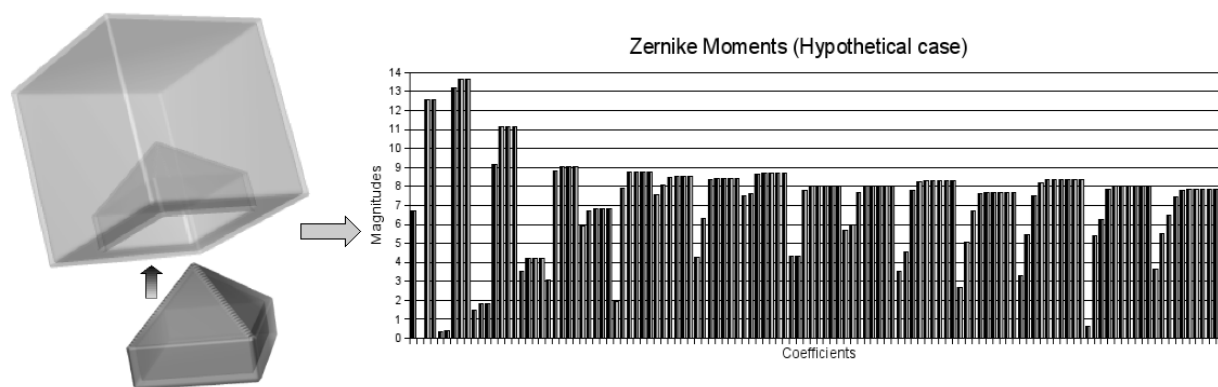


Supplemental Material 1



A hypothetical case where the surface of the triangular pyramid and the depression in the surface of the cube have a lock and key fit, *i.e.* complementary shapes. Magnitudes of the 3D Zernike descriptors for the surfaces of the triangular pyramid (black) and the complementary region of the cube surface (gray) are shown on the right side. 3DZD of the order $n = 20$ is used, which yields 121 coefficients. The spectrum is identical yielding a correlation coefficient of 1.0.

Supplemental Material 2

Comparison of Context shapes, ZDOCK, PATCHDOCK, and VDOCK on the bound-bound test set of ZDOCK Benchmark 2.0.

| Complex | Context Shapes | | ZDOCK | | PatchDock | | VDOCK | |
|---------|----------------|-------|-----------|-------|------------|-------|-------------|-------|
| | Rank | iRMSD | Rank | iRMSD | Rank | iRMSD | Rank | iRMSD |
| 1A2K | 40 | 1.08 | 570 | 2.41 | 300 | 1.47 | 826 | 1.2 |
| 1ACB | 8 | 2.32 | 6 | 0.82 | 10 | 1.6 | 15 | 1.5 |
| 1AHW | 7 | 1.2 | 56 | 1.18 | 40 | 1.55 | 1 | 1.21 |
| 1AK4 | 2925 | 2.08 | 3471 | 1.14 | - | - | 1024 | 2.18 |
| 1AKJ | 265 | 2.15 | 448 | 1.88 | - | - | - | - |
| 1ATN | 49 | 2.1 | 558 | 1.15 | - | - | - | - |
| 1AVX | 10 | 1.76 | 1 | 1.96 | 43 | 2.14 | 812 | 1.8 |
| 1AY7 | 193 | 1.23 | 46 | 1.68 | 24 | 2.07 | 46 | 1.31 |
| 1B6C | 11 | 1.78 | 24 | 1.69 | 40 | 1.92 | 44 | 0.94 |
| 1BGX | 1 | 1.96 | - | - | - | - | - | - |
| 1BJ1 | 1 | 1.05 | 3 | 1.42 | - | - | 58 | 0.73 |
| 1BUH | 61 | 1.55 | 393 | 1.43 | 83 | 1.14 | 2378 | 0.82 |
| 1BVK | 45 | 1.69 | 1087 | 1.43 | 131 | 2.12 | - | - |
| 1BVN | 1 | 1.55 | 10 | 1.24 | 1 | 0.75 | 1 | 0.71 |
| 1CGI | 1 | 1.37 | 1 | 1.12 | 1 | 1.08 | 1 | 0.61 |
| 1D6R | 4 | 1.68 | 35 | 1.04 | - | - | 27 | 0.52 |
| 1DE4 | 13 | 1.21 | 452 | 1.62 | - | - | - | - |
| 1DQJ | 67 | 1.65 | 19 | 2 | 83 | 1.71 | 391 | 0.67 |
| 1E6E | 1 | 1.58 | 58 | 2.06 | 2 | 2.29 | 11 | 1.45 |
| 1E6J | 1337 | 1.92 | 699 | 2.02 | 1706 | 1.43 | 104 | 0.62 |
| 1E96 | 1206 | 1.84 | - | - | 1767 | 1.44 | - | - |
| 1EAW | 1 | 1.41 | 1 | 1.75 | 1 | 0.99 | 2 | 0.81 |
| 1EER | 1 | 1.62 | - | - | 1 | 1.66 | 38 | 1.4 |
| 1EWY | 518 | 2.26 | - | - | 139 | 1.42 | 141 | 0.86 |
| 1EZU | 1 | 1.6 | - | - | 1 | 0.94 | 1 | 1.18 |
| 1F34 | 1 | 1.99 | - | - | 1 | 1.9 | 1 | 1.32 |

| Complex | Context Shapes | | ZDOCK | | PatchDock | | VDOCK | |
|----------------|-----------------------|------|--------------|------|------------------|------|--------------|------|
| 1F51 | 7 | 2.01 | - | - | 1 | 1.92 | 26 | 1.58 |
| 1FAK | 1997 | 1.7 | - | - | - | - | - | - |
| 1FC2 | 7 | 1.85 | 55 | 2.18 | 49 | 1.24 | 42 | 0.52 |
| 1FQJ | 12 | 1.94 | 120 | 1.94 | 248 | 1.48 | 15 | 1.78 |
| 1FSK | 9 | 2.06 | 19 | 1.7 | 218 | 1.57 | 454 | 0.94 |
| 1GCQ | 2 | 1.26 | 382 | 1.81 | - | - | 119 | 0.44 |
| 1GP2 | 53 | 1.86 | - | - | - | - | - | - |
| 1GRN | 1 | 1.84 | 7 | 2.26 | 3 | 1.45 | 1 | 1.08 |
| 1H1V | 14 | 2.37 | 1510 | 2.4 | - | - | - | - |
| 1HE1 | 1 | 1.44 | 7 | 1.67 | 1 | 1.06 | 15 | 0.81 |
| 1HIA | 2 | 1.07 | 1 | 1.7 | 14 | 1.19 | 3 | 1.5 |
| I2M | 6 | 1.36 | 14 | 1.8 | - | - | 1 | 1.48 |
| 1I4D | 104 | 1.42 | 793 | 2.08 | 167 | 1.05 | 695 | 1.68 |
| 1I9R | - | - | 1271 | 2.04 | - | - | 174 | 2.11 |
| 1IB1 | 2 | 1.48 | - | - | - | - | 5 | 2.04 |
| 1IBR | 1 | 2.05 | - | - | - | - | 336 | 1.34 |
| 1IJK | - | - | - | - | - | - | 115 | 0.9 |
| 1IQD | 14 | 1.19 | 55 | 1.83 | - | - | 1 | 0.35 |
| 1JPS | 2 | 1.26 | 23 | 2.3 | 96 | 1.87 | 21 | 0.98 |
| 1K4C | 5 | 0.88 | 30 | 1.16 | 337 | 1.53 | 274 | 0.88 |
| 1K5D | 2 | 2.06 | 10 | 2.11 | - | - | 2245 | 2.03 |
| 1KAC | - | - | 381 | 1.52 | - | - | - | - |
| 1KKL | 226 | 1.67 | - | - | - | - | 946 | 0.91 |
| 1KLU | 1108 | 1.8 | - | - | - | - | - | - |
| 1KTZ | 2280 | 1.41 | - | - | - | - | - | - |
| 1KXP | 3 | 2.17 | - | - | - | - | - | - |
| 1KXQ | 229 | 1.51 | 30 | 1.6 | 29 | 1.63 | 47 | 0.83 |
| 1M10 | - | - | 33 | 2.23 | - | - | - | - |
| 1MAH | 1 | 1.45 | 1 | 1.91 | 1 | 1.27 | 1 | 0.98 |
| 1ML0 | 569 | 1.91 | 75 | 1.94 | 7 | 0.58 | 157 | 1.48 |

| Complex | Context Shapes | | ZDOCK | | PatchDock | | VDOCK | |
|---------------------|-----------------------|------|--------------|------|------------------|------|--------------|------|
| 1MLC | 30 | 1.15 | 1205 | 1.37 | 516 | 1.79 | 1378 | 2.49 |
| 1N2C | 3 | 1.36 | - | - | - | - | - | - |
| 1NCA | 3 | 1.77 | 20 | 1.48 | - | - | 354 | 1.21 |
| 1NSN | - | - | - | - | - | - | 36 | 0.82 |
| 1PPE | 1 | 2.32 | 2 | 1.21 | 1 | 1.03 | 1 | 0.41 |
| 1QA9 | 972 | 1.3 | - | - | - | - | - | - |
| 1QFW | 1247 | 2.21 | 16 | 2.46 | - | - | 14 | 2.1 |
| 2QFW | 38 | 2.13 | 54 | 1.84 | - | - | 1 | 1.02 |
| 1RLB | 311 | 1.63 | - | - | 3143 | 2.32 | 561 | 0.49 |
| 1TMQ | 1 | 2.32 | 8 | 1.79 | 1 | 1.52 | 1 | 0.65 |
| 1UDI | 3 | 1.52 | 1 | 1.5 | 1 | 1.97 | 1 | 0.8 |
| 1VFB | 8 | 1.5 | - | - | - | - | 37 | 1.74 |
| 1WEJ | 496 | 1.25 | 1120 | 1.11 | - | - | 707 | 2.4 |
| 1WQ1 | 1 | 1.14 | 4 | 2.04 | 1 | 0.84 | 1 | 1.77 |
| 2BTF | 4 | 1.13 | 21 | 1.21 | 137 | 1.82 | 40 | 0.94 |
| 2JEL | 56 | 1.4 | 532 | 1.77 | 282 | 1.65 | 1 | 0.79 |
| 2MTA | 21 | 1.45 | 1447 | 2.26 | 115 | 1.71 | 92 | 0.88 |
| 2SIC | 4 | 1.36 | 9 | 1.19 | - | - | 1 | 0.88 |
| 2SNI | 2 | 1.27 | 4 | 2.5 | 13 | 2.1 | 4 | 1.27 |
| 7CEI | 123 | 1.9 | 5 | 2.18 | - | - | 54 | 1.08 |
| Summary | | | | | | | | |
| | Context Shapes | | ZDOCK | | PatchDock | | VDOCK | |
| Rank<100 | 53 | | 37 | | 28 | | 39 | |
| Rank<500 | 61 | | 43 | | 38 | | 50 | |
| Rank<1000 | 64 | | 48 | | 39 | | 56 | |
| Rank<2000 | 69 | | 54 | | 41 | | 58 | |
| | Context Shapes | | ZDOCK | | PatchDock | | VDOCK | |
| Win | 49 | | 10 | | 14 | | 20 | |

Values for the first three blocks, Context Shapes, ZDOCK, and PatchDock are taken from previously

published study (Shentu et al., Proteins 70: 1067-1073, 2008). Cases where either algorithm achieves a better rank are highlighted in bold. “-” indicates that the algorithm was unable to find a hit among the top 3600 predictions.

Supplemental Material 3

The weighting factors for the terms involved in the scoring function were obtained using a genetic algorithm. The training set consisted of 2000 predictions for each protein-protein docking pair (29 bound docking cases). The training samples were chosen randomly to include 75% of the hits ($i\text{RMSD} \leq 2.5\text{\AA}$) produced for each docking pair. Twenty such sets were created and the GA was trained 10 times over each set. Each GA run is performed using a population size of 30 and is repeated for 250 cycles. Fitness of the chromosome/individual is assessed based on the BEDROC value (Eqn. 12 in the main text). The evolutionary scheme uses self-adaptive mutation combined with a blend crossover operator. As, in all 200 (20x10) independent runs were tried, 200 sets of parameters were obtained. These values were then applied to the entire bound docking predictions (used in the training) and the parameter set yielding the highest BEDROC value were selected.

Figure below shows the final BEDROC values for the 200 different sets that were produced. The lowest and highest BEDROC values observed were 0.41 and 0.62, respectively, with a standard deviation of 0.05. For the BEDROC scores examined, 83 sets (41.5%) had values in the range 0.4-0.5, 115 (57.5%) in the range 0.5-0.6, and the remaining two had values greater than 0.6.

