

CONCISE ARTICLE



Cite this: *Med. Chem. Commun.*,
2015, 6, 810

Inhibition of insulin amyloid fibrillization by glyco-acridines: an *in vitro* and *in silico* study†

Quan Van Vuong,^{‡,a} Zuzana Bednarikova,^{‡,bc} Andrea Antosova,^b Pham Dinh Quoc Huy,^{ad} Katarina Siposova,^b Nguyen Anh Tuan,^e Mai Suan Li^{*ad} and Zuzana Gazova^{*b}

The formation of insulin amyloid fibrils leads to accumulation of protein aggregates at the sites of insulin injection and interferes with insulin delivery for treatment of diabetes. We investigated the ability of small molecules, aromatic glyco-acridine derivatives, to prevent insulin fibrillization. Fluorescence spectroscopy and atomic force microscopy have shown that glyco-acridines interfere with insulin aggregation and that their inhibitory activity depends on their structure. The binding free energies, estimated by all-atom molecular dynamics simulations, indicate that the non-polar interaction is the key factor controlling the binding affinity of glyco-acridine derivatives to insulin. We introduced, for the first time, geometrical descriptors that allowed us to distinguish the binding affinities of stereo-isomers. The binding free energies correlate with the distance between the planes of the acridine tricyclic core and the side parts in the unbound and bound states. In addition, the aromatic part of glyco-acridines is important for directing the ligand-dimer insulin interaction. Our findings may provide a basis for the development of new small molecule inhibitors for the treatment of amyloid-related diseases.

Received 7th January 2015,
Accepted 2nd February 2015

DOI: 10.1039/c5md00004a

www.rsc.org/medchemcomm

Introduction

Self-assembly of misfolded proteins into amyloid oligomers or fibrils is associated with a variety of human diseases known as systemic amyloidoses.¹ It has been shown that more than 20 human (poly)peptides are associated with about 40 different diseases including Alzheimer's disease, prion disorders, Parkinson's disease, and diabetes mellitus.^{2,3} Thus, it is of the most importance to understand the mechanisms of amyloid formation as well as to find ways to control or inhibit this process.

Considerable efforts have been devoted to the development of new efficient treatments for amyloidosis. Taking into

account the multistep character of protein aggregation, several approaches have been proposed.⁴ Direct inhibition of amyloid aggregation by small molecules appears to be one of the most promising methods.⁵ Small molecules can easily reach every tissue in the body and interfere with the process of amyloid aggregation.

It was previously shown that small aromatic molecules are efficient inhibitors of amyloid fibril formation *in vitro*^{6,7} and can prevent cell death in amyloidogenic cytotoxicity assays.⁸ Aromatic compounds, such as rottlerin, clotrimazole, or sulconazole, interfere with amyloidogenesis of prion proteins and lysozyme.^{9,10} Curcumin can inhibit aggregation of amyloid beta peptide¹¹ and α -synuclein.^{12,13} Paclitaxel and several small aromatic compounds showed the ability to inhibit insulin amyloid assembly *in vitro*.^{6,7,14} Polyphenols were also identified as effective amyloid inhibitors.^{13,15} We have previously demonstrated the structure-related activity of diverse aromatic acridine derivatives toward inhibition of aggregation of the lysozyme *in vitro*^{16,17} and *in silico*.¹⁸ The anti-aggregation activity of aromatic compounds is possibly related to their interactions with aromatic residues in the binding site or simply due to their non-polar interaction with the highly hydrophobic binding sites of proteins. Porat *et al.* suggested that aromatic interactions between the phenolic compounds and aromatic residues of amyloidogenic sequences may direct the inhibitor to the amyloidogenic core of the protein and thus facilitate the interference of the compound with amyloid

^a Institute for Computational Science and Technology, Tan Chanh Hiep Ward, District 12, Ho Chi Minh City, Vietnam. E-mail: masli@ipfpan.edu.pl

^b Department of Biophysics, Institute of Experimental Physics, Slovak Academy of Sciences, Watsonova 47, 040 01 Kosice, Slovakia. E-mail: gazova@saske.sk

^c Department of Biochemistry, Faculty of Chemistry, Safarik University, Moyzesova 11, 041 54 Kosice, Slovakia

^d Institute of Physics, Polish Academy of Sciences, Al. Lotników 32/46, 02-668 Warsaw, Poland

^e University of Science, VNUHCM, Ho Chi Minh City, Vietnam

† Electronic supplementary information (ESI) available: Additional Fig. S1–S10 and Tables S1–S4. This material is available free of charge via the Internet at <http://pubs.rsc.org>. See DOI: 10.1039/c5md00004a

‡ These authors contributed equally. The manuscript was written through the contributions of all authors. All authors have given approval to the final version of the manuscript.

aggregation.⁶ That would apply not only to phenolic compounds but also for aromatic compounds.

Insulin, a protein hormone produced by beta cells in the pancreas, plays a central role in glucose metabolism.¹⁹ Insulin is composed of chain A (21 residues) and chain B (30 residues), which are linked together by two inter-chain disulfide bonds.²⁰ Although the most stable form of insulin is a hexamer coordinated with 2–4 Zn²⁺ atoms *in vivo*, only less stable monomer and dimer states present the active forms.²¹ Amyloid deposits of insulin are observed at the site of frequent insulin injection in diabetes patients.²² Under certain destabilized conditions, such as low pH, high temperature, and the presence of salts or detergents, insulin has higher propensity to form amyloid aggregates.²³ The understanding and effective inhibition of the insulin amyloid aggregation would improve the storage and delivery of insulin for the treatment of diabetes.

In this study, we employed experimental and computational methods to investigate the interference of glyco-acridine derivatives with amyloid fibrillization of insulin. Experimental data obtained by AFM and ThT fluorescence assay indicate that glyco-acridines prevent amyloid fibrillization with IC₅₀ values in the μM range and their activity depends on their structure. The classical docking and molecular mechanics with generalized Born surface area (MM-GBSA) methods were used to understand the binding process at the atomic level. Our study confirmed the critical role of the aromatic ring in directing the inhibitor's interaction with the insulin dimer. We show that non-polar interactions dominate over the polar interactions and the glyco-acridine binding affinities correlate with the distance between the planes of the core and the side parts. These results point out the importance of compound geometry for distinguishing binding affinities of stereo-isomers.

Materials and methods

Chemicals

Human recombinant insulin (lyophilized powder, I2643-50MG), NaCl and Thioflavin T (ThT) were obtained from Sigma Chemical Company (St. Louis, MO). All other analytical reagent grade chemicals were purchased from either Sigma-Aldrich or Fluka. Glyco-acridine derivatives were synthesized at the Department of Organic Chemistry, Faculty of Science, Safarik University, Kosice. Glyco-acridines were dissolved in DMSO and freshly prepared before the measurements.

In vitro insulin amyloid fibrillization

Insulin amyloid aggregates were prepared by dissolving the lyophilized powder in 100 mM NaCl buffer (pH 1.6) to a final concentration of 10 μM. The solution was then incubated under constant stirring at 65 °C for 2 h. Formation of amyloid fibrils was monitored by characteristic changes in ThT fluorescence intensity and confirmed by atomic force microscopy.

ThT fluorescence assay

Amyloid aggregation of insulin was detected as a significant enhancement of the ThT fluorescence intensity. ThT was added to the 10 μM insulin samples to reach a final ThT concentration of 20 μM. Measurements were performed using a Synergy MX (BioTek) spectrophotometer in a 96-well plate. The excitation was set at 440 nm and the emission was recorded at 485 nm. The excitation and emission slits were adjusted to 9.0/9.0 nm and the top probe vertical offset was 6 nm.

Examination of the ability of glyco-acridines to affect insulin amyloid fibrillization

Glyco-acridine derivatives (at 60 μM final concentration) were added to 10 μM insulin solution at the beginning of the insulin amyloid fibrillization described above. The ability of glyco-acridines to inhibit formation of insulin amyloid aggregates was evaluated by ThT assay. The experiments were performed in triplicate and the reported values are the averages of measured values with standard deviation. In control experiments, the insulin was omitted to measure the fluorescence of glyco-acridine derivatives alone.

Time dependence of insulin amyloid aggregation – effect of glyco-acridines

Glyco-acridine derivatives (at 60 μM final concentration) were added to 10 μM native insulin solution to evaluate its effect on the process of insulin amyloid fibril formation. The mixture was exposed to the conditions triggering the aggregation process (pH 1.6, 65 °C and stirring at 1200 rpm). To determine the extent of insulin amyloid aggregation, ThT was added to the selected aliquots at indicated time intervals and fluorescence was measured by ThT assay. Every experiment was performed in triplicate and the reported values are the averages of measured data; the error bars represent the average deviation. The growth curves were obtained by fitting the average values with the non-linear least-square method.

Determination of IC₅₀

The IC₅₀ value (concentration of glyco-acridine derivative leading to 50% inhibition of insulin fibrillization) was determined by measuring the ability of glyco-acridine to inhibit formation of insulin amyloid aggregates within the compound concentration range from 10 pM to 1 mM at fixed 10 μM insulin concentration. The amount of amyloid aggregates was observed by ThT fluorescence assay. The fluorescence intensities observed in the presence of glyco-acridine derivatives were normalized to the fluorescence signal of amyloid fibrils alone. The final IC₅₀ value represents the average value obtained by fitting three independent concentration dependencies with the non-linear least-squares method. The volume of DMSO in samples was lower than 2% and had no effect on insulin fibrillization or stability of fibrils.

Attenuated total reflectance Fourier transform infrared (ATR-FT-IR) spectroscopy

ATR-FTIR spectra were obtained using a Nicolet 8700 Fourier Transform Infrared (FTIR) (Thermo Scientific) spectrometer equipped with Smart OMNI-Sampler (diamond crystal). The spectra were recorded for 680 μM insulin and glyco-acridine derivatives with the same concentration ratio (1:6) as in the ThT measurements. Each spectrum is an average of 256 interferograms recorded in the amide I region ($1740\text{--}1600\text{ cm}^{-1}$) with a resolution of 2 cm^{-1} . The buffer with 2% DMSO background was subtracted from the protein spectrum.

Atomic force microscopy

Samples of 10 μM insulin with 60 μM glyco-acridine derivatives were placed on a freshly cleaved mica surface by the drop casting method. After 5 min adsorption, samples were washed with ultrapure water and left to air dry. AFM images were taken in tapping mode using a Scanning Probe Microscope (Veeco di Innova, Bruker AXS Inc., Madison, USA) with an NCHV cantilever with a specific resistance of $0.01\text{--}0.025\Omega\text{ cm}$, antimony (n) doped Si, a radius of the tip curvature of 10 nm at a scan rate of $0.25\text{--}0.5\text{ kHz}$ and 512 pixels per line resolution (512×512 pixels/image). No smoothing or noise reduction was applied.

Docking method

The insulin dimer structure²⁴ was retrieved from the RSCB Protein Data Bank (PDB ID: 1GUJ). The structure of the protonated insulin dimer (at pH 2.1) was calculated using the PDB2PQR server²⁵ with the Amber force field. The PDBQT files for insulin and glyco-acridine derivatives were generated using Autodock Tools 1.5.4 (ref. 26 and 27) and used as the input for the Autodock Vina version 1.1,²⁸ which is more efficient than Autodock 4 (ref. 27) for docking glyco-acridine derivatives into insulin. The iterated local search global optimizer algorithm was employed with the Broyden–Fletcher–Goldfarb–Shanno method²⁹ for local optimization in Autodock Vina and the genetic algorithm was used for global optimization. To obtain accurate results, we set the exhaustiveness of global search equal to 1000. The maximum energy difference between the worst and best binding modes was set to 7. A total of 10 binding modes was generated with random starting positions of the glyco-acridine ligand and fully flexible torsion degrees of freedom. The center of grids was placed at the center of mass of the receptor with grid dimensions large enough to cover the whole receptor.

Molecular dynamics simulations

We used the MM-GBSA method to estimate the binding free energy of glyco-acridine derivatives to insulin. Insulin-ligand configurations obtained in the best docking mode were used as starting configurations for the molecular dynamics (MD) simulation at 300 K using the Amber99sb and gaff force fields for receptor and ligands, respectively. The simulation was performed with the help of the AMBER11 package.³⁰

RESP charges,³¹ calculated by Gaussian09 (ref. 32) (HF/6-31G*), were assigned to the ligands. The Generalized Born implicit solvent model^{33,34} (igb = 2) was used for simulation at pH 2.1. The SHAKE³⁵ algorithm was used to restrain all bonds involving hydrogen atoms. A time step of 2 fs was used. Non-bonded interactions were calculated with a cut-off of 2.5 nm. The MD simulation was kept running until a long enough stable trajectory was observed. Snapshots were taken every 100 ps for calculations of binding affinity by the MM-GBSA method.

MM-GBSA method

The MM-GBSA method,^{36,37} which is more accurate and reliable than the docking method, was used to estimate the binding free energy of glyco-acridine derivatives to insulin. In general, the MM-GBSA method calculates the binding free energy of the ligand to the receptor as:

$$\Delta G_{\text{bind}} = \Delta E_{\text{elec}} + \Delta E_{\text{vdw}} + \Delta G_{\text{sur}} + \Delta G_{\text{GB}} - T\Delta S \quad (1)$$

where ΔE_{elec} and ΔE_{vdw} are the contributions from the electrostatic and vdW interactions, respectively. ΔG_{sur} and ΔG_{GB} are the contributions of the non-polar and polar solvation energies. The entropic contribution ΔS is estimated using the normal mode approximation. Equilibrium snapshots collected from the MD run were used as the input files for MM-GBSA calculation. Energies were calculated by AMBER MM-GBSA utility³⁸ with the same force fields, charges and solvent model used as in the MD simulations, but no cut-off was used for the evaluation of non-bonded interactions. The default values were assigned to the rest of the parameters. The input files for MD and MM-GBSA calculations were prepared by utility sets³⁸ in AmberTools1.5.

Results and discussion

Several therapeutic approaches have been proposed for amyloid-related diseases. Based on the generally accepted evidence that amyloid oligomers are mostly involved in cell death, one of the approaches includes direct inhibition of the protein amyloid fibrillization. Therefore, the intensive research is aimed to identify new compounds able to prevent protein self-assembly into amyloid aggregates. A significant attention is devoted to small molecules with aromatic structures, which can specifically interact with amyloid aggregates and inhibit their formation. We have investigated the ability of nine glyco-acridine derivatives containing aromatic rings to inhibit amyloid fibrillization of insulin with the aim of understanding the mechanism of inhibition. We focused on the relationship between the binding affinities of compounds obtained using *in silico* methods and their inhibitory activities determined experimentally *in vitro*.

Glyco-acridine derivatives

The studied glyco-acridine derivatives contain of a planar tricyclic acridin-9-yl core and a side saccharide chain connected

by different linkers. Saccharide terminal units contain tetraacetylated or nonacetylated β -D-glucopyranosyl, β -D-galactopyranosyl or β -D-mannopyranosyl skeletons. Three types of linkers were used to attach saccharide terminals to the acridine core. Linkers have been prepared based on starting 9-hydrazinoacridine and 2,3,4,6-tetra-O-acetyl- β -D-glycopyranosylisothiocyanate synthons (Glu = 1, Gal = 2, Man = 3): thiosemicarbazides (T1, T1OH, T2), cyclic 1,3-thiazolidinones (C1, C2, C2OH and C3) and iso-thiosemicarbazides (I1 and I2). The structures of all screened compounds are shown in Fig. 1; the chemical compositions and formulas are presented in Table S1 in the ESI.[†] The diversity of glyco-acridine derivative structures allowed us to study the relationship between their chemical structure and ability to inhibit insulin amyloid fibrillization.

Impact of glyco-acridine derivatives on insulin aggregation

The effect of glyco-acridine derivatives on insulin fibrillization was studied using ThT fluorescence assay. The results obtained for 60 μ M concentration of glyco-acridine and 10 μ M insulin are presented in Fig. 2 (the fluorescence

intensities were normalized to the fluorescence signal detected for the insulin amyloid aggregates (Iagg) alone). The data clearly show significantly lower (lower than 50%) fluorescence intensities in the presence of glyco-acridine compounds when compared to fluorescence of untreated insulin aggregation. These results indicate that glyco-acridine derivatives have the ability to inhibit formation of insulin amyloid aggregates. The highest fluorescence intensities (about 45% of Iagg) were obtained for glyco-acridine derivatives C1, C2, T1, and T1OH. The lowest fluorescence intensities of about 18% of Iagg were observed for derivatives T2 and I1.

We used FTIR spectroscopy for examination of the content of secondary structure of native insulin and after its fibrillization in the absence and presence of glyco-acridines. The data are presented in Fig. 3. The FTIR spectrum of native insulin (violet dashed line) has the maxima corresponding to the high content of the α -helix structure (peak at 1651 cm^{-1}). The intensive increase of the β -sheet structure accompanying the formation of the peak at 1630 cm^{-1} together with the decrease of the peak intensity for α -helix structure was observed for insulin after amyloid fibrillization (violet solid line). The spectra obtained in the presence of glyco-acridines

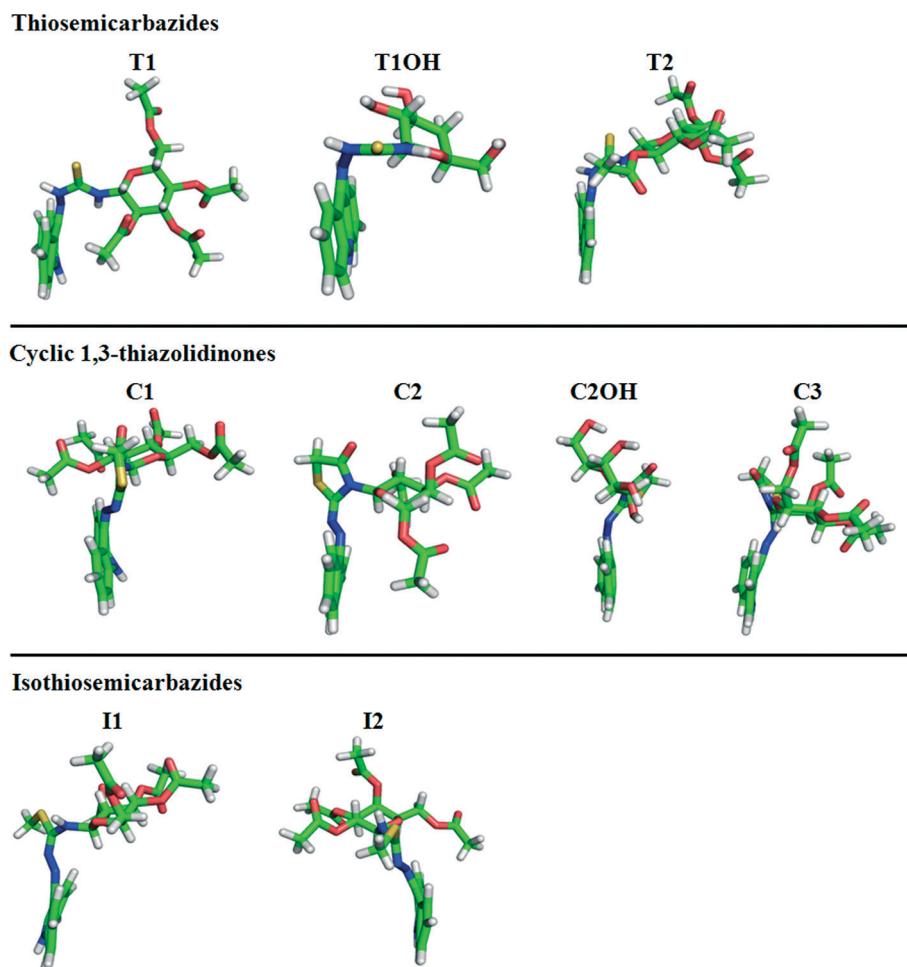


Fig. 1 The structural configuration of the studied glyco-acridine derivatives. The tricyclic acridine core is planar. The arrangement of the distinct linkers and side parts are presented.

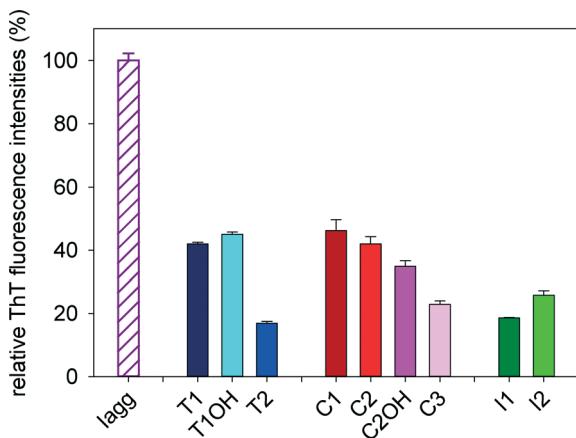


Fig. 2 ThT fluorescence intensities of insulin amyloid aggregates alone (lagg) and in the presence of 60 μ M glyco-acridine derivative. The extent of insulin aggregation was normalized to the control representing the fluorescence signal of the insulin amyloid fibrils (10 μ M) in the absence of the acridine compound (lagg, white hatched bar, taken as 100%). The single experiment was performed in triplicate samples. The error bars represent the standard deviation of three separate samples.

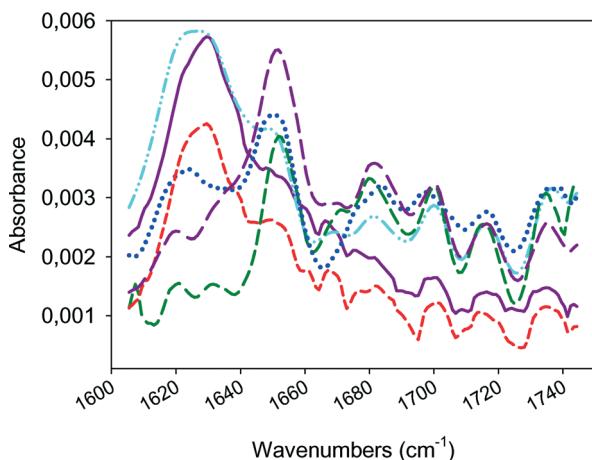


Fig. 3 FTIR spectra detected for native insulin (violet dashed line) and insulin after fibrillation alone (violet solid line) or in the presence of derivatives T1OH (cyan dash-dotted-dotted line), C1 (red short dashed line), T2 (blue dotted line) or I1 (green long dashed line). The ratio of insulin to glyco-derivative (1:6) is the same as in the ThT measurements. The concentration of the insulin was 680 μ M.

clearly indicate that the content of each type of secondary structure varies with the ability of the derivatives to inhibit insulin fibrillization. The high content of the β -sheet structure together with the lower content of the α -helix structure were observed for derivative T1OH (cyan line). In contrast to derivative I1 (green line), the spectrum is similar to the spectrum of native insulin with high content of the α -helix structure. These results indicate that the ability of T1OH to inhibit insulin fibrillization is lower than in the case of compound I1. The spectra obtained for T2 (blue line) and C1 (red line) show considerable contents of both the α -helix and β -sheet structures indicating that their inhibition activities are between derivatives T1OH and I1. We suggest that the

obtained FTIR spectra strongly support the data presented in Fig. 2.

Growth curves of insulin amyloid fibrillization

Next, we aimed to determine what step of insulin fibrillization is being inhibited by the interaction with glyco-acridine derivatives. Thus, the kinetics of the fibril formation in the presence of glyco-acridine derivatives was examined using the ThT fluorescence assay. The growth curves of the insulin fibrillization in the presence of 60 μ M glyco-acridines selected from each of the structural group are shown in Fig. 4.

The growth curve of insulin fibrillization represents nuclei-dependent kinetics typical for the formation of amyloid fibrils. The sigmoidal profile has the characteristic of a lag phase (formation of nuclei) lasting about two minutes followed by the steep increase of fluorescence intensity due to fibril polymerization. The steady-state fluorescence value was achieved after five minutes, suggesting that the formation of the mature fibrils was finished.

The presence of glyco-acridine derivatives changed the characteristics of the S-shape growth curves, namely the time of the lag phase, the slope of the curve and the steady-state fluorescence values in comparison to the insulin fibrillization occurring in the absence of glyco-acridine derivatives. For thiosemicarbazide derivatives, T1OH (cyan dash-dotted-dotted line) and T2 (blue dotted line) were observed to have longer lag phases and lower values for steady-state fluorescence intensities (more than 50% decrease). In the case of derivative C1 (red short dashed line), the lag phase was prolonged even further and the slope of the curve representing fibril growth was more modest. The characteristic of the growth curve for derivative I1 (green long dashed line) was completely different; a significantly longer lag phase (about 10 min) was followed by a rather sharp increase in

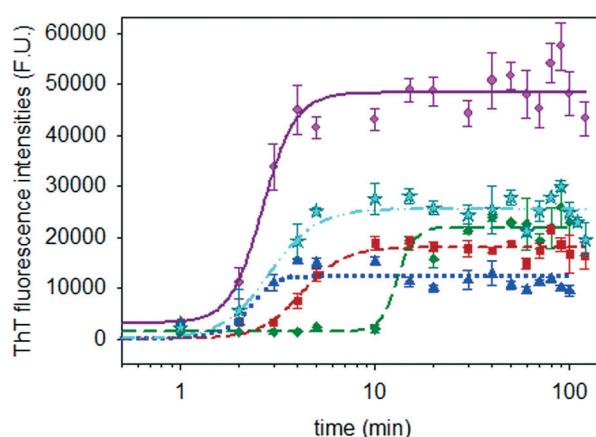


Fig. 4 Time course of insulin fibrillization (10 μ M) in the absence (violet circles) and in the presence of 60 μ M glyco-acridine derivatives T1OH (cyan stars), T2 (blue triangles), C1 (red squares) and I1 (green diamonds) monitored by ThT fluorescence assay. The error bars represent the standard deviation of three independent samples.

fluorescence intensity. These data suggest that glyco-acridine derivatives can slow nuclei formation and the subsequent formation of oligomers and fibrils.

We also examined the kinetics of insulin fibrillization by measuring the absorbance at 600 nm. The obtained data are presented in Fig. S1 in the ESI.[†] The kinetics are similar to the growth curves presented in Fig. 4 and thus strongly support the data obtained by ThT assay.

Atomic force microscopy (AFM)

The fast kinetics of insulin fibrillization was also confirmed by atomic force microscopy (Fig. 5). Formation of globular aggregates was observed after 1 min incubation with a tendency of prolongation into fibrils. Thin protofibrils elongated into fibrils were observed within 3 minutes, suggesting that the elongation phase was almost finished at the time. In the

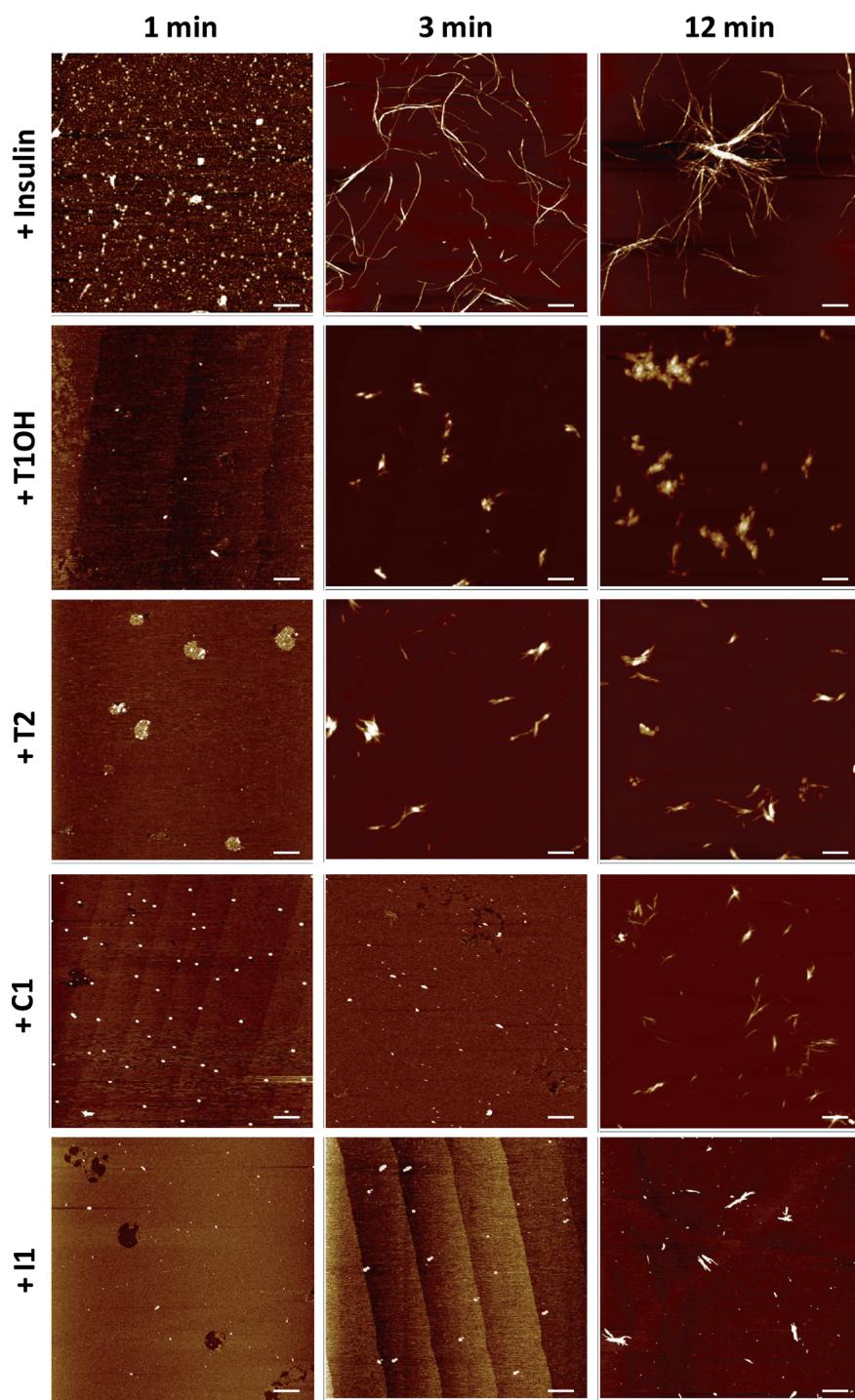


Fig. 5 AFM images of amyloid aggregates formed by 10 μM insulin alone and in the presence of 60 μM glyco-acridine derivatives T1OH, T2, C1 and I1 corresponding to 1 min, 3 min and 12 min time points on growth curves presented in Fig. 4. Bars represent 1 μm .

steady state, insulin formed long, straight and thin amyloid fibrils with a tendency to associate into bigger complexes.

The inhibitory effect of glyco-acridines was clearly demonstrated using this technique. Representative images of glyco-acridines belonging to each structural group are presented in Fig. 5. In the case of thiosemicarbazides T1OH and T2, a slightly prolonged lag phase was followed by the formation of significantly shorter fibrillar aggregates compared to structures observed for insulin alone after 3 min. In the steady state (after 12 min), we observed a huge amount of the mature fibrils. Interference of acridine derivatives C1 and I1 with the process of insulin amyloid fibrillization resulted in further prolongation of the lag phase. It corresponds with the presence of oligomers and globular aggregates 3 min after the start of the process of fibrillization. In the steady-state phase, the aggregates are much shorter and thinner even at 12 min. In the case of I1, the amount of aggregates was significantly reduced compared to that observed for other derivatives. Thus, visualization of the process of aggregation using AFM was clearly in agreement with the data obtained by ThT assay.

Determination of IC₅₀

IC₅₀ values (concentration of compound causing the half-maximal inhibition of insulin fibrillization) were determined for all experimentally studied glyco-acridine derivatives by ThT assay in the concentration range from 10 pM to 1 mM and fixed 10 µM insulin concentration. The relative fluorescence intensities (normalized to the fluorescence signal of insulin amyloid aggregates formed in the absence of glyco-acridines) for representative glyco-acridines from each structural group are provided in Fig. 6. The data clearly show that glyco-acridines inhibit the process of insulin aggregation in a dose-dependent manner. The IC₅₀ values, determined by fitting the average value curves with the nonlinear least-squares method, were within the micromolar range and are summarized in Table 1. The lowest IC₅₀ value corresponding to the most effective inhibitory effect was obtained for derivative I1 (7.22 µM). Higher IC₅₀ values, ranging from 22.4 µM to 39.79 µM, were determined for T1, T2, C2, C3 and I2. The least effective were derivatives C1 and T1OH with IC₅₀ values of 59.48 µM and 60.75 µM, respectively.

Docking of glyco-acridine derivatives

To gain insight into the inhibitory mechanism at the atomistic level, we first carried out the docking simulation. The glyco-acridine derivatives were docked into a dimer structure of insulin. In the best docking mode with the lowest binding energy, all ligands bound to the same binding pocket formed by residues G1, I2, V3, E4, Y19, C20, and N21 from chain A and L11, L15, F24, F25, Y26, T27, P28, and K29 from chain B of insulin (Fig. 7). The binding site belongs to one monomer because the dimer is not entirely symmetric with respect to the monomers. However, the second best docking position belonged to another monomer highlighted in blue. It should

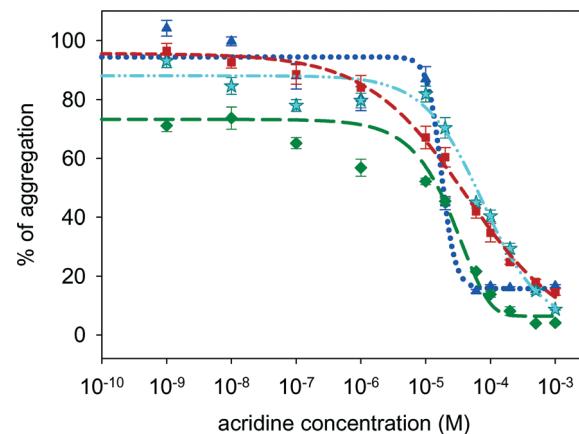


Fig. 6 The inhibition of insulin fibrillization induced by increasing concentrations of glyco-acridines detected by ThT assay; T1OH (cyan stars), T2 (blue triangles), C1 (red squares) and I1 (green diamonds). The degree of inhibition was plotted as a percentage of the control (insulin fibrillization in the absence of glyco-acridines). The fits were calculated as four-parameter logistic curves. Corresponding IC₅₀ values are listed in Table 1.

be noted that the fragment of residues 11–17 (LVEALYL) of chain B, which was considered as a crucial part in insulin fibrillization,^{39–42} is at the bottom of the binding pocket (in Fig. 7 highlighted in red).

It is apparent that the aromatic core of glyco-acridines binds to the binding pocket while the more hydrophilic sites are exposed to the solvent. This reasonable conformation of insulin–glyco-acridine complexes remains valid in an aqueous environment. The binding energies ΔE_{bind} of the ligands are listed in Table 1. The anti-correlation between binding energies and the experimentally determined IC₅₀ values (correlation coefficient $R = -0.76$, Fig. S2†) is not rational as, for instance, ligand I1 has the lowest IC₅₀, but its binding affinity is the worst. However, this poor correlation between experimental and docking results is not surprising because the docking method involves a number of crucial approximations. Thus, the docking approach failed to accurately estimate the binding energy presumably due to the crude approximations, in which the receptor is considered as a rigid object and the number of trial positions of ligand is limited.

To better understand the reason behind this correlation, we considered the binding site in more detail. As it is apparent from Fig. S3 in the ESI,† the binding pocket is rather flexible because it includes three gaps adjacent to the tangent of two insulin monomers. This great flexibility invalidates the docking method. However, the docking method did reliably predicted the binding site. As shown below, the more precise MM-GBSA method yielded the binding free energy and inhibition constant with great correlation.

Molecular dynamics simulations of glyco-acridine binding

The binding pocket and ligands were assumed to undergo major conformational changes to adapt to each other in

Table 1 The binding free energy estimated by the MM-GBSA method. Energy is measured in kcal mol⁻¹. ΔE_{bind} is the binding energy obtained by the docking method

Ligands	IC ₅₀ experimental (μM)	ΔE _{vdw}	ΔE _{elec}	ΔG _{GB}	ΔG _{sur}	-TΔS	ΔE _{vdw} + ΔG _{sur}	ΔE _{elec} + ΔG _{GB}	ΔG _{bind}	ΔE _{bind} docking
T1	39.8 ± 5.7	-36.32	-5.01	18.51	-3.93	18.60	-40.25	13.50	-8.16	-6.5
T1OH	60.8 ± 4.3	-21.63	-9.60	18.23	-2.39	13.20	-24.02	8.63	-2.19	-7.1
T2	22.4 ± 0.1	-44.95	-14.41	30.31	-4.88	19.58	-49.83	15.90	-14.34	-6.3
C1	59.5 ± 4.2	-41.20	-3.92	21.01	-4.79	20.59	-45.99	17.09	-8.31	-7.2
C2	34.6 ± 5.6	-46.51	-27.92	46.26	-5.60	21.68	-52.11	18.34	-12.09	-6.6
C2OH	14.9 ± 2.4	-44.96	-30.14	45.12	-5.19	16.80	-50.15	14.98	-18.36	-6.9
C3	30.8 ± 1.4	-56.01	-19.47	38.49	-5.45	19.41	-61.46	19.02	-23.03	-6.6
I1	7.2 ± 3.6	-52.33	-14.20	32.31	-5.32	15.66	-57.65	18.10	-23.88	-6.0
I2	33.3 ± 0.1	-48.56	-27.75	42.30	-5.25	23.04	-53.81	14.55	-16.22	-6.8

equilibrium. To simulate these structural dynamics, we performed the MD simulations using the configurations obtained in the docking best modes (Fig. 7, the configuration of each derivative is shown in Fig. S4†) as the starting configurations.

Equilibration. We carried out MD runs of 20–80 ns depending on the ligands (Fig. S5 in the ESI†). The time dependence of the receptor–ligand interaction and of Cα root mean square displacement (RMSD) of the receptor shows that the systems reach equilibrium (curves reach saturation) at different time frames. While T1 and I2 complexes reach equilibrium just after ≈7 ns, the T2 complex reached equilibrium at ≈65 ns (Fig. S5 in the ESI†). Snapshots were collected in equilibrium, *i.e.* after the arrow in Fig. S5,† and were used to estimate the binding free energies given in eqn (1).

Binding free energies and their correlation with IC₅₀. The values of ΔG_{bind} calculated by the MM-GBSA method are summarized in Table 1. They vary between -23.88 and -2.19 kcal mol⁻¹ depending on the ligands. As expected, the MM-GBSA method which is more precise than the docking method gives a good correlation between ΔG_{bind} and IC₅₀ ($R = 0.84$, Fig. 8). For example, the glyco-acridine derivative I1 has the lowest binding free energy and shows the strongest inhibition of insulin amyloid fibrillization.

The importance of the binding site dynamics. In sharp contrast with the docking results, the correlation between ΔG_{bind} and IC₅₀ correctly reflects the trend of having a higher binding affinity associated with a smaller inhibition constant, which implies the importance of the binding site dynamics. Thus, it confirms that the binding pocket and the ligand undergo major conformational changes to adapt to each other upon binding. For example, in the unbound state, the width and depth of the binding pocket are 3.2 Å and 3.8 Å, respectively (Fig. 9). In the presence of I1, the binding pocket becomes slightly wider, 3.5 Å, and much deeper, 13.0 Å. Fragment LVEALYL at the bottom of the binding pocket (Fig. 7) is expected to substantially change its conformation upon ligand binding. Because LVEALYL provides the basic structural organization of the spine of insulin fibrils,⁴⁰ the binding glyco-acridine derivatives might inhibit insulin fibrillization through conformational modulation of this peptide.

A lack of the key role of hydrogen bonding in binding affinity. We calculated the hydrogen bond (HB) interaction energy between ligands and the receptor to examine the influence of HBs on the binding affinity. The hydrogen bond energy, ΔE_{HB}, was estimated as the total interaction energy between donor (–NH, –OH) and acceptor (–O, –N) groups using Amber99sb and gaff force fields. The calculation was

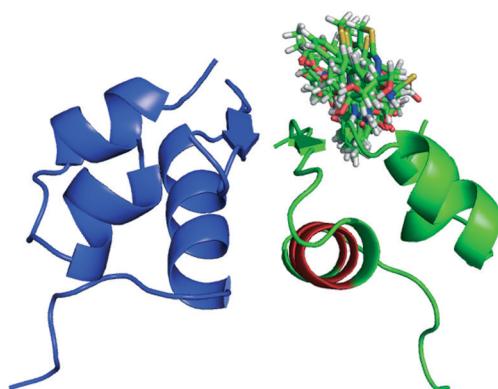


Fig. 7 The best docking positions of all glyco-acridine derivatives in complex with insulin dimer; one insulin monomer is in green and the second is in blue. Fragment LVEALYL at the bottom of the binding pocket is highlighted in red.

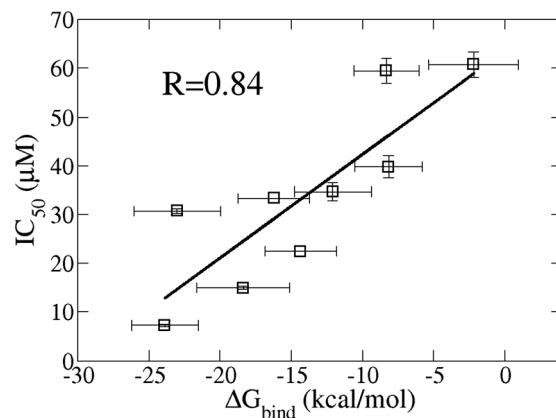


Fig. 8 Correlation of IC₅₀ values with binding energies ΔG_{bind} obtained by MM-GBSA. $R = 0.84$. Error bars of ΔG_{bind} were obtained by averaging over snapshots collected in equilibrium.

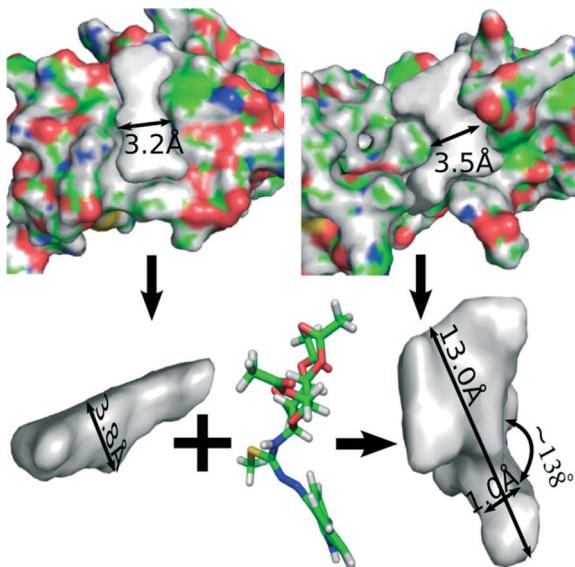


Fig. 9 Structure of the binding pocket in the absence and presence of the most active ligand I1. In the unbound state described by the PDB structure, the width and depth of the binding pocket are 3.2 Å and 3.5 Å, respectively. In the bound state (structure is the last snapshot from the MD run), these values level up to 3.5 Å and 13.0 Å, respectively.

performed using equilibrium snapshots collected from the MD runs for all ligands and the results are shown in Table S3 in the ESI†. The obtained hydrogen bond energies ΔE_{HB} show no correlation with the free binding energies ΔG_{bind} (correlation coefficient $R = -0.29$, Fig. S6 in the ESI†). Particularly, the HB energies of the two most active ligands, I1 and C2OH, are the lowest ones. Thus, the hydrogen bonding presumably has only a minor impact on the binding affinity of glyco-acridine derivatives to insulin.

The dominating role of the non-polar interaction energy. To get further insight into key factors determining binding affinity and inhibitory activity of glyco-acridine derivatives, the free binding energies obtained by the MM-GBSA method were decomposed for detail analysis. Basically, ΔG_{bind} consists of entropy ($-T\Delta S$) and the polar and non-polar interaction energies. The polar interaction energy is the sum of electrostatic interaction energies between the ligand with receptor and the solvent, $\Delta E_{\text{elec}} + \Delta G_{\text{GB}}$, while the non-polar interaction energy is the sum of the van der Waals (vdW) interaction energies of the ligand with receptor and the solvent, $\Delta E_{\text{vdw}} + \Delta G_{\text{sur}}$.

As it was previously demonstrated, the electrostatic effects play a key role in mediating the interaction between proteins as well as in the formation and growth of amyloid fibrils.⁴³ It was also proposed that binding of inhibitors can change the charges on the surface of proteins and influence interactions between proteins and thus suppress the process of fibril formation.^{42,44–46} Our results indicate that the polar and non-polar contributions varied widely from ligand to ligand. However, the non-polar interaction energy was always negative and dominates over the small and positive polar term (Table 1). The non-polar part showed high correlation with

the free binding energy ΔG_{bind} ($R = 0.9$, Fig. 10), while the anti-correlation between polar interaction energy and ΔG_{bind} is poorer ($R = -0.71$). Thus, the non-polar interactions favor ligand binding while the polar interactions disfavor ligand binding due to the high energy needed for desolvation. Thus, glyco-acridine derivatives bind to insulin mainly by non-polar interaction.

The entropy contribution was always positive due to association of ligand with receptor that reduces the freedom of the system. Although most of the ligands are stereo-isomers which have the same number of atoms, entropy contributions were between 13.2 and 23.04 kcal mol⁻¹. The results suggest that the ligand geometry is more important than the number of atoms and entropy may affect the ranking of glyco-acridine derivatives in the free binding energy as well as in their inhibitory activity.

Van der Waals interactions of the glyco-acridine tricyclic core and binding of derivatives to insulin. Because the data above indicated that non-polar interactions determine the ligand binding affinity, we next tried to narrow down what part of the ligand is the most important contributor. For this purpose, the vdW interactions between insulin and the acridine core, linker and side parts of the ligand were analyzed in detail (Table S4 in the ESI†). For all ligands with the exception of C2, the contribution of the core part was greater than those of linker and side chain. More importantly, only the core part showed a high correlation with the free binding energy ($R = 0.9$, Fig. S7 in the ESI†). We have estimated the vdW and electrostatic interactions of the core with aromatic rings of residues HIS, PHE, TYR, and TRP. As evident from Table S4 in the ESI†, the interaction between two aromatic

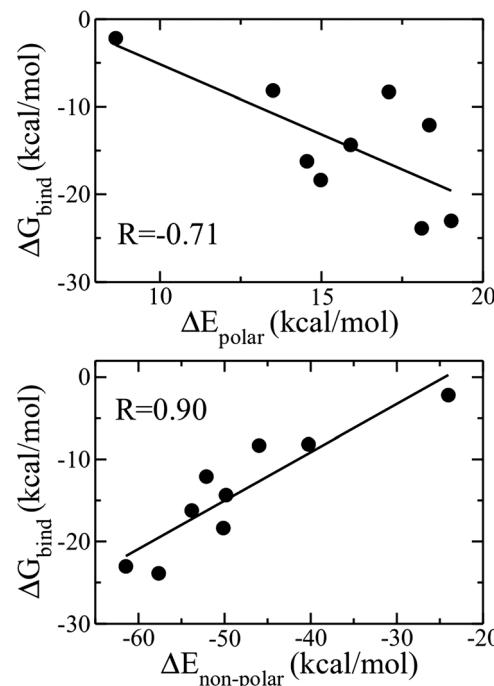


Fig. 10 Correlation between ΔG_{bind} and the polar (ΔE_{polar}) and non-polar ($\Delta E_{\text{non-polar}}$) interaction energies.

rings is repulsive due to their positive charges. In addition, there was no correlation between the aromatic interaction energy and the binding free energy (Fig. S8 in the ESI[†]) suggesting that although the hydrophobic aromatic tricyclic core of acridines strongly binds to insulin, the aromatic–aromatic interactions play a minor role.

Structural geometry of glyco-acridine derivatives and the binding affinity

The vdW interactions of the glyco-acridine core with the receptor define the differences in non-polar interaction as well as binding affinity. Because the core is an invariant part and some of the ligands are stereo-isomers, it remains unclear why the binding affinities of the studied ligands are so different. In order to understand this issue, we have investigated the structure of the ligands in more detail. The geometry of the glyco-acridine derivatives is characterized by the orientation of the core and the side chains linked together by a linker. We introduced two geometrical descriptors, D_{gxy} and D_{gyz} , representing the distance between the planes of the core and the side chains in the vertical oz direction and the horizontal ox direction, respectively (Fig. 11).

The bound state values of the geometrical descriptors were calculated using MD snapshots collected in equilibrium (Table 2). The distance from the side chain to the core in the vertical oz direction (D_{gxy}) was anti-correlated with the free binding energy ($R = -0.71$, Fig. 12), while a poorer correlation was observed for D_{gyz} ($R = 0.51$). Interestingly, there was also a high anti-correlation between the vdW interaction energies of the core with D_{gxy} ($R = -0.88$, Fig. S9 in the ESI[†]) and a lower correlation with distance in the horizontal ox direction, D_{gxy} ($R = 0.66$). The side chains oriented upward away from

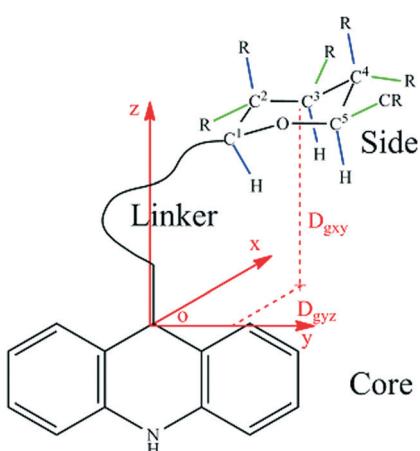


Fig. 11 Common structure of glyco-acridine derivatives with the planar tricyclic acridin-9-yl core and a side saccharide connected by a linker R which can be $-H$, $-OCOCH_3$ or $-OH$ depending on derivatives. Axial positions of the side part are in blue and equatorial positions are in green. D_{gxy} is the distance from the center of the cyclic six-membered ring of the side part to the oxy plane which is perpendicular to the bond between the core and the linker part. Origin o was placed on the C atom of this bond. D_{gyz} is the distance from the center of the cyclic six-membered ring of the side saccharide to the plane of the core represented by the oy plane.

Table 2 Geometric descriptors D_{gxy} and D_{gyz} of glyco-acridine derivatives in the bound and unbound (denoted by *) states

Ligands	D_{gxy} (Å)	D_{gyz} (Å)	D_{gxy}^* (Å)	D_{gyz}^* (Å)
T1	2.39	5.19	2.72	4.84
T1OH	1.34	3.83	1.32	3.80
T2	5.90	2.34	1.68	3.57
C1	3.78	1.88	3.23	2.34
C2	3.02	2.79	3.51	1.98
C2OH	3.33	2.50	3.68	2.23
C3	4.16	2.59	3.34	2.63
I1	5.51	2.30	3.99	2.69
I2	3.93	2.22	4.12	2.70

the core in the vertical oz direction favored the vdW interaction of the core as well as the binding free energy while those with the larger horizontal ox distance did not. This was because the core binds to the binding pocket at both sides of its planar tricyclic core. The side chain geometrically blocks the core from interacting with the binding pocket, reducing the binding affinity of the ligands. The upward vertical oz orientation of the side chain is the optimal orientation for binding.

To estimate distances D_{gyz} and D_{gxy} in the state where the ligand is not bound to insulin, we have performed 100 ns MD for each free ligand under the same conditions as for the receptor–ligand complex. Skipping the first 10 ns as time for equilibration, snapshots collected from the last 90 ns were used for the estimation of geometrical descriptors (Table 2, Fig. S10[†]). The calculated values of D_{gxy}^* in the unbound state is also anti-correlated with the free binding energy ($R = -0.66$, Fig. 12) pointing out the important role of ligand geometry in its binding affinity. As expected, the correlation level with geometrical descriptors in the unbound state is lower compared to that in the bound state (Fig. 12 and S9 in the ESI[†]). The correlation between ΔG_{bind} and D_{gxy}^* opens up a new possibility to obtain the information about the binding affinity by simulating only the dynamics of the free ligands. This kind of simulation is much less time consuming than the simulation of protein–ligand complexes.

The role of geometry and flexibility of the linker and side chains

Combining all of the computational results revealed that the planar hydrophobic tricyclic core binds to the binding pocket exposing the hydrophilic side chain to the polar solvent.

Both the ligand and the binding pocket undergo major conformational changes to adapt to each other favoring the vdW interactions of the tricyclic core of the ligand with the receptor. To accommodate the tricyclic core inside the binding pocket, the ligand has to re-orient the side chains upward from the core increasing D_{gxy} . Thus, ligand with high binding affinity should have a compact side chain and high flexibility.

Tetraacetylated β -D-galactopyranosyl and tetraacetylated β -D-mannopyranosyl side chains contain an acetyl group bonded to C^4 and C^2 at axial position as opposed to the

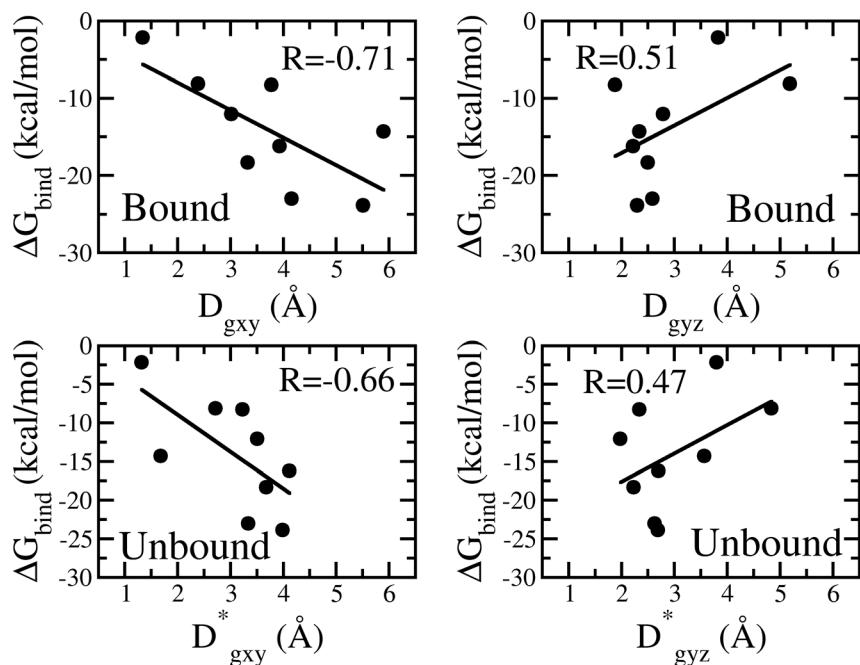


Fig. 12 Correlation between the free binding energies and geometrical descriptors of glyco-acridine derivatives in the bound and unbound states. * denotes the unbound state.

equatorial position in tetraacetylated β -D-glucopyranosyl (Fig. 11). The orientation of the acetyl group at these axial positions is upward far away from the core, reducing the geometric obstacle against the binding of the core. That explains why the binding affinity increases in ligands T1 and T2 and C1, C2, and C3, when the tetraacetylated β -D-glucopyranosyl in the side chain is replaced by tetraacetylated β -D-galactopyranosyl or tetraacetylated β -D-mannopyranosyl containing a side chain. Thus, ligand C3 is more active than ligand C2 because C² is closer to the core than C⁴ and the effect of axial position of C² is stronger than that of C⁴.

In addition to the geometry of the side chains, the relative orientation of the core and side chains is also determined by the geometry and flexibility of the linker. Among the three

types of linkers, cyclic 1,3-thiazolidinone, which includes a five-membered ring and only one rotatable bond (Fig. 13), is the most rigid linker and geometric descriptors, D_{gxy} and D_{gyz} , of the ligands from the C1, C2, and C3 groups, which contain this linker, vary between the bound and unbound states the least and remain small. It also explains the lower binding affinity of those ligands compared to the others with the same side chains. Thiosemicarbazide and iso-thiosemicarbazide linkers containing three and two rotatable bonds, respectively, are more flexible (Fig. 11), but both of them also contain an intra-hydrogen bond which reduces their flexibility. For both linkers, only the rotatable bond N–N, which does not change the length of the intra-hydrogen bond rotating around itself, keeps its flexibility intact.

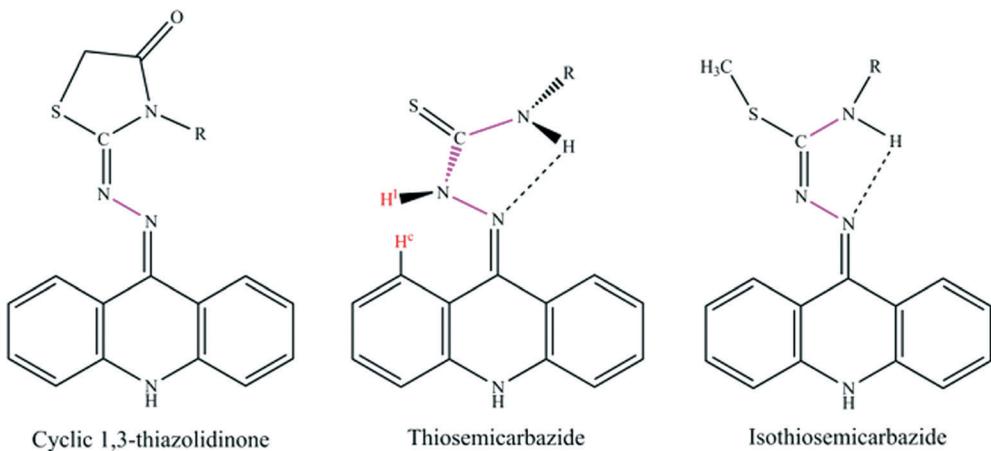


Fig. 13 Structures of three types of linkers. R is a side part; rotatable bonds of the linker are in purple.

In the case of thiosemicarbazides, there is repulsion between H^l of the linker and H^c of the core (Fig. 11), which prevent thiosemicarbazides from rotating around the bond N–N and thus to achieve an upward orientation away from the core. Thus, isothiosemicarbazides have more flexible linkers, and consequently, the binding affinity of I1 is higher than that of T1.

The smaller side chain is undoubtedly less geometrically limiting and thus increases the binding affinity. As a result, C2OH with a small non-acetylated β-D-galactopyranosyl side chain is one of the most active ligands superior to C1, C2, and C3 (Table 2). However, this is not true for T1OH which has the lowest binding affinity because T1OH distorted the flexible thiosemicarbazide linker with a small non-acetylated β-D-glucopyranosyl side chain. With a small side chain, the repulsion between H^l of the linker and H^c of the core can be minimized by rotating the linker around the N–N bond. The positioning of the side part closer to the core (Fig. 13) prevents the core insertion into the binding pocket of insulin.

Conclusions

The results obtained using *in vitro* experiments (ThT assay, FTIR spectroscopy and AFM) showed that glyco-acridines are able to interfere with insulin amyloid fibrillization with the half-maximal inhibition constant IC₅₀ in the μM range. The result of this interaction is the reduction of insulin amyloid fibrillization. *In silico* calculation presumed the importance of conformations of the side chain and the core of glyco-acridine derivatives as well as their linkers.

Our data indicate that the non-polar interactions together with the core of acridine derivatives are the key factors determining the ligand's free binding energy to insulin. The contribution of the linker and side chains to the inhibitory activity depends not only on their position relative to the core but also on their flexibility.

The experimentally determined inhibitory activities of glyco-acridines are in agreement with *in silico* results. The lowest inhibitory activities were obtained for glyco-acridine derivatives C2 and T1OH *in vitro*. The distorted linker connecting a small side chain (glyco-residue) to the acridine core of T1OH produces tight conformation of the side chain and core resulting in the lowest binding affinity. The highest inhibitory activities were observed for derivatives I1, C2OH, and T2, in which side chains and the glyco-acridine core are not close to each other, and thus, the core can interact with insulin.

The interesting task was to explain the differences in the inhibitory activities as well as binding affinities of the stereoisomeric glyco-acridine compounds. To address this problem, we have introduced geometrical descriptors D_{gxy} and D_{gyz} to characterize the three-dimensional orientations of different parts of glyco-acridines. The binding affinity was shown to be correlated with the distance between the planes of the core and the side chains in the bound state explaining different activities of stereo-isomers.

Based on the detailed analysis of the effect of side chain geometry and linker flexibility, we predict that the glyco-acridine derivatives with high binding affinity consist of i) a planar tricyclic acridin-9-yl core, ii) a flexible iso-thiosemicarbazide linker or a planar linker like cyclic 1,3-thiazolidin-4-one, which bonds to the side part of the oxygen atom of the carbonyl group (etherized) instead of the nitrogen atom and iii) a small side chain such as nonacetylated β-D-glucopyranosyl.

In addition, the crucial role of the aromatic part in directing the interaction of the inhibitor to receptor was confirmed by simulation. By elucidating the critical role of the planar aromatic core and other geometric factors affecting the binding free energy, our findings will be useful for design of potential lead compounds for amyloid-related diseases.

Derivatives I1, C2OH and T2 were the most effective inhibitors of insulin fibrillization with IC₅₀ values in the micromolar range. Thus, these glyco-acridines represent potential therapeutic agents for insulin-associated amyloidosis and, possibly, for amyloid-related diseases in general.

Acknowledgements

This work was supported by Project VEGA 0181, EU grant 26220220005 and 26110230097, grant UPJS-VVGS 98/13-14, Narodowe Centrum Nauki in Poland (grant no. 2011/01/B/NZ1/01622), Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 106-YS.02-2013.01 and the Department of Science and Technology at Ho Chi Minh City, Vietnam. The authors thank J. Imrich for synthesizing and providing the glyco-acridine derivatives.

References

- 1 F. Chiti and C. M. Dobson, *Annu. Rev. Biochem.*, 2006, 75, 333–366.
- 2 C. M. Dobson, *Nature*, 2003, 426, 884–890.
- 3 M. M. Varughese and J. Newman, *J. Biophys.*, 2012, 434289.
- 4 F. E. Cohen and J. W. Kelly, *Nature*, 2003, 426, 905–909.
- 5 T. Härd and C. Lendel, *J. Mol. Biol.*, 2012, 421, 441–465.
- 6 Y. Porat, A. Abramowitz and E. Gazit, *Chem. Biol. Drug Des.*, 2006, 67, 27–37.
- 7 Y. Hong, L. Meng, S. Chen, C. W. T. Leung, L.-T. Da, M. Faisal, D.-A. Silva, J. Liu, J. W. Y. Lam, X. Huang and B. Z. Tang, *J. Am. Chem. Soc.*, 2012, 134, 1680–1689.
- 8 D. Latawiec, F. Herrera, A. Bek, V. Losasso, M. Candotti, F. Benetti, E. Carlino, A. Kranjc, M. Lazzarino, S. Gustincich, P. Carloni and G. Legname, *PLoS One*, 2010, 5, e9234.
- 9 B. Y. Feng, B. H. Toyama, H. Wille, D. W. Colby, S. R. Collins, B. C. H. May, S. B. Prusiner, J. Weissman and B. K. Shochet, *Nat. Chem. Biol.*, 2008, 4, 197–199.
- 10 N. Sarkar, M. Kumar and V. K. Dubey, *Biochim. Biophys. Acta, Gen. Subj.*, 2011, 1810, 809–814.
- 11 S. T. Ngo and M. S. Li, *J. Phys. Chem. B*, 2012, 116, 10165–10175.

- 12 F. Yang, G. P. Lim, A. N. Begum, O. J. Ubeda, M. R. Simmons, S. S. Ambegaokar, P. P. Chen, R. Kayed, C. G. Glabe, S. A. Frautschy and G. M. Cole, *J. Biol. Chem.*, 2005, **280**, 5892–5901.
- 13 K. Ono and M. Yamada, *J. Neurochem.*, 2006, **97**, 105–115.
- 14 E. Kachooei, A. A. Moosavi-Movahedi, F. Khodagholi, F. Mozaffarian, P. Sadeghi, H. Hadi-Alijanvand, A. Ghasemi, A. A. Saboury, M. Farhadi and N. Sheibani, *J. Biochem.*, 2014, **155**, 361–373.
- 15 Z. Gazova, K. Siposova, E. Kurin, P. Mučaji and M. Nagy, *Proteins*, 2013, **81**, 994–1004.
- 16 Z. Gazova, A. Bellova, Z. Daxnerova, J. Imrich, P. Kristian, J. Tomascikova, J. Bagelova, D. Fedunova and M. Antalik, *Eur. Biophys. J.*, 2008, **37**, 1261–1270.
- 17 A. Antosova, B. Chelli, E. Bystrenova, K. Siposova, F. Valle, J. Imrich, M. Vilkova, P. Kristian, F. Biscarini and Z. Gazova, *Biochim. Biophys. Acta, Gen. Subj.*, 2011, **1810**, 465–474.
- 18 Q. Van Vuong, K. Siposova, T. T. Nguyen, A. Antosova, L. Balogova, L. Drajna, J. Imrich, M. S. Li and Z. Gazova, *Biomacromolecules*, 2013, **14**, 1035–1043.
- 19 M. Evans, P. M. Schumm-Draeger, J. Vora and A. B. King, *Diabetes, Obes. Metab.*, 2011, **13**, 677–684.
- 20 K. R. Wilhelm, K. Yanamandra, M. A. Gruden, V. Zamotin, M. Malisauskas, V. Casaite, A. Darinskas, L. Forsgren and L. A. Morozova-Roche, *Eur. J. Neurol.*, 2007, **14**, 327–334.
- 21 K. Giger, R. P. Vanam, E. Seyrek and P. L. Dubin, *Biomacromolecules*, 2008, **9**, 2338–2344.
- 22 S. Störkel, H. M. Schneider, H. Müntefering and S. Kashiwagi, *Lab. Invest.*, 1983, **48**, 108–111.
- 23 L. Nielsen, S. Frokjaer, J. F. Carpenter and J. Brange, *J. Pharm. Sci.*, 2001, **90**, 29–37.
- 24 J. L. Whittingham, D. J. Scott, K. Chance, A. Wilson, J. Finch, J. Brange and G. G. Dodson, *J. Mol. Biol.*, 2002, **318**, 479–490.
- 25 T. J. Dolinsky, J. E. Nielsen, J. A. McCammon and N. A. Baker, *Nucleic Acids Res.*, 2004, **32**, W665–W667.
- 26 M. F. Sanner, *J. Mol. Graphics Modell.*, 1999, **17**, 57–61.
- 27 G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson, *J. Comput. Chem.*, 2009, **30**, 2785–2791.
- 28 O. Trott and A. J. Olson, *J. Comput. Chem.*, 2010, **31**, 455–461.
- 29 C. G. Broyden, *J. Inst. Math. Its Appl.*, 1970, **6**, 76–90.
- 30 D. Case, T. A. Darden, T. E. Cheatham, C. Simmerling, J. Wang, R. Duke, R. Luo, M. Crowley, R. Walker, W. Zhang, K. M. Merz, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, X. Wu, S. Brozell, T. Steinbrecher, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, D. H. Mathews, M. G. Seetin, C. Sagué, V. Babin and P. Kollman, *Amber 11 OR*, University of California, San Francisco, 2010.
- 31 M. M. Franci and L. E. Chirlian, *Reviews in Computational Chemistry*, John Wiley & Sons, Inc., 2000, pp. 1–31.
- 32 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 Jr, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. 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, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09*, Gaussian Inc., Wallingford, CT, 2009.
- 33 V. Tsui and D. A. Case, *J. Am. Chem. Soc.*, 2000, **122**, 2489–2498.
- 34 V. Tsui and A. D. Case, *Theory and applications of the generalized Born solvation model in macromolecular simulations*, 2001.
- 35 S. G. Lambrakos, J. P. Boris, E. S. Oran, I. Chandrasekhar and M. Nagumo, *J. Comput. Phys.*, 1989, **85**, 473–486.
- 36 T. Hou, J. Wang, Y. Li and W. Wang, *J. Chem. Inf. Model.*, 2011, **51**, 69–82.
- 37 F. Godschalk, S. Genheden, P. Soderhjelm and U. Ryde, *Phys. Chem. Chem. Phys.*, 2013, **15**, 7731–7739.
- 38 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.
- 39 M. I. Ivanova, M. J. Thompson and D. Eisenberg, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 4079–4082.
- 40 M. I. Ivanova, S. A. Sievers, M. R. Sawaya, J. S. Wall and D. Eisenberg, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 18990–18995.
- 41 H.-L. Chiang, S. T. Ngo, C.-J. Chen, C.-K. Hu and M. S. Li, *PLoS One*, 2013, **8**, e65358.
- 42 J. Haas, E. Vöhringer-Martinez, A. Bögehold, D. Matthes, U. Hensen, A. Pelah, B. Abel and H. Grubmüller, *ChemBioChem*, 2009, **10**, 1816–1822.
- 43 A. K. Buell, P. Hung, X. Salvatella, M. E. Welland, C. M. Dobson and T. P. Knowles, *Biophys. J.*, 2013, **104**(5), 1116–1126.
- 44 A. K. Buell, A. Dhulesia, D. A. White, T. P. J. Knowles, C. M. Dobson and M. E. Welland, *Angew. Chem., Int. Ed.*, 2012, **51**, 5247–5251.
- 45 L. Di Michele, E. Eiser and V. J. Fodera, *J. Phys. Chem. Lett.*, 2013, **4**(18), 3158–3164.
- 46 V. Fodera, A. Zaccone, M. Lattuada and A. M. Donald, *Phys. Rev. Lett.*, 2013, **111**, 108105.