Support Information

Materials and methods

Enzymes, substrates and inhibitors

The synthesis of the amides and esters compounds was previously described by Barros *et al*. [1]. The mature human tissue kallikreins 5 and 7 were expressed as recombinant proteins from an insect cell/baculovirus expression system, as described previously [2]. The Fluorescence Resonance Energy Transfer (FRET) peptides Abz-KLYSSKQ-EDDnp (Abz, o-aminobenzoic acid; EDDnp, N-(2,4-dinitrophenyl)ethylenediamine) and Abz-KLRSSKQ-EDDnp were synthesized by solid phase synthesis methods, as described previously [3].

Kinetic assays

Stock solutions of the compounds were previously dissolved in a solution H_2O:DMSO 1:1 to the appropriate concentration prior to the assays, with the final DMSO concentration kept at no more than 5 % (v/v). The compounds were initially screened against human kallikreins 1, 5, 6 and 7 at an initial concentration of 100 μM. All reactions were performed in a buffer 50 mM Tris-HCl, pH 7.5. The enzymes were incubated previously during 5 minutes with the compound at 37°C, and the reactions were started by the addition of the substrate Abz-KLRSSKQ-Eddnp (KLK5) or Abz-KLYSSKQ-Eddnp (KLK7) at a final concentration of 20 μM. A spectrofluorimeter Hitachi F2500 was used to follow the hydrolysis of the appropriate fluorogenic substrate using an excitation wavelength of 320 nm and an emission wavelength of 420 nm. Control assays were performed without inhibitor (negative control).

The IC_{50} of an inhibitor was calculated from initial reaction rates in the absence and presence of the inhibitor over a range of 1 μM to 500 μM. The values represent means ± SD of at least three individual experiments. Analysis of all concentration-response curve data was done using GraFit 7 program [4].

The mechanism of the inhibition of the compounds against the KLK5/KLK7 was obtained by visual inspection of Lineweaver–Burk plots of all inhibitory data at different substrate concentrations and the K_i value was determined by the secondary plot of K_M/V_{max} vs. [I].

Molecular Docking

The crystal structure of KLK5 [5] complexed in the presence of the inhibitor leupeptin was used in the docking studies and were retrieved from the Protein Data Bank [6]. Docking of compounds **6** and **14** were performed with Autodock4.2.[7]. Polar and aromatic hydrogens were added, and Gastegier charges were computed by Autodock Tools on each atom of the ligand. The

AutoTors utility was used to define torsional degrees of freedom for the ligand. The grid box was centered in the macromolecule and the dimension of the grid was 51 × 51 × 51 $Å^3$ with the spacing between the grid points at 0.375 Å in order to include the entire binding site. Grid potential maps were calculated using the module AutoGrid 4.0. The Lamarckian genetic algorithm was used to perform docking simulation, with an initial population of 150 randomly placed individuals with a maximum number of 2500000 energy evaluations, 27000 generations, mutation rate of 0.02, a crossover rate of 0.8, and an elitism value of 1, were used. Then 100 docking runs were performed. Pseudo-Solis and Wets algorithm was used for local search method. Finally, the resulting docked conformations were clustered together on the basis of root-mean-square deviation (RMSD) tolerance of 2.0 Å and represented by most favorable free energy of binding.

Molecular dynamics simulation

The lowest energy pose for ligand **6** and **14** in complex with KLK5 structure was used as a starting configuration of an MD simulation. Missing atoms and hydrogens were added using the *tleap* program in Amber11 software [1]. Amber ff99SB force field [8] and GAFF force field [9,10] were assigned respectively to amino acid residues and to the inhibitors. The GAFF parameters were generated using *antechamber* and *parmchk* program, partial atomic charges were derived using AM1-BCC methodology [11]. Each complex structure was solvated in a TIP3P water cubic box extending at least 10 Å from the complex, and then neutralized by the addition of appropriate number of monovalent counterions. Each system was relaxed by two successive energy minimizations in *sander* program, firstly the heavy atoms of the solute was kept fixed with 10.0 $kcal.mol^{-1}.Å^{-2}$ force constant while the hydrogen atoms and waters were relaxed with the first 2500 steps being steepest descent before switching to the conjugate gradient algorithm for the remaining 2500 steps. Then all atoms were relaxed using another 2500 steps of steepest descent followed by2500 steps of conjugate gradient minimization. Starting from the minimized system, the complex was equilibrated by five successive steps. First, it was gradually heated from 0 K to 300 K during 100 ps with a restrain force constant of 10.0 $kcal.mol^{-1}.Å^{-2}$ applied to heavy atoms of the solute under NVT conditions. Next, a 100 ps MD simulation with a 10.0 $kcal.mol^{-1}.Å^{-2}$ restraint on the complex was carried out at a pressure of 1 atm and 300 K to equilibrate the density using Berendsen's barostat [12]. Finally, the complex was equilibrated through four additional 100 ps at 300 K and 1atm with a restrain force constant of 8.0, 6.0, 4.0 and 2.0 $kcal.mol^{-1}.Å^{-2}$, respectively, applied to heavy atoms of complex. The temperature was kept fixed at 300 K using a Langevin thermostat [13] with a collision frequency of 2 ps^{-1}. After the equilibration phase, a 10 ns simulation at constant pressure with a target temperature of 300 K and pressure of 1 atm was conducted in *pmemd* program. The SHAKE algorithm [14] was employed to constrain all hydrogen atoms and

the time step was set to 2.0 fs. Particle mesh Ewald (PME) [15] method was used to treat the long-range electrostatic interactions with default cutoff of 8.0 Å. Both energies and coordinates were saved every 5 ps, thus yielding a MD trajectory of 2000 frames which was used in the postproduction energetic analysis.

MM/GBSA calculations

For every snapshot, a free energy is calculated for each molecular species (complex, receptor, and ligand), and the binding free energy is calculated as the difference:

$$\Delta G_{bind} = G_{complex} - \left(G_{receptor} + G_{ligand}\right) \quad (1)$$

In MM/GBSA, the binding free energy (ΔG_{bind}) of the protein-ligand complex can be calculated as

$$\Delta G_{bind} = \Delta E_{MM} + \Delta G_{sol} - T\Delta S \quad (2)$$

$$\Delta E_{MM} = \Delta E_{internal} + \Delta E_{electrostatic} + \Delta E_{vdW} \quad (3)$$

$$\Delta G_{sol} = \Delta G_{GB} + \Delta G_{SA} \quad (4)$$

where ΔE_{MM}, ΔG_{sol} and $-T\Delta S$ are the changes of the gas phase molecular mechanics energy, the solvated free energy contribution, and the solute entropy estimated from the MD trajectory, respectively. ΔE_{MM} is further divided into a $\Delta E_{internal}$ (bond, angle, and dihedral energies), $\Delta E_{electrostatic}$ (electrostatic), and ΔE_{vdw} (van der Waals) energies. ΔG_{solv} is the sum of two contributions: ΔG_{GB} (polar solvation energy), and ΔG_{SA} (nonpolar solvation energy). The polar contribution is calculated using the GB model, while the nonpolar energy is estimated by solvent accessible surface area (SASA). Finally, the change in solute entropy during ligand association ($-T\Delta S$) is usually computed by normal-mode analysis [16].

In this work, a single trajectory approach was adopted in the calculation of the binding free energies between the receptor and inhibitors using 2000 snapshots taken from the MD trajectory, which represent the last 4ns of simulation. The MMPBSA.py script implemented in AmberTools 1.5 was used to perform the MM/GBSA calculations. All the required steps to estimate the binding free energy of the protein-ligand complexes are automatically performed with this script. In the MM/GBSA calculation, ΔG_{GB} was determined using a modified GB model developed by Onufriev *et. al.* (GB^{OBC}) [17]. The nonpolar contribution was determined on the basis of the solvent-accessible surface area (SASA) using the LPCO method [18], where the surface tension constant was set to 0.0072 $kcal.mol^{-1}.Å^{-2}$. In this work all changes in solute entropy were omitted, since we was interested only in the relative rank-ordering of ΔG_{bind} for each complex, and is generally assumed the entropy change will be constant for a set analogues resulting in no influence in their relative binding energies [19, 20].

References

(1) Barros, T.G.; Zorzanelli, B.C.; Pinheiro, S.; de Brito, M.A.; Tanuri, A.; da Costa, E.C.B.; Mohana-Borges, R.S.; Rodrigues, C.R.; de Souza, A.M.T.; Ferreira, V.F.; Muri, E.M.F. *Let. Org. Chem.*, **2012**, 9, 239-249.

(2) Bernett, M.J.; Blaber, S.I.; Scarisbrick, I.A.; Dhanarajan, P.; Thompson, S.M.; and Blaber, M. *J. Biol. Chem.*, **2002**, 277, 24562–24570.

(3) Hirata, I.; Cezari, M.H.C.; Nakaie, C.R.; Boshcov, P.; Ito, A. S.; Juliano, M.A.; Juliano, L. *Lett. Peptide Sci*. **1994**, 1, 299–308.

(4) Leatherbarrow, R. J. GraFit Version 7; Erithacus Software Ltd.: Horley, U.K., **2009**.

(5) Debela, M.; Goettig, P.; Magdolen, V.; Huber, R.; Schechter, N.M.; Bode, W. *J. Mol. Biol*. **2007**, 373, 1017-1031.

(6) Berman, H.M.; Westbrook J.; Feng Z.; Gilliland G.; Bhat T.N.; Weissig H.; Shindyalov I.N.; Bourne P.E. *Nucleic Acids Res*. **2000**, 28, 235-242.

(7) Morris, G.M.; Huey, R.; Lindstrom, W.; Sanner, M.F.; Belew, R.K.; Goodsell, D.S.; Olson, A.J. *J. Comput. Chem*. **2009**, 16, 2785- 2791.

(8) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling C. *Proteins* **2006**, 65, 712-725.

(9) Wang, J.M.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D.A. *J. Comput. Chem*. **2004**, 25, 1157-1174.

(10) Wang, J.M.; Wang, W.; Kollman, P.A.; Case, D.A. *J. Mol. Graph*. **2006**, 25, 247-260.

(11) Jakalian, A.; Bush, B.L.; Jack, D.B.; Bayly, C.I. *J. Comput. Chem.* **2000**, 21, 132-146.

(12) Berendsen, H.J.C.; Postma, J.P.M.; van Gunsteren, W.F.; Di Nola, A.; Haak, J.R. *J. Chem. Phys.* **1984**, 81, 3684-3690.

(13) Pastor, R.W.; Brooks, B.R.; Szabo, A. *Mol. Phys*. **1988**, 65, 1409-1419.

(14) Ryckaert, J.P.; Ciccotti, G.; Berendsen, H.J.C. *J. Comput. Phys.* **1977**, 23, 327-341.

(15) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys*. **1993**, 98, 10089-10092.

(16) Hou, T., Wang J., Li, Y., Wang, W. *J. Chem. Inf. Model*. **2011**, 51, 69-82.

(17) Onufriev, A.; Bashford, D.; Case, D.A. *Protein*, **2004**, 55, 383-394.

(18) Weiser, J.; Shenkin, P. S.; Still, W. C. *J. Comput. Chem.* **1999**, 20, 217-230.

(19) Weis, A.; Katebzadeh, K.; Soderhjelm, P.; Nilsson, I.; Ryde, U. *J. Med. Chem*. **2006**, 49, 6596-6606.

(20) Brown, S. P.; Muchmore, S. W. *J. Chem. Inf. Model*. **2007**, 47, 1493-1503.

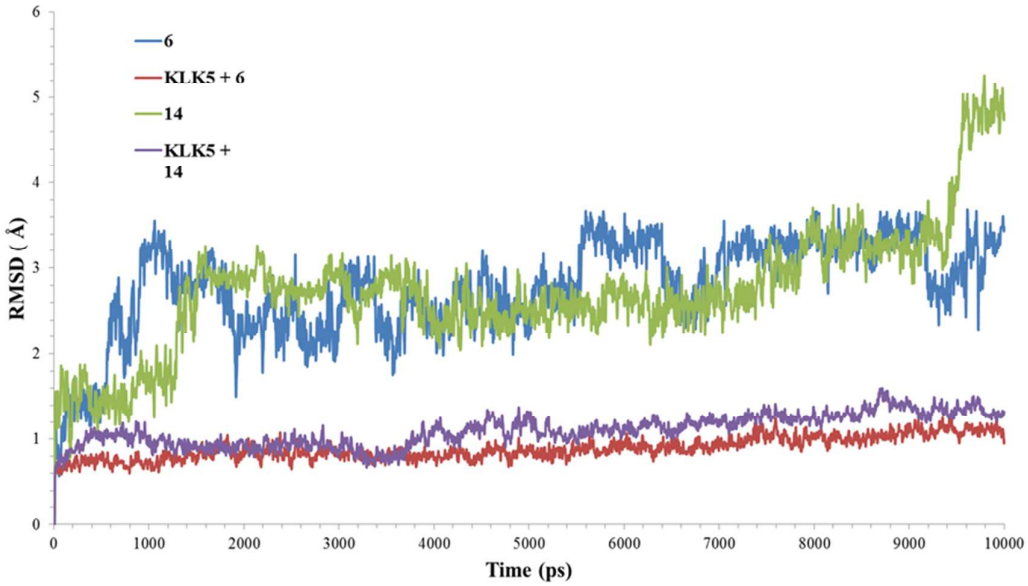
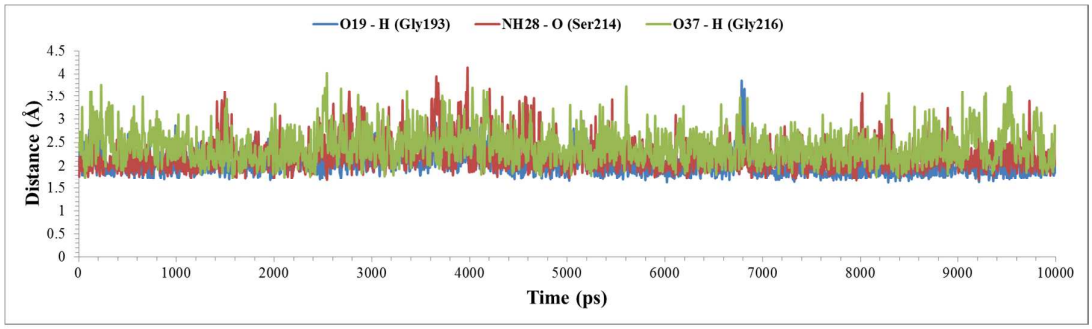
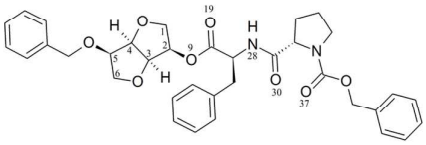
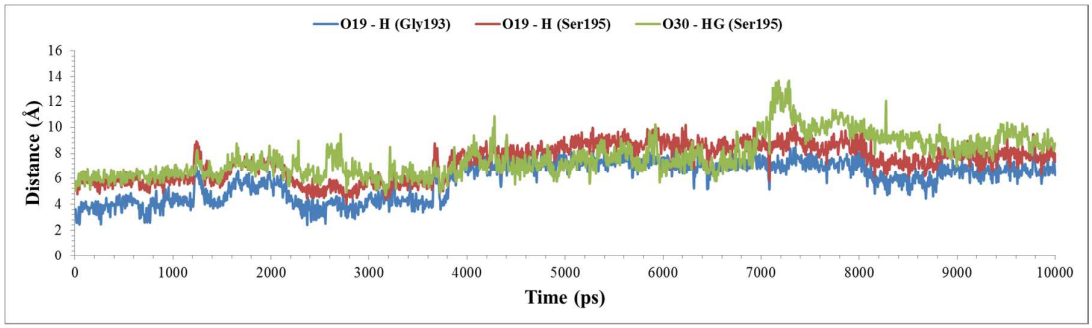


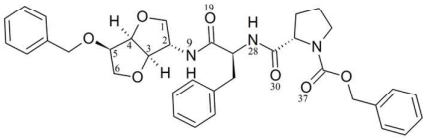
Figure S1. RMSD of the backbone atoms of KLK5 enzyme and heavy atoms of the inhibitors **6** and **14**.





6





14

Figure S2. Plots of the hydrogen bonds lengths *vs*. time for the MD simulation of KLK5 in complex with compound **6** and **14**.

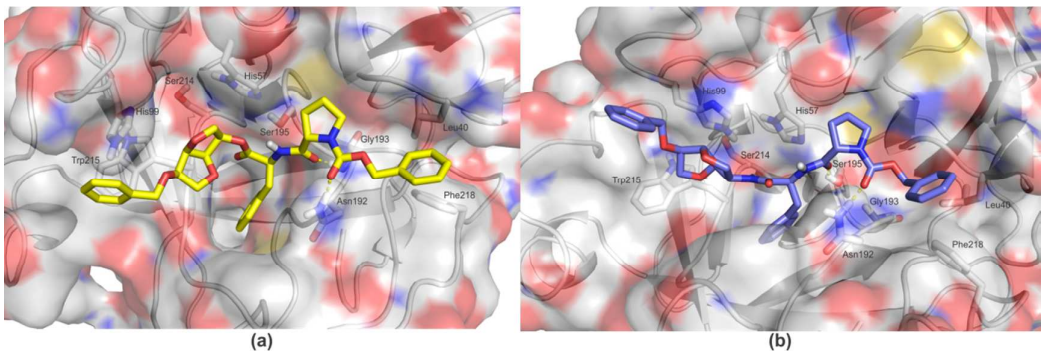


Figure S3. Predicted binding mode of inhibitor **6** (a) and **14** (b) in the active site of KLK7, illustrating the essential residues and the hydrogen bonds (yellow dash line). Inhibitors are shown in stick representation and the protein backbone as ribbons. Selected active site residues predicted to interact with the inhibitors are illustrated. The surface of the enzyme is shown colored by element.

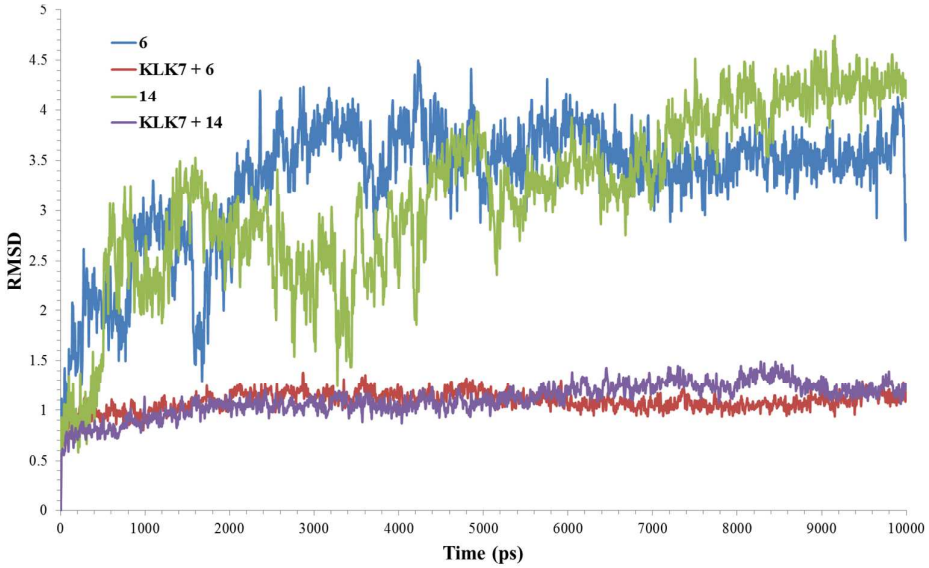


Figure S4. RMSD of the backbone atoms of KLK7 enzyme and heavy atoms of the inhibitors **6** and **14**.

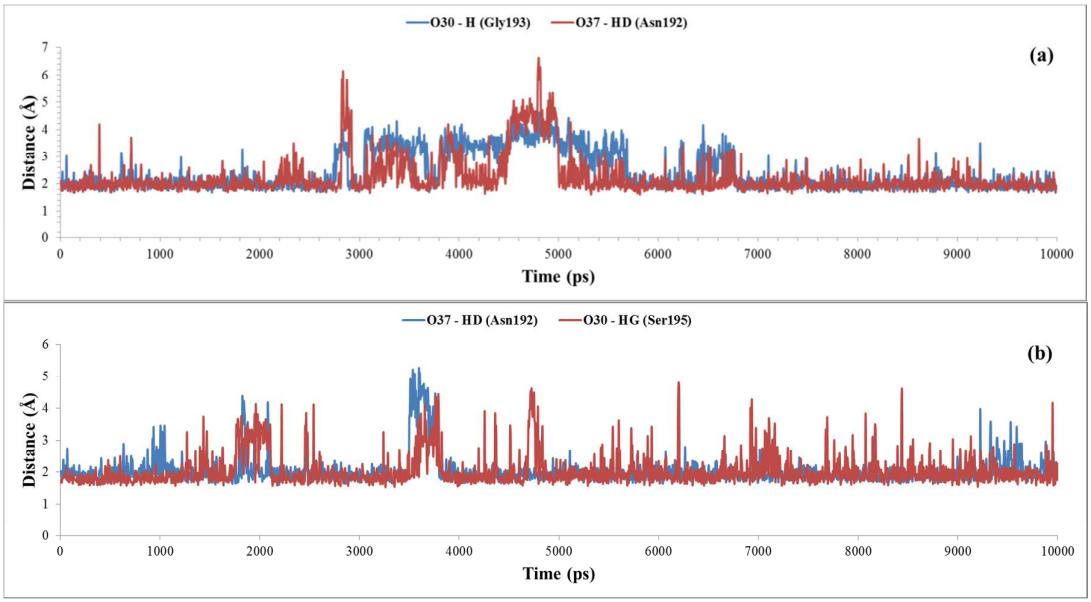


Figure S5. Plots of the hydrogen bonds lengths *vs*. time for the MD simulation of KLK7 in complex with compound **6** (a) and **14** (b).

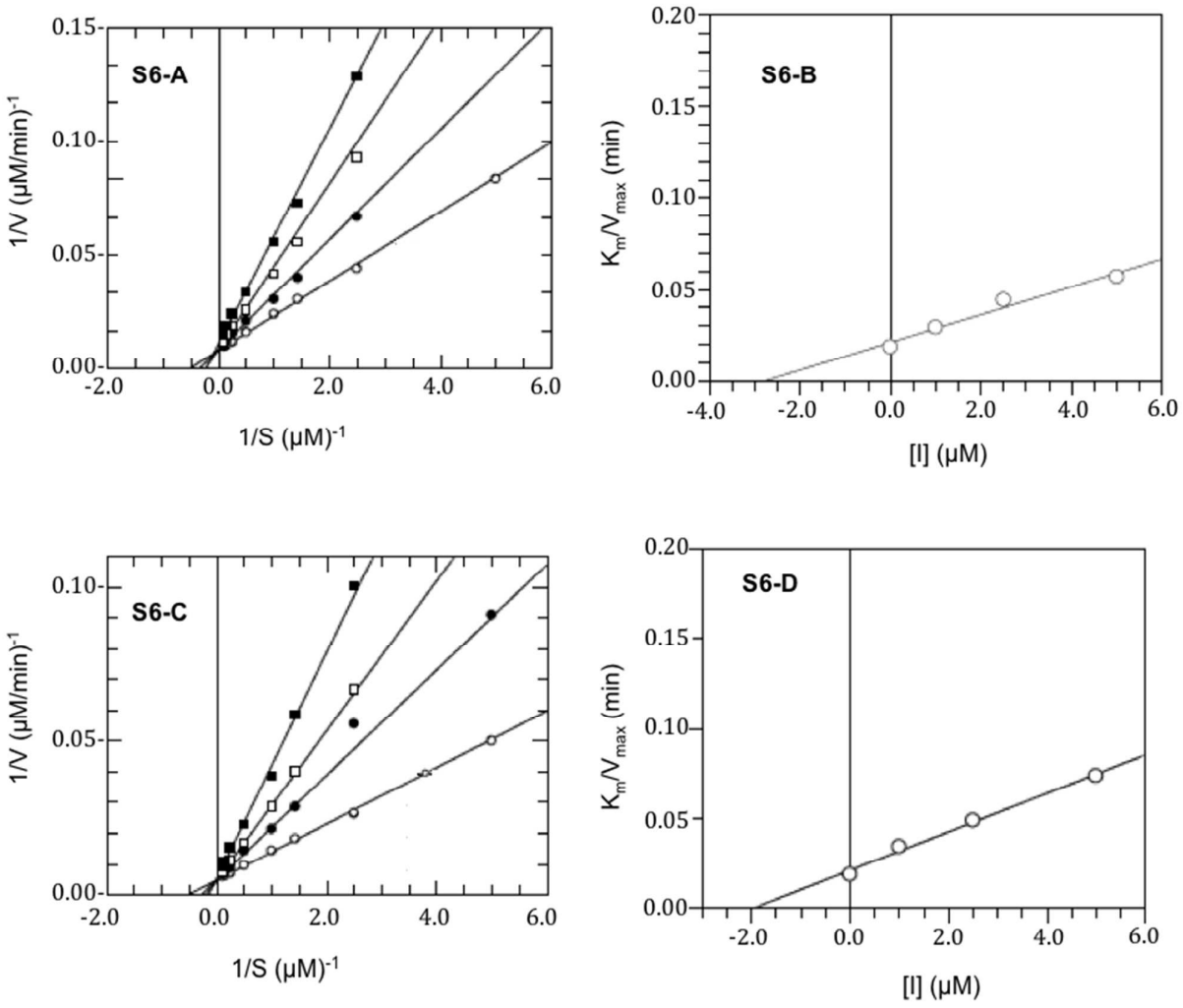


Figure S6. Representative graphs for the kinetic constants determination. The mechanism of inhibition was determined by the Lineweaver-Burk plots for the hydrolysis of FRET substrate Abz-KLYSSKQ-EDDnp by KLK7, in the presence of inhibitor **6** (S6A) and **7** (S6C). The Ki values were determined by the graph K_m/V_{max} vs [I] for the compounds **6** (S6B) and **7** (S6D). In S6A and S6C the solid lines represent the linear regression in fits obtained by software GraFit 7.0 in absence of inhibitor (○) and three different inhibitor concentrations (1.0 μM (●), 2.5 μM (□), and 5.0 μM (■).