

Supporting information for: Arylsulfonamides as Inhibitors for Carbonic Anhydrase: Prediction & Validation

Maurus Schmid,^{†,‡} Fabien W. Monnard,[†] Elisa S. Nogueira,[†] Thomas R. Ward,[†]
and Markus Meuwly^{*,‡}

*Department of Chemistry, University of Basel, Spitalstrasse 51, CH-4056 Basel, Switzerland, and
Department of Chemistry, University of Basel, Klingelbergstrasse 80, CH-4056 Basel, Switzerland*

E-mail: thomas.ward@unibas.ch, m.meuwly@unibas.ch

Contents

| | |
|------------------------------------|------------|
| Force Field Parametrization | S2 |
| General Aspects | S3 |
| Site-directed mutagenesis | S4 |
| Supplementary Figures | S6 |
| References | S10 |

*To whom correspondence should be addressed

[†]Department of Chemistry, University of Basel

[‡]Department of Chemistry, University of Basel

Force Field Parametrization

Force field parametrization was carried out by comparing with reference electronic structure calculations. First, the structure of the ligand was optimized with Density Functional Theory (DFT¹) using Gaussian² and the hybrid B3LYP functional.^{3,4} The atomic basis sets consisted of an effective core potential (LanL2DZ⁵) for the metal and the explicit 6-31G(d,p) basis for all remaining atoms.

All ligands required custom parameters for the force field. They were either taken from parameters for similar models through analogies or derived by fitting CHARMM force field terms⁶ to reproduce energy surface scans of the respective parameters calculated with DFT, as described above. To assist the parametrization, perl scripts using PerlMol⁷ were used to create the topology and necessary input files for the CHARMM calculations directly from the Gaussian output files. The fitting was performed using CHnolls⁸, the CHARMM interface for Inolls⁹. The charges for the ligand atoms were set to the Mulliken charges from the DFT optimizations (see above).

Two approaches can be envisaged for parametrizing the Zn(His)₃ centre in CAs: A "non-bonded"¹⁰ and a "bonded" one¹¹⁻¹³. Recently, a systematic derivation of force field parameters for the bonded model was reported by Lin et al.¹⁴ In the context of MM-GBSA, however, the fully bonded approach is not suitable because calculating a ligand-binding free energy ΔG requires the protein and the ligand to be non-covalently bound entities. Furthermore, test simulations with the fully non-bonded model failed to reproduce the tetrahedral geometry of the zinc centre. To circumvent this, a cationic dummy atom model was developed by Pang.¹⁵

Inspired by these previous efforts, a model consisting of a combination of bonded and non-bonded terms, was envisaged: The zinc center was parametrized with explicit bonds to the protein, but devoid of a bond to the arylsulfonamide ligand. This is motivated by the fact that the computation of ligand binding free energies (see below) is more straightforward with such a model. The zinc atom thus had a bond to each of the three histidine nitrogen atoms His₉₄N _{ϵ} , His₉₆N _{ϵ} , and His₁₁₉N _{δ} , but none to the nitrogen atom N_S of the sulfonamide. Some of the parameters for the zinc-histidine interactions from Lu and Voth¹², who parametrized an approach with four bonds to the zinc, proved

to be valid here as well and were used subsequently. The charge of the zinc atom was set to $q_{\text{Zn}} = +1e$, which was supported by a DFT calculation for a model system of the binding site consisting of 3 imidazoles and a arylsulfonamide, where the resulting Mulliken charge on zinc was $+1.04 e$. The parameters for the zinc center are given in Table S2.

To maintain a tetrahedrally coordinated zinc ion, as observed in the crystal structures, constraints on the zinc binding site were introduced using the Colvars module of NAMD. The 3 valence angles $\text{N}_S\text{ZnN}_{\epsilon,\alpha}$, where $\alpha = 1, 2, 3$ for the three His-nitrogen atoms, were constrained with a harmonic constraint $k_{\theta}(\theta - \theta_0)^2$ with $\theta_0 = 120^\circ$ with a rescaled force constant of $0.05 \text{ kcal/mol/deg}^2$. Further harmonic potentials were set on the 3 dihedral angles formed by these atoms and the sulphur S of the ligand. The dihedrals were constrained to the corresponding values observed in the crystal structure with a rescaled force constant of $0.00111 \text{ kcal/mol/deg}^2$. These values proved to yield a stable and conserved tetrahedral structure for the binding site throughout all simulations. The constraints were only used during the simulations and not for the subsequent analysis.

General Aspects

Materials and reagents were purchased from the highest commercially available grade and used without further purification. Esterase activity was recorded on a Tecan Safire spectrophotometer using NUNC 96-well plates. Human Carbonic Anhydrase isozyme II (hCA II) was expressed in 1 L shaking-flasks in an Infors HT Ecotron shaker and culture growth was checked by UV-Vis at 600 nm with a Varian Cary 50 Scan. Cell cultures were centrifuged either using a Heraeus Multifuge 4KR or a Heraeus Suprafuge 22 ($5346 \times g$). Cells were resuspended using Lab-Shaker (Adolf Kükner AG, Switzerland). SDS-PAGE gels were analyzed on a BioRad Gel Doc XR (Software: Quantity One 4.6.6). Sulfonamide affinity chromatography was performed using ÄKTAprime (Amersham Pharmacia Biotech, Software: PrimeView 5.0). Molecular weight was confirmed by MS ES-TOF (Bruker micrOTOF II, USA) and analyzed with Bruker Daltonics Data-Anakysis software.

Site-directed mutagenesis

Plasmid encoding hCA II and containing a T7 RNA polymerase promoter and an ampicillin resistance gene (pACA)¹⁶ was a generous gift from Carol Fierke, Michigan University.¹⁷ The construct of this plasmid has a serine residue at position 2 instead of an alanine, with no effect on protein expression or catalytic properties.

Primers were designed following the method described by Zheng et al.¹⁸ and tested *in silico* to minimize hairpin formation (Kibbe, www.basic.northwestern.edu/biotools/OligoCalc.html). Primers were obtained from Microsynth Balgach, Switzerland. PCR reactions were prepared by addition of 5 μ L 10x Pfu buffer, 2 μ L of 10 mM dNTP (final concentration 0.4 mM), 2.5 μ L DMSO (final concentration 5 %), 1.5 μ L Pfu Turbo polymerase, 1.5 μ L of 10 μ M primers (forward and reverse), 35 μ L H₂O to 1 μ L of template. The cycle conditions were: initial denaturation (95 °C, 5 min), followed by 16 cycles of 1 min at 95 °C; 1 min at 60 °C; 15 min at 68 °C. The final elongation was performed at 68 °C for 1 h. PCR products were analyzed by 2.4 % agarose gel electrophoresis.

The initial DNA template (wild type sequence) was digested by *DpnI* (4 h at 42 °C). 5 μ L of PCR product was used to transform ultra competent XL1-blue *E. coli* cells (produced in-house). Plasmids were purified using a Wizard Plus SV Miniprep DNA purification System (Promega - Switzerland) and were sequenced either by Starseq (Germany) or Microsynth (Switzerland).

Tables

Table S1: Primers sequences for Q-PCR

| Mutant | Forward primer (+) | Reverse primer (-) | Melting temp (°C) | Product length |
|--------|---|---|---------------------|----------------|
| L198A | CCTACCCAGGCT CAGCGACCACCC CTCCTCTTCTG | GAGGAGGGGTGG TCGCTGAGCCTG GGTAGGTCC | 72.7(+)/ 73.5(-) | 35(+)/33(-) |
| L198F | CCTACCCAGGCT CATTACCACCC CTCCTCTTCTG | GAGGAGGGGTGG TAAATGAGCCTG GGTAGGTCC | 70.2(+)/ 71.0(-) | 35(+)/33(-) |
| L198Q | CCTACCCAGGCT CACAGACCACCC CTCCTCTTCTG | GAGGAGGGGTGG TCTGTGAGCCTG GGTAGGTCC | 71.5(+)/ 71.9(-) | 35(+)/33(-) |

Table S2: MM Parameters for the Zinc moiety used in this study

| Atoms | | | | Force constant | equilibrium | |
|---------------------|------|-----|----------------------------|------------------------------|-------------|-------------|
| Bonds | | | | k | r_e | |
| Zn | NR2 | | 40 kcal/mol/Å ² | 2.24 Å | | |
| Angles ^a | | | | k | θ_e | |
| NR2 | Zn | NR2 | | 23 kcal/mol/rad ² | 109.5° | |
| CPH2 | NR2 | ZNB | | 20 kcal/mol/rad ² | 126.0° | |
| Dihedrals | | | | k | n | φ_0 |
| NR1 | CPH2 | NR2 | ZNB | 5.0 kcal/mol | 2 | 180.0° |
| CPH1 | CPH1 | NR2 | ZNB | 5.0 kcal/mol | 2 | 180.0° |
| HR1 | CPH2 | NR2 | ZNB | 5.0 kcal/mol | 2 | 180.0° |
| HR3 | PH1 | NR2 | ZNB | 5.0 kcal/mol | 2 | 180.0° |
| Improper Dihedrals | | | | k | n | φ_0 |
| ZNB | NR2 | NR2 | NR2 | 25 kcal/mol/rad ² | 0 | 0° |

^a parameters from¹²

Supplementary Figures

Figure captions

Figure S1 Variation of RMSD as a function of simulation time for **1** ⊂ hCA II.

Figure S2 MM-GBSA interaction energies of **1** ⊂ hCA II per frame with running averages over 100 steps.

Figure S3 Selected active site residues of **1** ⊂ hCA II. Residues colored by residue type (red: acidic, white: neutral, light blue: polar and dark blue basic residues).

Figures

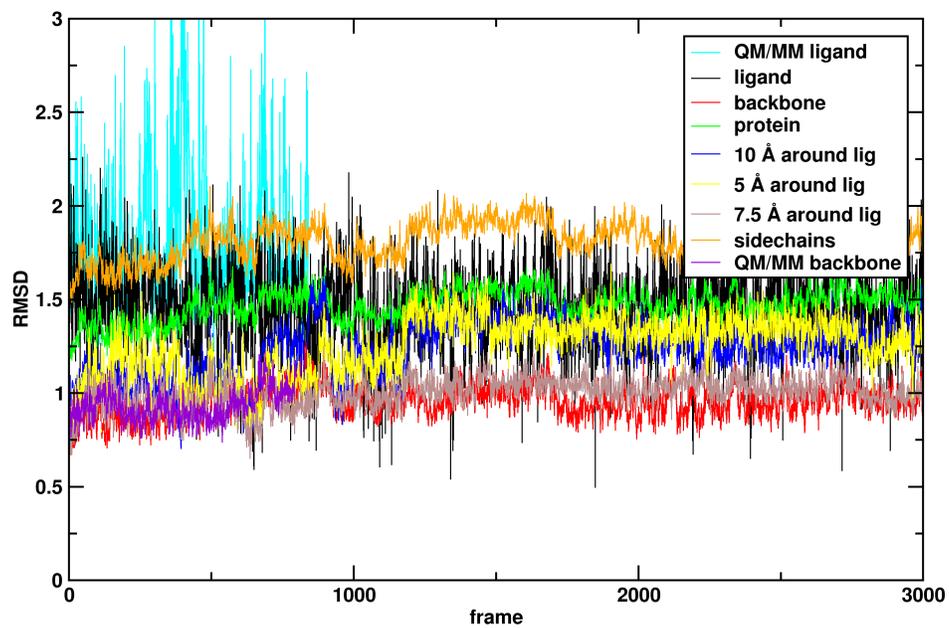


Figure S1: Variation of RMSD as a function of simulation time for **1** C hCA II.

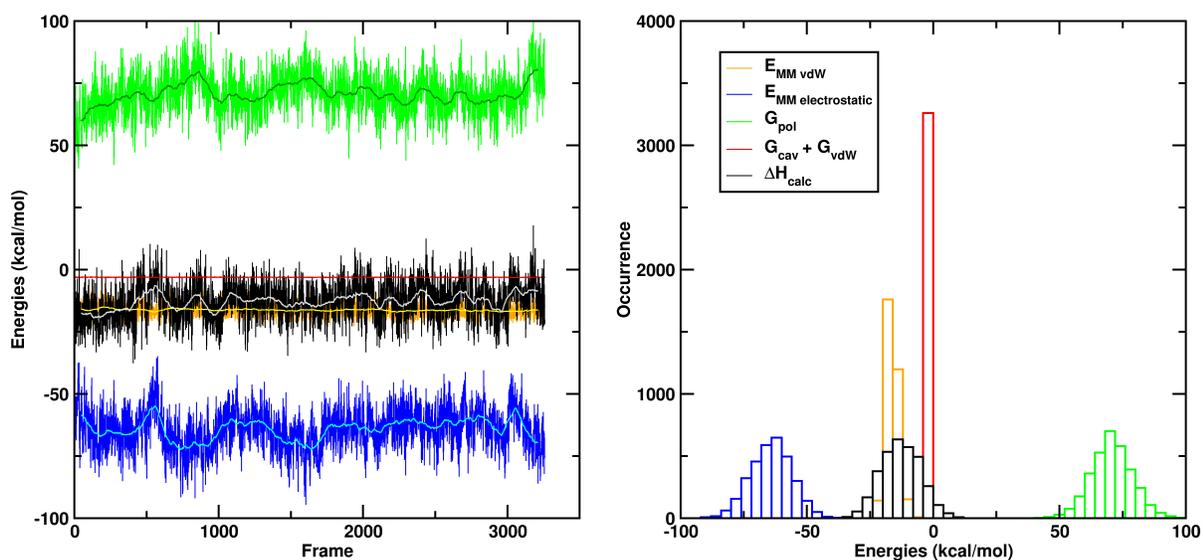


Figure S2: MM-GBSA interaction energies of **1** in hCA II per frame with running averages over 100 steps.

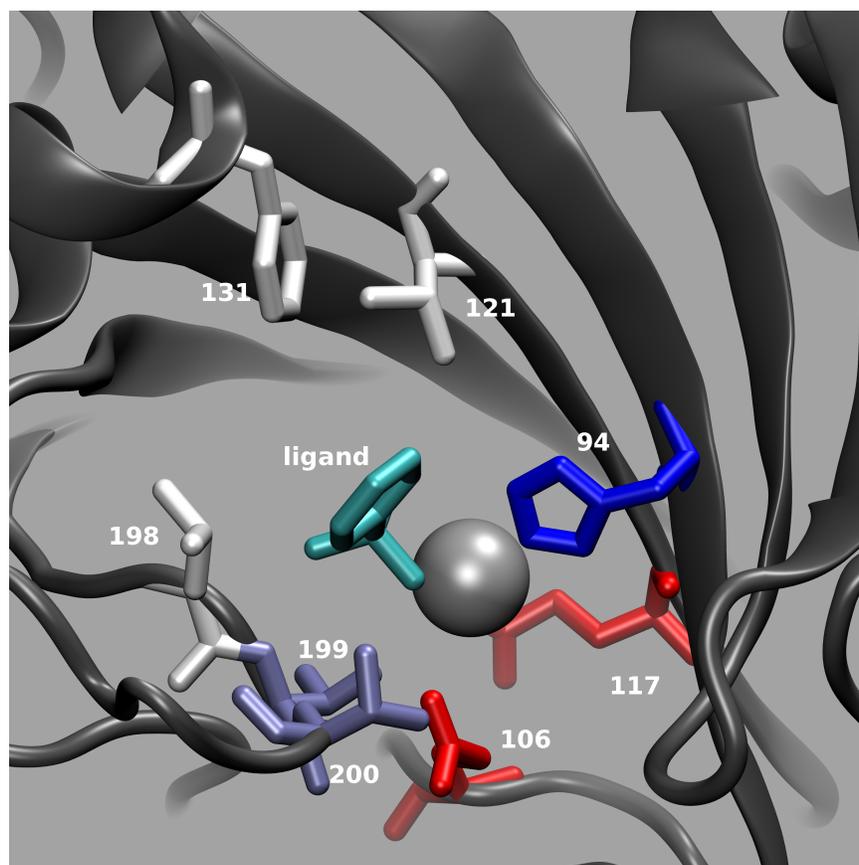


Figure S3: Selected active site residues of **1** in hCA II. Residues colored by residue type (red: acidic, white: neutral, light blue: polar and dark blue basic residues).

References

- (1) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864–B871.
- (2) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03, Revision B.01*.
- (3) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (4) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623–11627.
- (5) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 270–283.
- (6) MacKerell Jr., A. D.; Brooks III, C. L.; Nilsson, L.; Roux, B.; Won, Y.; Karplus, M. In *CHARMM: The Energy Function and Its Parameterization with an Overview of the Program*; v. R. Schleyer et al., P., Ed.; John Wiley & Sons: Chichester, 1998; Vol. 1, pp 271–277.
- (7) Tubert-Brohman, I. *The Perl Journal* **2004**, *8*, 3–5.
- (8) Devereux, M.; Meuwly, M. *J. Chem. Inf. Model.* **2010**, *50*, 349–357.

- (9) Law, M.; Hutson, J. *Comput. Phys. Commun.* **1997**, *102*, 252–268.
- (10) Stote, R. H.; Karplus, M. *Proteins Struct. Funct. Bioinf.* **1995**, *23*, 12–31.
- (11) Hoops, S. C.; Anderson, K. W.; Merz, K. M. *J. Am. Chem. Soc.* **1991**, *113*, 8262–8270.
- (12) Lu, D.; Voth, G. A. *Proteins Struct. Funct. Bioinf.* **1998**, *33*, 119–134.
- (13) Alterio, V.; Vitale, R. M.; Monti, S. M.; Pedone, C.; Scozzafava, A.; Cecchi, A.; De Simone, G.; Supuran, C. T. *J. Am. Chem. Soc.* **2006**, *128*, 8329–8335.
- (14) Lin, F.; Wang, R. *J. Chem. Theory Comput.* **2010**, *6*, 1852–1870.
- (15) Pang, Y.-P. *J. Mol. Model.* **1999**, *5*, 196–202.
- (16) Studier, F.; Moffatt, B. *J. Mol. Biol.* **1986**, *189*, 113–130.
- (17) Nair, S. K.; Calderone, T. L.; Christianson, D. W.; Fierke, C. A. *J. Biol. Chem.* **1991**, *266*, 17320–17325.
- (18) Zheng, L.; Baumann, U.; Reymond, J.-L. *Nucleic Acids Res.* **2004**, *32*, e115.