|  |  |
| --- | --- |
| **Table 1** Selection criteria applied to Binding DB for generating the negative dataset | |
| **Selection Criteria** | **Total examples** |
| Number of Compound-Protein Pairs in Binding-DB (June 2021) | 22,782,226 |
| Examples with single-chain protein targets | 2,169,607 |
| Examples with Binding Affinity measured in terms of | 490,605 |
| Examples with valid Uniport identifiers | 487,732 |
| Examples with valid SMILES strings | 485,550 |
| Examples with valid Inhibition constant values | 479,802 |
| Compound-protein pairs with | 4,355 |
| **Final Negative set after removal of duplicate examples** | **3,657** |