

# Postprocessing of Protein–Ligand Docking Poses Using Linear Response MM-PB/SA: Application to Wee1 Kinase Inhibitors

Kanin Wichapong,<sup>†,‡</sup> Michael Lawson,<sup>‡</sup> Somsak Pianwanit,<sup>†</sup> Sirirat Kokpol,<sup>†</sup> and Wolfgang Sippel<sup>\*,‡</sup>

Department of Chemistry, Faculty of Science, Chulalongkorn University, Bangkok, 10330, Thailand and  
Department of Pharmaceutical Chemistry, Martin-Luther-University Halle-Wittenberg,  
06120, Halle(Saale), Germany

Received May 31, 2010

Prediction of the binding strength of untested ligands is a central issue in structure-based drug design. In order to rapidly screen large compound databases, simple scoring schemes are often used in target-based virtual screening. The resulting scores often correlate poorly with biological affinities. More rigorous scoring methods, such as MM-PB/SA, correlate better with biological data by considering solvation effects and protein flexibility in the calculation of the binding free energy of a ligand. Here we describe the performance of a modified MM-PB/SA method on 222 Wee1 kinase inhibitors (48 pyridopyrimidine and 174 pyrrolocarbazole derivatives). Docking of these inhibitors into the available Wee1 kinase crystal structure yielded a consistent binding mode, and the derived MM-PB/SA models showed a significant correlation between calculated and experimental data ( $r^2$  values between 0.64 and 0.67). Further study of these models on external test sets of Wee1 kinase inhibitors and structurally related decoys showed that a model based on a single kinase–inhibitor conformation can discriminate the active inhibitors from decoys. We also tested whether the linear interaction energy method with continuum electrostatics (LIECE) yields comparable results to MM-PB/SA and whether the LIECE and MM-PB/SA models can be applied for virtual screening of compound libraries.

## 1. INTRODUCTION

Docking-based approaches are commonly implemented in prediction of the binding mode of a novel lead compound to a target protein. Although current docking-based approaches are usually able to reproduce the experimentally observed binding mode, accurate prediction of the binding affinity of these ligands is still challenging.<sup>1</sup> Simple scoring functions are fast but often not accurate enough to discriminate actives from weak actives or inactives in cases where protein flexibility and solvation effects must be considered.<sup>1</sup> More complex methods such as MM-PB/SA, MM-GB/SA, TI, and FEP account for protein flexibility and solvation effects but are also computationally demanding. Linear response models provide an intermediate method between the extremes of static and dynamic methods. One implementation of a linear response model is the linear interaction energy (LIE) method developed by Aqvist et al.,<sup>2–4</sup> in which binding free energies are calculated by averaging interaction energies between the protein and the ligand from MD simulations. The approach estimates the binding free energy of protein–ligand complexes by using the following equation

$$\Delta G = \alpha(\langle E^{\text{vdW}} \rangle_{\text{bound}} - \langle E^{\text{vdW}} \rangle_{\text{free}}) + \beta(\langle E^{\text{ele}} \rangle_{\text{bound}} - \langle E^{\text{ele}} \rangle_{\text{free}}) + \gamma \quad (1)$$

where  $\alpha$  and  $\beta$  are coefficients derived from empirical fitting to experimental data (and vary widely depending on the force

field and computational methods<sup>5,6</sup>),  $\gamma$  represents the absolute binding free energy,<sup>3</sup> and  $E^{\text{vdW}}$  and  $E^{\text{ele}}$  represent van der Waals and electrostatic interaction energies between the ligand and its environment. LIE requires several protein–ligand snapshots, which are obtained by MD<sup>4</sup> or Monte Carlo<sup>7</sup> simulations. The application of LIE for different protein targets has shown that it outperforms classical docking scores and can be used for prediction of novel compounds from classical lead series.<sup>5</sup> The method is limited by the fact that two computationally expensive MD systems have to be carried out (unbound ligand in solvent and solvated protein–ligand complex) for each ligand<sup>5</sup> and is thus only useful for calculating up to approximately 100 compounds.

In the case where a single ligand scaffold is being studied and protein flexibility can be ignored, the computationally expensive MD simulations of LIE can be replaced with energy minimizations of protein–ligand complexes. This method, which is known as linear interaction energy with continuum electrostatics (LIECE),<sup>8</sup> is fast enough to be suitable for virtual screening. LIECE calculates binding free energy using eqs 2–4.

A one-parameter model

$$\Delta G = \alpha \Delta E^{\text{vdW}} \quad (2)$$

A two-parameter model with continuum electrostatics

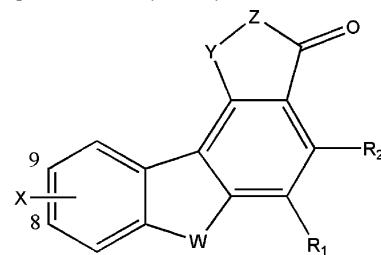
$$\Delta G = \alpha \Delta E^{\text{vdW}} + \beta \Delta E^{\text{ele}} \quad (3)$$

A three-parameter model with decomposed electrostatics

\* Corresponding author phone: +49-345-55-25040; fax: +49-345-55-27355; e-mail: wolfgang.sippel@pharmazie.uni-halle.de.

<sup>†</sup> Chulalongkorn University.

<sup>‡</sup> Martin-Luther-University Halle-Wittenberg.

**Chart 1.** Molecular Structures of the Pyrrolo[3,4-*c*]carbazole-1,3(2*H*,6*H*)-dione Derivatives Used in This Study

compd	W	X	Y	Z	R <sub>1</sub>	R <sub>2</sub>	activity		
							IC <sub>50</sub> ( $\mu$ M)	pIC <sub>50</sub>	$\Delta G_{\text{exp}}$
training set (144 compounds)									
3	NH	9-OH	CO	NH	H	I	2.3	5.63	-7.74
4	NH	8-OH	CO	NH	H	Ph	0.31	6.51	-8.93
5	O	9-OH	CO	NH	H	Ph	0.43	6.37	-8.74
6	S	9-OH	CO	NH	H	Ph	0.078	7.11	-9.75
8	NH	9-OH	CO	NH	Me	Ph	0.13	6.89	-9.45
9	NH	9-OH	CO	NH	Et	Ph	1.6	5.79	-7.95
10	NH	9-OH	CO	NH	Ph	Me	9.7	5.01	-6.88
11	NH	9-OH	CO	NH	Ph	Ph	2.3	5.63	-7.74
12	NH	9-OH	CO	NH	Ph	H	4.0	5.39	-7.41
13	NH	9-OH	CH <sub>2</sub>	NH	H	Ph	37	4.43	-6.08
14	NH	9-OH	CO	N-NH <sub>2</sub>	H	Ph	3.9	5.40	-7.42
15	NH	9-OH	CO	NH	H	2-ClPh	0.011	7.96	-10.92
16	NH	9-OMe	CO	NH	H	2-ClPh	0.64	6.19	-8.50
17	NMe	9-OH	CO	NH	H	2-ClPh	0.057	7.24	-9.94
19	NH	9-OH	CO	NH	H	2-FPh	0.33	6.48	-8.89
21	NH	9-OH	CO	NH	H	2-IPh	0.013	7.89	-10.82
22	NH	9-OH	CO	NH	H	2-MePh	0.15	6.82	-9.36
23	NH	9-OH	CO	NH	H	2-EtPh	0.51	6.29	-8.64
24	NH	9-OH	CO	NH	H	2-CF <sub>3</sub> Ph	0.58	6.24	-8.56
25	NH	9-OH	CO	NH	H	2-CH <sub>2</sub> OHPh	0.45	6.35	-8.71
26	NH	9-OH	CO	NH	H	2-CNPh	0.19	6.72	-9.22
27	NH	9-OH	CO	NH	H	2-COMePh	0.83	6.08	-8.35
28	NH	9-OH	CO	NH	H	2-Ph-Ph	0.57	6.24	-8.57
31	NH	9-OH	CO	NH	H	2-OEtPh	0.26	6.59	-9.04
32	NH	9-OH	CO	NH	H	2-SMePh	0.033	7.48	-10.27
33	NH	9-OH	CO	NH	H	2-SOMePh	0.22	6.66	-9.14
34	NH	9-OH	CO	NH	H	2-NO <sub>2</sub> Ph	0.047	7.33	-10.06
35	NH	9-OH	CO	NH	H	2-NH <sub>2</sub> Ph	0.21	6.68	-9.16
36	NH	9-OH	CO	NH	H	3-FPh	0.22	6.66	-9.14
37	NH	9-OH	CO	NH	H	3-ClPh	0.055	7.26	-9.96
38	NH	9-OH	CO	NH	H	3-MePh	0.23	6.64	-9.11
39	NH	9-OH	CO	NH	H	3-CH <sub>2</sub> OHPh	0.87	6.06	-8.32
40	NH	9-OH	CO	NH	H	3-CH <sub>2</sub> NH <sub>2</sub> Ph	4.4	5.35	-7.35
41	NH	9-OH	CO	NH	H	3-CNPh	0.18	6.74	-9.26
42	NH	9-OH	CO	NH	H	3-COMePh	4.3	5.36	-7.36
45	NH	9-OH	CO	NH	H	3-OMePh	0.62	6.21	-8.52
46	NH	9-OH	CO	NH	H	3-NO <sub>2</sub> Ph	0.30	6.52	-8.95
47	NH	9-OH	CO	NH	H	3-NH <sub>2</sub> Ph	0.070	7.15	-9.82
48	NH	9-OH	CO	NH	H	4-FPh	16	4.80	-6.58
49	NH	9-OH	CO	NH	H	4-ClPh	0.73	6.14	-8.42
52	NH	9-OH	CO	NH	H	4-CNPh	1.8	5.74	-7.88
53	NH	9-OH	CO	NH	H	4-COMePh	3.6	5.44	-7.47
54	NH	9-OH	CO	NH	H	4-OHPh	0.067	7.17	-9.85
55	NH	9-OH	CO	NH	H	4-OMePh	12	4.90	-6.75
57	NH	9-OH	CO	NH	H	4-SO <sub>2</sub> MePh	1.1	5.95	-8.18
58	NH	9-OH	CO	NH	H	4-NH <sub>2</sub> Ph	0.15	6.82	-9.36
59	NH	9-OH	CO	NH	H	2-Cl, 3-ClPh	0.028	7.55	-10.37
60	NH	9-OH	CO	NH	H	2-Cl, 3-OHPh	0.012	7.55	-10.87
61	NH	9-OH	CO	NH	H	2-Cl, 3-NH <sub>2</sub> Ph	0.021	7.92	-10.54
62	NH	9-OH	CO	NH	H	2-Cl, 4-OHPh	0.023	7.64	-10.48
63	NH	9-OH	CO	NH	H	4-Cl, 3-NH <sub>2</sub> Ph	0.024	7.62	-10.46
64	NH	9-OH	CO	NH	H	2-Cl, 5-ClPh	0.49	6.31	-8.66
66	NH	9-OH	CO	NH	H	2-Cl, 5-NH <sub>2</sub> Ph	0.020	7.70	-10.57
67	NH	9-OH	CO	NH	H	2-Cl, 6-ClPh	0.028	7.55	-10.37
68	NH	9-OH	CO	NH	H	2-Cl, 6-OHPh	0.045	7.35	-10.08
69	NH	9-OH	CO	NH	H	2-Cl, 6-OMePh	0.015	7.82	-10.74
70	NH	9-OH	CO	NH	H	2-Br, 4-NH <sub>2</sub> Ph	0.020	7.70	-10.57

Chart 1. Continued

compd	W	X	Y	Z	R <sub>1</sub>	R <sub>2</sub>	activity		
							IC <sub>50</sub> ( $\mu$ M)	pIC <sub>50</sub>	$\Delta G_{\text{exp}}$
71	NH	9-OH	CO	NH	H	2-Br, 6-BrPh	0.035	7.46	-10.23
72	NH	9-OH	CO	NH	H	2-Me, 3-MePh	0.27	6.57	-9.01
73	NH	9-OH	CO	NH	H	2-Me, 5-MePh	0.96	6.02	-8.26
74	NH	9-OH	CO	NH	H	2-Me, 6-MePh	0.075	7.12	-9.78
75	NH	9-OH	CO	NH	H	2-OMe, 4-NH <sub>2</sub> Ph	0.019	7.72	-10.60
76	NH	9-OH	CO	NH	H	2-OMe, 5-NH <sub>2</sub> Ph	0.11	6.96	-9.55
77	NH	9-OH	CO	NH	H	2-OMe, 6-OmePh	0.027	7.57	-10.39
78	NH	9-OH	CO	NH	H	2-OMe, 6-FPh	0.029	7.54	-10.34
79	NH	9-OH	CO	NH	H	2,6-diCl, 3-OHPh	0.018	7.74	-10.63
81	NH	9-OH	CO	NH	H	2-thienyl	0.14	6.85	-9.41
82	NH	9-OH	CO	NH	H	3-thienyl	0.042	7.38	-10.12
83	NH	9-OH	CO	NH	H	2-pyrrolyl	0.18	6.74	-9.26
84	NH	9-OH	CO	NH	H	3-pyrrolyl	0.038	7.42	-10.18
87	NEt	9-OH	CO	NH	H	2-ClPh	0.050	7.30	-10.02
88	N-n-Pr	9-OH	CO	NH	H	2-ClPh	0.063	7.20	-9.88
89	N-i-Pr	9-OH	CO	NH	H	2-ClPh	0.053	7.28	-9.98
90	N-n-Bu	9-OH	CO	NH	H	2-ClPh	0.059	7.23	-9.92
91	N(CH <sub>2</sub> ) <sub>2</sub> i-Pr	9-OH	CO	NH	H	2-ClPh	0.15	6.82	-9.36
92	N-n-pentyl	9-OH	CO	NH	H	2-ClPh	0.17	6.77	-9.29
93	NH	9-OH	CO	NH	H		0.097	7.01	-9.62
94	NH	9-OH	CO	NH	H	2,6-diClPh	0.028	7.55	-10.37
95	NMe	9-OH	CO	NH	H		0.14	6.85	-9.41
96	NMe	9-OH	CO	NH	H	2-ClPh	0.057	7.24	-9.94
97	N-n-Bu	9-OH	CO	NH	H	2-ClPh	0.059	7.23	-9.92
98	N(CH <sub>2</sub> ) <sub>2</sub> OH	9-OH	CO	NH	H		0.025	7.60	-10.43
99	N(CH <sub>2</sub> ) <sub>2</sub> OH	9-OH	CO	NH	H	2-ClPh	0.045	7.35	-10.08
101	N(CH <sub>2</sub> ) <sub>3</sub> OH	9-OH	CO	NH	H		0.2	6.70	-9.19
102	N(CH <sub>2</sub> ) <sub>3</sub> OH	9-OH	CO	NH	H	2-ClPh	0.009	8.05	-11.04
103	N(CH <sub>2</sub> ) <sub>3</sub> OH	9-OH	CO	NH	H	2,6-diClPh	0.007	8.15	-11.19
104	N(CH <sub>2</sub> ) <sub>3</sub> OH	9-OH	CO	NH	H	2-OMePh	0.03	7.52	-10.32
105	N(CH <sub>2</sub> ) <sub>2</sub> CONH <sub>2</sub>	9-OH	CO	NH	H		0.021	7.68	-10.54
106	N(CH <sub>2</sub> ) <sub>2</sub> CONH <sub>2</sub>	9-OH	CO	NH	H	2-ClPh	0.006	8.22	-11.23
107	N(CH <sub>2</sub> ) <sub>2</sub> CONH <sub>2</sub>	9-OH	CO	NH	H	2,6-diClPh	0.33	6.48	-8.89
108	N(CH <sub>2</sub> ) <sub>2</sub> CN	9-OH	CO	NH	H	2-ClPh	0.015	7.82	-10.74
109	N(CH <sub>2</sub> ) <sub>2</sub> COOME	9-OH	CO	NH	H	2-ClPh	0.03	7.52	-10.32
110	N(CH <sub>2</sub> ) <sub>3</sub> CN	9-OH	CO	NH	H	2-ClPh	0.033	7.48	-10.27
112	N(CH <sub>2</sub> ) <sub>2</sub> NMe <sub>2</sub>	9-OH	CO	NH	H		0.2	6.70	-9.19
113	N(CH <sub>2</sub> ) <sub>2</sub> NMe <sub>2</sub>	9-OH	CO	NH	H	2-ClPh	0.096	7.02	-9.63
114	N(CH <sub>2</sub> ) <sub>2</sub> NMe <sub>2</sub>	9-OH	CO	NH	H	2,6-diClPh	0.17	6.77	-9.29
116	N(CH <sub>2</sub> ) <sub>2</sub> Nmorpholide	9-OH	CO	NH	H	2-ClPh	0.064	7.19	-9.87
117	N(CH <sub>2</sub> ) <sub>2</sub> Nmorpholide	9-OH	CO	NH	H	2,6-diClPh	0.11	6.96	-9.55
118	N(CH <sub>2</sub> ) <sub>2</sub> Nimidazolide	9-OH	CO	NH	H		0.23	6.64	-9.11
119	N(CH <sub>2</sub> ) <sub>2</sub> Nimidazolide	9-OH	CO	NH	H	2-ClPh	0.092	7.04	-9.66
120	N(CH <sub>2</sub> ) <sub>2</sub> Nimidazolide	9-OH	CO	NH	H	2,6-diClPh	0.12	6.92	-9.50
121	N(CH <sub>2</sub> ) <sub>3</sub> NHMe	9-OH	CO	NH	H		0.28	6.55	-8.99
122	N(CH <sub>2</sub> ) <sub>3</sub> NHMe	9-OH	CO	NH	H	2-ClPh	0.069	7.16	-9.83
123	N(CH <sub>2</sub> ) <sub>3</sub> NHMe	9-OH	CO	NH	H	2,6-diClPh	0.11	6.96	-9.55
125	N(CH <sub>2</sub> ) <sub>3</sub> NMe <sub>2</sub>	9-OH	CO	NH	H	2-ClPh	0.1	7.00	-9.61
126	N(CH <sub>2</sub> ) <sub>3</sub> NMe <sub>2</sub>	9-OH	CO	NH	H	2,6-diClPh	0.14	6.85	-9.41
127	N(CH <sub>2</sub> ) <sub>3</sub> Nmorpholide	9-OH	CO	NH	H		0.29	6.54	-8.97
128	N(CH <sub>2</sub> ) <sub>3</sub> Nmorpholide	9-OH	CO	NH	H	2-ClPh	0.071	7.15	-9.81
129	N(CH <sub>2</sub> ) <sub>3</sub> Nmorpholide	9-OH	CO	NH	H	2,6-diClPh	0.064	7.19	-9.87
132	N(CH <sub>2</sub> ) <sub>3</sub> Nimidazolide	9-OH	CO	NH	H	2,6-diClPh	0.059	7.23	-9.92
133	N(CH <sub>2</sub> ) <sub>3</sub> N(4-Mepiperazine)	9-OH	CO	NH	H		0.3	6.52	-8.95
134	N(CH <sub>2</sub> ) <sub>3</sub> N(4-Mepiperazine)	9-OH	CO	NH	H	2-ClPh	0.082	7.09	-9.72
135	N(CH <sub>2</sub> ) <sub>3</sub> N(4-Mepiperazine)	9-OH	CO	NH	H	2,6-diClPh	0.062	7.21	-9.89
136	N(CH <sub>2</sub> ) <sub>3</sub> NHPh	9-OH	CO	NH	H		0.093	7.03	-9.65
139	N(CH <sub>2</sub> ) <sub>2</sub> CONH(CH <sub>2</sub> ) <sub>2</sub> NMe <sub>2</sub>	9-OH	CO	NH	H		0.17	6.77	-9.29
140	N(CH <sub>2</sub> ) <sub>2</sub> CONH(CH <sub>2</sub> ) <sub>2</sub> NMe <sub>2</sub>	9-OH	CO	NH	H	2-ClPh	0.035	7.46	-10.23
141	N(CH <sub>2</sub> ) <sub>2</sub> CONH(CH <sub>2</sub> ) <sub>2</sub> NMe <sub>2</sub>	9-OH	CO	NH	H	2,6-diClPh	0.014	7.85	-10.78
142	N(CH <sub>2</sub> ) <sub>2</sub> COOH	9-OH	CO	NH	H		0.023	7.64	-10.48
143	N(CH <sub>2</sub> ) <sub>2</sub> COOH	9-OH	CO	NH	H	2-ClPh	0.009	8.05	-11.04
144	N(CH <sub>2</sub> ) <sub>2</sub> COOH	9-OH	CO	NH	H	2,6-diClPh	0.39	6.41	-8.80
145	N(CH <sub>2</sub> ) <sub>2</sub> CONHSO <sub>2</sub> Me	9-OH	CO	NH	H	2-ClPh	0.012	7.92	-10.87
146	N(CH <sub>2</sub> ) <sub>2</sub> CONHSO <sub>2</sub> Ph	9-OH	CO	NH	H	2-ClPh	0.007	8.15	-11.19
147	N(CH <sub>2</sub> ) <sub>2</sub> C-tetrazole	9-OH	CO	NH	H	2-ClPh	0.021	7.68	-10.54
148	N(CH <sub>2</sub> ) <sub>2</sub> S-triazole	9-OH	CO	NH	H	2-ClPh	0.024	7.62	-10.46
149	N(CH <sub>2</sub> ) <sub>2</sub> SO-triazole	9-OH	CO	NH	H	2-ClPh	0.009	8.05	-11.04
152	N(CH <sub>2</sub> ) <sub>3</sub> C-tetrazole	9-OH	CO	NH	H	2-ClPh	0.016	7.80	-10.70
153	NH	8-(CH <sub>2</sub> ) <sub>4</sub> -NMe <sub>2</sub> , 9-OH	CO	NH	H	2-ClPh	0.049	7.31	-10.03

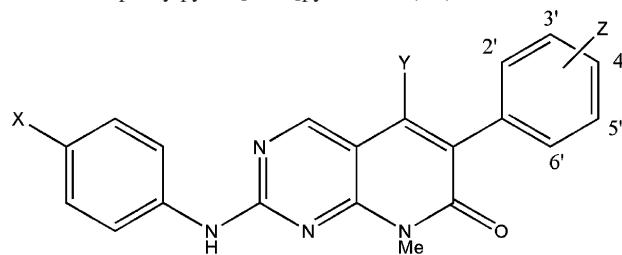
Chart 1. Continued

compd	W	X	Y	Z	R <sub>1</sub>	R <sub>2</sub>	activity		
							IC <sub>50</sub> (μM)	pIC <sub>50</sub>	ΔG <sub>exp</sub>
154	NH	8-(CH <sub>2</sub> ) <sub>4</sub> -Npyrrol, 9-OH	CO	NH	H	2-CIPh	0.05	7.30	-10.02
155	NH	8-(CH <sub>2</sub> ) <sub>4</sub> -Nmorph, 9-OH	CO	NH	H	2-CIPh	0.037	7.43	-10.20
156	NMe	8-(CH <sub>2</sub> ) <sub>4</sub> -NMe <sub>2</sub> , 9-OH	CO	NH	H	2-CIPh	0.034	7.47	-10.25
157	NMe	8-(CH <sub>2</sub> ) <sub>4</sub> -Npyrrol, 9-OH	CO	NH	H	2-CIPh	0.036	7.44	-10.22
159	N(CH <sub>2</sub> ) <sub>2</sub> OH	8-(CH <sub>2</sub> ) <sub>4</sub> -Npyrrol, 9-OH	CO	NH	H	2-CIPh	0.024	7.62	-10.46
160	N(CH <sub>2</sub> ) <sub>2</sub> OH	8-(CH <sub>2</sub> ) <sub>4</sub> -Nmorph, 9-OH	CO	NH	H	2-CIPh	0.019	7.72	-10.60
161	NH	8-O(CH <sub>2</sub> ) <sub>3</sub> -NMe <sub>2</sub> , 9-OH	CO	NH	H	2-CIPh	0.026	7.59	-10.41
163	NH	8-O(CH <sub>2</sub> ) <sub>3</sub> -Nmorph, 9-OH	CO	NH	H	2-CIPh	0.026	7.59	-10.41
164	NMe	8-O(CH <sub>2</sub> ) <sub>3</sub> -NMe <sub>2</sub> , 9-OH	CO	NH	H	2-CIPh	0.058	7.24	-9.93
165	NMe	8-O(CH <sub>2</sub> ) <sub>3</sub> -Npyrrol, 9-OH	CO	NH	H	2-CIPh	0.075	7.12	-9.78
167	N(CH <sub>2</sub> ) <sub>2</sub> OH	8-O(CH <sub>2</sub> ) <sub>3</sub> -NMe, 9-OH <sub>2</sub>	CO	NH	H	2-CIPh	0.018	7.74	-10.63
168	N(CH <sub>2</sub> ) <sub>2</sub> OH	8-O(CH <sub>2</sub> ) <sub>3</sub> -Npyrrol, 9-OH	CO	NH	H	2-CIPh	0.024	7.62	-10.46
169	N(CH <sub>2</sub> ) <sub>2</sub> OH	8-O(CH <sub>2</sub> ) <sub>3</sub> -Nmorph, 9-OH	CO	NH	H	2-CIPh	0.015	7.82	-10.74
170	NMe	8-S(CH <sub>2</sub> ) <sub>3</sub> -Npyrro, 9-OH	CO	NH	H	2-CIPh	0.02	7.70	-10.57
171	NMe	8-SO(CH <sub>2</sub> ) <sub>3</sub> -Npyrro, 9-OH	CO	NH	H	2-CIPh	0.033	7.48	-10.27
173	NMe	8-SO <sub>2</sub> NH(CH <sub>2</sub> ) <sub>2</sub> -Npyrro, 9-OH	CO	NH	H	2-CIPh	0.046	7.34	-10.07
174	NMe	8-CONH(CH <sub>2</sub> ) <sub>2</sub> -Npyrro, 9-OH	CO	NH	H	2-CIPh	0.015	7.82	-10.74
test set (30 compounds)									
1	NH	9-OH	CO	NH	H	Ph	0.097	7.01	-9.62
2	NH	9-OH	CO	NH	H	H	4.0	5.39	-7.41
7	NMe	9-OH	CO	NH	H	Ph	0.26	6.59	-9.04
18	O	9-OH	CO	NH	H	2-CIPh	0.033	7.48	-10.27
20	NH	9-OH	CO	NH	H	2-BrPh	0.023	7.64	-10.48
29	NH	9-OH	CO	NH	H	2-OHPh	0.060	7.22	-9.91
30	NH	9-OH	CO	NH	H	2-OMePh	0.024	7.62	-10.46
43	NH	9-OH	CO	NH	H	4-biphenyl	40	4.30	-6.04
44	NH	9-OH	CO	NH	H	3-OHPh	0.089	7.05	-9.68
50	NH	9-OH	CO	NH	H	4-MePh	3.3	5.48	-7.52
51	NH	9-OH	CO	NH	H	4-CH <sub>2</sub> OHPh	1.2	5.92	-8.13
56	NH	9-OH	CO	NH	H	4-SMe	29	4.50	-6.23
65	NH	9-OH	CO	NH	H	2-Cl, 5-OHPh	0.042	7.38	-10.12
80	NH	9-OH	CO	NH	H	2,6-diCl, 4-OHPh	0.049	7.31	-10.03
85	NH	9-OH	CO	NH	H	4-pyridyl	0.82	6.09	-8.35
86	NH	9-OH	CO	NH	H	3-pyridyl	0.58	6.24	-8.56
100	N(CH <sub>2</sub> ) <sub>2</sub> OH	9-OH	CO	NH	H	2,6-diClPh	0.008	8.10	-11.11
111	N(CH <sub>2</sub> ) <sub>3</sub> OMe	9-OH	CO	NH	H	2-CIPh	0.027	7.57	-10.39
115	N(CH <sub>2</sub> ) <sub>2</sub> Nmorpholide	9-OH	CO	NH	H	H	0.14	6.85	-9.41
124	N(CH <sub>2</sub> ) <sub>2</sub> NMe <sub>2</sub>	9-OH	CO	NH	H	H	0.36	6.44	-8.84
130	N(CH <sub>2</sub> ) <sub>2</sub> Nimidazolide	9-OH	CO	NH	H	H	0.11	6.96	-9.55
131	N(CH <sub>2</sub> ) <sub>2</sub> Nimidazolide	9-OH	CO	NH	H	2-CIPh	0.054	7.27	-9.97
137	N(CH <sub>2</sub> ) <sub>2</sub> NHPh	9-OH	CO	NH	H	2-CIPh	0.074	7.13	-9.79
138	N(CH <sub>2</sub> ) <sub>2</sub> NHPh	9-OH	CO	NH	H	2,6-diClPh	0.067	7.17	-9.85
150	N(CH <sub>2</sub> ) <sub>2</sub> SO <sub>2</sub> triazole	9-OH	CO	NH	H	2-CIPh	0.019	7.72	-10.60
151	N(CH <sub>2</sub> ) <sub>2</sub> COOH	9-OH	CO	NH	H	2-CIPh	0.013	7.89	-10.82
158	NMe	8-(CH <sub>2</sub> ) <sub>4</sub> -N-morph, 9-OH	CO	NH	H	2-CIPh	0.03	7.52	-10.32
162	NH	8-O(CH <sub>2</sub> ) <sub>3</sub> -N-pyrrol, 9-OH	CO	NH	H	2-CIPh	0.036	7.44	-10.22
166	NMe	8-O(CH <sub>2</sub> ) <sub>3</sub> -N-morph, 9-OH	CO	NH	H	2-CIPh	0.057	7.24	-9.94
172	NMe	8-SO <sub>2</sub> (CH <sub>2</sub> ) <sub>3</sub> -N-pyrrol, 9-OH	CO	NH	H	2-CIPh	0.16	6.80	-9.33

$$\Delta G = \alpha \Delta E^{\text{vdW}} + \beta_1 \Delta E^{\text{ele}} + \beta_2 \Delta G^{\text{sol}} \quad (4)$$

where  $E^{\text{vdW}}$  and  $E^{\text{ele}}$  correspond to van der Waals and electrostatic interactions,  $\Delta G^{\text{sol}}$  is the free energy of solvation, and  $\Delta G^{\text{ele}}$  is the sum of the electrostatic interaction between the ligand and the protein in the gas phase and the change in the solvation energy upon binding. A rigorous treatment of the solvation term within the continuum electrostatics approximation (such as the numerical solution of the Poisson–Boltzmann (PB) equation by the finite-difference technique)<sup>9,10</sup> is applied in LIECE. A LIECE model of 165 kinase inhibitors (CDK2, Lck, and p38 inhibitors) yielded the correct prediction of a test set of 128 EGFR and 37 EphB4 inhibitors.<sup>10</sup> LIECE has also been successfully applied for predicting novel inhibitors of West Nile Virus NS2B/NS3 protease<sup>11</sup> and β-secretase (BACE-1).<sup>12,13</sup>

Linear response models can also be applied to MM-PB/SA to generate LR-MM-PB/SA,<sup>14–16</sup> in which linear regression-based optimization of the coefficients for  $E^{\text{vdW}}$ ,  $E^{\text{ele}}$ , solvation energy, and solvent-accessible surface area (SASA) is implemented. When a single conformation derived from energy minimization of the protein–ligand complex is used to calculate the binding free energy, LR-MM-PB/SA calculations are similar in computation time to LIECE. Both models have also been shown in the literature to outperform docking-based scoring. One such example is a LR-MM-PB/SA model based on single conformation of Cathepsin B inhibitor complexes, which showed good correlation ( $r^2 = 0.92$ ) and internal predictivity ( $q^2 = 0.89$ ).<sup>16</sup> The notable difference between the two methods is that the original LIECE model does not

**Chart 2.** Molecular Structures of the 2-Anilino-6-phenylpyrido[2,3-d]pyrimidin-7(8H)-ones Derivatives Used in This Study

compd	X	Y	Z	activity		
				IC <sub>50</sub> ( $\mu$ M)	pIC <sub>50</sub>	$\Delta G_{\text{exp}}$
test set (38 compounds)						
175	O(CH <sub>2</sub> ) <sub>2</sub> N <i>Et</i> <sub>2</sub>	H	2',6'-diCl	0.165	6.78	-9.31
176	H	H	2',6'-diCl	2.6	5.59	-7.66
177	H	H	2',6'-diF	9.7	5.01	-6.88
179	H	H	2',6'-diMe	0.99	6.00	-8.24
180	H	H	2',6'-diCF <sub>3</sub>	41	4.39	-6.02
181	H	H	2',6'-diOH	27	4.57	-6.27
182	H	H	2'-Cl, 6'-F	2.4	5.62	-7.71
183	H	H	2'-Cl, 6'-Me	1.9	5.72	-7.85
186	H	H	2'-Me, 6'-Br	4.5	5.35	-7.34
189	H	H	2',6'-diCl, 3'-CH <sub>2</sub> OH	3.5	5.46	-7.49
190	H	H	2',6'-diCl, 3'-CH <sub>2</sub> NH <sub>2</sub>	31	4.51	-6.19
191	H	H	2',6'-diCl, 3'-COOH	3.2	5.49	-7.54
192	H	H	2',6'-diCl, 3'-CONH <sub>2</sub>	8.6	5.07	-6.95
193	H	H	2',6'-diCl, 3'-OH	0.074	7.13	-9.79
194	H	H	2',6'-diCl, 3'-NH <sub>2</sub>	2.6	5.59	-7.66
195	H	H	2',6'-diCl, 4'-Cl	8.6	5.07	-6.95
196	H	H	2',6'-diCl, 4'-OH	0.22	6.66	-9.14
198	H	H	2',6'-diCl, 4'-NH <sub>2</sub>	3.7	5.43	-7.45
199	H	H	2',6'-diCl, 4'-NHAc	36	4.44	-6.10
201	H	H	2',6'-diMe, 3',5'-diOMe	33	4.48	-6.15
202	H	H	2',6'-diMe, 3',5'-diOH	0.14	6.85	-9.41
203	CH <sub>2</sub> CONH <sub>2</sub>	H	2',6'-diCl	0.12	6.92	-9.50
205	(CH <sub>2</sub> ) <sub>4</sub> CONH <sub>2</sub>	H	2',6'-diCl	0.26	6.59	-9.04
206	OCH <sub>2</sub> CONH <sub>2</sub>	H	2',6'-diCl	0.25	6.60	-9.06
207	O(CH <sub>2</sub> ) <sub>2</sub> N <i>Et</i> <sub>2</sub>	H	2',6'-diMe	0.99	6.00	-8.24
208	O(CH <sub>2</sub> ) <sub>3</sub> COOH	H	2',6'-diCl	0.086	7.07	-9.70
209	(CH <sub>2</sub> ) <sub>3</sub> COO(CH <sub>2</sub> ) <sub>2</sub> Nmorph	H	2',6'-diCl	0.095	7.02	-9.64
210	(CH <sub>2</sub> ) <sub>3</sub> COO(CH <sub>2</sub> ) <sub>2</sub> N <i>Me</i> <sub>2</sub>	H	2',6'-diCl	0.124	6.91	-9.48
211	(CH <sub>2</sub> ) <sub>3</sub> COO(CH <sub>2</sub> ) <sub>2</sub> Npip	H	2',6'-diCl	0.142	6.85	-9.40
212	(CH <sub>2</sub> ) <sub>3</sub> COOH	H	2',6'-diCl	0.032	7.49	-10.29
213	CH <sub>2</sub> CH(NH <sub>2</sub> )COOH	H	2',6'-diCl	0.09	7.05	-9.67
214	(CH <sub>2</sub> ) <sub>3</sub> tetrazole	H	2',6'-diCl	0.069	7.16	-9.83
215	O(CH <sub>2</sub> ) <sub>2</sub> N <i>Et</i> <sub>2</sub>	H	2',6'-diCl, 3'-OH	0.15	6.82	-9.36
218	O(CH <sub>2</sub> ) <sub>3</sub> COOH	H	2',6'-diCl, 4'-OH	0.04	7.40	-10.15
219	H	Me	2'-Cl	0.41	6.39	-8.77
220	O(CH <sub>2</sub> ) <sub>2</sub> N <i>Et</i> <sub>2</sub>	Me	2'-Cl	0.55	6.26	-8.59
221	H	Me	2',6'-diMe	1.2	5.92	-8.13
222	O(CH <sub>2</sub> ) <sub>2</sub> N <i>Et</i> <sub>2</sub>	Me	2',6'-diMe	0.54	6.27	-8.60
test set (10 compounds)						
178	H	H	2',6'-diBr	0.41	6.39	-8.77
184	H	H	2'-Cl, 6'-OMe	3.4	5.47	-7.50
185	H	H	2'-Cl, 6'-OH	1.5	5.82	-7.99
187	H	H	2'-OMe, 6'-OH	11	4.96	-6.80
188	H	H	2',6'-diCl, 3'-Me	50	4.30	-5.90
197	H	H	2',6'-diMe, 4'-OH	0.58	6.24	-8.56
200	H	H	2',6'-diCl, 3',5'-diOH	0.14	6.85	-9.41
204	(CH <sub>2</sub> ) <sub>2</sub> CONH <sub>2</sub>	H	2',6'-diCl	0.19	6.72	-9.22
216	O(CH <sub>2</sub> ) <sub>3</sub> COOH	H	2',6'-diCl, 3'-OH	0.04	7.40	-10.15
217	O(CH <sub>2</sub> ) <sub>2</sub> N <i>Et</i> <sub>2</sub>	H	2',6'-diCl, 4'-OH	0.08	7.10	-9.74

contain the fitting term  $\gamma$  and has been validated as a general equation on a diverse set of inhibitors and targets.

Described here are LR-MM-PB/SA and LIECE models for a data set of Wee1 kinase inhibitors, in which a single protein–inhibitor conformation derived from minimizing docking solutions was used for generating the LR models. The LIECE model<sup>10</sup> generated originally for CDK2, Lck,

and p38 inhibitors was applied to evaluate whether the model predicts the binding free energies of Wee1 kinase inhibitors. Since the Wee1 kinase inhibitors show different scaffolds compared to the originally studied CDK2, Lck, and p38 inhibitors, it was interesting to evaluate the predictive ability of LIECE models for such an external test set.

**Table 1.** Statistical Values of the Training set of LR-MM-PB/SA Models<sup>a</sup>

model	no. of compds	outlier	statistical values			
			RMSE	r <sup>2</sup>	XRMSE	q <sub>LOO</sub> <sup>2</sup>
1	<b>144</b>		pyrrolocarbazole model	<b>0.61</b>	<b>0.64</b>	<b>0.64</b>
2	140	13, 48, 71, 107		0.55	0.67	0.57
3	38		pyridopyrimidine model	0.86	0.52	0.97
4	<b>35</b>	<b>177, 181, 199</b>		<b>0.65</b>	<b>0.67</b>	<b>0.76</b>
5	182		general model	0.74	0.61	0.77
6	<b>175</b>	<b>13, 48, 71, 107, 177, 181, 199</b>		<b>0.64</b>	<b>0.67</b>	<b>0.66</b>

<sup>a</sup> No. of compds = number of compounds. RMSE = root mean square of error. XRMSE = root mean square of error derived from leave-one-out cross validation.

## 2. COMPUTATIONAL METHODS

**2.1. Inhibitor Data Set.** One hundred seventy four pyrrolo[3,4-*c*]carbazole-1,3(2*H*,6*H*)-dione derivatives and 48 2-anilino-6-phenylpyrido[2,3-*d*]pyrimidin-7(8*H*)-ones derivatives developed and tested in uniform assay conditions as Wee1 kinase inhibitors by Denny et al.<sup>17–20</sup> were analyzed in the present work. The IC<sub>50</sub> values of the inhibitors were converted into binding free energies (ΔG<sub>exp</sub>) by using the equation ΔG<sub>exp</sub> = −RT ln(IC<sub>50</sub>). The structures of the inhibitors, the biological activities (IC<sub>50</sub> and pIC<sub>50</sub>), and the ΔG<sub>exp</sub> values are shown in Charts 1 and 2. The 174 pyrrolocarbazole compounds were split 144/30 into the training/test sets, and the 48 pyridopyrimidine derivatives were split 38/10 between training and test (Tables 2 and 3) by random selection.

**2.2. Ligand Preparation and Docking Study.** All ligands were generated using Sybyl 7.2,<sup>21</sup> energy minimized using

**Table 2.** Comparison between ΔG<sub>exp</sub> and Predicted ΔG<sub>pred</sub> (LR-MM-PB/SA model 1) of Pyrrolocarbazole Test Set

compd	ΔG <sub>exp</sub>	ΔG <sub>pred</sub>	residual
1	-9.62	-9.28	-0.34
2	-7.41	-9.47	2.06
7	-9.04	-9.53	0.49
18	-10.27	-9.89	-0.38
20	-10.48	-10.35	-0.13
29	-9.91	-9.16	-0.75
30	-10.46	-9.73	-0.72
43	-6.04	-8.27	2.24
44	-9.68	-9.00	-0.68
50	-7.52	-8.17	0.65
51	-8.13	-8.88	0.76
56	-6.23	-7.57	1.34
65	-10.12	-9.51	-0.62
80	-10.03	-8.48	-1.55
85	-8.35	-8.80	0.45
86	-8.56	-8.94	0.38
100	-11.11	-10.16	-0.95
111	-10.39	-10.59	0.20
115	-9.41	-9.13	-0.28
124	-8.84	-9.79	0.95
130	-9.55	-9.82	0.27
131	-9.97	-10.26	0.28
137	-9.79	-10.33	0.55
138	-9.85	-9.93	0.09
150	-10.60	-10.26	-0.34
151	-10.82	-9.97	-0.85
158	-10.32	-10.75	0.43
162	-10.23	-10.69	0.46
166	-9.94	-11.37	1.43
172	-9.33	-9.45	0.12

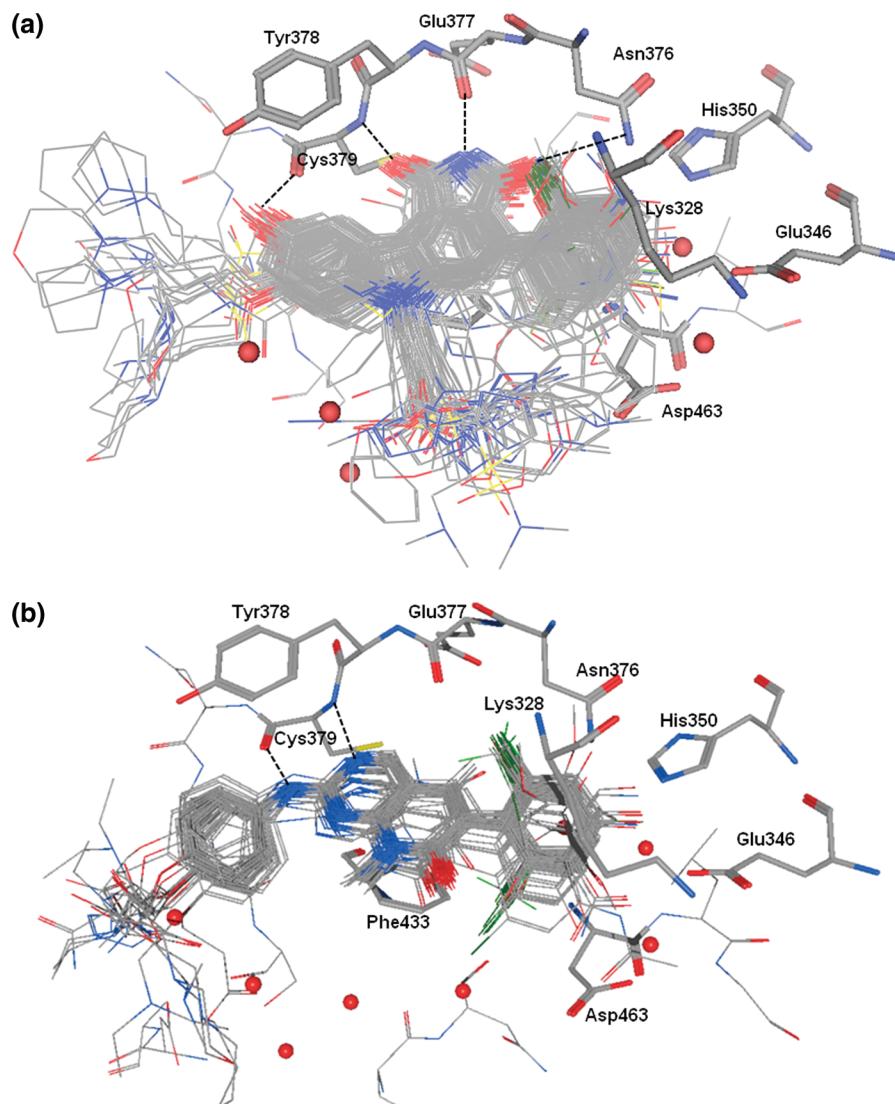
the MMFF94s force field<sup>22–28</sup> and the BFGS optimization method<sup>29–32</sup> with a convergence criteria of 0.05 kcal/mol, and protonated at pH 7.1 using protonate-3D in MOE2008.10.<sup>33</sup> Docking of pyrrolo[3,4-*c*]carbazole-1,3(2*H*,6*H*)-dione derivatives was carried out as described in our previous work,<sup>34</sup> in which we showed that GOLD<sup>35,36</sup> correctly predicted the binding mode of Wee1 kinase inhibitors. The compounds were docked into the X-ray structure of Wee1 kinase (PDB code 1X8B) using GOLD version 3.2 (which was validated on Wee1 kinase inhibitors in previous work<sup>35,36</sup>) with the binding site defined at Cys379 and a radius of 20 Å. We considered eight cocrystallized water molecules within the binding pocket using the water ‘toggle’ mode in GOLD.

**2.3. Refinement of Docking Poses.** All derived docking poses were energy minimized using a combination of steepest descent and conjugate gradient algorithm with a root-mean-square of gradient at 0.001 kcal/mol. AM1-BCC charges<sup>37</sup> were assigned for inhibitors, and the Amber99<sup>38</sup> force field was applied for the protein. The nonbonded cutoff was set at 16 Å. All heavy atoms of Wee1 kinase were tethered with a force constant of 100 kcal/mol, whereas the inhibitor atoms were relaxed during the energy minimization process. The minimized protein–inhibitor complex was subsequently used for calculating the van der Waals and electrostatic interaction energy as well as the solvation energy.

**2.4. Interaction Energy Calculation.** The minimized complexes were prepared for calculation of interaction energy using the LEaP module in AMBER10.<sup>39</sup> The parameters from the general Amber force field (GAFF)<sup>40</sup> were used for the ligands, and the Amber1999SB force field was used for the protein. van der Waals (E<sup>vdW</sup>) and electrostatic (E<sup>ele</sup>) interaction between ligand and the protein in the gas phase were calculated with an infinite cutoff using the SANDER

**Table 3.** Comparison between ΔG<sub>exp</sub> and ΔG<sub>pred</sub> (LR-MM-PB/SA model 4) of Pyridopyrimidine Test Set

compd	ΔG <sub>exp</sub>	ΔG <sub>pred</sub>	residual
178	-8.77	-7.67	-1.09
184	-7.50	-8.29	0.79
185	-7.99	-8.73	0.73
187	-6.80	-8.70	1.90
188	-5.90	-6.96	1.05
197	-8.56	-7.97	-0.59
200	-9.41	-8.49	-0.92
204	-9.22	-9.03	-0.19
216	-10.15	-9.60	-0.55
217	-9.74	-8.68	-1.06



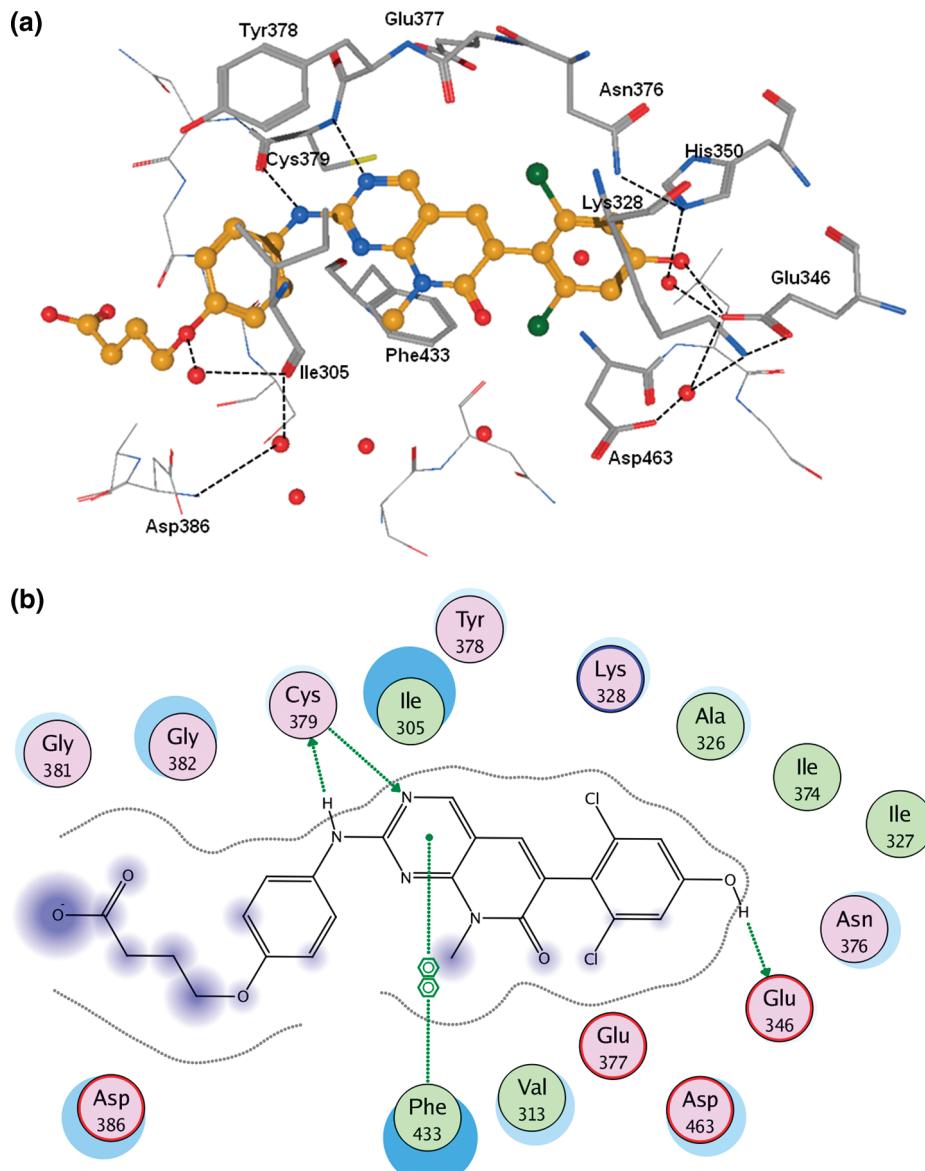
**Figure 1.** Docking poses of all pyrrolocarbazole (A) and pyridopyrimidine (B) derivatives shown at the binding pocket of Wee1 kinase. Hydrogen bonds are shown as dashed lines; cocrystallized water molecules are presented as red balls. Interacting amino acid residues are shown in ball and stick mode.

module in AMBER10. The electrostatic free energy of solvation ( $G^{\text{ele-sol}}$ ) was calculated with numerical solvation of the Poisson–Boltzmann (PB) equation as implemented in the *pbsa* program<sup>41</sup> in AMBER10. Default parameters were used for the PB solver (grid spacing of 0.5 Å, dielectric constants of 1.0 for solute and 80.0 for implicit PB solvent, and solvent probe radius of 1.4 Å). The nonelectrostatic free energy of solvation ( $G^{\text{nonele-sol}}$ ) was calculated using the linear function of the solvent-accessible surface area (SASA, eq 5),<sup>42</sup> where  $\gamma$  and  $b$  were set at the default values ( $\gamma = 0.00542 \text{ kcal/mol } \text{\AA}^2$  and  $b = 0.92 \text{ kcal/mol}$ ).

$$\Delta G^{\text{non-ele-sol}} = \gamma \text{SASA} + b \quad (5)$$

To evaluate the stability of the derived kinase–inhibitor complexes we carried out molecular dynamics simulations using AMBER10 and the force field setup described above. Exemplarily, the two most active inhibitors—pyrrolocarbazole 106 ( $\text{pIC}_{50} 8.22$ ) and pyridopyrimidine 212 ( $\text{pIC}_{50} 7.49$ )—were simulated for 9.5 ns. The complexes were soaked in a box of TIP3P water molecules with a margin of 10 Å. Prior to the free MD simulations, two steps of

relaxation were carried out; in the first step, we kept the protein–inhibitor complex fixed with a constraint of 500 kcal mol<sup>-1</sup> Å<sup>-1</sup>. In the second step, the inhibitor structures were relaxed for 0.5 ps, during which the protein atoms were restrained to the X-ray coordinates with a force constant of 500 kcal mol<sup>-1</sup> Å<sup>-1</sup>. In the final step, all restraints were removed and the complexes were relaxed for 1 ps. The temperature of the relaxed system was then equilibrated at 300 K through 20 ps of MD using 2 fs time steps. A constant volume periodic boundary was set to equilibrate the temperature of the system by the Langevin dynamics using a collision frequency of 10 ps<sup>-1</sup> and a velocity limit of 5 temperature units. During the temperature equilibration routine, the complex in the solvent box was restrained to the initial coordinates with a weak force constant of 10 kcal mol<sup>-1</sup> Å<sup>-1</sup>. The final coordinates of the temperature equilibration routine (after 20 ps) were then used to complete a 10 ns MD routine using 2 fs time steps, during which the temperature was kept at 300 K by the Langevin dynamics using a collision frequency of 1 ps<sup>-1</sup> and a velocity limit of 20 temperature units, and the pressure of the solvated system



**Figure 2.** (A) GOLD docking solution for one of the most potent inhibitors, compound 218. (B) Schematic representation of the interaction of compound 218 with the residues at the Wee1 binding site. Hydrogen bonds are indicated as arrows.

was equilibrated at 1 bar at a certain density in a constant pressure periodic boundary by an isotropic pressure scaling method employing a pressure relaxation time of 2 ps. The time step of the free MD simulations was 2 fs with a cutoff of 9 Å for the nonbonded interaction, and SHAKE was employed to keep all bonds involving hydrogen atoms rigid. Electrostatic interactions were computed using the particle mesh Ewald method.

**2.5. Binding Free Energy Calculation Using LR-MM-PB/SA.** Binding free energies were calculated by applying a three-parameter (eq 4) LR-MM-PB/SA method as shown previously and a four-parameter (eq 6) LR-MM-PB/SA method in which  $\Delta G^{\text{sol}}$  is the summation of  $\Delta G^{\text{ele-sol}}$  and  $\Delta G^{\text{nonele-sol}}$

$$\Delta G = \alpha \Delta E^{\text{vdW}} + \beta \Delta E^{\text{ele}} + \beta_1 \Delta G^{\text{ele-sol}} + \beta_2 \Delta G^{\text{non-ele-sol}} + \gamma \quad (6)$$

The coefficients ( $\alpha$ ,  $\beta$ ,  $\beta_1$ , and  $\beta_2$ ) of LR-MM-PB/SA models were generated by empirical fitting with the experimental

binding free energy ( $\Delta G_{\text{exp}}$ ). We did not include the entropy contribution in our calculations since we analyzed the relative binding free energies of series of similar molecules. Additionally, entropy calculations are computationally expensive, and it has been reported elsewhere<sup>10</sup> that modern entropy calculation methods do not improve the accuracy of binding free energy calculations.

**2.6. Binding Free Energy Calculation Using LIECE.** LIECE models from Caflish et al.<sup>10</sup> based on the CHARMM22 force field and the PBEQ method within CHARMM were applied for predicting the binding free energies of the Wee1 kinase inhibitors. The van der Waals and electrostatic interaction free energy were calculated by subtracting the values of the isolated components from the energy of the complex. The protocol described in section 2.3 was used for energy minimization of the ligands, and the Amber1999SB force field was used for minimizing the protein structure. An entropy term calculation was not included in the LIECE model.

### 3. RESULTS AND DISCUSSION

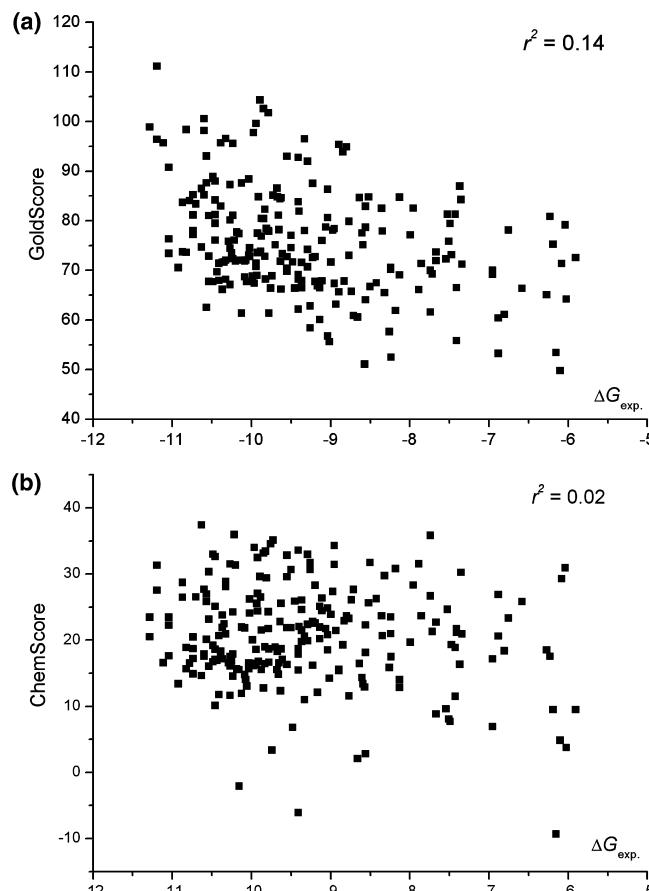
**3.1. Binding Mode of Pyridopyrimidine Derivatives at Wee1.** Since the GOLD docking program was able to reproduce the binding mode of cocrystallized pyrrolocarbazole Wee1 inhibitors,<sup>34</sup> we used GOLD to dock compounds with a pyridopyrimidine scaffold, which revealed a single binding mode for most compounds. Superimposition of the selected docking solutions of the pyridopyrimidines is shown in Figure 1B and revealed a conserved binding mode for pyridopyrimidines, including H-bond interaction with the backbone atoms of Cys379,  $\pi$ -cation interactions between the phenyl ring and the side chains of His350 and Lys328, at the hydrophobic pocket (Figure 1B), and  $\pi$ - $\pi$  stacking between the pyridine ring of inhibitors and Phe433. The most active inhibitors (e.g., compound 218) showed a further H-bond interaction between the OH group at position 4 of the phenyl ring and Glu346 (Figure 2A and 2B). The cocrystallized water molecules located at the binding pocket were involved in H-bond interaction between ligand and enclosed residues such as Ile305 (Figure 2A).

To evaluate the stability and flexibility of the kinase–inhibitor complexes we carried out MD simulations of two exemplarily chosen Wee1 kinase inhibitors. Pyrrolocarbazole 106 and pyridopyrimidine 212, which are the most active compounds in each subset, were simulated for 9.5 ns in complex with Wee1 kinase. In both cases the protein–inhibitor complex was stable and the main interactions, i.e., hydrogen bonds with the hinge region, were retained over the whole simulation time. Only the flexible side chains of compounds 106 and 212 adopt slightly different conformations during the MD simulation, which can be indicated in the rmsd plot shown in Figure S1 (Supporting Information). The low rmsd values observed for both ligands are a good indicator for the high stability of the kinase–inhibitor complexes.

Neither GoldScore nor ChemScore showed significant correlation between  $\Delta G_{\text{exp}}$  and  $\text{pIC}_{50}$  or  $\text{IC}_{50}$  values ( $r^2 = 0.14$  for GoldScore, Figure 3A, and  $r^2 = 0.02$  for ChemScore, Figure 3B), which was not surprising since docking scores often correlate poorly with biological affinities. In addition, we tested whether other scoring functions perform better on the Wee1 kinase data set. Using the X-Score function developed by Wang et al.<sup>43</sup> for the purpose of predicting binding affinities we observed a slightly better correlation coefficient of  $r^2 = 0.19$ . A similar value was observed using the ASP score within GOLD ( $r^2 = 0.19$ ). All scoring functions tested yielded models which are not applicable for structure-based optimization of Wee1 kinase inhibitors.

**3.2. LR-MM-PB/SA-Based Prediction.** The LR-MM-PB/SA method was used to generate a model for the pyrrolocarbazole subset, a model for the pyridopyrimidine subset, and a general model for both subsets. When the four-parameter method described in eq 6 was tested, the coefficients of the term  $\Delta G^{\text{nonele-sol}}$  were found to not significantly contribute to the models, so we reverted back to the three-parameter model.

The LR-MM-PB/SA model (model 1) for the pyrrolocarbazole subset was generated using 144 compounds as the training set (the same set as in our previous publication<sup>34</sup>). Relevant statistics for model 1 were  $r^2 = 0.64$ , root-mean-square of error (RMSE) = 0.61 kcal/mol,  $q_{\text{LOO}}^2 = 0.62$  and leave-one-out cross-validated root-mean-square of error



**Figure 3.** Correlation between  $\Delta G_{\text{exp}}$  and GoldScore (A) as well as ChemScore (B) for all 222 Wee1 kinase inhibitors.

(XRMSE) = 0.64 kcal/mol (Table 1). Compounds 13, 48, 71, and 107 showed residual values ( $\Delta G_{\text{exp}} - \Delta G_{\text{est.}}$ ) that were twice that of the RMSE value of model 1. Since compounds 13, 48, and 107 were also found to be outliers in a receptor-based 3D-QSAR model of the same data set,<sup>34</sup> a reduced model was generated by removing the three compounds and compound 71. The refined model (model 2, eq 7) showed only a slight improvement ( $r^2 = 0.67$ , RMSE = 0.55 kcal/mol,  $q_{\text{LOO}}^2 = 0.65$ , and XRMSE = 0.57 kcal/mol).

$$\Delta G = 0.10877\Delta E^{\text{vdW}} + 0.06786\Delta E^{\text{ele}} + 0.06239\Delta G^{\text{sol}} - 10.91832 \quad (7)$$

The LR-MM-PB/SA model for the pyridopyrimidine subset was built in the same way as that described for the pyrrolocarbazole subset. A three-parameter model was initially generated using 38 compounds as the training set (Chart 2). The correlation between observed and calculated data (model 3) yielded an  $r^2$  of 0.52, RMSE of 0.86 kcal/mol,  $q_{\text{LOO}}^2$  of 0.41, and XRMSE = 0.91 kcal/mol. The model for the pyridopyrimidine was thus less accurate than the pyrrolocarbazole model. Compounds 177, 181, and 199 were found to have residual values higher than 2.0 times that of the RMSE value and removed from the data set, and their docking results were visually analyzed in an attempt to understand the shortcomings of our model. In the case of compounds 177 and 181, it was not clear why a fluoro substitution at position 2 (compound 177) or a hydroxyl group (compound 181) results in lower activities, and our

calculated binding free energies were not consistent with the biological data. This effect is unlikely due to the electron-withdrawing group, as several compounds in the training set contain an electron-withdrawing group ( $-\text{Cl}$ ,  $-\text{F}$ ,  $-\text{Br}$ , or  $-\text{OH}$ ) at the same position but exhibited high biological activity. Finally, a refined model (model 4, eq 8) was generated from the remaining 35 compounds and had a much higher correlation than the unrefined model ( $r^2 = 0.67$ , RMSE = 0.65 kcal/mol,  $q_{\text{LOO}}^2 = 0.58$ , and XRMSE = 0.76 kcal/mol).

$$\Delta G = 0.09395\Delta E^{\text{vdW}} - 0.00187\Delta E^{\text{ele}} + 0.00726\Delta G^{\text{sol}} - 8.60966 \quad (8)$$

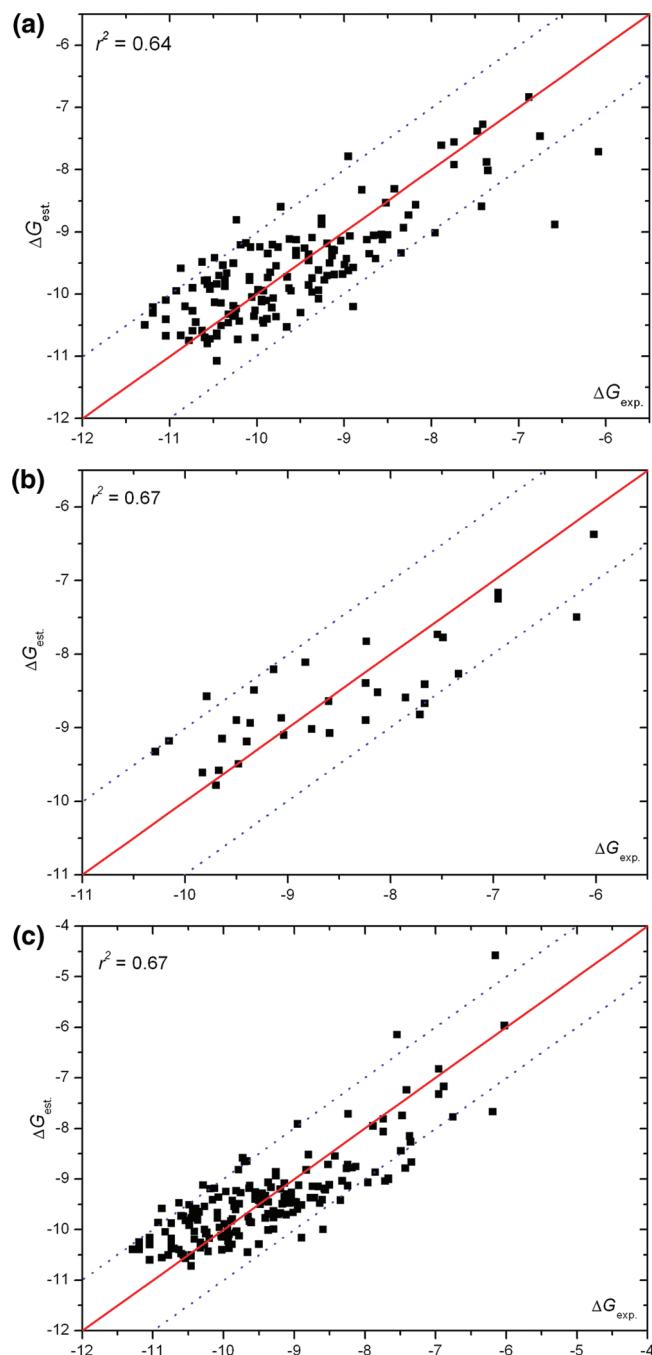
A unified model was then built based on the 182 compounds from the previous model, and the seven previously identified outliers were subsequently removed. The statistical values of the final model 6 were  $r^2 = 0.67$ , RMSE = 0.64 kcal/mol,  $q_{\text{LOO}}^2 = 0.64$ , and XRMSE = 0.66 kcal/mol. The  $r^2$  value indicates a good correlation between  $\Delta G_{\text{exp}}$  and  $\Delta G_{\text{est}}$ . In addition, the obtained  $q_{\text{LOO}}^2$  value suggests the robustness of the model for predicting further compounds. The derived unified model is shown in eq 9

$$\Delta G = 0.10505\Delta E^{\text{vdW}} - 0.05267\Delta E^{\text{ele}} + 0.05176\Delta G^{\text{sol}} - 10.74762 \quad (9)$$

The correlation between  $\Delta G_{\text{exp}}$  and  $\Delta G_{\text{est}}$  derived from LR-MM-PB/SA models 1, 4, and 6 are shown in Figure 4A, 4B, and 4C, respectively. A good correlation was observed in all three models, and most compounds show low residual values in the range of  $\pm 1.0$  kcal/mol (indicated by the dashed lines in Figure 4A, 4B, and 4C).

The LR-MM-PB/SA models (models 1, 4, and 6) were then examined with an external test set of 30 pyrrolocarbazoles and 10 pyridopyrimidines. The  $\Delta G_{\text{exp}}$  and  $\Delta G_{\text{pred}}$  and residual values of each compound of the test set are listed in Tables 2, 3 and 4, respectively.  $\Delta G_{\text{exp}}$  shows a good correlation with  $\Delta G_{\text{pred}}$ , yielding  $r^2$  values of 0.54 for the pyrrolocarbazole subset (Figure 5A), 0.42 for the pyridopyrimidine subset (Figure 5B), and 0.57 for the unified test set (Figure 5C). With one exception (compound 187, Figure 5B), the compounds of this model yielded low residual ( $\Delta G_{\text{exp}} - \Delta G_{\text{pred}}$ ) values, which suggests the model was reliable and robust.

**3.3. Linear Interaction Energy Method with Continuum Electrostatics (LIECE).** The calculated one-parameter models (models 7–13, shown in Table 5) using different  $\alpha$  values taken from the work of Caflish et al.<sup>10</sup> had a similar correlation coefficient ( $r^2 = 0.39$ ) and RMSE value (0.96 kcal/mol) to that of the LR-MM-PB/SA models. The high similarity of the results of models 7–13 is due to the small deviation of the  $\alpha$  coefficient among the different models (0.2338–2.770). In general, all one-parameter models (model 7–13) yielded lower correlation coefficients and higher RMSE values as observed for the LR-MM-PB/SA models. Interestingly, the correlation increases significantly when the electrostatic term is taken into account. For example, the two-parameter model (models 14–20), in which the electrostatic free energy ( $\Delta G^{\text{ele}}$ ) is included, yielded  $r^2$  values in the range of 0.40–0.54 (RMSE values between 0.84 and 0.95 kcal/mol). The three-parameter models (models



**Figure 4.** Correlation between  $\Delta G_{\text{exp}}$  and  $\Delta G_{\text{est}}$  of the compounds (A) in the pyrrolocarbazole training set (LR-MM-PB/SA model 1), (B) in the pyridopyrimidine training set (LR-MM-PB/SA model 4), and (C) in the unified training set (LR-MM-PB/SA model 6).

21–27), in which  $\Delta G^{\text{ele}}$  is decomposed into two terms (electrostatic interaction and solvation energy), showed slightly lower  $r^2$  (0.24–0.50) and higher RMSE values (0.87–1.07 kcal/mol).

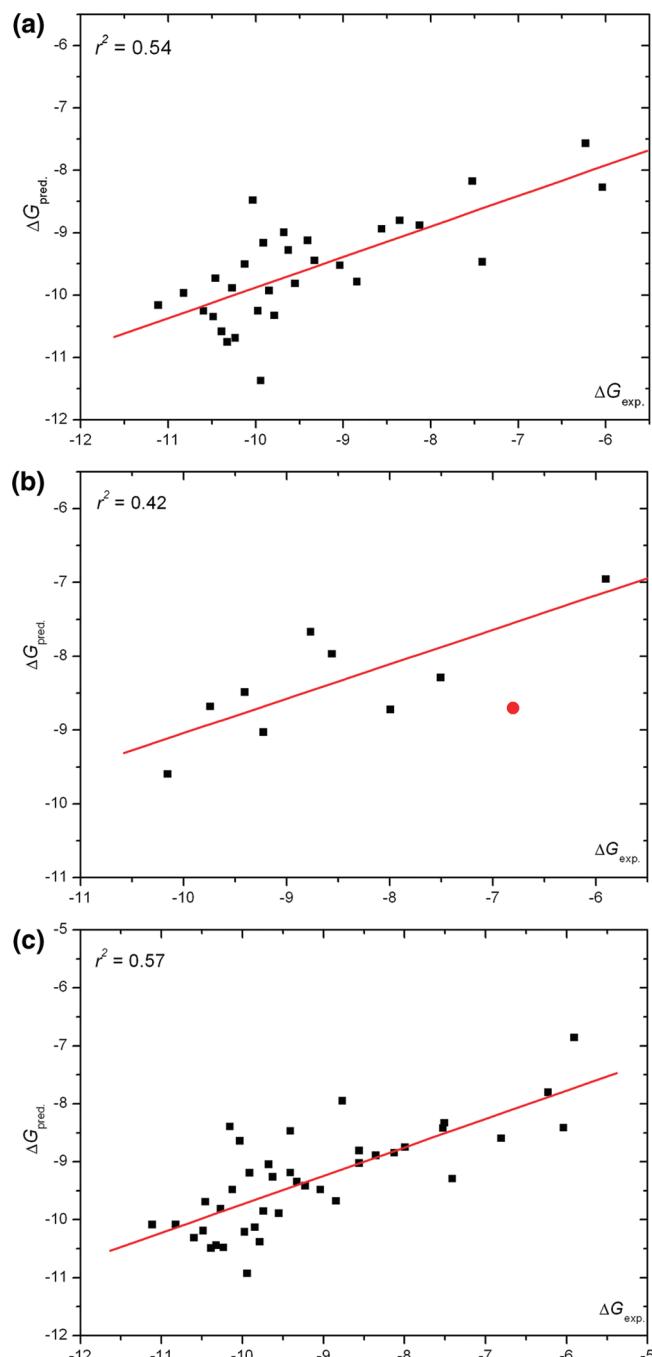
By applying the LIECE models for predicting the binding free energies of the 222 Wee1 kinase inhibitors, we observed similar accuracy as reported in the original work of Caflish et al.<sup>10</sup> The two-parameter model (model 20) based on three different kinase inhibitor data sets gave a slightly higher correlation coefficients ( $r^2 = 0.50$ ) than the three-parameter model (model 27,  $r^2 = 0.48$ ) generated from the same data set of kinase inhibitors. The highest correlation ( $r^2 = 0.54$ ) between  $\Delta G_{\text{pred}}$  and  $\Delta G_{\text{exp}}$  was observed using model 17, a

**Table 4.** Comparison between  $\Delta G_{\text{exp}}$  and Predicted  $\Delta G_{\text{pred}}$  (LR-MM-PB/SA model 6) of the Unified Test Set

compd	$\Delta G_{\text{exp}}$	$\Delta G_{\text{pred}}$	residual
1	-9.62	-9.27	-0.36
2	-7.41	-9.30	1.89
7	-9.04	-9.48	0.45
18	-10.27	-9.81	-0.45
20	-10.48	-10.19	-0.29
29	-9.91	-9.19	-0.72
30	-10.46	-9.69	-0.76
43	-6.04	-8.42	2.38
44	-9.68	-9.05	-0.63
50	-7.52	-8.42	0.90
51	-8.13	-8.85	0.72
56	-6.23	-7.81	1.58
65	-10.12	-9.48	-0.64
80	-10.03	-8.64	-1.39
85	-8.35	-8.89	0.54
86	-8.56	-9.03	0.47
100	-11.11	-10.09	-1.02
111	-10.39	-10.49	0.11
115	-9.41	-9.19	-0.21
124	-8.84	-9.68	0.84
130	-9.55	-9.89	0.34
131	-9.97	-10.21	0.24
137	-9.79	-10.38	0.60
138	-9.85	-10.13	0.29
150	-10.60	-10.31	-0.28
151	-10.82	-10.08	-0.74
158	-10.32	-10.44	0.12
162	-10.23	-10.48	0.25
166	-9.94	-10.93	0.99
172	-9.33	-9.34	0.02
178	-8.77	-7.95	-0.81
184	-7.50	-8.33	0.83
185	-7.99	-8.75	0.76
187	-6.80	-8.60	1.79
188	-5.90	-6.86	0.96
197	-8.56	-8.81	0.25
200	-9.41	-8.47	-0.93
204	-9.22	-9.42	0.19
216	-10.15	-8.40	-1.76
217	-9.74	-9.85	0.11

two-parameter model using two data sets (CDK2 and Lck inhibitor data set). This may be due to the structural similarity of the Wee1, CDK2, and Lck inhibitors. A two-parameter model based on three kinases (model 20) was generated using inhibitors covering more chemical space of the kinase inhibitors than model 17 and found to have a similar  $r^2$  (0.50) to model 17. Model 20 was thus deemed the most robust and reliable model for prediction of relative binding free energies of other kinase inhibitors. Although the correlation between the  $\Delta G_{\text{exp}}$  and  $\Delta G_{\text{pred}}$  of Wee1 kinase inhibitors using model 20 showed a good correlation ( $r^2 = 0.50$ , RMSE = 0.86 kcal/mol), most  $\Delta G_{\text{pred}}$  values derived from model 20 were positive and had high residual values (Figure 6). This may have been because the force field used was different from that of the original study. Our results suggest that the LIECE models can be applied to rank compounds retrieved from virtual screening of Wee1 kinases and provide a further validation for the robustness of this method.

**3.4. Virtual Screening.** The significant correlation between the experimental values and the predicted binding free energies of these models suggested the models could be useful for directing a virtual screening experiment. Enrichment factors (EFs)<sup>44</sup> were thus calculated to determine the ability of these derived models to discriminate true hits from



**Figure 5.** Correlation between  $\Delta G_{\text{exp}}$  and  $\Delta G_{\text{pred}}$  of the compounds (A) in the pyrrolocarbazole test set, (B) in the pyridopyrimidine (red point indicates the compound which gave high residual value), and (C) in the unified model.

decoys of the databases. The EF<sup>43</sup> for each data set was calculated according to following formula

$$\text{EF} = \{N_{\text{total}}/N_{\text{sampled}}\} \{\text{hits}_{\text{sampled}}/\text{hits}_{\text{total}}\}$$

where  $N_{\text{total}}$  and  $N_{\text{samples}}$  represent the total number of compounds in the database (1000 compounds) and the number of screened compounds, respectively, and  $\text{hits}_{\text{sampled}}$  and  $\text{hits}_{\text{total}}$  are the number of found actives and the total number of actives (20 compounds). Selection of compounds from the set of known active inhibitors of Wee1 kinase can lead to artificially high EF values. Thus, to avoid this bias, different databases containing different active compounds were used to calculate the EF value, and the average value

**Table 5.** Statistical Values of the Derived LIECE Models

model	data set of inhibitor of which protein used for generating model	A	$\beta$ or $\beta_1$	$\beta_2$	$r^2$	RMSE	$q_{\text{LOO}}^2$	XRMSE
one-parameter model								
7	CDK2	0.2338			0.39	0.96	0.37	0.97
8	Lck	0.277			0.39	0.96	0.37	0.97
9	p38	0.2377			0.39	0.96	0.37	0.97
10	CDK2+ Lck	0.251			0.39	0.96	0.37	0.97
11	CDK2+ p38	0.2383			0.39	0.96	0.37	0.97
12	Lck+ p38	0.2513			0.39	0.96	0.37	0.97
13	CDK2+ Lck+ p38	0.2463			0.39	0.96	0.37	0.97
two-parameter model								
14	CDK2	0.2866	0.0525		0.51	0.85	0.50	0.86
15	Lck	0.2735	0.0046		0.40	0.95	0.39	0.96
16	p38	0.2699	0.0264		0.46	0.90	0.45	0.90
17	CDK2+ Lck	<b>0.3072</b>	<b>0.0657</b>		<b>0.54</b>	<b>0.84</b>	<b>0.53</b>	<b>0.85</b>
18	CDK2+ p38	0.2632	0.0235		0.45	0.91	0.44	0.91
19	Lck+ p38	0.3033	0.0508		0.51	0.86	0.50	0.87
20	CDK2+ Lck+ p38	<b>0.2898</b>	<b>0.0442</b>		<b>0.50</b>	<b>0.86</b>	<b>0.49</b>	<b>0.88</b>
three-parameter model								
21	CDK2	0.2395	0.075	0.0294	0.41	0.94	0.40	0.95
22	Lck	0.2446	0.1528	0.0076	0.27	1.05	0.25	1.05
23	p38	0.1827	0.1584	-0.0013	0.24	1.07	0.21	1.08
24	CDK2+ Lck	0.3118	0.044	0.062	0.50	0.87	0.47	0.88
25	CDK2+ p38	0.219	0.0439	0.0006	0.36	0.98	0.34	0.99
26	Lck+ p38	0.2939	0.1186	0.0584	0.43	0.92	0.41	0.93
27	CDK2+ Lck+ p38	<b>0.2961</b>	<b>0.0325</b>	<b>0.0454</b>	<b>0.48</b>	<b>0.88</b>	<b>0.47</b>	<b>0.89</b>

of EF was also considered. The EFs were considered at 5% and 10% of the screened databases. Note that the maximum values of EFs are 20 and 10 at the 5% and 10% levels, respectively. When these maximum values are obtained, it means that all active compounds are recovered at the considered fraction of the screened database.

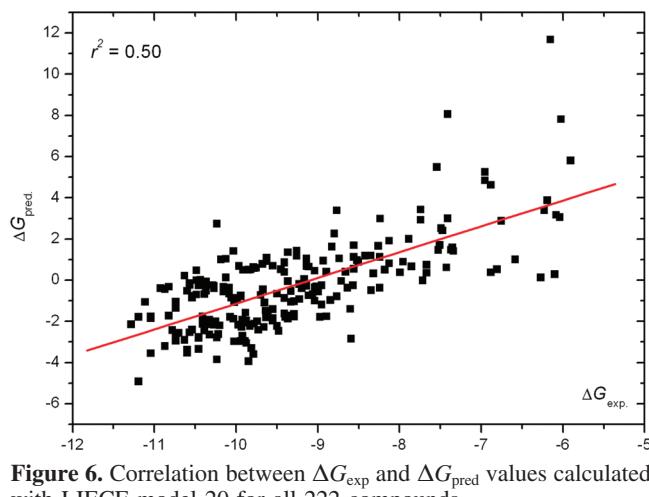
The first database contained 980 randomly selected compounds from the Chembridge Kinaset collection (Chembridge Corp., San Diego, CA), a kinase library of 11 000 compounds selected thorough a ligand-based pharmacophore search. The second data set contained 980 compounds randomly selected from the directory of useful decoys (DUD).<sup>45</sup> Here we took all decoy sets available for kinase inhibitors which includes EGFR, FGFR1, P38, PDGFR, SRC, TK, and VEGFR2. Then, 6 different sets of active Wee1 kinase inhibitors were filled with each of these two databases, leading to 12 final data sets of compounds. One data set contains the 20 most active inhibitors, and the other

5 data sets were generated by randomly selecting 20 actives from the large data set of 222 Wee1 inhibitors. In total, 12 different data sets, each of them containing 1000 compounds, were applied for the enrichment study. The DUD provides a well-defined, unbiased set of active and decoy compounds, and the chemical structures of the Kinaset library have been designed using known kinase inhibitor scaffolds. Discriminating actives from inactive compounds with similar chemical structure is a challenge and nontrivial task in virtual screening. Therefore, the chosen compounds represent suitable data sets to evaluate the performance of LR-MM-PB/SA and LIECE models.

The EFs are summarized in Table 6. Results revealed that most of the top 20 actives were recovered within 5% of the databases by using the LR-MM-PB/SA model 6. The LIECE model 20 was able to identify 16 active compounds within the first 5% of the screened databases. At 10% of the screened databases, it was found that all active compounds were recovered by the LR-MM-PB/SA model 6 and 19 compounds were received by LIECE model 20. In general, the LR-MM-PB/SA models yielded slightly better EFs compared to the LIECE models when considering all 12 data sets.

The virtual screening performance is additionally assessed by the ability to distinguish known active compounds (P) from the selected decoys (N). For each compound in the sorted row, the true positive rate (TPR) and the false positive rate (FPR) were calculated. Solutions that score better than or equal to that particular compound are defined as positive solutions. Active compounds within the range of positive solutions are true positives (TP), and decoys within the range of defined positive solutions are false positives (FP). TPR and FPR are calculated according to these equations

$$\text{TPR} = \frac{\text{TP}}{\text{P}} \quad \text{and} \quad \text{FPR} = \frac{\text{FP}}{\text{P}}$$



**Figure 6.** Correlation between  $\Delta G_{\text{exp.}}$  and  $\Delta G_{\text{pred.}}$  values calculated with LIECE model 20 for all 222 compounds.

**Table 6.** Enrichment Factors (EFs)<sup>a</sup>

	LR-MM-PB/SA (model 6)		LIECE (model 20)	
	5%	10%	5%	10%
Chembridge Kinaset enriched with Wee1 inhibitors				
top 20 actives	20.0	10.0	16.0	9.5
data set 1	13.0	9.0	13.0	8.0
data set 2	15.0	9.0	11.0	7.5
data set 3	15.0	8.5	13.0	8.5
data set 4	16.0	9.5	14.0	9.0
data set 5	16.0	9.5	13.0	7.5
average	<b>15.8</b>	<b>9.3</b>	<b>13.3</b>	<b>8.3</b>
DUD enriched with Wee1 inhibitors				
top 20 actives	19.0	10.0	16.0	9.5
data set 1	12.0	8.5	13.0	8.0
data set 2	12.0	9.0	11.0	7.5
data set 3	15.0	8.5	12.0	8.5
data set 4	15.0	9.0	13.0	9.0
data set 5	15.0	9.5	13.0	7.5
average	<b>14.7</b>	<b>9.1</b>	<b>13.0</b>	<b>8.3</b>

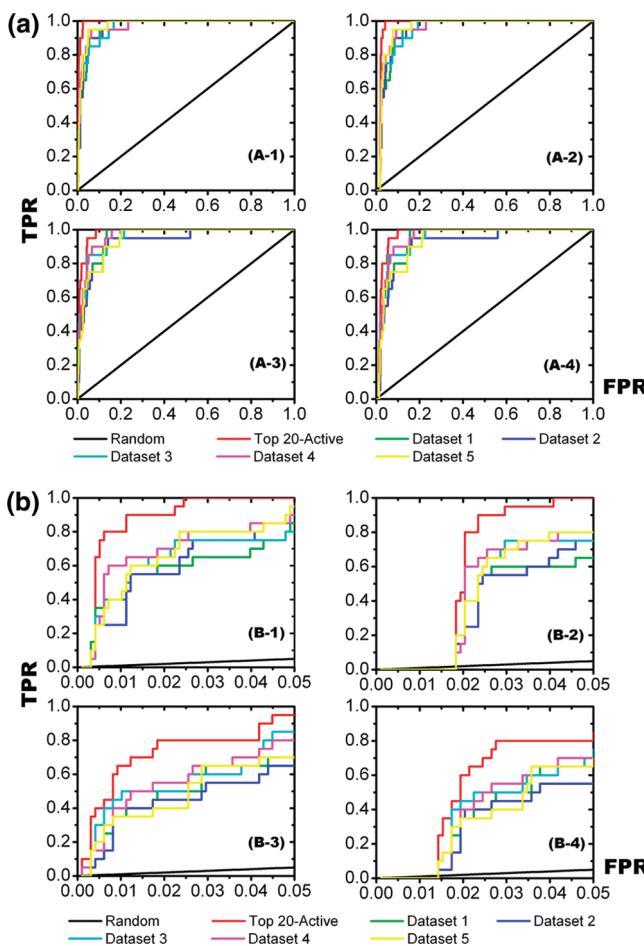
<sup>a</sup> Twenty active inhibitors (top 20 and five randomly selected inhibitor collections) and 980 decoys from the DUD decoy set and the Chembridge Kinaset compound collection were considered. EFs were calculated at 5% and 10% of screened database using LR-MM-PB/SA model 6 and LIECE model 20.

The receiver operator characteristic (ROC) diagrams resulting from plotting the TPR and FPR values are shown in Figure 7A-1–A-4. Ideally, ROC curves show a steep early ascent, almost parallel to the y axis, and then, close to the maximal value for y, continue parallel to the x axis. Such a behavior can be seen in all data sets as demonstrated in Figure 7A-1–A-4. Moreover, for practical reasons, the enrichment within the few top-ranking solutions is of great interest; economic demands allow only the processing of a limited number of compounds. Therefore, the focus on the top-ranking 5% of solutions of these ROC curves are additionally displayed in Figure 7B-1–B-4. Using the linear models (LR-MM-PB/SA and LIECE model) for virtual screening, we could show that these models produce very steep ascending of the ROC curves as displayed in Figure 7A-1–B-4. In general, the LR-MM-PB/SA and LIECE models show a good performance for ranking and separating active compounds from decoys.

#### 4. CONCLUSIONS

We applied molecular docking and binding free energy calculations to rationalize the binding mode of pyridopyrimidine and pyrrolocarbazole derivatives with Wee1 kinase. A conserved binding mode was observed, which included two H bonds to the backbone, interactions with a hydrophobic pocket, and  $\pi-\pi$  stacking with a phenylalanine of the protein. Analysis of the stability and flexibility of the kinase–inhibitor complexes by means of MD simulations showed that the inhibitor retained the binding mode observed in the docking study. However, concerning the docking results, no correlation was observed between either GoldScore or ChemScore and the experimental activities of the inhibitors. Transferable linear response methods such as LR-MM-PB/SA and LIECE were then applied to develop a better predictor of experimental activities.

To establish the LR-MM-PB/SA model, we took the docking poses of the Wee1 kinase inhibitors and mini-



**Figure 7.** (A) ROC curves obtained for different data sets: (A-1) Chembridge Kinaset decoys enriched with 6 different data sets of active Wee1 kinase inhibitors (calculated using LR-MM-PB/SA model 6), (A-2) DUD decoys enriched with 6 different data sets of active Wee1 kinase inhibitors (calculated with LR-MM-PB/SA model 6), (A-3) Chembridge Kinaset decoys enriched with 6 different data sets of active Wee1 kinase inhibitors (calculated with LIECE model 20), and (A-4) DUD decoys enriched with 6 different data sets of active Wee1 kinase inhibitors (calculated with LIECE model 20). TPR = true positive rate, FPR = false positive rate. (B) Focus on the first 5% of the ROC curves. Same data set collections are used as explained and shown in Figure 7A.

mized the molecules within the protein binding pocket. The derived models (for pyrrolocarbazole, pyridopyrimidine, and the unified data set) gave significant and predictive models, indicated by high correlation coefficients as well as high leave-one-out cross-correlation coefficients. The obtained models also effectively predicted the binding free energies of external test set compounds. Although the methodology of LR-MM-PB/SA is related to that of scoring functions since it is also a linear equation of the interaction energies, it performs better because of a more sophisticated calculation of the electrostatic term and generation of a unique linear equation of the interaction term for every protein.

The LIECE models originally derived for inhibitors of CDK2, Lck, and p38 were applied to predict the experimental binding affinities of Wee1 kinase inhibitors. The predicted biological values of the training set molecules showed high residual values, which may be because we used a different force field from the original publication. Despite these high

residual values, the significant correlation ( $r^2 = 0.50$ ) observed for the LIECE model suggests that it could be possible to use this approach for calculating the relative binding free energies of other kinase inhibitors. The derived LR-MM-PB/SA and LIECE models were able to separate active inhibitors from decoys, and most of the actives were identified among the first 10% of the database screened and are ranked correctly by both models. ROC curves generated using LR-MM-PB/SA and LIECE models showed a good performance for ranking and separating active compounds from decoys.

### ACKNOWLEDGMENT

The authors would like to thank the Department of Pharmaceutical Chemistry, Martin-Luther University Halle-Wittenberg, Halle (Saale), Germany, and Computational Chemistry Unit Cell (CCUC), Department of Chemistry, Faculty of Science, Chulalongkorn University, Bangkok, Thailand, for support in hardware and software. K.W. also thanks the Thailand Research Fund through the the Royal Golden Jubilee Ph.D. Program (Grant Number PHD/0044/2548), Center for Petroleum, Petrochemicals, and Advanced Materials, Chulalongkorn University, and Deutscher Akademischer Austausch Dienst (DAAD) for financial support. M.L. thanks the German Fulbright Kommission and IIE for financial support.

**Supporting Information Available:** rmsd plots of inhibitors (pyrrolocarbazole compound 106 and pyridopyrimidine compound 212) during the MD simulation time. This material is available free of charge via the Internet at <http://pubs.acs.org>.

### REFERENCES AND NOTES

- (1) Klebe, G. Virtual ligand screening: strategies, perspectives and limitations. *Drug Discovery Today* **2006**, *11*, 580–594.
- (2) Aqvist, J.; Marelius, J. The linear interaction energy method for predicting ligand binding free energies. *Comb. Chem. High Throughput Screening* **2001**, *4*, 613–626.
- (3) Carlsson, J.; Boukharta, L.; Aqvist, J. Combining docking, molecular dynamics and the linear interaction energy method to predict binding modes and affinities for non-nucleoside inhibitors to HIV-1 reverse transcriptase. *J. Med. Chem.* **2008**, *51*, 2648–2656.
- (4) Aqvist, J.; Medina, C.; Samuelsson, J. E. A new method for predicting binding affinity in computer-aided drug design. *Protein Eng.* **1994**, *7*, 385–391.
- (5) Stjernschantz, E.; Marelius, J.; Medina, C.; Jacobsson, M.; Vermeulen, N. P.; Oostenbrink, C. Are automated molecular dynamics simulations and binding free energy calculations realistic tools in lead optimization? An evaluation of the linear interaction energy (LIE) method. *J. Chem. Inf. Model.* **2006**, *46*, 1972–1983.
- (6) van Lipzig, M. M.; ter Laak, A. M.; Jongejan, A.; Vermeulen, N. P.; Wamelink, M.; Geerke, D.; Meerman, J. H. Prediction of ligand binding affinity and orientation of xenoestrogens to the estrogen receptor by molecular dynamics simulations and the linear interaction energy method. *J. Med. Chem.* **2004**, *47*, 1018–1030.
- (7) Jones-Hertzog, D. K.; Jorgensen, W. L. Binding affinities for sulfonamide inhibitors with human thrombin using Monte Carlo simulations with a linear response method. *J. Med. Chem.* **1997**, *40*, 1539–1549.
- (8) Huang, D.; Caflisch, A. Efficient evaluation of binding free energy using continuum electrostatics solvation. *J. Med. Chem.* **2004**, *47*, 5791–5797.
- (9) Warwicker, J.; Watson, H. C. Calculation of the electric potential in the active site cleft due to alpha-helix dipoles. *J. Mol. Biol.* **1982**, *157*, 671–679.
- (10) Kolb, P.; Huang, D.; Dey, F.; Caflisch, A. Discovery of kinase inhibitors by high-throughput docking and scoring based on a transferable linear interaction energy model. *J. Med. Chem.* **2008**, *51*, 1179–1188.
- (11) Ekonomiuk, D.; Su, X. C.; Ozawa, K.; Bodenreider, C.; Lim, S. P.; Yin, Z.; Keller, T. H.; Beer, D.; Patel, V.; Otting, G.; Caflisch, A.; Huang, D. Discovery of a non-peptidic inhibitor of west nile virus NS3 protease by high-throughput docking. *PLoS Negl. Trop. Dis.* **2009**, *3*, e356.
- (12) Huang, D.; Lüthi, U.; Kolb, P.; Cecchini, M.; Barberis, A.; Caflisch, A. In silico discovery of beta-secretase inhibitors. *J. Am. Chem. Soc.* **2006**, *128*, 5436–5443.
- (13) Huang, D.; Lüthi, U.; Kolb, P.; Edler, K.; Cecchini, M.; Audetat, S.; Barberis, A.; Caflisch, A. Discovery of cell-permeable non-peptide inhibitors of beta-secretase by high-throughput docking and continuum electrostatics calculations. *J. Med. Chem.* **2005**, *48*, 5108–5111.
- (14) Zhou, Z.; Madura, J. D. Relative free energy of binding and binding mode calculations of HIV-1 RT inhibitors based on dock-MM-PB/GS. *Proteins* **2004**, *57*, 493–503.
- (15) Zhou, Z.; Bates, M.; Madura, J. D. Structure modeling, ligand binding, and binding affinity calculation (LR-MM-PBSA) of human heparanase for inhibition and drug design. *Proteins* **2006**, *65*, 580–592.
- (16) Zhou, Z.; Wang, Y.; Bryant, S. H. Computational analysis of the cathepsin B inhibitors activities through LR-MMPBSA binding affinity calculation based on docked complex. *J. Comput. Chem.* **2009**, *30*, 2165–2175.
- (17) Palmer, B. D.; Smail, J. B.; Newcastle, G. W.; Dobrusin, E. M.; Kraker, A.; Moore, C. W.; Steinkampf, R. W.; Denny, W. A. Structure-activity relationships for 2-anilino-6-phenylpyrido[2,3-d]pyrimidin-7(8H)-ones as inhibitors of the cellular checkpoint kinase Wee1. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 1931–1935.
- (18) Palmer, B. D.; Thompson, A. M.; Booth, R. J.; Dobrusin, E. M.; Kraker, A. J.; Lee, H. H.; Lunney, E. A.; Mitchell, L. H.; Ortwin, D. F.; Smail, J. B.; Swan, L. M.; Denny, W. A. 4-Phenylpyrrolo[3,4-c]carbazole-1,3(2H,6H)-dione inhibitors of the checkpoint kinase Wee1. Structure-activity relationships for chromophore modification and phenyl ring substitution. *J. Med. Chem.* **2006**, *49*, 4896–4911.
- (19) Smail, J. B.; Baker, E. N.; Booth, R. J.; Bridges, A. J.; Dickson, J. M.; Dobrusin, E. M.; Ivanovic, I.; Kraker, A. J.; Lee, H. H.; Lunney, E. A.; Ortwin, D. F.; Palmer, B. D.; Quin, J., III; Squire, C. J.; Thompson, A. M.; Denny, W. A. Synthesis and structure-activity relationships of N-6 substituted analogues of 9-hydroxy-4-phenylpyrrolo[3,4-c]carbazole-1,3(2H,6H)-diones as inhibitors of Wee1 and Chk1 checkpoint kinases. *Eur. J. Med. Chem.* **2008**, *43*, 1276–1296.
- (20) Smail, J. B.; Lee, H. H.; Palmer, B. D.; Thompson, A. M.; Squire, C. J.; Baker, E. N.; Booth, R. J.; Kraker, A.; Hook, K.; Denny, W. A. Synthesis and structure-activity relationships of soluble 8-substituted 4-(2-chlorophenyl)-9-hydroxypyrido[3,4-c]carbazole-1,3(2H,6H)-diones as inhibitors of the Wee1 and Chk1 checkpoint kinases. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 929–933.
- (21) Syby7.2, Tripos Inc.: South Hanley, St. Louis, MO, 2007.
- (22) Halgren, T. A. Merck Molecular Force Field: I. Basis, Form, Scope, Parameterization and Performance of MMFF94. *J. Comput. Chem.* **1996**, *17*, 490–519.
- (23) Halgren, T. A. Merck Molecular Force Field. II. MMFF94 van der Waals and Electrostatic Parameters for Intermolecular Interactions. *J. Comput. Chem.* **1996**, *17*, 520–552.
- (24) Halgren, T. A. Merck Molecular Force Field. III. Molecular Geometries and Vibrational Frequencies. *J. Comput. Chem.* **1996**, *17*, 553–586.
- (25) Halgren, T. A.; Nachbar, R. B. Merck Molecular Force Field. IV. Conformational Energies and Geometries for MMFF94. *J. Comput. Chem.* **1996**, *17*, 587–615.
- (26) Halgren, T. A. Merck Molecular Force Field. V. Extension of MMFF94 Using Experimental Data, Additional Computational Data and Empirical Rules. *J. Comput. Chem.* **1996**, *17*, 616–641.
- (27) Halgren, T. A. MMFF VI. MMFF94s Option for Energy Minimization Studies. *J. Comput. Chem.* **1999**, *20*, 720–729.
- (28) Halgren, T. A. MMFF VII. Characterization of MMFF94, MMFF94s, and Other Widely Available Force Fields for Conformational Energies and for Intermolecular-Interaction Energies and Geometries. *J. Comput. Chem.* **1999**, *20*, 730–748.
- (29) Broyden, C. G. The convergence of a class of double-rank minimization algorithms: I. General considerations. *J. Inst. Math. Appl.* **1970**, *6*, 76–90.
- (30) Fletcher, R. A new approach to variable metric methods. *Comput. J.* **1970**, *13*, 317–322.
- (31) Goldfarb, D. A family of variable-metric methods derived by variational means. *Math. Comp.* **1970**, *24*, 23–26.
- (32) Shanno, D. F. Conditioning of quasi-Newton methods for function minimization. *Math. Comp.* **1970**, *24*, 647–656.
- (33) MOE 2008.10; Chemical Computing Group, Inc.: Montreal, Canada, 2008.
- (34) Wichapong, K.; Lindner, M.; Pianwanit, S.; Kokpol, S.; Sippl, W. Receptor-based 3D-QSAR studies of checkpoint Wee1 kinase inhibitors. *Eur. J. Med. Chem.* **2009**, *44*, 1383–1395.

- (35) Jones, G.; Willett, P.; Glen, R. C. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J. Mol. Biol.* **1995**, *245*, 43–53.
- (36) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727–748.
- (37) Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: II. Parameterization and Validation. *J. Comput. Chem.* **2002**, *23*, 1623–1641.
- (38) Wang, J.; Cieplak, P.; Kollman, P. A. How Well Does a Restrained Electrostatic Potential (RESP) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules. *J. Comput. Chem.* **2000**, *21*, 1049–1074.
- (39) Case, D. A.; Darden, T. A.; Cheatham, I. T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Sestin, M. G.; Sagui, C.; Babin, V.; Kollman, P. A. *AMBER 10*; University of California, San Francisco., 2008.
- (40) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.
- (41) Luo, R.; David, L.; Gilson, M. K. Accelerated Poisson-Boltzmann calculations for static and dynamic systems. *J. Comput. Chem.* **2002**, *23*, 1244–1253.
- (42) Sitkoff, D.; Sharp, K. A.; Honig, B. Accurate Calculation of Hydration Free Energies Using Macroscopic Solvent Models. *J. Phys. Chem.* **1994**, *98*, 1978–1988.
- (43) Wang, R.; Lai, L.; Wang, S. Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 11–26.
- (44) Pearlman, D. A.; Charifson, P. S. Improved scoring of ligand-protein interactions using OWFEG free energy grids. *J. Med. Chem.* **2001**, *44*, 502–511.
- (45) Huang, N.; Shoichet, B. K.; Irwin, J. J. Benchmarking sets for molecular docking. *J. Med. Chem.* **2006**, *49*, 6789–67801.

CI1002153