

**Table S1a. Model performance with graph attention network based on dataset3**

<b>Test data</b>	<b>RMSE</b>	<b>MAE</b>	<b>SD</b>	<b>R</b>
<b>PDBbind v.2016 core set</b>	1.7594	1.414	1.7226	0.61
<b>PDBbind v.2013 core set</b>	1.848	1.5282	1.8062	0.598

Graph attention model used graph attention network instead of graph convolution network.

**Table S1b. Model performance with capsule network based on dataset3**

<b>Test data</b>	<b>RMSE</b>	<b>MAE</b>	<b>SD</b>	<b>R</b>
<b>PDBbind v.2016 core set</b>	1.7562	1.4214	1.7206	0.612
<b>PDBbind v.2013 core set</b>	1.8338	1.5244	1.796	0.604

Capsule network model used capsule network instead of graph gather layer.