



Shantheya Balasupramaniam (Autor)  
**Validierung der Leistungsfähigkeit von  
Bewertungsfunktionen zur Vorhersage von  
Bindungsaffinitätsdifferenzen auf Basis von Ligand-  
Paar-Datensätzen**



<https://cuvillier.de/de/shop/publications/8006>

Copyright:  
Cuvillier Verlag, Inhaberin Annette Jentsch-Cuvillier, Nonnenstieg 8, 37075 Göttingen,  
Germany  
Telefon: +49 (0)551 54724-0, E-Mail: [info@cuvillier.de](mailto:info@cuvillier.de), Website: <https://cuvillier.de>



# Inhaltsverzeichnis

<b>Abbildungsverzeichnis .....</b>	<b>XV</b>
<b>Tabellenverzeichnis .....</b>	<b>XVII</b>
<b>Abkürzungen und Symbole .....</b>	<b>XIX</b>
<b>1. Einleitung .....</b>	<b>1</b>
1.1    Bewertungsfunktionen.....	3
1.1.1    Einsatzgebiete .....	3
1.1.2    Anforderungen und Validierungsgrundlagen .....	7
1.1.3    Klassen von Bewertungsfunktionen .....	8
1.1.3.1    Kraftfeldbasierte Bewertungsfunktionen.....	8
1.1.3.2    Empirische Bewertungsfunktionen .....	9
1.1.3.3    Wissensbasierte Bewertungsfunktionen.....	10
1.1.3.4    Bewertungsfunktionen des Maschinellen Lernens .....	11
1.1.3.5    Hybride und Targetspezifische Bewertungsfunktionen .....	11
1.1.4    Consensus Scoring.....	12
1.1.5    Weitere theoretische Ansätze zur Vorhersage der Bindungsaffinität .....	12
1.2    Bindungsaffinität .....	13
1.3    PDBbind-Datenbank .....	14
1.3.1    Aufbau der PDBbind-Datenbank .....	14
1.4    Matched Molecular Pairs .....	17
<b>2. Zielsetzung.....</b>	<b>19</b>
<b>3. Material und Methoden.....</b>	<b>23</b>
3.1    Untersuchte Bewertungsfunktionen .....	23
3.1.1    ASP.....	23
3.1.2    ChemScore.....	24
3.1.3    ChemPLP.....	25
3.1.4    GoldScore .....	25
3.1.5    London dG .....	26
3.1.6    ASE.....	27
3.1.7    Affinity dG.....	27
3.1.8    Alpha HB .....	28
3.1.9    GBVI/WSA dG .....	28



3.2	3-D-Matched Molecular Pair-Datensatz.....	30
3.2.1	Aufbau und Aufbereitung des 3-D-MMP-Datensatzes.....	30
3.2.2	Vorbereitung des 3-D-MMP-Datensatzes .....	32
3.2.3	Scoring des 3-D-MMP-Datensatzes .....	34
3.2.3.1	Scoring durch die MOE-Bewertungsfunktionen.....	34
3.2.3.2	Scoring durch die GOLD-Bewertungsfunktionen .....	35
3.2.4	Messgrößen und Einflussfaktoren des 3-D-MMP-Datensatzes.....	36
3.2.4.1	Bestimmung der Bindungsaffinitätsdifferenz ( $\Delta$ AFFINITY) .....	36
3.2.4.2	Bestimmung der Score-Differenz ( $\Delta$ SCORE) .....	37
3.2.4.3	Bestimmung des CONSENSUS der $\Delta$ SCORE-Werte .....	38
3.2.4.4	Bestimmung der Differenz in der Anzahl schwerer Atome ( $\Delta$ HA) .....	38
3.2.4.5	$\Delta$ HA-bezogene und HA-identische Komplex-Paare des 3-D-MMP-Datensatzes.....	39
3.2.5	Subgruppenanalyse und $\sigma_{\text{krit}}$ -Wert .....	40
3.3	Related Ligand Pair-Datensatz.....	43
3.3.1	Aufbau und Aufbereitung des RLP-Datensatzes.....	43
3.3.2	Vorbereitung des RLP-Datensatzes .....	47
3.3.3	Scoring des RLP-Datensatzes .....	48
3.3.4	Messgrößen und Einflussfaktoren des RLP-Datensatzes .....	48
3.3.4.1	Bestimmung von $\Delta$ AFFINITY, $\Delta$ HA und $\Delta$ SCORE.....	49
3.3.4.2	$\Delta$ HA-bezogene und HA-identische Komplex-Paare des RLP-Datensatzes .....	49
3.3.5	Subgruppenanalyse und $\sigma_{\text{krit}}$ -Wert .....	51
3.4	Positivkontrolle.....	55
3.4.1	CSAR Benchmark .....	55
3.4.2	Verwendung des CSAR-NRC-Datensatzes bei GBVI/WSA dG .....	57
3.4.3	Durchführung der Positivkontrolle.....	58
3.5	Anforderungen an Validierdatensätze .....	60
<b>4.</b>	<b>Ergebnisse und Auswertung .....</b>	<b>63</b>
4.1	Ergebnisse des 3-D-MMP-Datensatzes .....	64
4.1.1	Vorhersagegenauigkeit der $\Delta$ AFFINITY-Werte.....	64
4.1.1.1	Vorhersagegenauigkeit der $\Delta$ AFFINITY-Werte der WF-KOMPLEX-PAARE .....	65
4.1.1.2	Vorhersagegenauigkeit der $\Delta$ AFFINITY-Werte der WH-KOMPLEX-PAARE .....	68
4.1.1.3	Vorhersagegenauigkeit der $\Delta$ AFFINITY-Werte der WF-GO-KOMPLEX-PAARE .....	70
4.1.1.4	Vorhersagegenauigkeit der $\Delta$ AFFINITY-Werte der WH-GO-KOMPLEX-PAARE .....	73
4.1.1.5	Vergleich der Vorhersagegenauigkeiten der $\Delta$ AFFINITY-Werte .....	75



4.1.2	Vorhersage der $\Delta$ HA-bezogenen sowie HA-identischen $\Delta$ AFFINITY-Werte.....	76
4.2	Ergebnisse des RLP-Datensatzes .....	81
4.2.1	Vorhersagegenauigkeit der $\Delta$ AFFINITY-Werte .....	81
4.2.2	Vorhersage der $\Delta$ HA-bezogenen sowie HA-identischen $\Delta$ AFFINITY-Werte.....	85
4.3	Ergebnisse der Positivkontrolle .....	93
<b>5.</b>	<b>Diskussion.....</b>	<b>97</b>
5.1	$\Delta$ AFFINITY-Vorhersagegenauigkeiten .....	97
5.1.1	Einfluss des Wassers und der Geometrieoptimierung.....	98
5.1.2	Einfluss der Molekülgröße.....	101
5.2	Subgruppenanalyse .....	104
5.3	Consensus Scoring .....	105
5.4	Weitere beitragende Faktoren .....	106
5.5	Fazit zum 3-D-MMP- und RLP-Datensatz .....	108
<b>6.</b>	<b>Zusammenfassung und Ausblick.....</b>	<b>111</b>
<b>7.</b>	<b>Summary .....</b>	<b>115</b>
<b>8.</b>	<b>Anhang .....</b>	<b>119</b>
8.1	Beispiel einer gold.conf-Datei .....	174
8.2	Auslesen der Scores der GOLD-Bewertungsfunktionen (get_goldscores.m).....	178
8.3	Zuweisung der Scores (assign_scores.m) .....	179
8.4	Berechnung der $\Delta$ AFFINITY-Vorhersagegenauigkeit (MOE) (prediction_moe.m).....	179
8.5	Berechnung der $\Delta$ AFFINITY-Vorhersagegenauigkeit (GOLD) (prediction_gold.m).....	180
8.6	Bestimmung der Subgruppen des RLP-Datensatzes (generate_subgroups.m).....	181
8.7	Bestimmung des $\Delta$ AFFINITY-Vektors für den CONSENSUS (consensus_vector_affinity.m).....	181
8.8	Bestimmung des CONSENSUS der $\Delta$ SCORE-Werte (prediction_consensus.m) .....	182
8.9	Bestimmung der Relationen zwischen Molekülgröße und Bindungsaffinität (relationen_BA_HA_Score_diss.m).....	182
8.10	Berechnung der Fingerprints (calculate_fingerprint.py) .....	184
8.11	Berechnung der Tanimoto-Ähnlichkeit (calculate_tanimoto.m) .....	186
8.12	Berechnung des RDKit Fingerprints und der Tanimoto-Ähnlichkeit (Auszug aus calculate_RDKit_Tanimoto.py).....	186
<b>9.</b>	<b>Literaturverzeichnis .....</b>	<b>189</b>