Intramolecular hydrogen bond expectations in medicinal chemistry

Fabrizio Giordanetto[a], Christian Tyrchan[b], Johan Ulander[c]

[a]D.E. Shaw Research, 120W 45th Street, New York, NY 10036, USA

[b]Medicinal Chemistry, Respiratory and Inflammation, Innovative Medicines and Early Development Biotech Unit, Astrazeneca, Mölndal, Pepparedsleden 1, SE-431 83, Mölndal, Sweden

[c]Medicinal Chemistry, Cardiovascular and Metabolic Diseases, Innovative Medicines and Early Development Biotech Unit, Astrazeneca, Mölndal, Pepparedsleden 1, SE-431 83, Mölndal, Sweden

Supplementary Information

1. Materials and Methods

Molecular topologies with highest frequency of IMHB occurrence (>70%), as described by Kuhn *et al.*[1] (Table S1) and selected monoatomic transformations (Table S2) served to define the corresponding SMIRKS[2] for the sought matched molecular pairs (MMP). These definitions were employed to query a subset of ChEMBL[3] (i.e., compounds with molecular weight < 750 Da) using KNIME[4] and the matches and associated biological data returned for further analysis. MMPs were removed if any of the following conditions applied: 1) the transformed compound was still capable to form an IMHB due to structural symmetry or additional IMHB features in its topology, 2) the transformation resulted in a net charge change for the compound, 3) the associated biological data was not of the "binding" type as from ChEMBL definitions[3], was missing, labelled as inconclusive or undefined, originated from single concentration response experiments or high throughput screening assays. Numerical analysis and visualization of the resulting data set was performed using Dotmatics' Vortex[5].

2. Supplementary Tables

IMHB pair (pseudo ring)	Topology	% H-bond occurrence in CSD
C=O···H-N (6-membered)	cC3aC3a	93
C=O···H-N (6-membered)	aC3cC3a	89.5
C=O···H-N (6-membered)	cNaC3a	96
C=O···H-N (6-membered)	aNaC3c	93.5
C=O···H-N (6-membered)	aNaC3a	85.3
C=O···H-O (6-membered)	aC3cC3a	84.7
N···H-N (6-membered)	aNaC3a	80.9
C-O···H-N (5-membered)	aC4aC3a	74.5
C-O···H-N (5-membered)	aC3cC3a	71.2
C-O···H-N (6-membered)	aC3cC3aC3a	87.9
C-O···H-N (6-membered)	aC3cC3aC3c	70

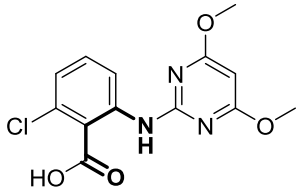
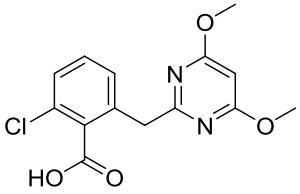
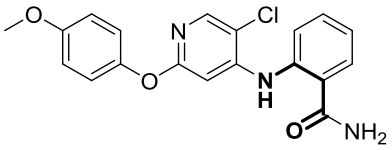
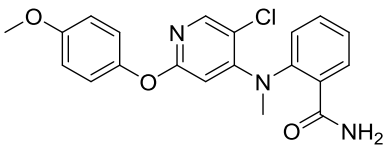
Table S1. Topologies with the highest occurrence of IMHB, based on Kuhn et al.[1]

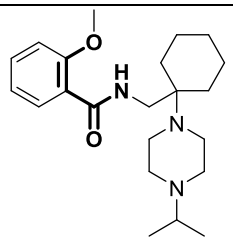
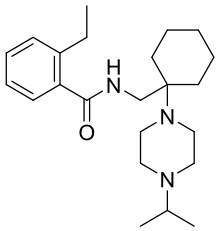
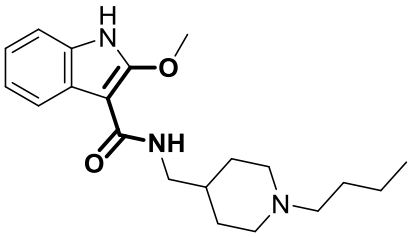
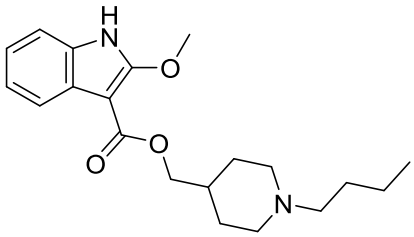
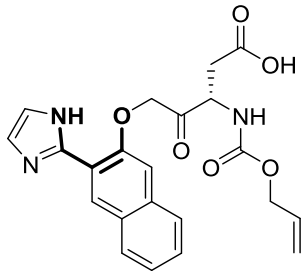
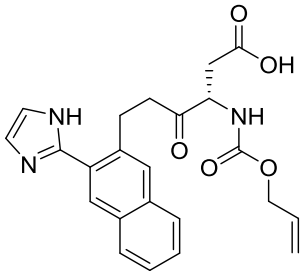
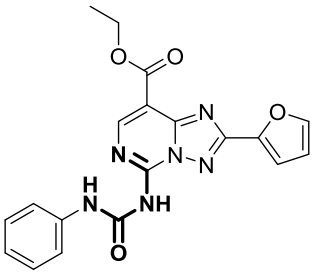
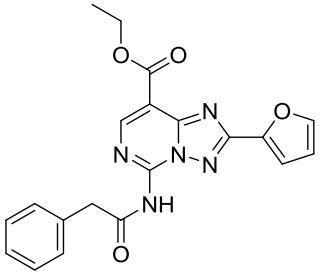
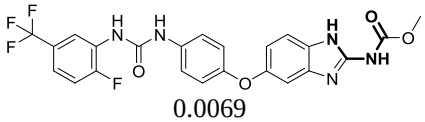
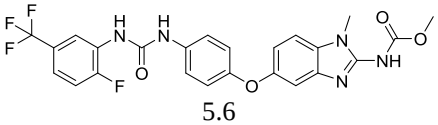
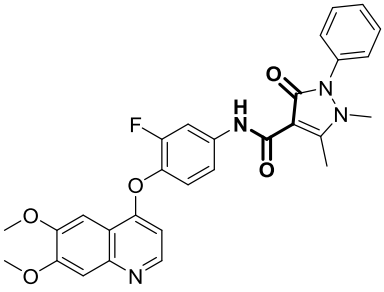
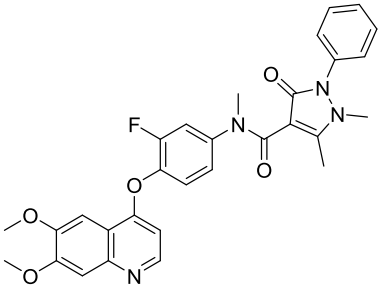
SMIRKS	Topology	Transform	IMHB Pair (HBA-HBD)	Pseudo Ring Size
[$([O]([#6X4])!@[#6X3]@[#6X3]!@[#7!H0]):2]>>[#6:2]	aC3cC3a	HBA-CH	Alkoxy-NH	5
[$([#7!H0]!@[#6X3]@[#6X3]!@[O][#6X4]):5]>>[#6:5]	aC3cC3a	HBD-CH	Alkoxy-NH	5

[$([#7H1]!@[#6X3]@[#6X3]!@[O][#6X4]):5]>>[#7:5][CH3]	aC3cC3a	HBD-Me	Alkoxy-NH	5
[$([#7H1]!@[#6X3]@[#6X3]!@[O][#6X4]):5]>>[#8:5]	aC3cC3a	HBD-O	Alkoxy-NH	5
[$([#7!H0]!@[#6X3]@[#6X3]!@[#6]=[O]):5]>>[#6:5]	aC3cC3a	HBD-CH	Carbonyl-NH	6
[$([#7H1]!@[#6X3]@[#6X3]!@[#6]=[O]):5]>>[#7:5][CH3]	aC3cC3a	HBD-Me	Carbonyl-NH	6
[$([#7H1]!@[#6X3]@[#6X3]!@[#6]=[O]):5]>>[#8:5]	aC3cC3a	HBD-O	Carbonyl-NH	6
[$([#8H1]!@[#6X3]@[#6X3]!@[#6]=[O]):5]>>[#6:5]	aC3cC3a	HBD-CH	Carbonyl-OH	6
[$([#8H1]!@[#6X3]@[#6X3]!@[#6]=[O]):5]>>[#8:5][CH3]	aC3cC3a	HBD-Me	Carbonyl-OH	6
[$([O]([#6X4])!@[#6X3]@[#6X3]!@[#6X3]!@[#7!H0]):2]>>[#6:2]	aC3cC3aC3a	HBA-CH	Alkoxy-NH	6
[$([#7!H0]!@[#6X3]!@[#6X3]@[#6X3]!@[O][#6X4]):6]>>[#6:6]	aC3cC3aC3a	HBD-CH	Alkoxy-NH	6
[$([#7H1]!@[#6X3]!@[#6X3]@[#6X3]!@[O][#6X4]):6]>>[#7:6][CH3]	aC3cC3aC3a	HBD-Me	Alkoxy-NH	6
[$([#7H1]!@[#6X3]!@[#6X3]@[#6X3]!@[O][#6X4]):6]>>[#8:6]	aC3cC3aC3a	HBD-O	Alkoxy-NH	6
[$([O]([#6X4])!@[#6X3]@[#6X3]!@[#6X3]@[#7!H0]):2]>>[#6:2]	aC3cC3aC3c	HBA-CH	Alkoxy-NH	6
[$([#7!H0]@[#6X3]!@[#6X3]@[#6X3]!@[O][#6X4]):6]>>[#6:6]	aC3cC3aC3c	HBD-CH	Alkoxy-NH	6
[$([#7H1]@[#6X3]!@[#6X3]@[#6X3]!@[O][#6X4]):6]>>[#7:6][CH3]	aC3cC3aC3c	HBD-Me	Alkoxy-NH	6
[$([#7H1]@[#6X3]!@[#6X3]@[#6X3]!@[O][#6X4]):6]>>[#8:6]	aC3cC3aC3c	HBD-O	Alkoxy-NH	6
[$([O]([#6X4])!@[#6X4]!@[#6X3]!@[#7!H0]):2]>>[#6:2]	aC4aC3a	HBA-CH	Alkoxy-NH	5
[$([#7!H0]!@[#6X3]!@[#6X4]!@[O][#6X4]):5]>>[#6:5]	aC4aC3a	HBD-CH	Alkoxy-NH	5
[$([#7H1]!@[#6X3]!@[#6X4]!@[O][#6X4]):5]>>[#7:5][CH3]	aC4aC3a	HBD-Me	Alkoxy-NH	5
[$([#7H1]!@[#6X3]!@[#6X4]!@[O][#6X4]):5]>>[#8:5]	aC4aC3a	HBD-O	Alkoxy-NH	5
[$([#7!H0]!@[#6X3]!@[#7X3]!@[#6]=[O]):5]>>[#6:5]	aNaC3a	HBD-CH	Carbonyl-NH	6
[$([#7H1]!@[#6X3]!@[#7X3]!@[#6]=[O]):5]>>[#7:5][CH3]	aNaC3a	HBD-Me	Carbonyl-NH	6
[$([#7H1]!@[#6X3]!@[#7X3]!@[#6]=[O]):5]>>[#8:5]	aNaC3a	HBD-O	Carbonyl-NH	6
[$([#7X2H0]@[#6X3]!@[#7X3]!@[#6X3]!@[#7!H0]):2]>>[#6:2]	aNaC3a	HBA-CH	Nhet-NH	6
[$([#7!H0]!@[#6X3]!@[#7X3]!@[#6X3]@[#7X2H0]):5]>>[#6:5]	aNaC3a	HBD-CH	Nhet-NH	6
[$([#7H1]!@[#6X3]!@[#7X3]!@[#6X3]@[#7X2H0]):5]>>[#7:5][CH3]	aNaC3a	HBD-Me	Nhet-NH	6
[$([#7H1]!@[#6X3]!@[#7X3]!@[#6X3]@[#7X2H0]):5]>>[#8:5]	aNaC3a	HBD-O	Nhet-NH	6
[$([#7!H0]@[#6X3]!@[#7X3]!@[#6]=[O]):5]>>[#6:5]	aNaC3c	HBD-CH	Carbonyl-NH	6

[$([#7H1]@[#6X3]!@[#7X3]!@[#6]=[O]):5]>>[#7:5][CH3]	aNaC3c	HBD-Me	Carbonyl-NH	6
[$([#7H1]@[#6X3]!@[#7X3]!@[#6]=[O]):5]>>[#8:5]	aNaC3c	HBD-O	Carbonyl-NH	6
[$([#7X2H0]@[#6X3]!@[#7X3]!@[#6X3]@[#7!H0]):2]>>[#6:2]	aNaC3c	HBA-CH	Nhet-NH	6
[$([#7!H0]@[#6X3]!@[#7X3]!@[#6X3]@[#7X2H0]):5]>>[#6:5]	aNaC3c	HBD-CH	Nhet-NH	6
[$([#7H1]@[#6X3]!@[#7X3]!@[#6X3]@[#7X2H0]):5]>>[#7:5][CH3]	aNaC3c	HBD-Me	Nhet-NH	6
[$([#7H1]@[#6X3]!@[#7X3]!@[#6X3]@[#7X2H0]):5]>>[#8:5]	aNaC3c	HBD-O	Nhet-NH	6
[$([#7!H0]!@[#6X3]!@[#6X3]@[#6]=[O]):5]>>[#6:5]	cC3aC3a	HBD-CH	Carbonyl-NH	6
[$([#7H1]!@[#6X3]!@[#6X3]@[#6]=[O]):5]>>[#7:5][CH3]	cC3aC3a	HBD-Me	Carbonyl-NH	6
[$([#7H1]!@[#6X3]!@[#6X3]@[#6]=[O]):5]>>[#8:5]	cC3aC3a	HBD-O	Carbonyl-NH	6
[$([#7!H0]!@[#6X3]!@[#7X3]@[#6]=[O]):5]>>[#6:5]	cNaC3a	HBD-CH	Carbonyl-NH	6
[$([#7H1]!@[#6X3]!@[#7X3]@[#6]=[O]):5]>>[#7:5][CH3]	cNaC3a	HBD-Me	Carbonyl-NH	6
[$([#7H1]!@[#6X3]!@[#7X3]@[#6]=[O]):5]>>[#8:5]	cNaC3a	HBD-O	Carbonyl-NH	6

Table S2. SMIRKS definition of chemical transformations used to identify IMHB MMPs.

Topology IMHB Pair Pseudo Ring Size Transform	**Bioactivity Type**	**IMHB Compound Bioactivity Value**	**"No IMHB" Compound Bioactivity Value**	**Δlog (Bioactivity)**
aC3cC3a Carbonyl-NH 6 HBD-CH	ALS IC_{50} (μM)[6]	 >100	 0.182	2.74
aC3cC3a Carbonyl-NH 6 HBD-Me	JNK1 IC_{50} (μM)[7]	 0.008	 >20	-3.39

aC3cC3aC3a Alkoxy-NH 6 HBA-CH	HCN1 IC_{50} (μM)[8]	 0.032	 >199	-3.80
aC3cC3aC3a Alkoxy-NH 6 HBD-O	5-HT4R IC_{50} (μM)[9]	 0.079	 0.0001	2.89
aC3cC3aC3c Alkoxy-NH 6 HBA-CH	CASP1 K_i (μM)[10]	 0.090	 >100	-3.04
aNaC3a Nhet-NH 6 HBD-CH	A2A AR K_i (μM)[11]	 0.0048	 47.2	-3.94
aNaC3c Carbonyl-NH 6 HBD-Me	TIE-2 IC_{50} (μM)[12]	 0.0069	 5.6	-2.90
cC3aC3a Carbonyl-NH 6 HBD-Me	c-Met K_i (μM)[13]	 0.001	 >20	-4.30

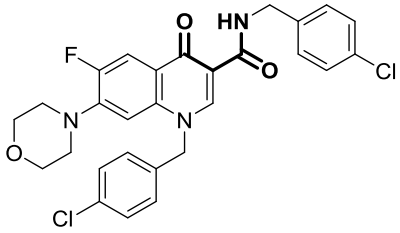
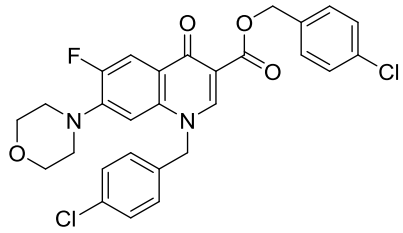
cC3aC3a Carbonyl-NH 6 HBD-O	NS5B IC_{50} (nM)[14]	 16	 0.02	2.90

Table S3. Representative IMHB MMPs outliers defined as Δlog(Bioactivity) > 1.5 * interquartile range. IMHB topologies and atomic pairs are highlighted in bold.

1. References

(1) Kuhn, B.; Mohr, P.; Stahl, M. Intramolecular Hydrogen Bonding in Medicinal Chemistry. *J. Med. Chem.* **2010**, *53* (6), 2601–2611.

(2) Leach, A. R.; Bradshaw, J.; Green, D. V.; Hann, M. M.; Delany, J. J. Implementation of a System for Reagent Selection and Library Enumeration, Profiling, and Design. *J. Chem. Inf. Comput. Sci.* **1999**, *39* (6), 1161–1172.

(3) Gaulton, A.; Bellis, L. J.; Bento, A. P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B.; Overington, J. P. ChEMBL: A Large-Scale Bioactivity Database for Drug Discovery. *Nucleic Acids Res.* **2011**, gkr777.

(4) Berthold, M. R.; Cebron, N.; Dill, F.; Gabriel, T. R.; Kötter, T.; Meinl, T.; Ohl, P.; Sieb, C.; Thiel, K.; Wiswedel, B. KNIME: The Konstanz Information Miner. In *Data Analysis, Machine Learning and Applications*; Preisach, C., Burkhardt, P. D. H., Schmidt-Thieme, P. D. L., Decker, P. D. R., Eds.; Studies in Classification, Data Analysis, and Knowledge Organization; Springer Berlin Heidelberg, 2008; pp 319–326.

(5) *Vortex, 2016.08; Dotmatics Limited, The Old Monastery, Windhill, Bishops Stortford, Herts CM23 2ND UK, 2016.*

(6) Nezu, Y.; Wada, N.; Saitoh, Y.; Takahashi, S.; Miyazawa, T. Synthesis and Herbicidal Activity of Pyrimidinyl Salicylic and Thiosalicylic Acids. *J. Pestic. Sci.* **1996**, *21* (3), 293–303.

(7) Song, X.; Chen, W.; Lin, L.; Ruiz, C. H.; Cameron, M. D.; Duckett, D. R.; Kamenecka, T. M. Synthesis and SAR of 2-Phenoxypyridines as Novel c-Jun N-Terminal Kinase Inhibitors. *Bioorg. Med. Chem. Lett.* **2011**, *21* (23), 7072–7075.

(8) McClure, K. J.; Maher, M.; Wu, N.; Chaplan, S. R.; Eckert III, W. A.; Lee, D. H.; Wickenden, A. D.; Hermann, M.; Allison, B.; Hawryluk, N.; Breitenbucher, J. G.; Grice, C. A. Discovery of a Novel Series of Selective HCN1 Blockers. *Bioorg. Med. Chem. Lett.* **2011**, *21* (18), 5197–5201.

(9) Gaster, L. M.; Joiner, G. F.; King, F. D.; Wyman, P. A.; Sutton, J. M.; Bingham, S.; Ellis, E. S.; Sanger, G. J.; Wardle, K. A. N-[(1-Butyl-4-Piperidinyl)methyl]-3,4- Dihydro-2H-[1,3]oxazino[3,2-A]indole- 10-Carboxamide Hydrochloride: The First Potent and Selective 5-HT4 Receptor Antagonist Amide with Oral Activity. *J. Med. Chem.* **1995**, *38* (24), 4760–4763.

(10) Mjalli, A. M. M.; Zhao, J. J.; Chapman, K. T.; Thornberry, N. A.; Peterson, E. P.; MacCoss, M.; Hagmann, W. K. Inhibition of Interleukin-1β Converting Enzyme by N-Acyl-Aspartyl Aryloxymethyl Ketones. *Bioorg. Med. Chem. Lett.* **1995**, *5* (13), 1409–1414.

(11) Federico, S.; Ciancetta, A.; Porta, N.; Redenti, S.; Pastorin, G.; Cacciari, B.; Klotz, K. N.; Moro, S.; Spalluto, G. Scaffold Decoration at Positions 5 and 8 of 1,2,4-Triazolo[1,5-c]Pyrimidines to Explore the Antagonist Profiling on Adenosine Receptors: A Preliminary Structure–Activity Relationship Study. *J. Med. Chem.* **2014**, *57* (14), 6210–6225.

(12) Hasegawa, M.; Nishigaki, N.; Washio, Y.; Kano, K.; Harris, P. A.; Sato, H.; Mori, I.; West, R. I.; Shibahara, M.; Toyoda, H.; Wang, L.; Nolte, R. T.; Veal, J. M.; Cheung, M. Discovery of Novel Benzimidazoles as Potent Inhibitors of TIE-2 and VEGFR-2 Tyrosine Kinase Receptors. *J. Med. Chem.* **2007**, *50* (18), 4453–4470.

(13) Norman, M. H.; Liu, L.; Lee, M.; Xi, N.; Fellows, I.; D'Angelo, N. D.; Dominguez, C.; Rex, K.; Bellon, S. F.; Kim, T.-S.; Dussault, I. Structure-Based Design of Novel Class II c-Met Inhibitors: 1. Identification of Pyrazolone-Based Derivatives. *J. Med. Chem.* **2012**, *55* (5), 1858–1867.

(14) Kumar, D. V.; Rai, R.; Brameld, K. A.; Somoza, J. R.; Rajagopalan, R.; Janc, J. W.; Xia, Y. M.; Ton, T. L.; Shaghafi, M. B.; Hu, H.; Lehoux, I.; To, N.; Young, W. B.; Green, M. J. Quinolones as HCV NS5B Polymerase Inhibitors. *Bioorg. Med. Chem. Lett.* **2011**, *21* (1), 82–87.