

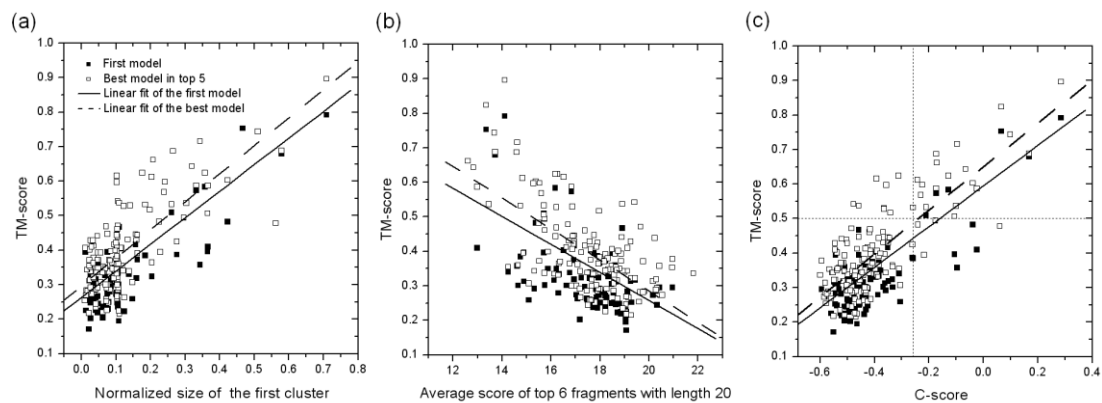
***Ab Initio* Structure Prediction for *Escherichia coli*: Towards  
Genome-wide Protein Structure Modeling and Fold Assignment**

Dong Xu and Yang Zhang

**SUPPLEMENTAL INFORMATION**

**Table S1.** Domain parsing results of 61 Hard multiple domain *E. coli* sequences.

UniProt Entry	Domains	UniProt Entry	Domains
PTPD_ECOLI	1-100,101-263	DIND_ECOLI	1-165,166-274
FLGI_ECOLI	1-159,160-365	FLK_ECOLI	1-200,201-331
FTSX_ECOLI	1-107,108-352	HEMX_ECOLI	1-193,194-393
HTRL_ECOLI	1-146,147-285	LGT_ECOLI	1-197,198-291
MALM_ECOLI	1-221,222-306	PTND_ECOLI	1-220,221-283
MLAA_ECOLI	1-154,155-251	MOTA_ECOLI	1-190,191-295
NDPA_ECOLI	1-176,177-335	PQIB_ECOLI	1-153,154-409,410-546
RECT_ECOLI	1-91,92-269	RFC_ECOLI	1-190,191-388
PTHB_ECOLI	1-228,229-319	TTDA_ECOLI	1-188,189-303
UPPP_ECOLI	1-169,170-273	WCAD_ECOLI	1-268,269-405
YADD_ECOLI	1-160,161-300	YAFZ_ECOLI	1-140,141-273
YAHE_ECOLI	1-168,169-287	Y AHL_ECOLI	1-83,84-271
YBFL_ECOLI	1-103,104-287,288-378	YBJX_ECOLI	1-137,138-330
YCJF_ECOLI	1-163,164-353	YDAU_ECOLI	1-132,133-285
YDCC_ECOLI	1-100,101-287,288-378	YDFE_ECOLI	1-127,128-255
YDIU_ECOLI	1-284,285-478	YEEA_ECOLI	1-59,60-352
YEER_ECOLI	1-82,83-242,243-510	YFAA_ECOLI	1-128,129-375,376-562
YFAD_ECOLI	1-149,150-299	YFBP_ECOLI	1-166,167-282
YFCI_ECOLI	1-158,159-296	YFDF_ECOLI	1-103,104-162,163-352
YFDI_ECOLI	1-116,117-270,271-443	YFJH_ECOLI	1-146,147-318
YFJN_ECOLI	1-166,167-357	YFJQ_ECOLI	1-142,143-273
YFJW_ECOLI	1-256,257-374,375-567	YGGM_ECOLI	1-145,146-335
YHGA_ECOLI	1-131,132-292	YHHI_ECOLI	1-102,103-219,320-378
YHIJ_ECOLI	1-151,152-277,278-540	YHIL_ECOLI	1-166,167-282,283-535
YICC_ECOLI	1-77,78-287	YIHF_ECOLI	1-50,51-295,296-476
YJCZ_ECOLI	1-156,157-292	YJEJ_ECOLI	1-182,183-289
YJIP_ECOLI	1-160,161-306	YJJP_ECOLI	1-76,77-256
YLBF_ECOLI	1-101,102-271	YNCI_ECOLI	1-99,100-289,290-374
YNEE_ECOLI	1-204,205-304	YNEK_ECOLI	1-213,214-371
YNJI_ECOLI	1-159,160-346	YPJB_ECOLI	1-90,91-263
YTFI_ECOLI	1-127,128-312		



**Figure S1.** TM-score of the final models versus parameters calculated from the QUARK simulations on the 145 benchmark proteins. The solid and open squares are the first and the best in top 5 models, respectively. (a) Normalized size of the first cluster. (b) Average score of top 6 fragments with length 20. (c) C-score.