A Water Soluble Pd$_2$L$_4$ Cage for Selective Binding of Neu5Ac

Schaapkens, X.; van Sluis, R.N.; Bobylev, E.O.; Reek, J.N.H.; Mooibroek, T.J.

DOI
10.1002/chem.202102176

Publication date
2021

Document Version
Final published version

Published in
Chemistry-A European Journal

License
CC BY-NC

Citation for published version (APA):
A Water Soluble Pd$_2$L$_4$ Cage for Selective Binding of Neu5Ac


Abstract: The sialic acid N-acetylanimerinic acid (Neu5Ac) and its derivatives are involved in many biological processes including cell-cell recognition and infection by influenza. Molecules that can recognize Neu5Ac might thus be exploited to intervene in or monitor such events. A key obstacle in this development is the sparse availability of easily prepared molecules that bind to this carbohydrate in its natural solvent; water. Here, we report that the carbohydrate binding pocket of an organic soluble [Pd$_2$L$_4$]$^{4+}$ cage could be equipped with guanidinium-terminating dendrons to give the water soluble [Pd$_2$L$_4$][NO$_3$]$_2$ cage 7. It was shown by means of NMR spectroscopy that 7 binds selectively to anionic monosaccharides and strongest to Neu5Ac with $K_a = 24$ M$^{-1}$. The cage had low to no affinity for the thirteen neutral saccharides studied. Aided by molecular modeling, the selectivity for anionic carbohydrates such as Neu5Ac could be rationalized by the presence of charge assisted hydrogen bonds and/or the presence of a salt bridge with a guanidinium solubilizing arm of 7. Establishing that a simple coordination cage such as 7 can already selectively bind to Neu5Ac in water paves the way to improve the stability, affinity and/or selectivity properties of M$_2$L$_4$ cages for carbohydrates and other small molecules.

Introduction

Sialic acids, a class of $\alpha$-keto acid sugars, have been found on the distal ends of cell surface glycoconjugates and play a wide variety of biological roles, especially in cell-to-cell recognition.[1] The most common member is N-acetylanimerinic acid (Neu5-

Ac), which is utilized by influenza or other viruses to enter mammalian cells.[2] Neu5Ac is also found at the end of tetrasaccharide Sialyl Lewis$^x$ 1 (sLe$^x$, Figure 1a) and clinical studies showed the importance of sLe$^x$ in leukocyte adhesion deficiency,[3] inflammatory response,[4] (in vitro) fertilization,[5] coronavirus binding,[6] and cancer metastasis.[7] An example is the study of molecular sensors for sLe$^x$, which facilitates extravasation of cancer cells out of the blood stream (metastasis), displaying leukocyte mimicry.[8] While Neu5Ac is typically O-linked to other molecules as $\alpha$-anomer (axial carboxylate),[9] which is utilized by influenza or other viruses to enter mammalian cells.[2] Neu5Ac is also found at the end of tetrasaccharide Sialyl Lewis$^x$ 1 (sLe$^x$, Figure 1a) and clinical studies showed the importance of sLe$^x$ in leukocyte adhesion deficiency,[3] inflammatory response,[4] (in vitro) fertilization,[5] coronavirus binding,[6] and cancer metastasis.[7] An example is the study of molecular sensors for sLe$^x$, which facilitates extravasation of cancer cells out of the blood stream (metastasis), displaying leukocyte mimicry.[8] While Neu5Ac is typically O-linked to other molecules as $\alpha$-anomer (axial carboxylate),[9]

Figure 1. a) Sialyl Lewis$^x$ 1 (sLe$^x$) bound to the E-selectin binding site as found in PDB entry 1G1T with charge assisted HBs highlighted in magenta[11]. b) Two artificial carbohydrate receptors that bind to Neu5Ac in part due to charge assisted HBs and/or the presence of a salt bridge with guanidinium solubilizing arms. c) Left: macrocycle 4 selective for GlcNAc-OMe. R$^4$ = carboxylate terminating dendrimer.$^{15}$ Right: Pt$_2$L$_4$ cage 5 selective for disaccharides where L = anthracene functionalized dipyrindil ligand and R$^2$=O(CH$_2$)$_3$OMe.$^{16}$ d) Pd$_2$L$_4$ cages 6 (X=BF$_3$)$^{10}$ and 7 (X=NO$_3$)$^{10}$ where L = isophthalamide linked dipyrindil ligand and R$^2$ a solubilizing group.

[a] X. Schaapkens, R. N. van Sluis, E. O. Bobylev, Prof. Dr. J. N. H. Reek, Dr. T. J. Mooibroek
Van ’t Hoff Institute for Molecular Sciences
University of Amsterdam
Science Park 904, 1098 XH, Amsterdam (The Netherlands)
E-mail: t.j.mooibroek@uva.nl
Supporting information for this article is available on the WWW under https://doi.org/10.1002/chem.202102176
© 2021 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
the chemical antecedent to such linkages is cytidine-5'-monophospho-β-Neu5Ac.[10] Unlinked Neu5Ac is predominantly present as the β-anomer in solution and has a rich in vivo chemistry.[11,12] Molecules that can bind selectively to Neu5Ac and its derivatives might thus be exploited to understand, monitor, or intervene in a range of biological processes.

Inspiration for the development of Neu5Ac binders can be drawn from selectins, a subclass of lectins (carbohydrate binding proteins). As is illustrated in Figure 1a for crystal structure 1GT (human E-selectin), there is a high degree of interaction complementarity in the binding mode with sLeX.[11] Notably, the anionic Neu5Ac fragment of sLeX forms a salt-bridge with an arginine residue and the galactose fragment has strong charge assisted hydrogen bonds (HBs) with the same arginine residue (highlighted in magenta).

The beneficial effect of employing charge assisted HBs to bind sialic acid derivatives has been mimicked by the artificial borane part and binding is further enhanced by the presence of the nearby guanidinium group. As is illustrated in Figure 1a for crystal structure drawn from selectins, a subclass of lectins (carbohydrate binding proteins). As is illustrated in Figure 1a for crystal structure 1GT (human E-selectin), there is a high degree of interaction complementarity in the binding mode with sLeX.[11] Notably, the anionic Neu5Ac fragment of sLeX forms a salt-bridge with an arginine residue and the galactose fragment has strong charge assisted hydrogen bonds (HBs) with the same arginine residue (highlighted in magenta).

We thus wondered what the binding properties of a cage such as 6 would be in aqueous solution, in particular for anionic carbohydrates like Neu5Ac. To this end, the solubility handles of the ligands in 6 were replaced by guanidinium-terminating dendrons to make the [Pd4L6]4+ cage 7 (Figure 1d). Herein, we report that 7 has selective affinity for anionic sugars, particularly for Neu5Ac, and that 7 has very low to no affinity for common neutral mono- and disaccharides.

Results and Discussion

The synthesis of the ligand precursor to cage 7 (penta nitric acid salt 13) is shown in Scheme 1. The starting trimesic pentafenourophenyl (PFP) ester 8 and amine 9 were synthesized according to literature procedures[13] and then coupled to each other to form bis-PFP ester 10 in 62% yield by using a previously reported protocol.[14,15] Subsequently, the remaining PFP esters of 10 were substituted by 3-aminopyridine to afford 11.[16] Deprotection of the Boc groups of 11 followed by basification and treatment with bis-boc-pyrazolocarboxamide afforded hexa-boc guanidine 12 in 69% yield.[17] The desired guanidinium ligand 13 could be obtained in 74% yield after treatment of 12 with 1 M nitric acid in a water/1,4-dioxane solvent mixture. The pyridyl rings in 13 could be selectively deprotonated by the addition of two equivalents of sodium hydride. The subsequent addition of a Pd(NO3)2 solution (0.55 eq.) gave cage 7 in a quantitative yield based on 1H NMR.

For example, macrocycle 4 (Figure 1c, left)[13] is highly selective for GlcNAc-β-OMe (Kθ ≈ 18,000 M−1 in water) by encapsulating the carbohydrate by regular HBs and CH···π interactions.[18] However, such covalent macrocycles are not selective for Neu5Ac or related anionic carbohydrates.[16c–e] This can be rationalized by the presence of anionic dendrimers (R+) used to solubilize the hydrophobic binding pockets.

Contrarily, coordination cages based on a dipyreryl ligand (L) an a square planar d8 metal (M, for example Pd2+ or Pt2+) are positively charged and are known to have affinity for anionic guests.[18] Another advantage of such coordination cages is the reversibility of the pyridine-metal bond. This allows for non-productive oligomerization products to become intermediates towards the desired macrocycle, thus evading low-yielding macrocyclization reactions needed in the synthesis of covalent cages. Recently, two examples of coordination cages with the structure [M6L6]4+ have been reported with affinity for carbohydrates.[14,15] As is exemplified in Figure 1c (right), [Pt4L4]4+ cage 5 is based on dipyreryl ligands separated by anthracene moieties. This cage bound selectively to D-sucrose (Kθ ≈ 1,000 M−1) by virtue of shape-complementarity and multiple CH···π interactions between the carbohydrate and the polyaromatic cavity of 5.[16] A similar [M6L6]4+ cage was reported (6 in Figure 1d) where the dipyreryl ligands are separated by isoptylamides, similar to macrocycle 4.[10] Cage 6 had an organic solubility handle and could be studied in CD2Cl2 containing 10% DMSO-d6, where selectivity towards n-octyl-β-glucoside (Kθ ≈ 51 M−1) versus n-octyl-β-galactoside (Kθ ≈ 29 M−1) was observed.

For example, macrocycle 4 (Figure 1c, left)[13] is highly selective for GlcNAc-β-OMe (Kθ ≈ 18,000 M−1 in water) by encapsulating the carbohydrate by regular HBs and CH···π interactions.[18] However, such covalent macrocycles are not selective for Neu5Ac or related anionic carbohydrates.[16c–e] This can be rationalized by the presence of anionic dendrimers (R+) used to solubilize the hydrophobic binding pockets.

Contrarily, coordination cages based on a dipyreryl ligand (L) an a square planar d8 metal (M, for example Pd2+ or Pt2+) are positively charged and are known to have affinity for anionic guests.[18] Another advantage of such coordination cages is the reversibility of the pyridine-metal bond. This allows for non-productive oligomerization products to become intermediates towards the desired macrocycle, thus evading low-yielding macrocyclization reactions needed in the synthesis of covalent cages. Recently, two examples of coordination cages with the structure [M6L6]4+ have been reported with affinity for carbohydrates.[14,15] As is exemplified in Figure 1c (right), [Pt4L4]4+ cage 5 is based on dipyreryl ligands separated by anthracene moieties. This cage bound selectively to D-sucrose (Kθ ≈ 1,000 M−1) by virtue of shape-complementarity and multiple CH···π interactions between the carbohydrate and the polyaromatic cavity of 5.[16] A similar [M6L6]4+ cage was reported (6 in Figure 1d) where the dipyreryl ligands are separated by isoptylamides, similar to macrocycle 4.[10] Cage 6 had an organic solubility handle and could be studied in CD2Cl2 containing 10% DMSO-d6, where selectivity towards n-octyl-β-glucoside (Kθ ≈ 51 M−1) versus n-octyl-β-galactoside (Kθ ≈ 29 M−1) was observed.
Information, an alternative route to prepare ligand 13 was unproductive due to reaction compatibility issues between PFP-esters and the Boc-linked NH of diBoc-protected guanidines.}

The synthesis of 7 from 13 could also be followed in D$_2$O by $^1$H NMR, as shown by the stacked spectra in Figure 2a (top). Upon addition of 2 equivalents of NaOH, the resonances belonging to the pyridyl ring in 13 were found significantly upfield, which is in line with deprotonation of the pyridyl nitrogens.

Subsequent stepwise addition of Pd(NO$_3$)$_2$ resulted in the disappearance of dipyril ligand signals with the proportional appearance of a new set of resonances.

The emerging Pd-complex and its parent ligand thus appear to be in slow exchange relative to the NMR time scale. All proton resonances of the resulting well-defined spectrum could be identified and are consistent with cage formation (see bottom of Figure 2a). In particular the large downfield shifts of proton resonances such as a (8.41—8.86) and d (8.77—9.74) are indicative of pyridyl-palladium coordination.[21]

The 2D DOSY NMR of this sample reveals that the diffusion constant (D) of 7 ($\log(D) = -9.83$) is substantially larger than that of the neutralized ligand ($\log(D) = -9.57$) which is also consistent with cage formation (see bottom of Figure 2a). Furthermore, the isotope distribution of a species with largest monoisotopic mass $m/z = 476.1798$ (Figure 2b) measured with CSI HRMS is in agreement with a simulated distribution of [7(NO$_3$)$_3$Cl]$^{7+}$ with largest monoisotopic mass of $m/z = 476.1823$. The modelled molecular formula of [7(NO$_3$)$_3$Cl]$^{7+}$ includes some deuterium and Cl$^-$ because the solution was measured from a D$_2$O sample in undeuterated solvent containing trace amounts of salts (see Supporting Information and Figures S74–S90 for details).

With the water soluble [Pd$_2$L$_4$]$_{16}^{16-}$ cage 7 in hand, the binding affinity for the carbohydrates listed in Table 1 was investigated by $^1$H NMR titration experiments in D$_2$O. Titrations with charge neutral carbohydrates 14–26 to about 140 mM only resulted in minor near-linear peaks shifts of some resonances of 7 with $\Delta \delta_{\max} \approx 0.02$ p.p.m. on average (see Figures S92–S104). These shifts could not result from the dilution of 7, as a dilution study in the concentration range used during titrations revealed that all resonances remained stationary (see Figure S91). Attempts to fit these shifts to a binding model was not feasible and could only be roughly modelled (not fitted) to binding with an affinity around or below the detection limit of $\approx 3$ M$^{-1}$. We thus interpret these shifts as resulting from very weak binding of $\leq 3$ M$^{-1}$ (entry 1, Table 1), spanning only the very start of possible binding curves.

Addition of neutralized solutions of 27–29, did result in significant non-linear shifting of the resonances of 7, with clear signs of saturation (Figures S105–S107). These shifts could be fitted accurately to a 1:1 binding model resulting in the binding constants listed in entries 2–4 of Table 1.

Selected spectra of the titration with 29 are shown in Figure 3a. With increasing concentration of Neu5Ac 29, aro-

![Figure 2](image-url)

**Figure 2.** a) Top: stepwise formation of 7 by deprotonation of the pyridyls in 13 (top two spectra) followed by stepwise addition of Pd(NO$_3$)$_2$ as a 10 mM solution. Bottom: overlay of the 2D DOSY NMR spectra of 7 and 13 with deprotonated pyridyl rings. b) representation of 7 with labelled protons and CSI HRMS isotope distribution with indicated highest isotopic mass as measured from a D$_2$O solution (top, blue) and simulated (bottom, green).

<table>
<thead>
<tr>
<th>Entry</th>
<th>Guest</th>
<th>$K_a$ (M$^{-1}$)[h]</th>
<th>Correlation of fit ($r^2$) [b]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14–26</td>
<td>$\leq 3$</td>
<td>–</td>
</tr>
<tr>
<td>2</td>
<td>27</td>
<td>6.6 ± 0.2</td>
<td>0.9974</td>
</tr>
<tr>
<td>3</td>
<td>28</td>
<td>8.3 ± 0.3</td>
<td>0.9609</td>
</tr>
<tr>
<td>4</td>
<td>29</td>
<td>24.0 ± 0.2</td>
<td>0.9981</td>
</tr>
</tbody>
</table>

[a] Obtained by curve fitting $^1$H NMR resonance shifts using HypNMR.[22]  
[b] Correlation of fit computed over all fitted data points.

Table 1. Overview of binding studies performed with 7 and the structures of titrants 14–29.
matic signals a, b, d, f and g shifted downfield and broadened slightly. In the presence of a large excess of 29, a minor species with lower symmetry arose, marked with blue asterisks in Figure 3a. This species disappeared after heating at 60°C (‘A.H.’), while the major species of [7–29] persisted and another minor symmetrical species appeared (marked with diamonds). This new set of resonances was nearly identical to a solution of deprotonated ligand 13 and neutralized 29 at the same concentration as present in the titration (shown at the top of Figure 3a and assigned with subscript ‘L’). The minor symmetrical species (diamonds) present after heating is thus probably ligand bound to 29, with the ligand originating from cage decomposition (likely driven by Pd-plating). To quantify binding of the major symmetrical species 7, the peak shifting could be fitted to a 1:1 model using HypNMR[22] as shown in Figure 3b. This fitting gave the association constant ($K_a$) of 24.0 ± 0.2 M$^{-1}$ listed in entry 4 of Table 1 with an excellent goodness of fit of $r^2$ = 0.9981.

To further understand binding of 7 with 29, selective 1D nuclear Overhauser effect (nOe) NMR spectra were recorded after heating at 60°C, as shown in Figure 4b. For reference purposes, the $^1$H NMR spectrum of 29 is shown in Figure 4a together with an assignment that is based on a full structure elucidation of 29 (see Section S3 for full details). As is shown in Figure 4b, clear nOe signals were observed between irradiated cage protons d, f/a and c/g and carbohydrate protons c3, c9 and c11. Proton d of 7 also shifted the most when binding to 29 (see Figure 3). The much weaker nOe with proton b suggest that these are further away from the C–H protons of Neu5Ac 29. It is worth pointing out that protons c3, c9 and c11 are located on different sides of 29 and that c3 and c9 are even opposite to each other (see Figure 4a). The fact that nOe’s were observed to these signals thus implies that 29 is bound inside 7.

The titration data (Table 1) clearly show that 7 is selective for anionic monosacharides, in particular for Neu5Ac 29, and that 7 binds about equally well to glucoronate 27 and galacturonate 28. To rationalize these observations, some molecular models were generated of cage 7 bound to the β-anomers of anionic carbohydrates 27, 28 and 29 using density functional theory (DFT, see Section S4 for details).

Shown in the top of Figure 5a is the space filling representation of a model of 7 that fully encapsulates glucoronate 27, with OH-3 protruding from one of the portals. Interestingly, as is shown in the bottom of Figure 5a, OH-3 is not involved in a HB to an amide, while all the other hydroxyl groups are. Rather, OH-3 is involved in the only (and weak) charge assisted HB involving the pyridyl C–H (blue arrow). The carboxylate is furthermore held in place by four amidic HBs with H–O distances in the range of 2.0–2.3 Å. This four-pronged charge assisted HB interaction might rationalize the selectivity of 7 for the anionic glucoronate 27 over the neutral glucose 14. The model of 7 with galacturonate 28 is very similar.
to the [7-27] model, which is evident from the overlaid HB patterning at the bottom of Figure 5a (see also Figure S116). Here too the carboxylate is held in place by four amide HBs and OH-3 is the only hydroxyl that has no HB to an amide and only weakly to a pyridyl C (blue arrow). This absence of an amide HB with OH-3 in both 27 and 28 might explain the lack in selectivity observed between these two carbohydrates. As can be seen in the top of Figure 5b, modelling the binding of 7 to Neu5Ac 29 resulted in a complex where the guanidinium arm (C2) protrudes through one of the four portals of 7. The glyceryl (C7-C9) and acetyl (C10, C11) fragments point out of two other portals (see Figure S117 for details). Presumably, the carboxylate and hydroxyl group on C2 of 29 are too large to be accommodated in the interior of 7 in the same manner as modelled for the carboxylates in 27 and 28 (Figure 5a). Moreover, the guanidinium arm that was incorporated in this model of 7 formed a strong salt-bridge with the carboxylate of 29. This is particularly evident from the HB pattern in the [7-29] model shown in the bottom of Figure 5b. Highlighted in magenta are the three strong charge assisted HBs with H--O distances in the range of 1.8–2.1 Å, which resemble the charge assisted HBs found in PDB entry 1G1T (Figure 1a). Additionally, there are three amide HBs and four relatively strong HBs involving pyridyl C--H’s with H--O≈2.2 Å (blue arrows). The salt-bridge formation, together with the encapsulation of 29 via three of the four portals of 7, the three amide HBs and the four HBs with pyridyl C--H’s observed in the model can offer a rationale for the ~3–4 fold selectivity observed for 29 over 27/28 (see Table 1).

**Figure 5.** Models of 7 bound to anionic carbohydrates calculated by DFT/ωB97X-D/6-31G*. a) top: space filling representation of 7 bound to 27; bottom: hydrogen bonding pattern of the [7-27] model overlaid with the [7-29] model. The C-atoms of 27 and 28 are labelled in red and the distances (in Å) are averages for both models. See Figure S116 for more details; b) top: space filling model of Neu5Ac 29 bound to a model of 7 with a guanidinium arm incorporated; bottom: hydrogen bonding pattern of [7-29] with distances in Å and C-atoms of 29 labelled in red. The structure is similar to a structure without such a guanidinium arm, see Figure S117.

**Conclusions**

A previous reported [Pd4L4]3+ cage was rendered water soluble by introducing guanidinium solubility groups on the ligand to form the [Pd4L4]3+ cage. Titration of 7 with common carbohydrates revealed selectivity for anionic carbohydrates. While the binding affinities towards neutral mono- and disaccharides was near or under the detection limit of K ≈ 3 M⁻¹, the binding affinities for anionic carbohydrates 27, 28 and 29 were found to be 6.6, 8.3 and 24 M⁻¹, respectively. Cage 7 is thus at least 8 times more selective for 29 than for neutral saccharides. This selectivity could be rationalized based on DFT modelling and likely originates from complementary ion-pair formation between the anionic sugars and cationic 7 by various (charge assisted) HBs, much like those found in nature.

The selectivity for anionic carbohydrates in aqueous solution is rare, and has not yet been reported with similar isophthalamide macrocyclic structures held together by either covalent or coordination bonds. While the affinities found are on the lower end, this study establishes the principle that a coordination cage such as 7 can selectively bind to anionic carbohydrates in water. This finding thus paves the way for further improvements of coordination cages to enhance their stability, affinity and/or selectivity properties. Ultimately, such developments could lead to selective synthetic lectins for Neu5Ac and its α/β-derivatives that can be used in biological studies such as Western blotting, targeted drug delivery.

**Acknowledgements**

This research was financially supported by the Netherlands Organization for Scientific Research (NWO) with VIDI grant number 723.015.006.

**Conflict of Interest**

There are no conflicts to declare.

**Keywords:** carbohydrate recognition · host-guest systems · molecular recognition · saccharides · self-assembly

