**Table S1**. Dataset of 214 protein-ssDNA complexes identified by their NDB and PDB IDs.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **NDB ID** | **PDB ID** | **NDB ID** | **PDB ID** | **NDB ID** | **PDB ID** |
| NA2119 | 4HQB | NA2115 | 4HQX | NA0891 | 3Q0A |
| NA2201 | 4IQV | NA2028 | 4H5Q | NA0594 | 3NGZ |
| NA2202 | 4IQW | NA1877 | 4FPV | NA0595 | 3NH0 |
| NA2415 | 4KI2 | NA1757 | 4ESV | NA0596 | 3NH1 |
| NA2128 | 4I27 | NA2001 | 4A75 | NA0895 | 3Q22 |
| NA2129 | 4I28 | NA2002 | 4A76 | NA0898 | 3Q23 |
| NA2130 | 4I29 | NA1520 | 3V9S | NA0899 | 3Q24 |
| NA2131 | 4I2A | NA1521 | 3V9U | NA0122 | 3A5U |
| NA2132 | 4I2B | NA1522 | 3V9W | NA0565 | 3N1I |
| NA2133 | 4I2C | NA1523 | 3V9X | NA0566 | 3N1J |
| NA2134 | 4I2E | NA1524 | 3V9Z | NA0567 | 3N1K |
| NA2135 | 4I2F | NA1525 | 3VA0 | NA0568 | 3N1L |
| NA2136 | 4I2G | NA1526 | 3VA3 | 2KN7 | 2KN7 |
| NA2137 | 4I2H | NA1883 | 4A8F | NA0603 | 3NGO |
| NA2322 | 4JRP | NA1884 | 4A8K | NA0369 | 3KQH |
| NA2323 | 4JRQ | NA1885 | 4A8M | NA0370 | 3KQK |
| NA2324 | 4JS4 | NA1886 | 4A8Q | NA0371 | 3KQL |
| NA2325 | 4JS5 | NA1887 | 4A8S | NA0372 | 3KQN |
| NA2247 | 4J1J | NA1888 | 4A8W | NA0373 | 3KQU |
| NA1506 | 3VAF | NA1889 | 4A8Y | NA0289 | 3KJO |
| NA1512 | 3VAG | 2LTT | 2LTT | NA0290 | 3KJP |
| NA1532 | 3VAH | NA1402 | 3UPU | NA0058 | 3I2O |
| NA1533 | 3VAI | NA1375 | 3UDG | NA0059 | 3I3M |
| NA1534 | 3VAJ | NA1472 | 3ULP | NA0060 | 3I49 |
| NA1535 | 3VAK | NA1593 | 3VDY | PD1283 | 3H15 |
| NA1536 | 3VAL | NA1384 | 3VKE | PD1273 | 3GP8 |
| NA1537 | 3VAM | NA1248 | 3SZ5 | PD1274 | 3GPL |
| NA1812 | 4FF1 | NA1564 | 4A15 | NA0014 | 2VTB |
| NA1813 | 4FF2 | NA0978 | 3AUO | PD1138 | 3D2W |
| NA1814 | 4FF3 | NA1323 | 3U3Y | NA0339 | 2VYE |
| NA1816 | 4FF4 | NA1355 | 3U6F | PD1083 | 3C2P |
| NA2084 | 4HID | 2L45 | 2L45 | PD1084 | 3C3L |
| NA2083 | 4HIK | 2L46 | 2L46 | PD1086 | 3C46 |
| NA2082 | 4HIM | NA1373 | 3UGO | PD1153 | 3DLB |
| NA2080 | 4HIO | NA1374 | 3UGP | PD1155 | 3DLH |
| NA2079 | 4HJ5 | NA1376 | 3ZVM | PD1028 | 2Z70 |
| NA2078 | 4HJ7 | NA1039 | 3R8F | PD1108 | 3CMU |
| NA2077 | 4HJ8 | NA1040 | 3RA0 | PD1109 | 3CMW |
| NA2076 | 4HJ9 | NA0556 | 3MXI | PD1021 | 2QFJ |
| NA2075 | 4HJA | NA0558 | 3MXM | PD1045 | 3B39 |
| NA1993 | 4GOP | NA0883 | 3PX7 | PD1010 | 2PY5 |
| NA2114 | 4HQU | NA1014 | 2Y35 | PD0966 | 2OK0 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| PD0915 | 2O19 | PD0681 | 1ZZJ | 1L1V | 1L1V |
| PD0940 | 2O54 | PD0517 | 1S6M | PD0278 | 1KIX |
| PD0941 | 2O59 | PD0596 | 1XF2 | PD0266 | 1K8G |
| PD0942 | 2O5C | PD0597 | 1XJV | PD0206 | 1I8M |
| PD0943 | 2O5E | PD0599 | 1XHZ | PD0203 | 1I7D |
| PD0948 | 2O4I | PD0603 | 1XI1 | PD0150 | 1F0V |
| PD0959 | 2OA8 | PD0485 | 1RFF | PD0144 | 1EYG |
| PD0890 | 2NMV | PD0486 | 1RFI | PD0106 | 2KZZ |
| PD0785 | 2FR4 | PD0487 | 1RG1 | PD0072 | 1D8Y |
| PD0855 | 2DWL | PD0488 | 1RG2 | PR0016 | 1D9D |
| PD0856 | 2DWM | PD0489 | 1RGT | PD0080 | 1D9F |
| PD0726 | 2ES2 | PD0490 | 1RGU | PD0087 | 2UP1 |
| PD0858 | 2I5S | PD0491 | 1RH0 | PD0064 | 1QSL |
| PD0850 | 2I0Q | PD0493 | 1RC8 | PD0074 | 1OTC |
| PD0675 | 1ZTG | PD0494 | 1RPZ | PD0043 | 2PJR |
| PD0746 | 2D7D | PD0495 | 1RRC | PD0014 | 2KFN |
| PD0669 | 1ZM5 | PD0496 | 1OMH | PD0015 | 2KFZ |
| PD0766 | 2FDC | PD0497 | 1OSB | PDE0129 | 1UAA |
| PD0767 | 2FD8 | PD0498 | 1QX0 | PDE0137 | 1KFS |
| PD0768 | 2FDF | PD0499 | 1QZH | PDE0138 | 1KRP |
| PD0769 | 2FDG | PD0500 | 1QZG | PDE0136 | 1KSP |
| PD0770 | 2FDH | PD0501 | 1PGZ | PDO001 | 1JMC |
| PD0771 | 2FDI | PD0502 | 1PO6 | PDE090 | 1NOY |
| PD0772 | 2FDK | PD0503 | 1MW8 | PDE026 | 1LAU |
| PD0745 | 2F55 | PD0504 | 1OVF | PDE023 | 1RBJ |
| PD0688 | 2A0I | PD0505 | 1NOP | PDE0117 | 1RCN |
| PD0695 | 2A6O | PD0506 | 1M07 | PDE013 | 1HUT |
| PD0715 | 2AXY | PD0507 | 1KEG | PDA001 | 1CBV |
| PD0680 | 1ZZI | PD0508 | 1KDH | 1HVO | 1HVO |
|  |  | PDE0116 | 1RTA |  |  |

**Table S2.** Domain-based non-redundant ssDNA-binding domains in the SP group.

|  |  |  |
| --- | --- | --- |
| **CATH Domain ID** | **Domain Definition** | **CATH Classification** |
| 1cbvH01 | 1cbv:H (1-121) | 2.60.40.10 |
| 1cbvL01 | 1cbv:L (1-113) | 2.60.40.10 |
| 1otcA01 | 1otc:A (37-210) | 2.40.50.140 |
| 1otcA02 | 1otc:A (211-326) | 2.40.50.140 |
| 1pgzA01 | 1pgz:A (8-94) | 3.30.70.330 |
| 1qzgA00 | 1qzg:A | 2.40.50.140 |
| 1rbjA00 | 1rbj:A | 3.10.130.10 |
| 1xjvA02 | 1xjv:A (151-299) | 2.40.50.140 |
| 1zziA00 | 1zzi:A | 3.30.1370.10 |
| 1osbA00 | 1osb:A | TBC |
| 2i0qB00 | 2i0q:B | 2.40.200.10 |
| 2o19A04 | 2o19:A (289-416) | 1.10.290.10 |
| 2o5cA01 | 2o5c:A (1-156) | 3.40.50.140 |
| 2o5cA02 | 2o5c:A (157-216,489-659) | 1.10.460.10 |
| 4a75A00 | 4a75:A | 2.40.50.140 |
| 3c2pA07 | 3c2p:A (609-803) | TBC |
| 3c2pA08 | 3c2p:A (804-882,978-1052) | 1.20.140.110 |
| 3d2wA00 | 3d2w:A | 3.30.70.330 |
| 3q0aB01 | 3q0a:B (5-311,883-916) | TBC |
| 3ugoA00 | 3ugo:A | 1.20.120.1810 |
| 3v9xA00 | 3v9x:A | 3.30.420.10 |
| 3c2pA02 | 3c2p:A (312-338,411-454,546-562,917-977) | 3.30.70.2440 |

TBC: To Be Classified, pending cases in CATH.

**Table S3.** Domain-based non-redundant ssDNA-binding domains in the NS group.

|  |  |  |
| --- | --- | --- |
| **CATH Domain ID** | **Domain Definition** | **CATH Classification** |
| 1jmcA01 | 1jmc:A (183-296) | 2.40.50.140 |
| 1jmcA02 | 1jmc:A (297-420) | 2.40.50.140 |
| 1lauE00 | 1lau:E | 3.40.470.10 |
| 1noyB02 | 1noy:B (107-338) | 3.30.420.10 |
| 1uaaA01 | 1uaa:A (2-107,185-273) | 3.40.50.300 |
| 1uaaA02 | 1uaa:A (108-181) | 1.10.10.160 |
| 1uaaB03 | 1uaa:B (274-376,533-640) | 3.40.50.300 |
| 2kn7A00 | 2kn7:A | 1.10.150.20 |
| 2lttA00 | 2ltt:A | TBC |
| 2pjrA04 | 2pjr:A (387-543) | 1.10.486.10 |
| 3a5uA00 | 3a5u:A | 2.40.50.140 |
| 3cmwA01 | 3cmw:A(37-268,1001-1036) | 3.40.50.300 |
| 3dlbA02 | 3dlb:A (21-96) | 3.30.70.2620 |
| 3dlbA03 | 3dlb:A (178-266) | 2.170.260.50 |
| 3gp8A01 | 3gp8:A (153-297) | TBC |
| 3gp8A02 | 3gp8:A (298-492,703-716) | 3.40.50.300 |
| 3gp8A03 | 3gp8:A (493-571,641-702) | 3.40.50.300 |
| 3gplA04 | 3gpl:A (572-640) | 2.30.30.940 |
| 3h15A00 | 3h15:A | 2.40.50.140 |
| 2d7dA01 | 2d7d:A (3-89,117-148,320-338,388-410) | 3.40.50.300 |
| 2d7dA02 | 2d7d:A (90-116,248-319,339-387) | TBC |
| 3kqhA01 | 3kqh:A (189-326) | 3.40.50.300 |
| 3kqhA02 | 3kqh:A (327-430,452-483) | 3.40.50.300 |
| 3kqlA03 | 3kql:A (431-451,484-625) | 1.10.820.10 |
| 3n1iA00 | 3n1i:A | 2.30.31.10 |
| 3sz5A00 | 3sz5:A | 3.90.320.10 |
| 3ulpC00 | 3ulp:C | 2.40.50.140 |
| 3zvmA01 | 3zvm:A (143-337) | 3.40.50.1000 |
| 4a15A04 | 4a15:A (410-615) | 3.40.50.300 |
| 4a8mQ02 | 4a8m:Q (37-91,518-606) | 1.10.490.60 |
| 4fpvA00 | 4fpv:A | 3.60.10.10 |
| 4h5qB00 | 4h5q:B | TBC |
| 4hqbA00 | 4hqb:A | TBC |
| 4i2fA03 | 4i2f:A (302-440) | 3.30.460.10 |
| 4j1jA02 | 4j1j:A (133-235) | 1.10.472.180 |
| 4j1jC01 | 4j1j:C (2-132) | 1.20.142.20 |
| 4jrpA01 | 4jrp:A (7-201) | 3.30.420.10 |
| 4jrpB02 | 4jrp:B (214-313,314-357) | 3.30.1520.20 |
| 4jrpB03 | 4jrp:B (358-420) | 1.20.1280.70 |
| 3dlhB05 | 3dlh:B (320-466) | 3.40.50.2300 |
| 3dlhA01 | 3dlh:A (3-20,97-177,267-315,580-584) | 3.30.530.60 |
| 3dlhA04 | 3dlh:A (316-319,467-579,585-685) | 3.30.420.10 |

TBC: To Be Classified, pending cases in CATH.

**Table S4.** The non-redundant holo-apo pair dataset.

|  |  |  |  |
| --- | --- | --- | --- |
| **Dataset** | **Holo domain ID** | **Holo domain definition** | **Apo homolog** |
| SP | 1cbvL01 | 1cbv:L (1-113) | 1nbv:L (1-113) |
| 1pgzA01 | 1pgz:A (8-94) | 1l3k:A (8-94) |
| 1rbjA00 | 1rbj:A | 1kf5:A |
| 1zziA00 | 1zzi:A | 1zzk:A |
| 2fr4B01 | 2fr4:B (1-120) | 1xf3:B (1-120) |
| 2o19A01 | 2o19:A (1-156) | 1d6m:A (1-156) |
| 2o19A02 | 2o19:A (157-216,489-658) | 1d6m:A (157-216,489-653) |
| 2o19A04 | 2o19:A (289-416) | 1d6m:A (289-416) |
| 3c2pA02 | 3c2p:A (312-338,411-454,546-  562,917-977) | 2po4:A (312-338,411-454,546-  562,917-977) |
| 3c2pA07 | 3c2p:A (609-803) | 2po4:A (609-803) |
| 3c2pA08 | 3c2p:A (804-882,978-1052) | 2po4:A (804-882,978-1052) |
| 3q0aB01 | 3q0a:B (5-311,883-916) | 2po4:A (5-311,883-916) |
| 3ugoA00 | 3ugo:A | 1ku2:A |
| 4a75A00 | 4a75:A | 3ulj:A |
| NS | 1jmcA01 | 1jmc:A (183-296) | 1fgu:A (183-296) |
| 1jmcA02 | 1jmc:A (297-420) | 1fgu:A (297-420) |
| 1lauE00 | 1lau:E | 1udg:A |
| 1noyB02 | 1noy:B (107-338) | 1noz:B (107-338) |
| 2kn7A00 | 2kn7:A | 2aq0:A |
| 2lttA00 | 2ltt:A | 2ltd:A |
| 2pjrA01 | 2pjr:A (11-115,194-282) | 1pjr:A (11-115,194-282) |
| 2pjrA03 | 2pjr:A (283-385,546-548) | 1pjr:A (283-385,546-548) |
| 2pjrA04 | 2pjr:A (387-543) | 1pjr:A (387-543) |
| 2pjrB01 | 2pjr:B (571-646) | 1pjr:A (571-646) |
| 3a5uA00 | 3a5u:A | 1x3e:A |
| 3cmwA09 | 3cmw:A (4037-4268) | 2reb:A (37-268) |
| 3gp8A01 | 3gp8:A (153-297) | 3e1s:A (153-297) |
| 3gp8A02 | 3gp8:A (298-492,703-716) | 3e1s:A (298-492,703-716) |
| 3gp8A03 | 3gp8:A (493-571,641-702) | 3e1s:A (493-571,641-702) |
| 3gplA04 | 3gpl:A (572-640) | 3e1s:A (572-640) |
| 3h15A00 | 3h15:A | 3ebe:A |
| 3kqhA01 | 3kqh:A (189-326) | 5e4f:A (189-326) |
| 3kqhA02 | 3kqh:A (327-430,452-483) | 5e4f:A (327-430,452-483) |
| 3kqlA03 | 3kql:A (431-451,484-625) | 5e4f:A (431-451,484-625) |
| 3n1iA00 | 3n1i:A | 3n1h:A |
| 3sz5A00 | 3sz5:A | 3syy:A |
| 3zvmA01 | 3zvm:A (143-337) | 3zvl:A (143-337) |
| 4a8mQ02 | 4a8m:Q (37-91,518-606) | 1hhs:A (37-91,518-606) |
| 4h5qB00 | 4h5q:B | 4h5m:B |
| 4i2fA03 | 4i2f:A (302-440) | 1jms:A (302-440) |
| 4jrpA01 | 4jrp:A (7-201) | 3c95:A (7-201) |
| 4jrpB03 | 4jrp:B (358-420) | 3c95:A (358-420) |
| 4jrqB02 | 4jrq:B (214-313,314-359) | 3c95:A (214-313,314-359) |