Table S1. Statistics on the confidence and quality of models included in eRepo-ORP. Columns labeled "All data" report mean values ±standard deviation and the number of cases for all DrugBank and Orphanet structures as well as PDB→DrugBank and DrugBank→Orphanet alignments, whereas "Pocket matches" correspond to those pairs of DrugBank and Orphanet proteins producing statistically significant pocket alignments.

Dataset	Score[e]	All data		Pocket matches	
		mean ±std.	*# of cases*	*mean ±std.*	*# of cases*
DrugBank[a]	GDT-score	0.71 ±0.17	348	0.75 ±0.15	282
	Pocket conf.	0.87 ±0.18		0.87 ±0.19	
Orphanet[b]	GDT-score	0.68 ±0.15	922	0.69 ±0.14	761
	Pocket conf.	0.82 ±0.22		0.82 ±0.22	
PDB→DrugBank[c]	TM-score	0.65 ±0.14	2,012	0.68 ±0.18	744
	TC	0.49 ±0.27		0.65 ±0.30	
DrugBank→Orphanet[d]	TM-score	0.26 ±0.08	320,714	0.27 ±0.10	18,145
	eMS-score	0.13 ±0.20		0.91 ±0.11	

[a] Structure models generated for DrugBank proteins by eThread with binding sites predicted by eFindSite. [b] Structure models generated for Orphanet proteins by eThread with binding sites predicted by eFindSite. [c] Alignment between DrugBank structures and holo-templates selected from the PDB to position drugs within binding sites. [d] Pairs of DrugBank and Orphanet proteins. [e] GDT-score is estimated by ModelEvaluator, pocket confidence is estimated by eFindSite for the top-ranked drug-biding sites, TM-score measures the global structure similarity, TC is the Tanimoto coefficient reported by kcombu, eMS-score is the pocket similarity reported by eMatchSite.