
\[ ^3J_{\text{HH}} = A \cos^2 \phi + B \cos \phi + C, \] parameters: \( A = 7.9, B = -1.55, C = 1.35 \)

<table>
<thead>
<tr>
<th>Residue</th>
<th>D-Val-3</th>
<th>D-Dab-4</th>
<th>D-Dab-5</th>
<th>D-Dab-7</th>
</tr>
</thead>
<tbody>
<tr>
<td>( ^3J_{\text{HNH}\alpha} ) [Hz]</td>
<td>8.4</td>
<td>8.8</td>
<td>8.8</td>
<td>8.1</td>
</tr>
</tbody>
</table>

Table S4. Structure determination of neoglycopeptide 2 in solution. \( ^3J_{\text{H\alpha H\beta}} \) coupling constants as determined from DQF-COSY spectra indicate that the \( \chi_1 \) dihedral angles of D-Dab residues 2, 4, 5, and 7 do not adopt \( g^- \) rotamers and Glu-8 does not adopt the \( g^+ \) rotamer. These rotamers were excluded by \( \chi_1 \) torsion angle restraints during molecular dynamics simulated annealing within AMBER 8.

<table>
<thead>
<tr>
<th>Residue</th>
<th>D-Dab-2</th>
<th>D-Dab-4</th>
<th>D-Dab-5</th>
<th>D-Dab-7</th>
<th>Glu-8</th>
</tr>
</thead>
<tbody>
<tr>
<td>( ^3J_{\text{H\alpha H\beta}} ) [Hz]</td>
<td>12</td>
<td>12</td>
<td>10</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>( ^3J_{\text{H\alpha H\beta}'} ) [Hz]</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>allowed ( \chi_1 )</td>
<td>120 ± 120</td>
<td>120 ± 120</td>
<td>120 ± 120</td>
<td>120 ± 120</td>
<td>–120 ± 120</td>
</tr>
</tbody>
</table>
**Figure S5.** Structure determination of neoglycopeptide 2 in solution. Long-range (i.e. non-sequential) distance constraints are shown as blue arrows. $\phi$ dihedral angles restrained by direct $J$-coupling restraints are marked with red circles. $\chi_1$ Dihedral angle restraints are marked with green circles.

**Figure S6.** Structure determination of neoglycopeptide 2 in solution. Ramachandran plot of the final ensemble of 20 conformers. The conformers can be grouped into three families of conformers, I, II, and III, according to backbone rmsd after superposition.
Figure S7. Structure determination of neoglycopeptide 2 in solution: Ensemble of the dominant family of conformers of 2. (A) Peptide backbone structures, (B) ensemble turned by 90° around the vertical axis. (C) Orientations of sugar-bearing side chains, (D) ensemble turned by 90° around the vertical axis (blue D-Dab-2, red D-Dab-4, green D-Dab-5, magenta D-Dab-7).
$^1$H NMR of 5 (CDCl$_3$, 250.1 MHz)

$^{13}$C NMR of 5 (CDCl$_3$, 150.9 MHz)
$^1$H NMR of 15 (CDCl$_3$, 400.1 MHz)*

$^{13}$C NMR of 15 (CDCl$_3$, 100.6 MHz)

* The set of signals with lower intensity originates from a minor isomer due to cis-trans isomerization of the carbamate moiety, as was shown by observation of corresponding exchange signals in ROESY spectra.
$^1$H NMR of 16 (CDCl$_3$, 400.1 MHz)

$^{13}$C NMR of 16 (CDCl$_3$, 100.6 MHz)
$^1$H NMR of 17 (CDCl$_3$, 600.1 MHz)*

* The set of signals with lower intensity originates from minor isomers due to cis-trans isomerization of the carbamate moieties, as was shown by observation of corresponding exchange signals in ROESY spectra (cf. below).

ROESY NMR of 17 (CDCl$_3$, 600.1 MHz). Orange cross peaks (with the same sign as the diagonal) indicate chemical exchange whereas cyan cross peaks originate from dipolar cross relaxation (spatial vicinity).
$^{13}$C NMR of 17 (CDCl$_3$, 150.9 MHz)
$^1$H NMR of 18 (CDCl$_3$, 600.1 MHz)*

* The set of signals with lower intensity originates from minor isomers due to cis-trans isomerization of the carbamate moieties, as was shown by observation of corresponding exchange signals in ROESY spectra.

$^{13}$C NMR of 18 (CDCl$_3$, 150.9 MHz)
$^1$H NMR of 19 (CDCl$_3$, 600.1 MHz)*

* The set of signals with lower intensity originates from minor isomers due to cis-trans isomerization of the carbamate moieties, as was shown by observation of corresponding exchange signals in ROESY spectra.

$^{13}$C NMR of 19 (CDCl$_3$, 150.9 MHz)
$^1$H NMR of 20 (D$_2$O, 600.1 MHz)

$^{13}$C NMR of 20 (D$_2$O, 150.9 MHz)
$^1$H NMR of 21 (D$_2$O, 400.1 MHz)

$^{13}$C NMR of 21 (D$_2$O, 100.6 MHz)
$^1$H NMR of 22 (D$_2$O, 600.1 MHz)

$^{13}$C NMR of 22 (D$_2$O, 150.9 MHz)
$^1$H NMR of 23 (D$_2$O, 600.1 MHz)

$^{13}$C NMR of 23 (D$_2$O, 150.9 MHz)
$^1$H NMR of 24 (D$_2$O, 600.1 MHz)

$^{13}$C NMR of 24 (D$_2$O, 150.9 MHz)
$^1$H NMR of 28 (DMSO-$d_6$, 600.1 MHz)*

* The set of signals with lower intensity originates from minor isomers due to cis-trans isomerization of the carbamate moieties, as was shown by observation of corresponding exchange signals in ROESY spectra (cf. below).

ROESY NMR of 28 (DMSO-$d_6$, 600.1 MHz). Orange cross peaks (with the same sign as the diagonal) indicate chemical exchange whereas cyan cross peaks originate from dipolar cross relaxation (spatial vicinity).
$^{13}$C NMR of 28 (DMSO-$d_6$, 150.9 MHz)