Supporting Information

Benzoxaboroles: new potent inhibitors of the carbonic anhydrases of the pathogenic bacterium Vibrio cholerae.

Alessandro Bonardi, Alessio Nocentini, Roberta Cadoni, Sonia del Prete, Pascal Dumy, Clemente Capasso, Paola Gratteri, Claudiu T. Supuran and Jean-Yves Winum

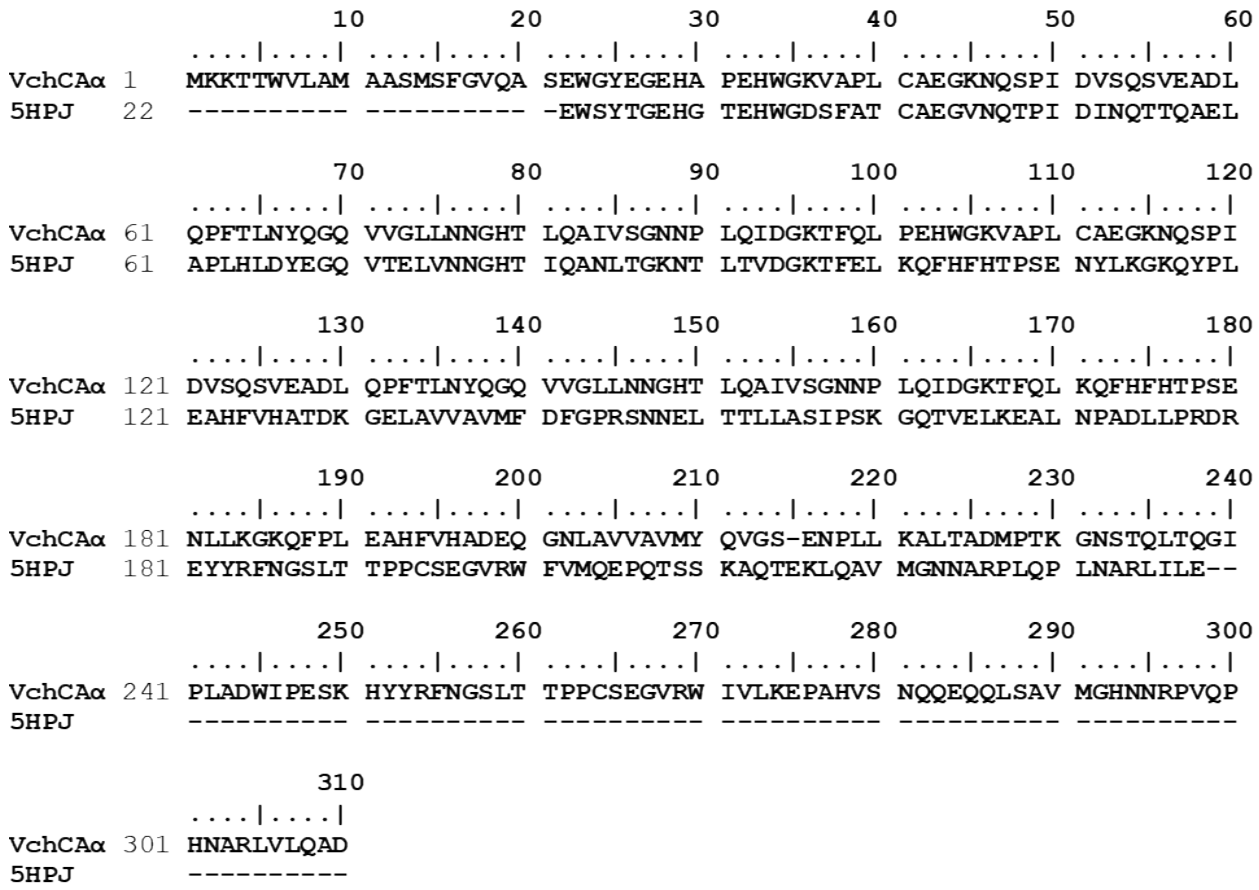

Figure S1. Sequence alignment of VchCAα with the template α-CA from *Photobacterium profundum* (pdb 5HPJ).

```
                    10         20         30         40         50         60
               ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
VchCAβ_A 1     MPEIKQLFEN NSKWSESIKA ETPEYFAKLA KGQNPDFLWI GCADSRVPAE RLTGLYSGEL
VchCAβ_B 1     MPEIKQLFEN NSKWSESIKA ETPEYFAKLA KGQNPDFLWI GCADSRVPAE RLTGLYSGEL
6D2N     3     --ALQQLFEN NVRWAEAIKQ EDPDFFAKLA RQQTPEYLWI GCSDARVPAN EIVGMLPGDL

                    70         80         90        100        110        120
               ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
VchCAβ_A 61    FVHRNVANQV IHTDLNCLSV VQYAVDVLQV KHIIVCGHYG CGGVTAAIDN PQLGLINNWL
VchCAβ_B 61    FVHRNVANQV IHTDLNCLSV VQYAVDVLQV KHIIVCGHYG CGGVTAAIDN PQLGLINNWL
6D2N     61    FVHRNVANVV LHTDLNCLSV IQFAVDVLKV KHILVTGHYG CGGVRASLHN DQLGLIDGWL

                   130        140        150        160        170        180
               ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
VchCAβ_A 121   LHIRDYYLKH REYLDQMPAE D-RSDKLAEI NVAEQVYNLA NSTVLQNAWE RGQAVEVHGF
VchCAβ_B 121   LHIRDYYLKH REYLDQMPAE D-RSDKLAEI NVAEQVYNLA NSTVLQNAWE RGQAVEVHGF
6D2N     121   RSIRDLAYEY REHLEQLPTE EERVDRLCEL NVIQQVANVS HTSIVQNAWH RGQSLSVHGC

                   190        200        210        220
               ....|....| ....|....| ....|....| ....|....| ...
VchCAβ_A 180   VYGIEDGRLE YLGVRCASRS AVEDNYHKAL EKILNPNHRL LCR
VchCAβ_B 180   VYGIEDGRLE YLGVRCASRS AVEDNYHKAL EKILNPNHRL LCR
6D2N     181   IYGIKDGLWKNLNVTVSG--- ---------L DQLP-PQYRL SPL
```

Figure S2. Sequence alignment of type-I VchCAβ with the template type-I β-CA from *Pseudomonas aeruginosa* (pdb 6D2N).

```
                    10         20         30         40         50         60
               ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
VchCAγ_A 1     SSIRSYKGIV PKLGEGVYVD SSAVLVGDIE LGDDASIWPL VAARGDVNHI RIGKRTNIQD
VchCAγ_B 1     SSIRSYKGIV PKLGEGVYVD SSAVLVGDIE LGDDASIWPL VAARGDVNHI RIGKRTNIQD
VchCAγ_C 1     SSIRSYKGIV PKLGEGVYVD SSAVLVGDIE LGDDASIWPL VAARGDVNHI RIGKRTNIQD
3TIO     1     DVLHPYRDLF PQIGQRVMID DSSVVIGDVR LADDVGIWPL VVIRGDVHYV QIGARTNIQD

                    70         80         90        100        110        120
               ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
VchCAγ_A 61    GSVLHVTHKN AENPNGYPLC IGDDVTIGHK VMLHGCTIHD RVLVGMGSIV LDGAVIENDV
VchCAγ_B 61    GSVLHVTHKN AENPNGYPLC IGDDVTIGHK VMLHGCTIHD RVLVGMGSIV LDGAVIENDV
VchCAγ_C 61    GSVLHVTHKN AENPNGYPLC IGDDVTIGHK VMLHGCTIHD RVLVGMGSIV LDGAVIENDV
3TIO     61    GSMLHVTHKS SYNPDGNPLT IGEDVTVGHK VMLHGCTIGN RVLVGMGSIL LDGAIVEDDV

                   130        140        150        160        170        180
               ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
VchCAγ_A 121   MIGAGSLVPP GKRLESGFLY MGSPVKQARP LSDKERAFLV KSSSNYVQSK NDYLNDVKTV
VchCAγ_B 121   MIGAGSLVPP GKRLESGFLY MGSPVKQARP LSDKERAFLV KSSSNYVQSK NDYLNDVKTV
VchCAγ_C 121   MIGAGSLVPP GKRLESGFLY MGSPVKQARP LSDKERAFLV KSSSNYVQSK NDYLNDVKTV
3TIO     121   MIGAGSLVPQ NKRLESGYLY LGSPVKQIRP LSDEEKAGLR YSANNYVKWK DEYL------


               ..
VchCAγ_A 181   RE
VchCAγ_B 181   RE
VchCAγ_C 181   RE
3TIO           --
```

Figure S3. Sequence alignment of VchCAγ with the template γ-CA from *Escherichia coli* (pdb 3TIO).

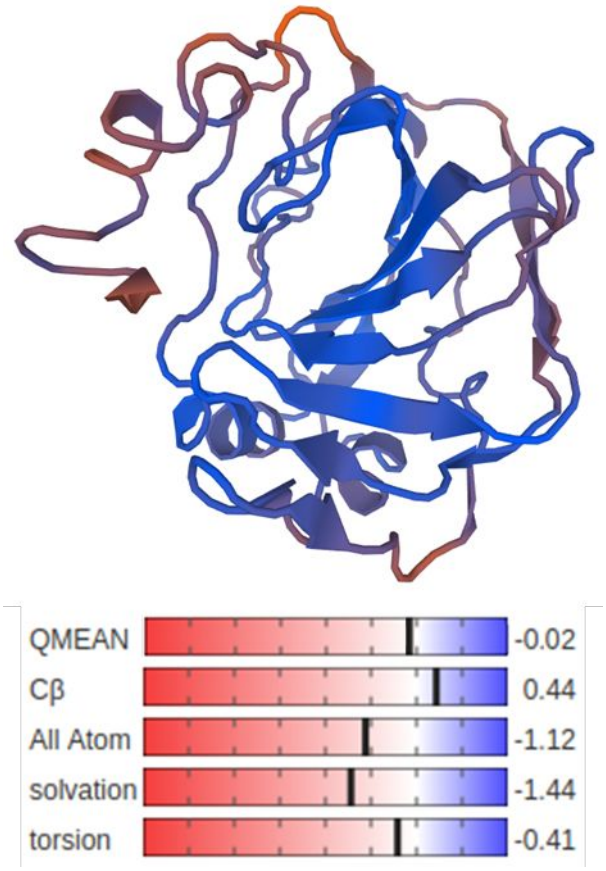

Figure S4. 3D representation of the homology model of VchCAα and related parameters calculated with Swiss model.

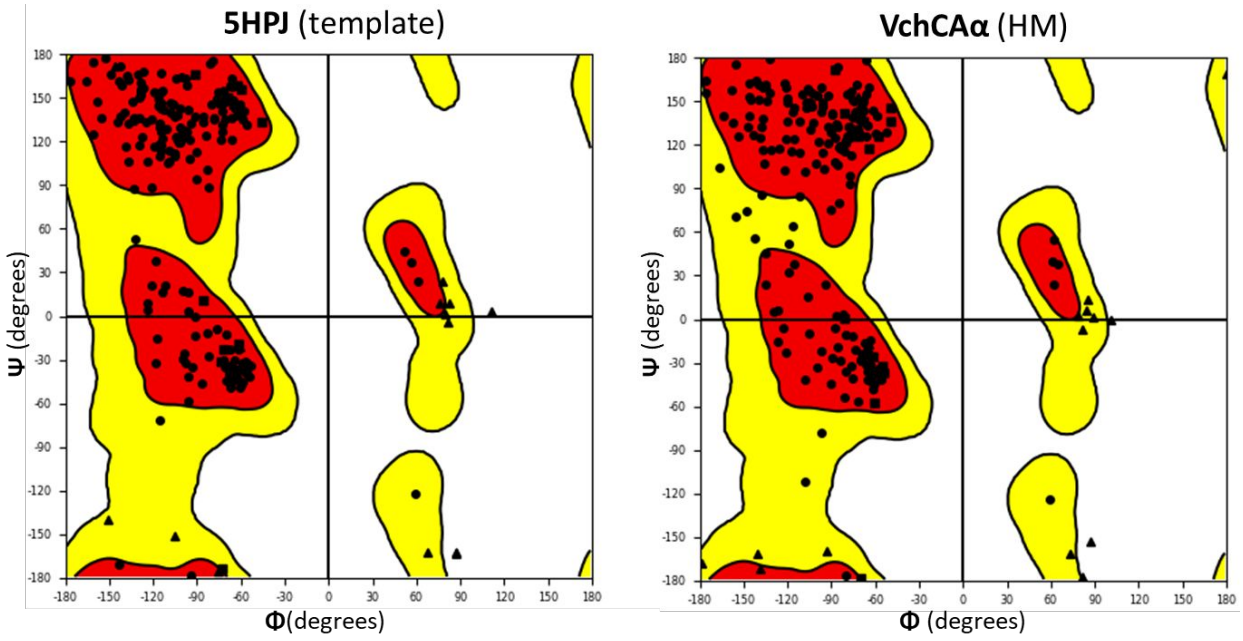

Figure S5. Ramachandran Plot of 5HPJ (template) and VchCAα (HM).

Table S1. Structural parameters of 5HPJ (template) and VchCAα (HM).

	5HPJ	VchCAα
MolProbity score	1.09	1.25
Clash score	2.96	0.90
Ramachandran Favoured	98.14 %	92.13 %
Ramachandran Outliers	0.00 %	0.46 %
Rotamer Outliers	0.53 %	0.00 %
C-Beta Deviations	0	0
Bad Bonds	0/1768	0/1748
Bad Angles	1/2415	23/2385
Cis Prolines	3/13	4/16
QMEAN	0.46	-0.02
Cβ	-0.37	0.44
All Atom	0.47	-1.12
solvation	-0.65	-1.44
torsion	0.81	-0.41

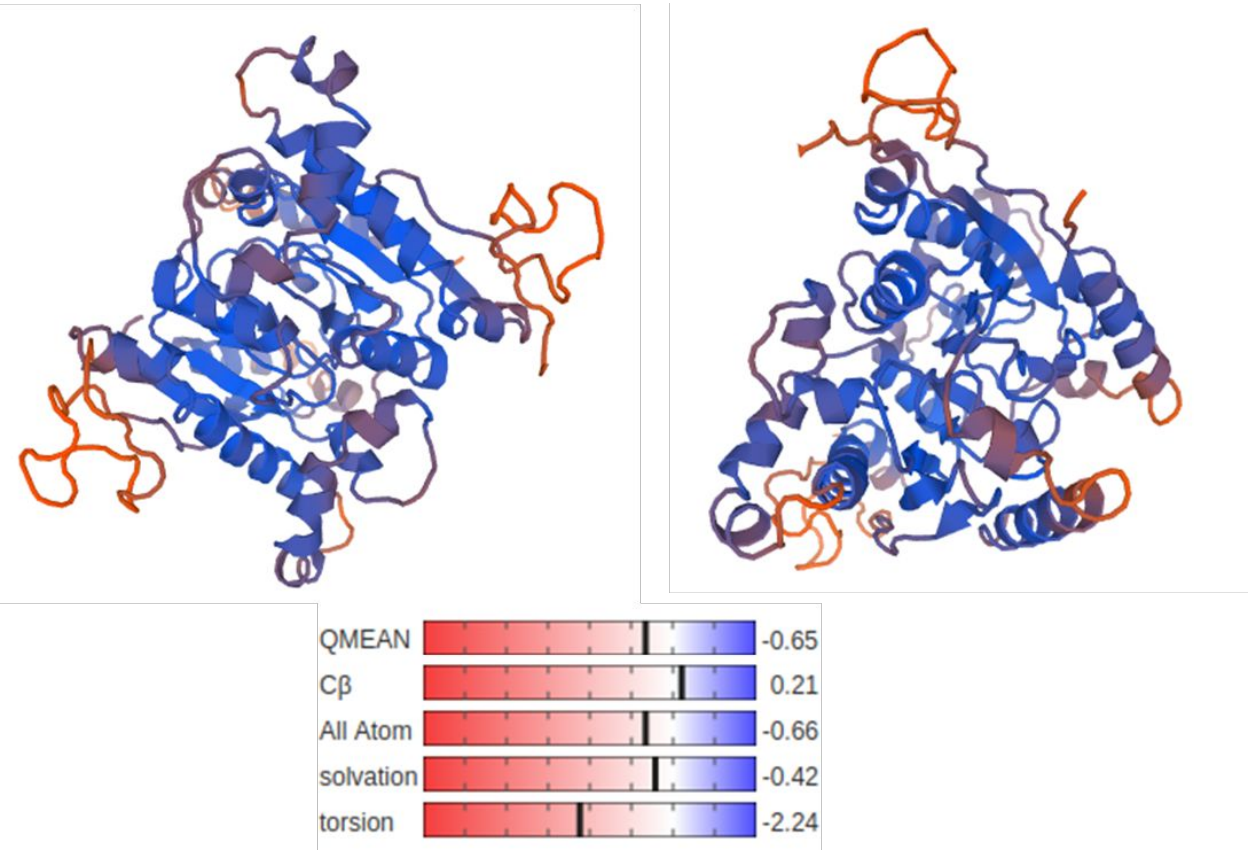

Figure S6. 3D representation of the homology model of type I VchCAβ and related parameters calculated with Swiss model.

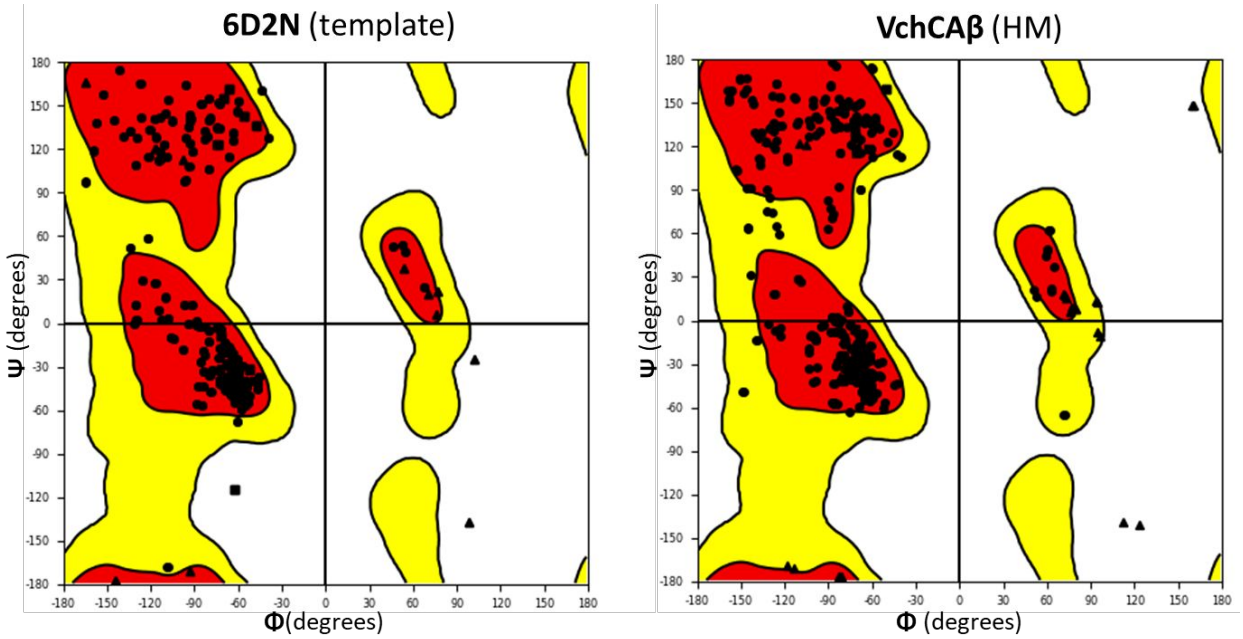

Figure S7. Ramachandran Plot of 6D2N (template) and VchCAβ (HM).

Table S2. Structural parameters of 6D2N (template) and VchCAβ (HM).

	6D2N	**VchCAβ**
MolProbity score	1.60	1.78
Clash score	3.95	1.99
Ramachandran Favoured	97.10 %	93.41 %
Ramachandran Outliers	0.97 %	0.00 %
Rotamer Outliers	2.21 %	3.16 %
C-Beta Deviations	0	0
Bad Bonds	0/3390	0/3634
Bad Angles	2/4630	38/4934
QMEAN	-0.34	-0.65
Cβ	0.06	0.21
All Atom	0.89	-0.66
solvation	0.41	-0.42
torsion	-0.56	-2.24

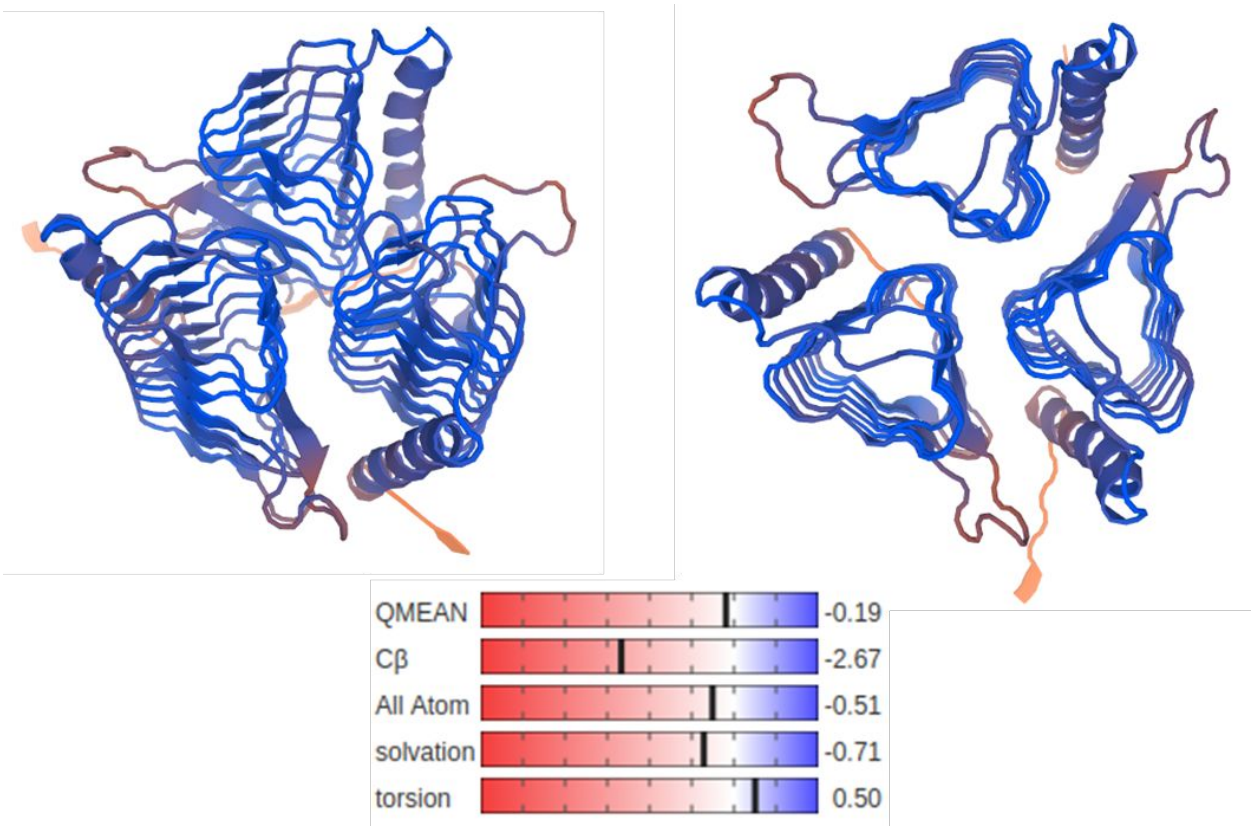

Figure S8. 3D representation of the homology model of VchCAγ and related parameters calculated with Swiss model.

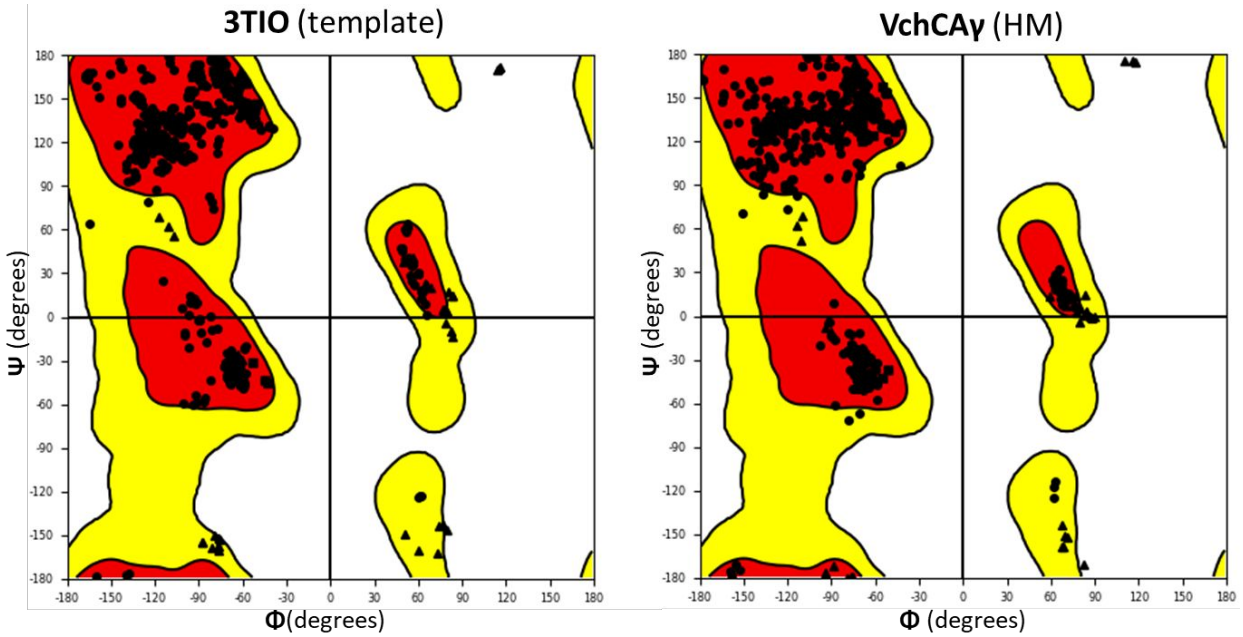

Figure S9. Ramachandran Plot of 3TIO (template) and VchCAγ (HM).

Table S3. Structural parameters of 6D2N (template) and VchCAγ (HM).

	3TIO	**VchCAγ**
MolProbity score	1.70	1.11
Clash score	7.28	0.25
Ramachandran Favoured	97.68 %	95.96 %
Ramachandran Outliers	0.00 %	0.19 %
Rotamer Outliers	2.00 %	2.05 %
C-Beta Deviations	0	3
Bad Bonds	0/8280	0/4024
Bad Angles	1/11244	46/5449
Cis Prolines	6/47	3/24
QMEAN	1.01	-0.19
Cβ	-1.49	-2.67
All Atom	-0.14	-0.51
solvation	0.38	-0.71
torsion	1.11	0.50

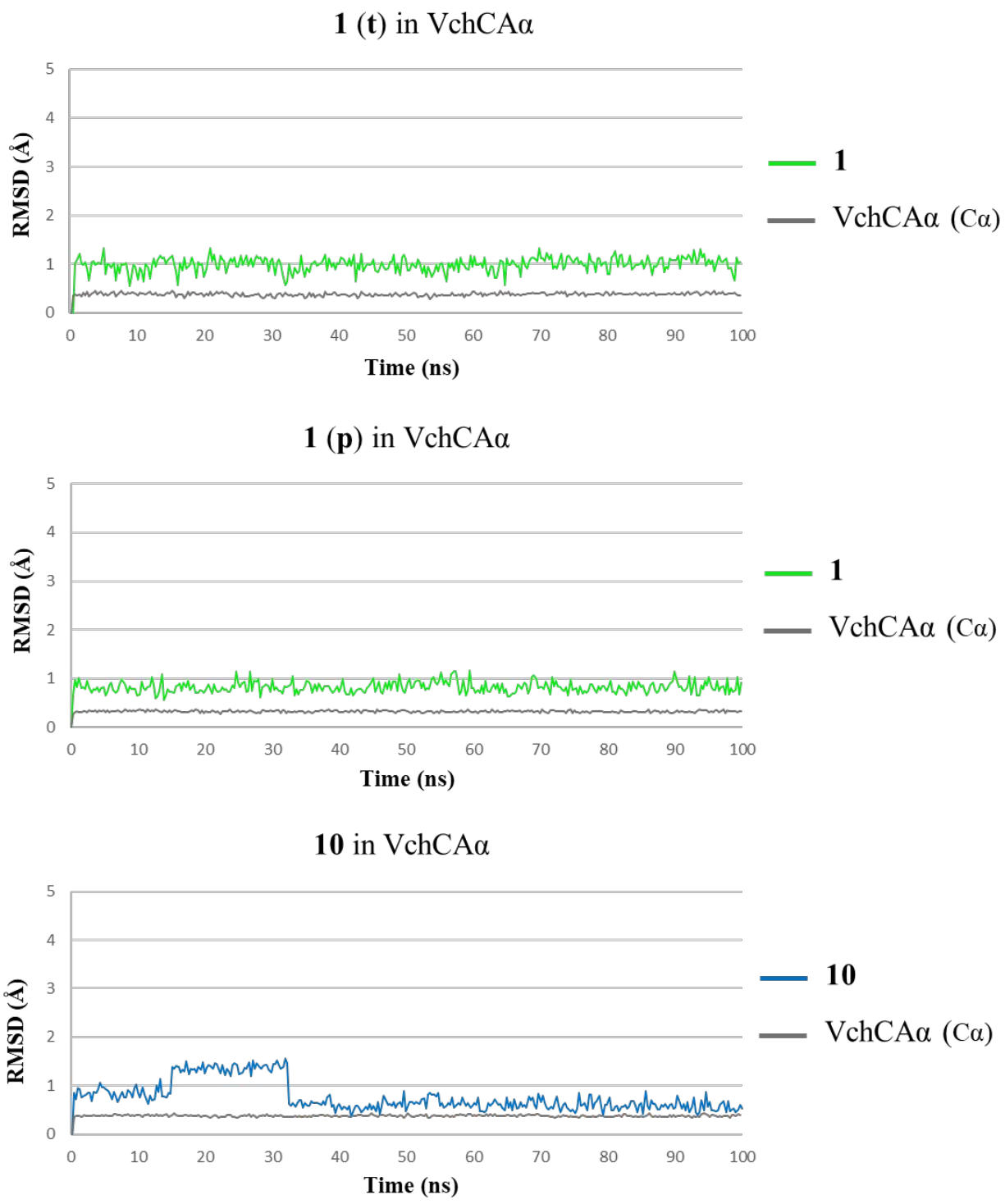

Figure S10. Ligand and protein RMSD (Å) *vs* time (100 ns) for the MD of the binding poses of benzoxaboroles **1** (**t** and **p**), and **10** with VchCAα.

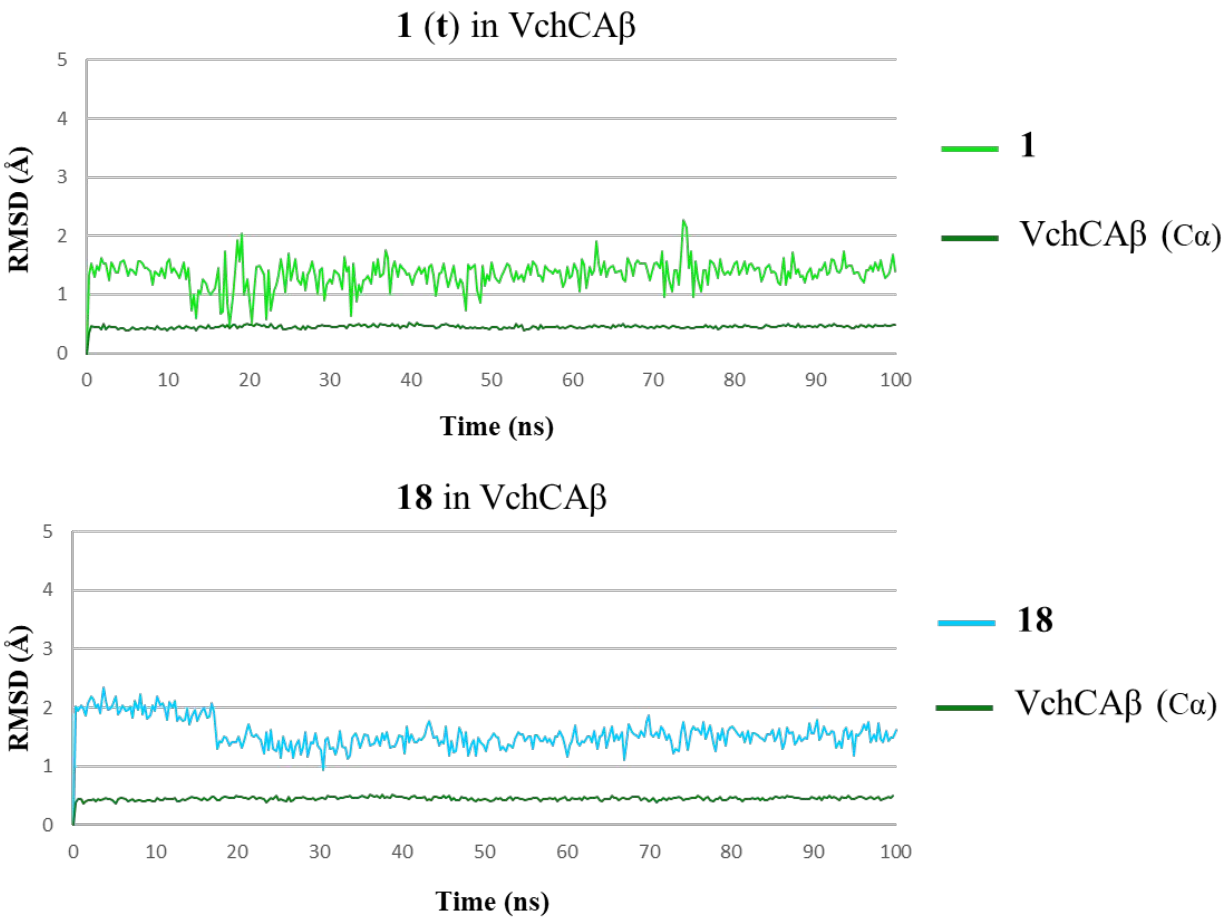

Figure S11. Ligand and protein RMSD (Å) *vs* time (100 ns) for the MD of the binding poses of benzoxaboroles **1** (**t**) and **18** with VchCAβ.

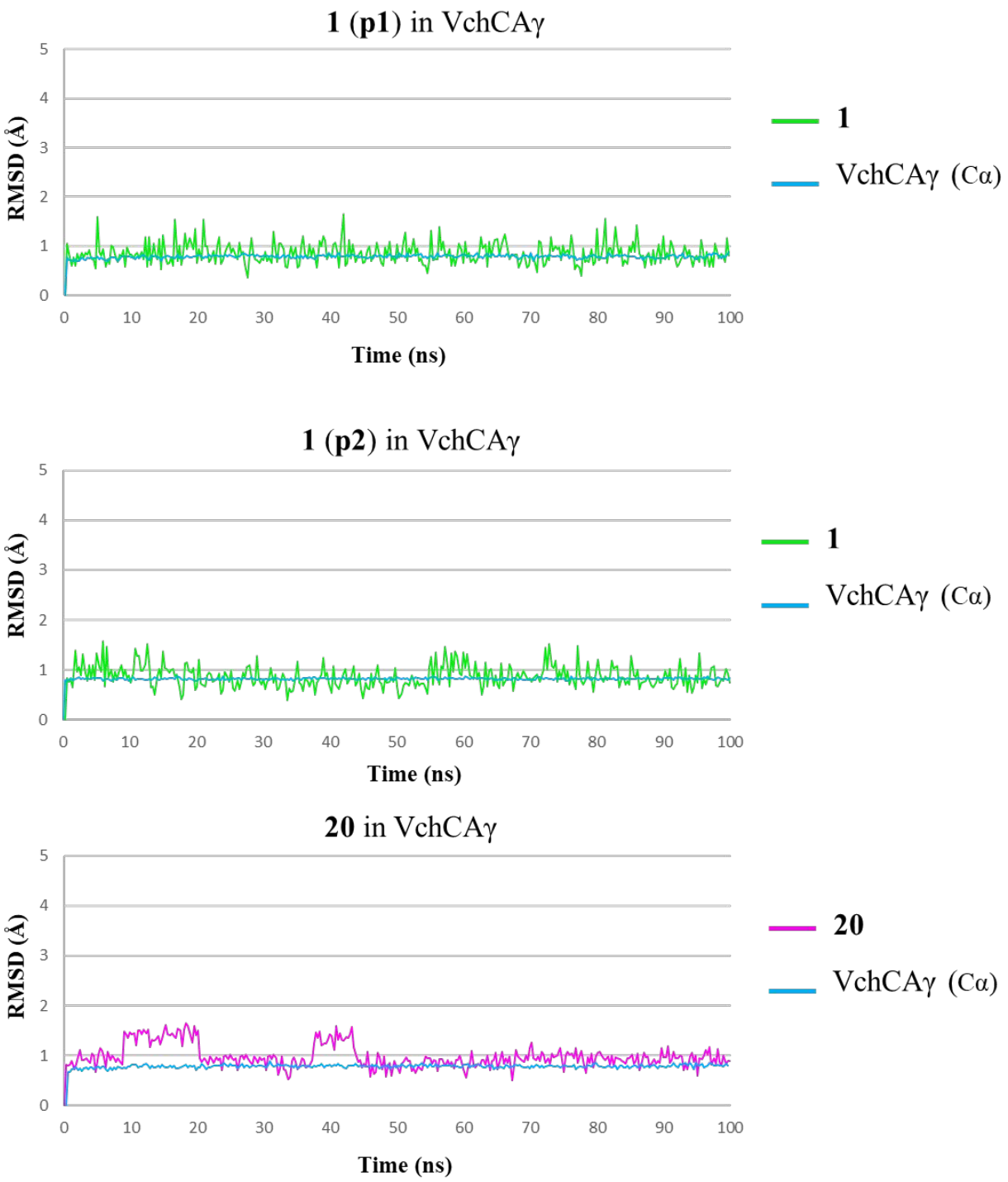

Figure S12. Ligand and protein RMSD (Å) *vs* time (100 ns) for the MD of the binding poses of benzoxaboroles **1** (**p1** and **p2**), and **20** with VchCAγ.

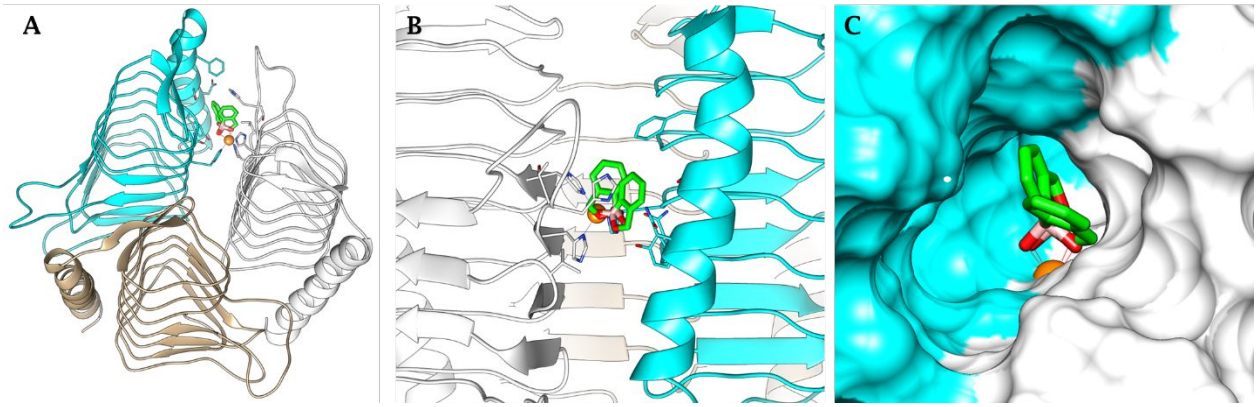

Figure S13. (A) Upper and (B) transversal and (C) surface view of VchCAγ in adduct with superimposed **p1** and **p2** orientations of compound **1**.

Materials and Methods

Chemistry

Benzoxaborole **1** and tavaborole were commercially available. Compounds **2–23** were previously reported by this group [*Chem. Commun.* **2016**, *52(80)*, 11983-11986.]

Carbonic anhydrase assays

An Sx.18Mv-R Applied Photophysics (Oxford, UK) stopped-flow instrument has been used to assay the catalytic activity of various CA isozymes for CO_2 hydration reaction [J. Biol. Chem. 1971, 246, 2561]. Phenol red (at a concentration of 0.2 mM) was used as indicator, working at the absorbance maximum of 557 nm, with 20 mM Tris (pH 7.4 for VchCAα and hCA II and 8.3 for VchCAβ and VchCAγ) as buffer, and 20 mM $NaClO_4$ (for maintaining constant the ionic strength), following the initial rates of the CA-catalyzed CO_2 hydration reaction for a period of 10–100 s. The CO_2 concentration ranged from 1.7 to 17 mM for the determination of the kinetic parameters and inhibition constants. For each inhibitor, at least six traces of the initial 5–10% of the reaction have been used for determining the initial velocity. The uncatalyzed rates were determined in the same manner and subtracted from the total observed rates. Stock solutions of inhibitor (0.1 mM) were prepared in distilled-deionized water and dilutions up to 0.01 nM were done thereafter with the assay buffer. Inhibitor and enzyme solutions were preincubated together for 30 min at room temperature prior to assay, in order to allow for the formation of the E–I complex. The inhibition constants were obtained by nonlinear least-squares methods using PRISM 3 and the Cheng-Prusoff equation, as reported earlier, and represent the mean from at least three different determinations [Bioorg. Chem. 2019, 93, 103336; J. Enzyme Inhib. Med. Chem. 2020, 35, 59-64; J. Enzyme Inhib. Med. Chem. 2019, 34(1), 510-518.]. VchCAα, VchCAβ, and VchCAγ were recombinant proteins obtained in-house as reported earlier [J. Enzyme Inhib. Med. Chem. 2014, 29(1), 23-7.; Bioorg. Med. Chem. 2016, 24(5), 1115-20.; 11-15; Acta Crystallogr. D Biol. Crystallogr. 2015, 71(Pt 12), 2449-56.; Bioorg. Med. Chem. Lett. 2016, 26(8), 1941-6.; Bioorg. Med. Chem. 2016, 24(16), 3413-7.].

Molecular Modelling

The primary sequences of VchCAα, VchCAβ, and VchCAγ were retrieved from the UniProt Consortium. The crystal structure of α-CA from *Photobacterium profundum* (PDB 5HPJ [PLoS One. 2016, 11, e0168022.]), β-CA from *Pseudomonas aeruginosa* (PDB 6D2N [ChemMedChem. 2018, 13, 2024-2029], and γ-CA homologous protein from *Escherichia coli* (PDB 3TIO [Acta Crystallogr. D Biol Crystallogr. 2012, 68, 920-6.] were used as templates in the homology modeling procedure. Multiple models were generated using the Prime module of Schrödinger [Prime, v.5.5] and the SwissModel platform (Nucleic Acids Res. 2018, 46, W296-W303) and submitted to loop refinements and quality evaluation procedures (Figures S4-S9 and Tables S1-S3, Supporting Information). The best scored structures of VchCAα, VchCAβ and the VchCAγ crystal structure (retrieved as PDB 5CXK [Acta Crystallogr. D Biol. Crystallogr. 2015, 71(Pt 12), 2449-56.]) were prepared using the Protein Preparation Wizard tool implemented in the Schrödinger suite [Schrödinger Suite Release 2019-1, Schrödinger, LLC, New York, NY, 2019: (a) Prime, v.5.5; (b) Maestro v.11.9; (c) Epik, v.4.7; (d) Impact, v.8.2; (e) Macromodel v.12.3. (f) Glide, v.8.2]. The energy minimization protocol with a root mean square deviation (RMSD) value of 0.30 Å was applied using force field OPLS3e. The ligand structures in the $B(OH)_2^-$ form were submitted to QM geometry optimization (B3LYP/6-31G^{*+}) and ESP charges calculation with the Jaguar module (v.10.3) of Schrödinger. The software Glide was used for docking [Glide, v.8.2]. Grids were centered on the centroids of the zinc-coordinating residues and ligands were docked using QM-computed charges and the standard precision mode (SP). The best poses for each compound were re-docked and scored for its binding free energies by the Prime MM-GBSA protocol [Prime, v.5.5], using QM charges and a VSGB solvation model. Additionally, they were submitted to a MD simulation using Desmond [Desmond v.5.7] and the OPL3e force field. Specifically, the system was solvated in an orthorhombic box using TIP4PEW water molecules, extended 15 Å away from any protein atom. It was neutralized adding chlorine and sodium ions. The simulation protocol included a starting relaxation step followed by a final production phase of 100 ns. In particular, the relaxation step comprised the following: (a) a stage of 100 ps at 10 K retaining the harmonic restraints on the solute heavy atoms (force constant of 50.0 kcal mol−1 Å−2) using the NPT ensemble with Brownian dynamics; (b) a stage of

12 ps at 10 K with harmonic restraints on the solute heavy atoms (force constant of 50.0 kcal mol−1 Å−2), using the NVT ensemble and Berendsen thermostat; (c) a stage of 12 ps at 10 K and 1 atm, retaining the harmonic restraints and using the NPT ensemble and Berendsen thermostat and barostat; (f) a stage of 12 ps at 300 K and 1 atm, retaining the harmonic restraints and using the NPT ensemble and Berendsen thermostat and barostat; (g) a final 24 ps stage at 300 K and 1 atm without harmonic restraints, using the NPT Berendsen thermostat and barostat. The final production phase of MD was run using a canonical NPT Berendsen ensemble at temperature 300 K. During the MD simulation, a time step of 2 fs was used while constraining the bond lengths of hydrogen atoms with the M-SHAKE algorithm. The atomic coordinates of the system were saved every 100 ps along the MD trajectory. Protein and ligand RMSD, ligand torsions evolution and occupancy of intermolecular hydrogen bonds and hydrophobic contacts were computed along the production phase of the MD simulation with the Simulation Interaction Diagram tools implemented in Maestro.