A Phage Display-Derived Peptide Binds To Human CD206
And Modeling Reveals A New Binding Site In The Receptor
Eliana Karina Asciutto, Sergei Kopanchuk, Anni Lepland, Lorena
Simón-Gracia, Carlos Alemán, Tambet Teesalu, and Pablo Scodeller

J. Phys. Chem. B, Just Accepted Manuscript • DOI: 10.1021/acs.jpcb.8b11876 • Publication Date (Web): 15 Feb 2019

Downloaded from http://pubs.acs.org on February 16, 2019

Just Accepted

“Just Accepted” manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides “Just Accepted” as a service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. “Just Accepted” manuscripts appear in full in PDF format accompanied by an HTML abstract. “Just Accepted” manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are citable by the Digital Object Identifier (DOI®). “Just Accepted” is an optional service offered to authors. Therefore, the “Just Accepted” Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the “Just Accepted” Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these “Just Accepted” manuscripts.
A Phage Display-Derived Peptide Binds To Human CD206 And Modeling Reveals A New Binding Site In The Receptor

Eliana K. Asciutto\textsuperscript{a*}, Sergei Kopanchuk\textsuperscript{b}, Anni Lepland\textsuperscript{c}, Lorena Simón-Gracia\textsuperscript{c}, Carlos Aleman\textsuperscript{f}, Tambet Teesalu\textsuperscript{c,d,e*}, Pablo Scodeller\textsuperscript{c*}

\textit{a) School of Science and Technology, National University of San Martin (UNSAM) and CONICET. Campus Migueletes, 25 de Mayo y Francia, Buenos Aires, Argentina.}
\textit{b) Institute of Chemistry, University of Tartu, Ravila 14, Tartu, 50411, Estonia.}
\textit{c) Laboratory of Cancer Biology, Institute of Biomedicine and Translational Medicine, University of Tartu, Ravila 14B, Tartu, 50411, Estonia.}
\textit{d) Cancer Research Center, Sanford-Burnham-Prebys Medical Discovery Institute, 10901 North Torrey Pines Road, La Jolla, 92037 California, USA.}
\textit{e) Center for Nanomedicine and Department of Cell, Molecular and Developmental Biology, University of California, Santa Barbara Santa Barbara, 93106 California, USA.}
\textit{f) Departament d’Enginyeria Química and Barcelona Research Center in Multiscale Science and Engineering, EEBE, Universitat Politècnica de Catalunya, C/ Eduard Maristany, 10-14, 08019, Barcelona, Spain}

*: Corresponding authors: easciutto@unsam.edu.ar, tteesalu@sbpdiscovery.org, pablo.david.scodeller@ut.ee

\textbf{ABSTRACT}
We recently identified a tumor homing peptide (mUNO, sequence: “CSPGAK”) that specifically interacts with mouse CD206 to target CD206/MRC1-expressing tumor-associated macrophages in mice. Here, we report studies on the binding of mUNO to human recombinant CD206 (hCD206) and on modeling the mUNO/hCD206 interaction by computational analysis. Fluorescence anisotropy analysis demonstrated that fluorophore-labeled mUNO interacts with hCD206. Microsecond timescale molecular dynamic simulations and docking predictions showed that mUNO binds to a newly identified epitope between C-type lectin domains 1 and 2. The physical mechanisms that contribute to the docking interactions of mUNO include electrostatic interactions, aromatic interactions, and hydrogen bonds. We also demonstrate selectivity of FAM-mUNO for CD206\textsuperscript{+} cultured human macrophages. The peptide mUNO appears to be the first ligand capable of interacting with this epitope of hCD206, to which no ligands have been reported. Our study has implications for targeting human M2-like tumor-associated macrophages, a subpopulation of immune cells with major protumoral role.
INTRODUCTION

Tumor-associated macrophages skewed towards the M2 phenotype (M2-TAMs) play major roles in several malignant tumors\(^1\), and are recognized as important targets to diagnose and treat cancer\(^2\). M2-TAMs express higher levels of anti-inflammatory cytokines, angiogenic factors, and proteases than M1-like macrophages. M2-TAMs contribute to cancer relapse\(^3\) and promote tumor growth by driving angiogenesis\(^4\) and immunosuppression\(^5\). M2-TAMs are identified based on the expression of certain marker receptors, such as CD163 and CD206. Because of the key role of M2 TAMs in tumor progression and treatment resistance, the amount of M2 macrophages in tumor tissue and in sentinel lymph nodes can be used as a prognostic factor and measure of therapeutic outcome. \(\gamma\)-Tilmanocept, a mannose-based ligand of lectin domain of human CD206 (hCD206), is a FDA-approved radiolabeled compound for lymphatic mapping in early cancer patients\(^2\). \(\gamma\)-Tilmanocept labels M2 macrophages in lymph nodes that drain from a primary tumor, a prognostic factor of early cancer dissemination. \(\gamma\)-Tilmanocept is administered locally in the sentinel lymph nodes, thus missing potentially important malignant foci in other parts of the body. The reason for local administration of \(\gamma\)-Tilmanocept is that hCD206 and other mannose binding proteins are promiscuously expressed: CD206 is, in addition to tumor M2-TAMs, expressed in normal organs such as liver\(^6,7\). Moreover, other mannose receptors are expressed in the nonmalignant tissues, such as skin and intestinal and genital mucosa\(^8,9\). Lately, a peptide with anti-inflammatory properties was reported to bind hCD206\(^10\), however the binding is not expected to be specific for M2 macrophages, as the peptide also interacts with CD47, RelB and Sirp-\(\alpha\). Recently, we identified a peptide named “mUNO” (CSPGAK) and its cyclic version CSPGAKVRC that specifically targets mCD206 expressed on MRC1/CD206-expressing TAMs (MEMs) across a spectrum of solid tumors of diverse origin and shows low accumulation in liver\(^11\). The peptide was identified by in vivo peptide phage display screens in mice bearing metastatic breast tumors. Using fluorescence anisotropy assays, we showed that mUNO interacted with recombinant mouse CD2016 (mCD206). Affinity targeting with mUNO allowed to guide fluorescent payloads inside the MEMs. BLAST analysis revealed that mUNO sequence includes SPGAK pentapeptide motif that is present in all CD206-binding collagens.

mUNO specificity to mCD206 can be understood through the study of peptide-receptor interactions. These interactions are affected by many factors, therefore, an accurate description must consider not only collective motions but also local flexibility of the peptide as well as the receptor to predict binding energies\(^12\). This can be achieved using all-atom molecular dynamics (MD) simulations. Such approach, in which all degrees of freedom and fine interaction details are explicitly considered, allows understanding how structural fluctuations and the dominant modes of motion affect the binding process.

Here we report cell-free and cellular binding studies of mUNO to hCD206 and explore, using computational methods, the mechanisms of the interaction and map the potential binding sites.

EXPERIMENTAL AND THEORETICAL METHODS
Materials
All peptides were purchased from TAG Copenhagen (Denmark). FAM-mUNO is FAM-X-CSPGAK-COOH, where X is an aminohexaonic acid linker and the peptide has a free C-terminus. Unlabeled mUNO is X-CSPGAK-COOH. FAM-CPMTDNE is FAM-X-CPMTDNE-COOH, where X is an aminohexaonic acid linker and the peptide has a free C-terminus.

In vitro binding studies
Fresh blood was obtained from University of Tartu Hospital. Blood was diluted 1:1-1:1.5 in PBS in 50mL tubes and mixed carefully. Ten mL of Ficoll–PaqueTM Plus (GE Healthcare) at 37°C was added to a new tube, 35mL of diluted blood was slowly layered over Ficoll without mixing. Tubes were centrifuged for 35 minutes at 400g at 20°C in a swinging-bucket rotor without brake. The upper layer containing plasma and platelets was aspirated and white layer containing peripheral blood mononuclear cells (PBMC) was transferred to a new 50mL tube. Thirty mL of ice cold buffer (PBS + 2mM EDTA-Na2) was added to PBMC, mixed and centrifuged at 300g at 20°C for 10 minutes. Supernatant was aspirated and 40mL of ice cold buffer was added and centrifuged at 200g at 20°C for 15 minutes to get rid of platelets still remaining inside the PBMC mixture. Supernatant was aspirated and cells were resuspended in 20mL of ice cold buffer for cell counting. For cell counting equal amounts of cells and 0.4% Trypan blue (Smart Mix) was mixed and counted with TC10TM Automated Cell Counter (BioRad). Mixture was centrifuged at 300g at 20°C for 10 min. Supernatant was aspirated and cells resuspended in ice cold buffer (80µL of buffer per 10⁷ total cells). Twenty µL of CD14 Microbeads (MACS Miltenyi Biotec) were added per 10⁷ total cells. Beads and cells were mixed and incubated for 15 min at 4°C, gently mixing every 5 min. Meanwhile the LS column (MACS Miltenyi Biotec) was placed in the magnetic field MACS Separator and washed with 3mL of ice cold buffer. After 15 min incubation mixture of cells and beads was added to the column and left to pass through with gravity. Three washing steps were performed, 3mL of ice cold buffer per each wash. Column was removed from separator, 5mL of ice cold buffer was added and cells were flushed by pushing the plunger that comes with the column. Cells were collected to a new 15mL tube and prepared for cell counting.

In 24-well plates with glass coverslips, 3.5x10⁵ CD14⁺ cells were seeded per well in 500µL RPMI with 10% FBS and penicillin/streptomycin. Forty-eight hours later 250µL of media were aspirated and replaced with 250µL of fresh full media. Ninety-six hours after the seeding, 250µL of media were aspirated and replaced with 250µL of fresh full media, and IL-4 was added to a final concentration of 40ng/mL (human recombinant human IL-4, carrier free, from Biolegend, Cat# 574002). Cells were then left for 48 hours and then FAM-mUNO and FAM-CPMTDNE were added to the well containing 500 µL of media (final concentration of 3nM), incubated for 45 min at 37°C, washed 2 times with full media, washed once with PBS, fixed with 4% PFA for 10 min at room temperature, and permeabilized with Triton. Cells were immunostained with rabbit anti-FAM (Invitrogen, Catalog# A889, dilution 1/250) and mouse anti-CD206 (Biolegend, catalog # 321102), dilution 1/200) overnight at 4°C, followed by incubation with goat anti-rabbit Alexa546 (Invitrogen, catalog # A11010, dilution 1/400)
and goat anti-mouse Alexa647 (Invitrogen, catalog # A21235, dilution 1/300), and counterstained with DAPI. The coverslips were mounted on a glass slide and imaged with a LSM-710 Zeiss Confocal microscope.

Fluorescence anisotropy
The fluorescence anisotropy measurements methodology was modified from the method developed for characterization of melanocortin receptors\textsuperscript{13}. The stocks of peptide (FAM-mUNO and mUNO) in mQ water were stored at −20 °C and diluted with incubation buffer (IB) (10 mM Na-HEPES, 150 mM NaCl, 1 mM CaCl₂, 0.1% Pluronic F-127, pH 7.4) on the day of experiment. The concentration of fluorescent ligands was determined by absorbance of FAM (ε₄₉₅ = 75000 M⁻¹ cm⁻¹). The recombinant human macrophage mannose receptor rhMMR/CD206 (R&D Systems, catalog number: 2534-MR-050) and recombinant mouse CD163 (R&D Systems, catalog number: 7435-CD-050) were reconstituted at 1 mg/ml in IB, aliquoted and stored at −20 °C. In all experiments, FA signal measurements were done under pseudo first-order conditions with 100 nM singly labelled FAM peptides and 0.4 µM proteins. Black 384-well round bottom polystyrene NBS surface microplates (Corning, Product No. 3676) that were found to give optimal results for our assays (low background of fluorescence and low adsorption of ligands onto plastic surface) were used in all FA experiments. Assays were carried out in kinetic mode in a total volume of 30 µl at 25 °C on the PHERAstar (BMG Labtech, Germany) microplate reader using an optical module with excitation and emission filters of 485 nm (slit 10 nm) and 520 nm (slit 10 nm), respectively. Dual emission detection mode allows simultaneous recording of intensities that are parallel ($I_{∥}$) and perpendicular ($I_{⊥}$) to the plane of excitation light. Sensitivities of channels (G factor) were corrected with gain adjustment of the photomultiplier tubes (PMTs) using fluorescein as a standard. The background fluorescence (receptor and buffer components in the absence of fluorescent ligand) was subtracted independently from all intensity channels. FA signals at time t after initiation of binding reaction was calculated as parameters r(t) from the equation:

$$r(t) = \frac{I(t)_{∥} - I(t)_{⊥}}{I(t)_{∥} + 2 \cdot I(t)_{⊥}}$$

Homology models of mannose receptor hCD206
Two models were generated to describe the extracellular portion of the ectodomain of hCD206 using the Swiss-Model automated comparative protein modeling server\textsuperscript{14}. Model 1 (M1) was built using Endo180 as template protein (PDB:5EW6) and a second model, Model 2 (M2) is based on the recently published crystal structure of the CysR-CTLD2 fragment of CD206 (PDB:5XTS)\textsuperscript{15}. Both models generated a folded structure for hCD206 with a total of 475/469 residues plus a Ca²⁺ metal ion for M1 and M2 respectively. For both models, missing residues were added with The Swiss Model software\textsuperscript{16, 17, 18} (Figure S4).

The modeled portion of hCD206 consists of four domains: a cysteine rich domain (CysR), a fibronectin type II domain (FNII), and two C-type lectin domains (CTLDs). Both models have, in the CysR domain, a
β trefoil fold-like structure formed by pairs of beta hairpins, alpha helices and short 3\textsubscript{10} helices. FNII domain is formed mainly by beta sheets and 3\textsubscript{10} helices, and lectin domains are more ordered with a total of 5 pure alpha helices connected by beta sheets. Although slight differences were found between these two preliminary models, we observed that once they were solvated and temperature and pressure equilibration was performed, the models became very similar (Figure S5). Differences were found only at the CysR domain, the most flexible domain, and because of these differences, we decided to work with the two models.

**CD206-mUNO Molecular Dynamic Simulations**

Models 1 and 2 for hCD206 and four mUNO peptides were combined using tLeap on AmberTools17\textsuperscript{19}. Missing hydrogen atoms were added in this step. Na\textsuperscript{+} and Cl\textsuperscript{-} ions were added to neutralize the total charge of the complex and to model a NaCl solution at 150 mM concentration. The system was solvated with TIP3P water molecules\textsuperscript{20}, resulting in a 105 x 117 x 141 Å cubic box, containing a total of 45499 water molecules. The ff14SB force field\textsuperscript{21} was used to describe the interactions of the complex. Molecular Dynamics simulations were carried out using The Amber 16 Molecular dynamics package\textsuperscript{19}. The system: complex (hCD206 and 4 mUNOs) in NaCl solution was first minimized using 1000 steps of the steepest descent algorithm followed by 4000 steps of the conjugate gradient algorithm. The system was then heated from 0K to 298K during 0.5 ns under the Langevin equilibration scheme using a collision frequency of 2.0 ps\textsuperscript{-1}, and under constant volume periodic boundary conditions. Next, a short (400 ps) run was performed using constant pressure periodic boundary condition with an average pressure of 1 atm. Isotropic position scaling was used to keep the pressure constant with a relaxation time of 1 ps. Next, 30 ns equilibrations were performed. Once the system was equilibrated, production runs were launched. Three independents 100 ns conventional production runs were performed for each model in an NVT ensemble, keeping the temperature constant at 298 K with Langevin dynamics. To treat non-bonded interactions a cutoff of 10 Å was used. Long range interactions were described using The Particle Mesh Ewald (PME) method\textsuperscript{22}. All bonds containing hydrogen atoms were constrained by the SHAKE algorithm\textsuperscript{23}. The time step for numerical integration was 2 fs.

Each independent run was started with different velocities. To ensure enough difference between initial configurations, each run was first heated to a different temperature (between 298 and 305 K), and then decreased to 298K.

Next, four independent aMD simulations were performed \textsuperscript{24}, using boost potentials applied to the total energy and an extra boost to dihedral angles with: 

\[ E_{\text{dih}} = 3.5 \times N_{\text{res}} + V_{\text{dih},\text{avg}}, \quad \alpha_{\text{dih}} = \left(\frac{1}{3}\right) \times N_{\text{res}}, \quad E_{\text{tot}} = V_{\text{total,avg}} + \left(\frac{1}{3}\right) \times N_{\text{atoms}}, \quad \alpha_{\text{tot}} = \left(\frac{1}{3}\right) \times N_{\text{atoms}} \]

aMD simulations were started from the previous 100 ns conventional MD runs at 298K.

Using the methodology just explained in the paragraph above, three extra conventional MD simulations 1.1 µs long were run with mUNO placed in the high affinity positions identified in the first part of this work.
Docking predictions
Molecular Docking was performed using Autodock Vina\textsuperscript{25}, where the peptide is treated as a flexible molecule with active rotatable bonds. Additionally, the three most populated cluster for hCD206 from aMD simulations were used as input coordinates. The search space is big enough to include the entire receptor and leave enough space for the peptide to fit in. This kind of search is known as "blind docking". Autodock Vina uses an empirical scoring function to assign scores to the reported binding poses.

Cluster analysis.
MD ensembles were generated using all trajectories from simulations as described before. From these ensembles the most populated states were extracted via a cluster analysis of the full pool of configurations. Cluster analysis was performed using a distance metric coordinate, RMSD and the average-linkage algorithm\textsuperscript{26}.

Binding Free Energy Calculations
MMGBSA calculations\textsuperscript{27} were performed from short MD trajectories 1 ns long each, using 200 snapshots each, for the free energy calculation. Within the MMGBSA approach, the binding process is decomposed in a solvation event and a binding in vacuum event, following a thermodynamic cycle. Solvation free energies were calculated by solving the Generalized Born (GB) equation for each of the receptor, ligand, and complex. Hydrophobic contributions were added via an empirical term. Free energy change of binding in vacuum was obtained by calculating the average interaction energy between hCD206 and mUNO. Explicit water molecules and ions were removed from the trajectory to avoid time consuming solvent-solvent interactions calculations. Finally, free energy of binding of the complex was calculated using Eq. (1):

\[ \Delta G_{\text{bind,solv}} = \Delta G_{\text{bind,vac}} + \Delta G_{\text{solv,complex}} - (\Delta G_{\text{solv,ligand}} + \Delta G_{\text{solv,receptor}}) \]  

(Eq.1)

Note that we are assuming (as commonly done in MMGBSA) changes in entropy to be negligible in comparison with changes in energy.

Electrostatic Potential
Electrostatic Potential for hCD206 was calculated using APBS software\textsuperscript{28}. The package finds the numerical solution of the Poisson-Boltzmann equation, a continuum model for describing electrostatic interactions between molecules in solution.

RESULTS
mUNO binds to hCD206
In our previous study, we established that mUNO can interact with mCD206\textsuperscript{11}. Whereas mouse and human CD206 proteins are 82.1\% identical, it was not shown that the peptide is capable of interacting with hCD206 – an aspect relevant to potential translational applications of the mUNO peptide.
Therefore, we performed experimental binding studies of FAM-mUNO on recombinant hCD206 and on control proteins using fluorescence anisotropy (FA) - a cell-free, solution-based assay. The hCD206, when incubated with FAM-mUNO, caused time-dependent increase of anisotropy (Figure 1, black curve) indicating interaction. This increase was not from the peptide alone as the peptide alone showed a minimal increase (Figure 1, green curve). In contrast, FAM-mUNO did not interact with CD163 (Figure 1, red curve), as the increase in anisotropy seen in this case was the same as that of the peptide alone. CD163 is another scavenger receptor, which like CD206 is also a marker of M2-TAMs and contains a cysteine-rich domain. Furthermore, when hCD206 was preincubated with an excess of unlabeled mUNO, a minimal increase of anisotropy was seen from FAM-mUNO (Figure 1, blue curve), indicating that the binding of FAM-mUNO to hCD206 is from the peptide and not from the FAM label, or the aminohexanoic acid linker used in the mUNO synthesis.

Figure 1. Binding of FAM-mUNO to hCD206. Change in fluorescence anisotropy of FAM-mUNO during incubation with human recombinant CD206 (black curve) or with CD163 (red curve). The concentration of FAM-mUNO was 100nM and the concentration of CD206 and CD163 was 0.4µM. The anisotropy for the FAM-mUNO alone is shown in the green curve. Unlabeled mUNO (100µM) was coincubated with FAM-mUNO and CD206 (blue curve), showing that binding of FAM-mUNO was from the peptidic portion and not from the linker or the FAM.
**The most populated mUNO conformation is a U-shaped structure.**

A cluster analysis of the simulated trajectory using an RMSD metric determined that mUNO ensemble at 298 K is composed of three main conformations: a bend, a turn, and helical structure. The most populated cluster has an occupation of 82% and consists of a bent coil structure forming a U shape with the hydrophobic P3 and G4 at the center of the arc (Figure 2, left panel). The peptide end-to-end average distance in this conformation is about 10 Å. The second populated conformation has a global shape that is very similar to the first node. The main difference is that P3 has a different orientation, pointing inwards.

Time evolution of secondary structure reveals that mUNO switches between these two structures easily, which makes sense since it only involves a P3 rotation (Figure 2, right panel).

![Figure 2: Evolution of secondary structure from MD simulation of mUNO (right panel). Conformation of the most populated node for mUNO from a cluster analysis (left panel)](image)

Coordinates determined for the most populated cluster for mUNO were used as starting coordinates for all the following analysis.

**Estimation of binding sites from molecular dynamic simulations.**

The complex: solvated hCD206 + four mUNO peptides was simulated in two rounds: a first round of 100 ns conventional molecular dynamics (cMD), followed by a second round of 200 ns accelerated molecular dynamics (aMD). Four of these protocols were performed, starting from different coordinates, giving a total simulation time of several microseconds. From aMD ensembles we extracted the regions where mUNO spends most of the time. These regions are high occupancy sites and therefore are an indicator of high affinity sites. Figure 3 left (red blobs) shows averaged 3D mUNO occupancy maps in all aMD simulations. Highest mUNO occupancy is observed at lectin domains, mainly in the region between the two lectin domains. Another region with high occupancy is the end of FNII domain, just where the protein bends. The residues in close contact (less than 3Å) with at least one mUNO peptide have also been identified and are shown in Figure 3.
Next, hCD206+mUNO ensembles generated by aMD were clustered using an RMSD criterion. The three most populated conformations for hCD206 were used as input configurations for blind docking predictions. We used hCD206 most populated conformations from aMD ensembles as input for a docking calculation because, in this way, the flexibility of hCD206 is fully considered. As the binding sites were not initially known, a blind search was performed using Autodock Vina. A total of 28 docking poses were obtained and classified according to the hCD206 domain to which they belong and to the calculated docking score. 82% of the docked found poses fell in the lectin domains (Table S1), including poses between lectin domains CTLD1 and CTLD2 (pose referred to as CTLD1/2. This result is fully compatible with that obtained in the simulations as these docking poses cover regions that were extensively explored by mUNO in both cMD and aMD simulations.

**Ordering best docking poses using MMGBSA**

Binding energies of the top ten docking poses from Table S1 were refined using the Molecular Mechanics Generalized Boltzmann Surface Area method (MMGBSA)\(^\text{27}\). Calculated free energies and energy decomposition corresponding to the three poses with lower binding energies are reported in Table S2. From MMGBSA results, we analyzed the three most favorable poses belonging to different domains: pose “C1”, a pose in which mUNO is located in the long loop connecting the two lectin domains CTLD1 and CTLD2, resulting in a remarkably stable complex; a pose “C2” in the CysR domain and a pose “C3” in the FNII domain. The dominant interaction is electrostatic for the three poses. It is interesting to note that the electrostatic and solvation terms behave differently in poses C2 and C3. In C3, electrostatic interactions are the dominant interactions while the solvation term gives a very unfavorable contribution. In C2, the difference between the electrostatic and solvation term is less dramatic than in C3. Also, VDW interaction in C1 is considerably greater than in the other two poses: C2 and C3. This result is compatible with the sampling obtained from aMD simulations.

The three binding sites found so far C1, C2, and C3 were re-entered into a second cMD stage to
evaluate their stability. New cMD simulations, starting from C1, C2, and C3 were performed. As
stability control, we ran a fourth cMD simulation with mUNO translated 2 Å upwards from the C1
pose (C1T). The displacement of mUNO from its original position was tracked in each simulation to
control the stability of each potential binding pose. Simulations that started from C1, C2, and C3
revealed highly stable complexes. In the stability control simulation C1T, mUNO immediately
dissociated from the receptor. After 70 ns a binding event attempt at the C1 pose takes place, and
finally at 85 ns binding at C1 is once more observed (Figure S1A).

To confirm that the binding observed is specific to the aminoacid sequence of mUNO, we
performed simulations with two control peptides: the peptide CPMTDNE (previously shown not to
bind to CD20611) and a scrambled version of mUNO, AKPCGS. The cMD simulations with these two
peptides starting from position C1 showed that CPMTDNE (Figure S1B green line) and AKPCGS
(Figure S1B, red line) move freely inside box, as opposed to mUNO (Figure S1B blue line).

Once the stability of the three reported high affinity sites was established, we continued the
dynamics for C1, C2, C3 past to the microsecond time scale. From these three long simulations it
was found that the most stable pose is C1, between the long loop connecting CTLD1 with CTLD2.
In the simulation started from this pose, mUNO spends 80% of the total simulated time in or very
close to the reported pose (Figure 4). The three simulations end with mUNO close to the C1 pose
(light blue background in Figure 4).

![Figure 4: Distance of mUNO's center of mass coordinates relative to a residue of CD206 in the center of lectin domains (residue 286), close to binding pose C1 (CTLD1/2), for the simulation with mUNO placed initially in C1 (blue line), with mUNO placed initially in C2 (CysR, green line), and mUNO placed initially in C3 (FNII, purple line). Domains are illustrated with their respective colors at the right panel.]

**Binding poses involve different physical interaction mechanisms**

The top three predicted binding sites correspond to: 1) mUNO located at the long loop between
the two lectin domains, pose C1 (Figure 5A, left panel); 2) mUNO at the CysR domain, with P78,
W83, and Y89 in close contact with the ligand, pose C2 (Figure 5B, left panel); and 3) mUNO
located just below the FNII domain, besides the hydrophobic surface formed by T167, T197, N203,
pose C3 (Figure 5C, left panel).

Binding site C1 consists of mUNO embedded at the long loop connecting the two lectin domains. This region has been previously identified as a strong electron-negative environment\(^ {29} \), to the extent that a Na\(^+ \) ion was assigned to that region, close to N348 and R295. To investigate the role of electrostatic interactions involved in C1 pose, we used Adaptive Poisson-Boltzman Solver (APBS)\(^ {28} \) to calculate the electrostatic potential at the surface of the receptor. We found that the long loop and the inner region of CTLD2 exerts a negative potential, explaining the strong attraction to the positively charged mUNO peptide (Figure 5A right panel). Binding pose C1 falls on a strategic region of the extracellular portion of CD206, as the long loop connecting the first two lectin domains is known to interact strongly with external long loops forming homodimers\(^ {29} \). It is important to note that this binding site in pose C1 is different from the carbohydrate binding site\(^ {30} \), showing that mUNO and mannose target different sites on the protein. Indeed, fluorescence anisotropy showed that an excess of mannose did not interfere with FAM-mUNO binding (Figure S2), confirming that mUNO and mannose target different sites. Hydrogen bonds also contribute significantly to binding at position C1. The following H-bonds are formed: K6-S288 with 88% occupancy, A5-R295 with 86% occupancy, G4-T444 with 79% occupancy, C1-F350 with 64% occupancy, K6-T446 with 53% occupancy, and G4-T446 with 50% occupancy. Occupancy was calculated as described in methods, and only H-bonds that were observed for more than half of the simulation time (occupancy > 50%) were reported here. H-bonds for pose C1 are shown in higher detail in Figure S3.
Figure 5: Docking clustered conformations and MMGBSA refinement reveal three top docking poses in CTLD1/2, CysR, and FNII domains. The three poses (A-C, left panels) with lowest binding energy as obtained from a MMGBSA refinement. The electrostatic potential felt by mUNO on hCD206 surface calculated with APBS is illustrated, near poses C1, C2, and C3 (A-C, right panels). The sign of the potential is represented in RGB scale color, from blue (positive) to red (negative). Binding pose reported for 4-SO$_4$-GalNAc is also indicated in the figure.

In pose C2, a double aromatic interaction between P3-W83 and P3-P78 keeps mUNO bound in this highly flexible portion of hCD206 (Figure 5B, left panel). The Proline residue of mUNO, P3, stacks simultaneously with W83 on one side and forming an edge-to face aromatic interaction with P78 at the other side. This type of pi stacking is widely observed in supramolecular chemistry. CysR
domain in CD206 is the region where glycoproteins bind; hence we investigated if mUNO binding pose C2 resembles well-known glycoprotein binding sites. We found that our determined C2 pose did not match with the well conserved binding pocket for 4-SO₄-GalNAc. This was expected for electrostatic reasons, as the total charge of mUNO is positive (+1), opposed to the negative sulfated glycoprotein. The calculated electrostatic potential on hCD206 surface showed that the binding region reported for 4-SO₄-GalNAc falls on a positive region of the electrostatic potential (Figure 5B, right panel), whereas mUNO chooses the groove at CysR domain where the electrostatic potential is negative. No disulfide bonds between mUNO and the receptor were observed.

In pose C3 (Figure 5C, left panel), polar residues T197, T198, and T199 form a surface where the N terminal of mUNO (S2 and P3) is supported. This region is very similar to the one in Fibronectin suggested to provide a binding site for collagen, i.e. the region limited at the top by Y193. The nature of the interaction for this binding pose at FNII domain seems to be mainly electrostatic (Figure 5C, right panel).

**Structural changes on hCD206 induced by mUNO**

MD simulations also revealed high flexibility in hCD206’s structure. It was observed that when mUNO peptides are present the curvature of the axis joining the CysR, the FNII and the two CTLD domains tends to close (Figure 6). A quantification of this closing, showing histograms of end-to-end distance for each ensemble (mUNO bound and unbound), revealed an increase in the opening of about 30 % for the unbound structure. On the other hand, no significant changes in hCD206 structure were observed in the lectin domains region between the bound and unbound structures.
Figure 6. Molecular dynamics reveals closing of the bent U shape when mUNO is present. (A) hCD206 structure (blue) and mUNO bound hCD206 structures (red) from aMD simulations showing the closing of the bent U shape for mUNO bound hCD206. Residues used to calculate end to end distance are represented as spheres. (B) Reweighted histogram of end to end distance from aMD simulations.

Selectivity of FAM-mUNO to CD206+ human macrophages

To evaluate applications of mUNO, we assessed the ability of mUNO to carry conjugated payload selectively to human CD206+ macrophages. Human CD206+ macrophages were obtained by stimulating monocytes isolated from human blood with Interleukin-4 (IL-4). This resulted in a mixed population of cells where 45% of cells were CD206+. Such a mixed population of CD206+ and CD206- cells is representative of the macrophage population in tumors. Using this in-vitro model of human M2 TAMs, we evaluated binding of FAM-mUNO and FAM-CPMTDNE control peptide. FAM-mUNO showed preference for CD206+ cells, as 86% of FAM-mUNO targeted cells were CD206+ (Figure 7A), whereas only 44% of FAM-CPMTDNE targeted cells were CD206+(Figure 7B), suggesting that only mUNO has a preference for CD206+ macrophages.
Figure 7: FAM-mUNO is selective to CD206+ cultured human macrophages. CD14+ cells were isolated from human blood and plated on coverslips in full RPMI medium. Four days later, human Interleukin-4 (IL-4) was added to a final concentration of 40ng/mL and left for 48 h. FAM-mUNO(A) or FAM-CPMTDNE (B) were added to a final concentration of 3nM and incubated at 37°C for 45 min. Cells were then washed, fixed with 4% PFA in PBS and stained for FAM and CD206. The percentage of (FAM and CD206)+ cells over the FAM+ cells was obtained by counting the cells from 6 different frames from two independent experiments (C). Scale bar represents 50µm.

DISCUSSION

MRC1/CD206-expressing tumor associated macrophages (MEMs) are major promoters of tumor immunosuppression, angiogenesis, metastasis, and contribute to resistance to chemotherapies and to programmed cell death protein-1 (PD1) checkpoint blockade. Identification of peptides or small molecules capable of selectively targeting human MRC1/CD206 is of great interest for diagnosing and treating solid tumors. We recently identified the mCD206 binding peptide mUNO and its cyclic version (UNO) that showed specific in vivo targeting and internalization in MEMs, either when coupled to a fluorescent tag or to a nanoparticle payload. Here, we evaluated interaction of mUNO with hCD206, determined the domains of hCD206 targeted by the peptide, provided a quantitative estimation of the binding energy, and determined the nature of the interactions that contribute to the binding.

Our previous studies demonstrated interaction of mUNO with recombinant mouse CD206 protein. Here we used fluorescence anisotropy analysis to evaluate whether the binding of mUNO to CD206 is conserved between the mouse and human, or species-specific. Our observation that mUNO is capable of binding to human recombinant CD206, is in line with peptide/receptor interaction being conserved between species. We also observed that mUNO was able to carry a conjugated model payload (FAM) selectively to primary human CD206+ macrophages. This increases translational relevance of studies on mUNO-mediated targeting of hCD206-positive M2 TAMs and warrants further studies on development of mUNO guided therapeutics and imaging agents.

The mannose receptor family is composed of five endocytic proteins: CD206, Endo180, CD280, CD205, and PLA2R. The members of the family have a similar structure, and their extracellular domains contain eight C-type lectin-like domains (CTLDs), a fibronectin type II domain, and a cysteine rich domain. When this work was initiated, there was no high-resolution structure of CD206. Thus, we used the structure of Endo180, a homologue of CD206, as the template to build a model for hCD206 (Model 1). During the course of this investigation, the crystalline...
structure of the first four domains of the hCD206 was published (PDB 5XTS) and we included it in our work (Model 2). Accelerated molecular dynamics combined with docking predictions revealed poses in the three domains. mUNO bound with highest affinity in a pose located at the long loop connecting lectin domains CTLD1 and CTLD2 (C1). Importantly, simulations with two control peptides showed no binding to hCD206. High affinity poses were also found at the FNII domain (C3) and the CysR domain (C2). Different physical mechanisms were identified to be involved in the binding of different poses. Although electrostatic interactions play a significant role in all identified binding poses, H-bonds between mUNO and the receptor make a significant contribution in the binding pose C1. Moreover, our molecular dynamics simulations showed that mUNO triggered a 30° closing of the L-shaped structure; this closing might secure the binding, by packing mUNO-binding domains closer together. The conformation of the L-shape has already been proposed to regulate ligand binding in CD206 and Endo180. Thus, if mUNO alters the conformation of the L-shape, the peptide could also have a biological activity.

The sequence of mUNO, CSPGAK, contains a motif present in all CD206-binding collagens, SPGAK. Two recent studies reported that both CTLD and FNII domains participate in the binding of collagen and denatured collagen to CD206 and Endo180. In prior studies, binding of collagen to CD206 and Endo180 was attributed solely to the FNII domain. Considering the interaction observed in our study to FNII in a very similar region to the one that provides a binding site for collagen in fibronectin, the binding observed in our study to CTLD1/2, and the homology of mUNO to SPGAK; it is reasonable to speculate that the SPGAK motif plays a role in the interaction between collagen and CD206.

An FDA-approved hCD206 binder, Manocept, used in detection of early cancer, is based on mannose. The fact that mUNO does not share a binding site with mannose may allow it to bypass other mannose binding proteins that exist in the body to render the mUNO more selective towards CD206 than mannose-based ligands. Another binder of CD206 is the peptide RP-182 from Riptide Biosciences, but it also binds to other proteins (RelB, Sirp-α, CD47).

CONCLUSIONS

Extensive computational analysis revealed binding of mUNO to a previously untargeted epitope of CD206, found between CTLD1 and CTLD2. Moreover, in vitro binding studies showed selectivity of FAM-mUNO to CD206+ human macrophages. Our study has implications for targeting human M2 TAMs. This study also serves as the starting point for designing mUNO-like peptides of higher affinity and longer blood half-life, or low molecular weight analogs of mUNO which could be orally available and resistant to enzymatic degradation.

SUPPORTING INFORMATION. The following figures are available as supporting information: Figure S1. Reported binding poses C1, C2, and C3 are highly stable; Figure S2. mUNO and mannose do not share binding site; Figure S3: H-bonds formed in C1 binding pose; Figure S4. Homology models for CD206; Figure S5. Comparison of M1 and M2 colored by RMSD values from blue to red (0 to 1); Table S1. Docking poses obtained with AutoDock Vina; Table S2: Binding energies calculated using MM-GBSA.
ACKNOWLEDGEMENTS
This work was supported by Agencia Nacional de Promoción Científica y Tecnológica grant (PICT-2015-1706, to E. K. Asciutto), Estonian Research Council grant PSG38 (to P. Scodeller), European Regional Development Fund (Project No. 2014-2020.4.01.15-0012, to T. Teesalu), Estonian Research Council grant IUT20-17 (to A. Rinken), European Research Council grant GLIOGUIDE from European Regional Development Fund (to T. Teesalu), Estonian Research Council grant PRG230 (to T. Teesalu), and European Regional Development Fund Mobilitas Plus postdoctoral fellowship MOBJD11 (to L. Simón-Gracia). We thank Ago Rinken for manuscript proofreading.

REFERENCES


(12) Elber, R.; Karplus, M. Multiple Conformational States of Proteins: A Molecular Dynamics


