

Supplementary Material

Supplementary Tables

Supplementary Table I

The list of protein chains of DBP-123 dataset (PDB id_Chain id)

1A02_J	1EFA_A	1IGN_A	1L3L_A	1P7D_A	1TDZ_A	2BNW_A	2O5C_A	3C2I_A	3FSI_A
1A31_A	1EGW_A	1J1V_A	1LLM_C	1P7H_L	1W0U_A	2BSQ_E	2PYJ_A	3CMW_A	3G9M_A
1A3Q_A	1EWN_A	1J3E_A	1LMB_3	1PP7_U	1WB9_A	2E52_A	2QOJ_Z	3COQ_A	3GFI_A
1AIS_B	1EYG_A	1JE8_A	1LQ1_A	1QRV_A	1X9M_A	2EX5_A	2R1J_L	3CVU_A	3GNA_A
1B01_A	1FLO_A	1JEY_A	1MNN_A	1QZH_A	1XO0_A	2EZV_A	2RBA_A	3DFX_A	3GOX_A
1BC8_C	1FZP_B	1JFI_A	1MUS_A	1R7M_A	1XYI_A	2FR4_B	2RBF_A	3DLB_A	3H0D_A
1BDT_A	1GD2_E	1JFI_B	1NKP_A	1REP_C	1YRN_A	2H7G_X	2VLA_A	3DVO_A	
1BF5_A	1GDT_A	1JKO_C	1ODH_A	1RH6_A	1YTF_C	2HDD_A	2VOA_A	3ERE_D	
1BL0_A	1GU4_A	1K78_A	1OE4_A	1RIO_H	1ZS4_A	2HEO_A	2W7N_A	3EYI_A	
1CEZ_A	1H6F_A	1KX5_A	1ORN_A	1RXW_A	1ZX4_A	2I06_A	2Z3X_A	3F27_D	
1CF7_B	1HLV_A	1KX5_B	1OSB_A	1SA3_A	2A07_F	2IRF_G	3BM3_A	3F2B_A	
1D02_A	1I3J_A	1KX5_C	1OZJ_A	1SKN_P	2A3V_A	2IS6_A	3BRD_A	3FDE_A	
1DC1_A	1IC8_A	1KX5_D	1P71_A	1TC3_C	2AXY_A	2O4A_A	3BS1_A	3FDQ_A	

Supplementary Table II

The list of protein chains of HOLO-83 and APO-83 datasets (PDB id_Chain id)

HOLO-83	APO-83								
1A73_A	1EVX_A	1MJO_A	1CMC_A	1ZRF_A	1G6N_A	2IH2_A	1AQJ_A	2YVH_A	2YVE_A
1B3T_A	1VHI_A	1NOP_A	1MU7_A	2AC0_A	2OCJ_A	2IHN_A	1TFR_A	3B39_A	1EQN_A
1DFM_A	1ES8_A	1ODG_A	1VSR_A	2BZF_A	1CI4_A	2ITL_A	2ITJ_A	3BEP_A	3D1G_A
1EMH_A	3FCI_A	1QNA_A	1VOK_A	2C5R_F	2BNK_B	2IVK_A	1OUO_A	3BKZ_A	3I3Q_A
1EON_A	1RVE_A	1QPI_A	2VKE_A	2C7P_A	1HMY_A	2NQJ_A	2NQH_A	3C25_A	3BVQ_A
1ESG_A	1BAM_A	1R8D_A	1JBG_A	2DNJ_A	3DNI_A	2OAA_A	2OA9_A	3C46_A	2PO4_A
1F4K_A	1J0R_A	1SAX_A	1OKR_A	2DP6_A	2D3Y_A	2ODI_A	2ODH_A	3CLC_A	3G5G_A
1FOK_A	2FOK_A	1SXQ_A	2BGU_A	2DTU_A	1IH7_A	2OFI_A	2OFK_A	3E6C_C	3E6D_A
1GXP_A	1GXQ_A	1T9I_A	2O7M_A	2E1C_A	1RI7_A	2Q2U_A	1P8L_A	3EI2_B	3EI3_B
1H9D_A	1E50_R	1TRO_A	3WRP_A	2FCC_A	2END_A	2QSH_A	2QSF_A	3GX4_X	3GVA_A
1IAW_A	1EV7_A	1U8R_A	1FX7_A	2FIO_A	2FIP_A	2RGR_A	1BJT_A	3GZ6_A	3GZ5_A
1JEY_B	1JEQ_B	1UUT_A	1M55_A	2FKC_A	1YNM_A	2VE9_A	2VE8_A	3I8D_A	1ZBF_A
1JJ4_A	1F9F_A	1V15_A	1V13_A	2FR4_A	1XF4_L	2VJV_A	2A6M_A	3PVI_A	1K0Z_A
1JT0_A	3BT9_A	1WTE_A	1WTD_A	2G1P_A	2ORE_F	2VS7_A	1B24_A	6CRO_A	5CRO_A
1JX4_A	3FDS_A	1XPX_A	1MIJ_A	2GIG_A	2AUD_A	2W36_A	3HD0_A	9ICW_A	1BPPE_A
1K3X_A	1Q39_A	1Y8Z_A	1XV5_A	2HAX_A	1C9O_A	2W42_A	1W9H_A		
1M3Q_A	1KO9_A	1Z63_A	1Z6A_A	2HHV_A	1XWL_A	2WIW_B	2WCZ_B		

Supplementary Table III

The list of protein chains of PDNA-62 for performance comparison between our method and DISPLAR (PDB id_Chain id)

1A02_F	1BF5_A	1CMA_A	1GCC_A	1HWT_D	1MDY_A	1PAR_B	1REP_C	1UBD_C	2DRP_D
1A02_J	1BHM_A	1D02_A	1GDT_A	1IF1_A	1MEY_F	1PDN_C	1SRS_A	1XBR_A	2GLI_A
1A02_N	1BL0_A	1D66_A	1HCQ_A	1IGN_A	1MHD_A	1PER_L	1SVC_P	1YRN_A	2HDC_A
1A74_A	1C0W_B	1DP7_P	1HCR_A	1IHF_A	1MNM_A	1PNR_A	1TC3_C	1YRN_B	3CRO_L
1AAY_A	1CDW_A	1ECR_A	1HDD_C	1IHF_B	1MNM_C	1PUE_E	1TF3_A	1YSA_C	
1AZQ_A	1CF7_A	1FJL_A	1HLO_A	1J59_A	1MSE_C	1PVL_B	1TRO_A	1YUI_A	
1B3T_A	1CJG_A	1GAT_A	1HRY_A	1LMB_4	1OCT_C	1PYI_A	1TSR_A	2BOP_A	

Supplementary Table IV

The P-values of the comparison of packing density at different cutoff distance of definition of DNA-binding residues and packing density. The average packing density values of DNA-binding and non-binding residues were obtained on DBP-123. The P-values were computed from the Kolmogorov-Smirnov test.

Interface threshold (Å)	PD threshold (Å) ^a	Average PD in binding residues	Average PD in non-binding residues	P-value
3.5	5.5	1.73	1.87	2.2E-05
	6.0	2.78	3.03	4.4E-06
	6.5	3.89	4.18	6.1E-06
4.0	5.5	1.74	1.87	1.4E-05
	6.0	2.77	3.04	9.0E-09
	6.5	3.91	4.19	7.6E-06
4.5	5.5	1.73	1.88	4.5E-06
	6.0	2.77	3.04	4.0E-11
	6.5	3.91	4.19	6.1E-07
5.0	5.5	1.73	1.88	5.9E-07
	6.0	2.78	3.05	2.8E-11
	6.5	3.92	4.20	3.2E-08

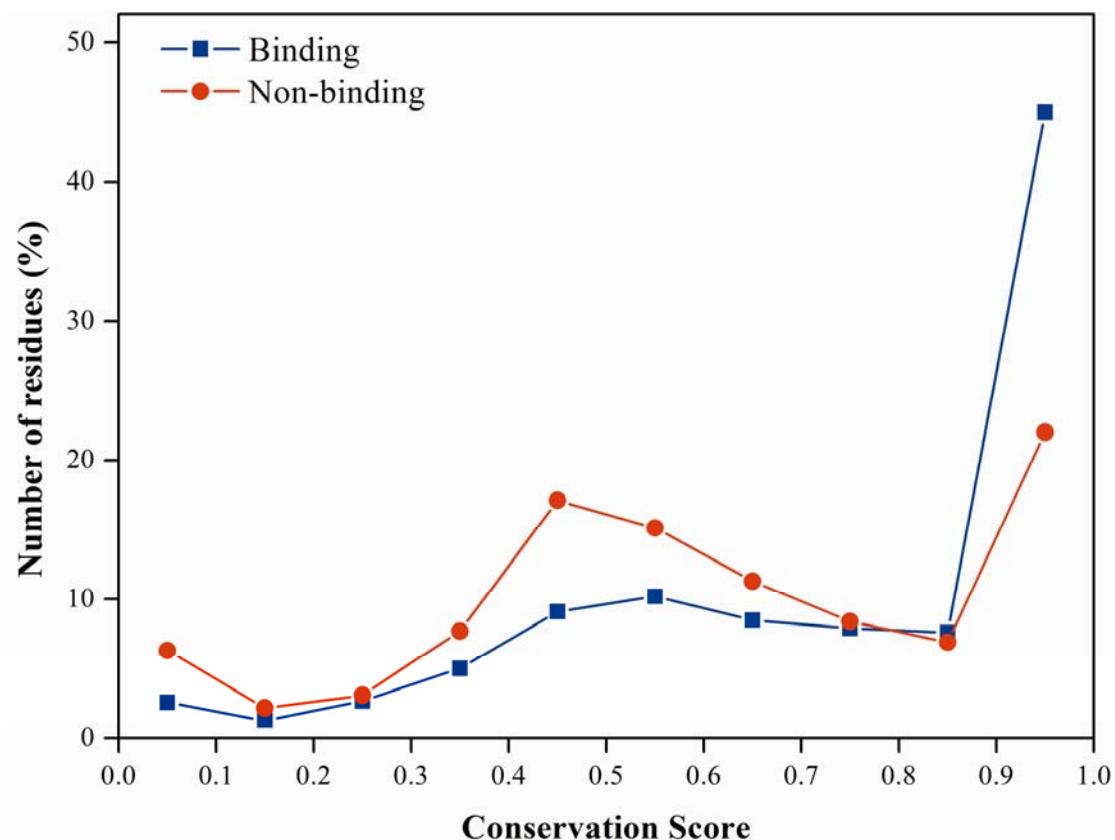
^aPD stands for pack density.

Supplementary Table V

Performance Comparison of the structural windows based on nearest spatial surface neighbors in our method and the structural window derived from closest spatial neighbors adopted by DISPALR.¹ These comparisons were performed on the data sets of DBP-123, HOLO-83 and APO-83, respectively.

Dataset	Window	SN (%)	SP (%)	PR(%)	F₁	AUC
DBP-123	All	74.1±6.5	73.9±3.7	35.8±6.5	0.479±0.068	0.808
	Surface	76.7±4.3	74.9±3.7	37.5±6.7	0.500±0.068	0.822
HOLO-83	All	67.0±20.7	76.2±14.3	38.7±19.9	0.459±0.180	0.791
	Surface	68.9±18.5	77.9±13.8	41.1±20.0	0.483±0.177	0.807
APO-83	All	67.5±20.4	76.7±15.1	38.7±20.2	0.457±0.181	0.790
	Surface	70.1±18.0	77.8±15.1	41.1±19.8	0.482±0.174	0.810

Supplementary Figures



Supplementary Figure 1

The relative frequency of DNA-binding and non-binding residues on the different range of conservation scores on DBP-123.

References

1. Tjong H, Zhou HX. DISPLAR: an accurate method for predicting DNA-binding sites on protein surfaces. Nucleic Acids Res 2007;35(5):1465-1477.