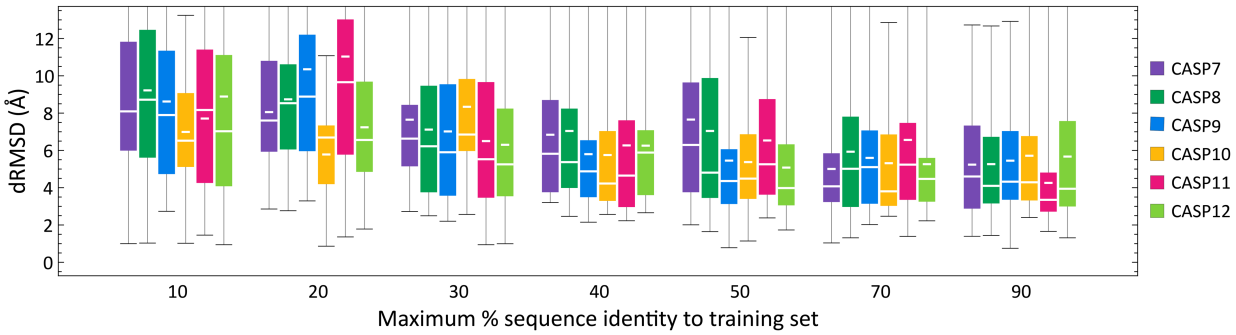

Figure S1. RGN performance based on percent sequence identity, related to Figure 3.
Distribution of RGN dRMSDs of ProteinNet validation sets grouped by maximum % sequence identity to training set and broken down by each CASP (medians are wide white lines, means are short white lines).

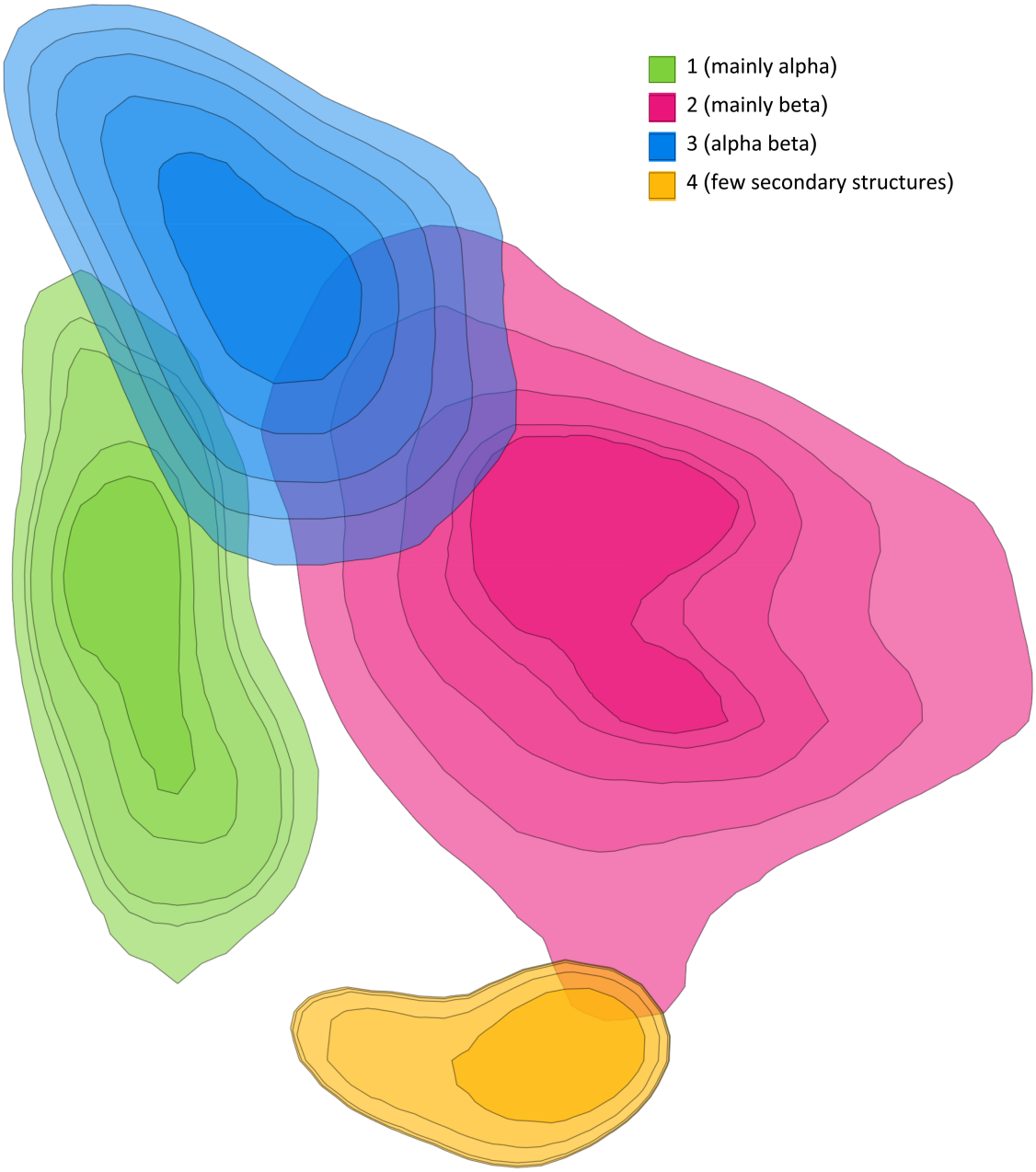

Figure S2. RGN representation of CATH classes, related to Figure 5.
Contour plots of the topmost CATH classes projected onto RGN latent space.

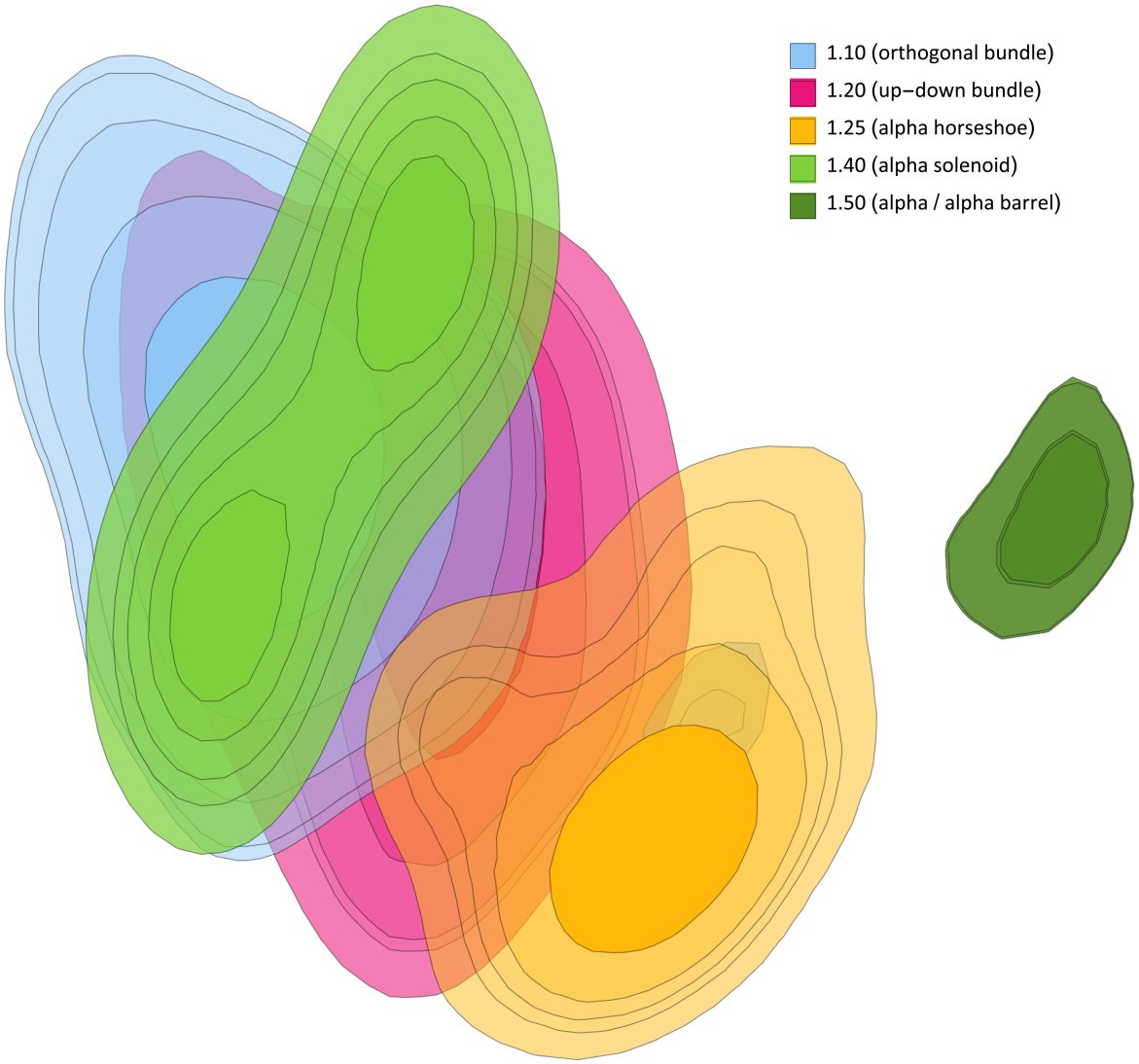

Figure S3. RGN representation of "Mainly Alpha" CATH classes, related to Figure 5.
Contour plots of subcategories in the "Mainly Alpha" CATH class projected onto RGN latent space.

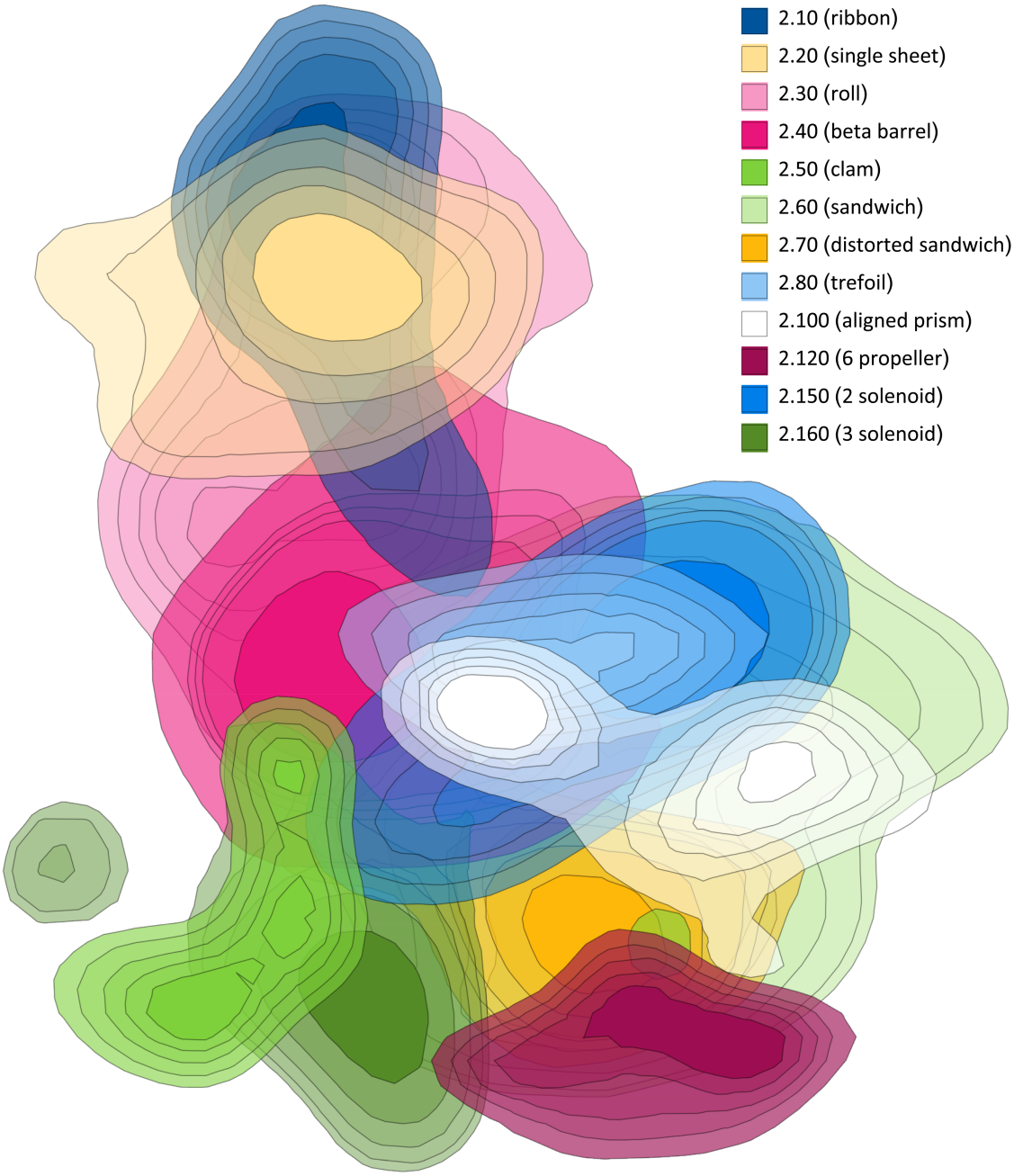

Figure S4. RGN representation of "Mainly Beta" CATH classes, related to Figure 5.
Contour plots of subcategories in the "Mainly Beta" CATH class projected onto RGN latent space.

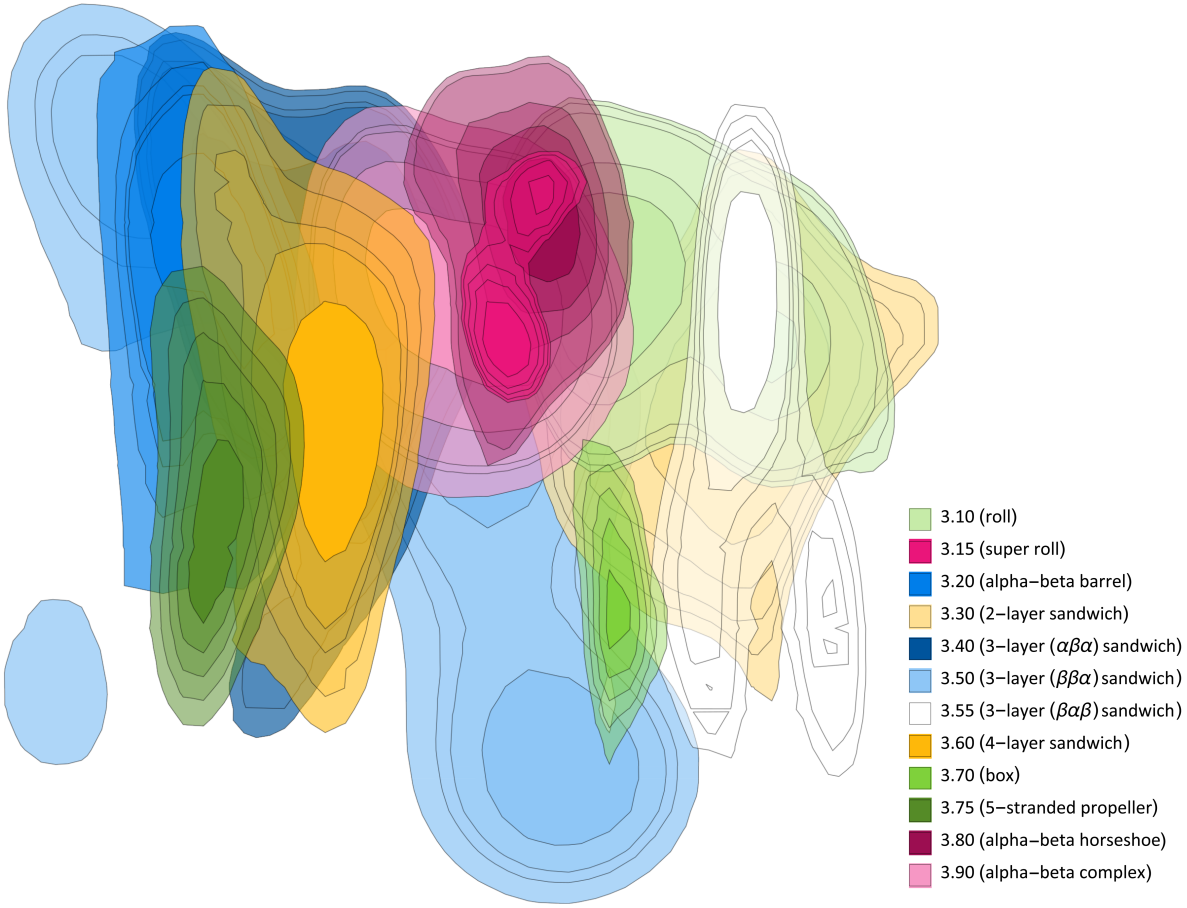

Figure S5. RGN representation of "Alpha Beta" CATH classes, related to Figure 5.
Contour plots of subcategories in the "Alpha Beta" CATH class projected onto RGN latent space.

		FM (novel folds) test set (Å)						TBM (known folds) test set (Å)					
		CASP12	CASP11	CASP10	CASP9	CASP8	CASP7	CASP12	CASP11	CASP10	CASP9	CASP8	CASP7
Training set	PN7	+0.9	+0.3	+1.1	+1.0	+1.8	0	+1.7	+1.8	+0.9	+1.5	+0.4	0
	PN8	+0.6	+0.2	+1.2	+0.3	0		+1.4	+1.0	+0.2	+0.9	0	
	PN9	0	+0.7	+0.8	0			+0.6	+0.6	0	0		
	PN10	+0.5	+1.2	0				+0.6	0	0			
	PN11	+0.2	0					+0.1	0				
	PN12	0						0					

Table S1. Effect of dataset size on RGN accuracy, related to Table 1.

RGNs trained on ProteinNet (PN) training set X were tested on all CASP test sets subsequent to X (e.g. RGN trained on ProteinNet 7 was tested on CASP 8-12) to assess the effect of data set size on model accuracy. Numbers shown are differences in average dRMSD (lower is better) relative to RGNs trained and tested on matching data sets (i.e. trained on ProteinNet X and tested on CASP X.)

	FM (novel folds) category (TM score)						TBM (known folds) category (TM score)					
	CASP7	CASP8	CASP9	CASP10	CASP11	CASP12	CASP7	CASP8	CASP9	CASP10	CASP11	CASP12
RGN	0.27	0.36	0.28	0.25	0.28	0.29	0.49	0.50	0.48	0.48	0.47	0.43
1st Server	**0.33**	**0.37**	**0.32**	**0.30**	**0.29**	**0.35**	**0.72**	**0.72**	**0.71**	**0.69**	**0.66**	**0.70**
2nd Server	0.30	0.33	0.32	0.29	0.27	0.33	0.71	0.70	0.71	0.68	0.66	0.70
3rd Server	0.29	0.31	0.30	0.27	0.26	0.31	0.71	0.70	0.70	0.68	0.65	0.70
4th Server	0.27	0.25	0.29	0.27	0.25	0.31	0.70	0.69	0.70	0.68	0.64	0.68
5th Server	0.24	0.24	0.28	0.26	0.22	0.30	0.68	0.69	0.70	0.67	0.64	0.68

Table S2. Comparative accuracy of RGNs using TM score, related to Table 1.

The average TM score (higher is better, range is between 0 and 1) achieved by RGNs and the top five servers at each CASP is shown for the novel folds **(left)** and known folds **(right)** categories. Numbers are based on common set of structures predicted by top 5 servers during each CASP. A different RGN was trained for each CASP, using the corresponding ProteinNet training set containing all sequences and structures available prior to the start of that CASP.

ProteinNet 7	Iteration	1,000	5,000			
	dRMSD (Å)	14	13.6			
ProteinNet 8	Iteration	1,000	5,000	20,000	50,000	
	dRMSD (Å)	13.4	13.2	12.6	12	
ProteinNet 9	Iteration	1,000	5,000	20,000	50,000	100,000
	dRMSD (Å)	13	12.7	12.2	11.2	10.3
ProteinNet 10	Iteration	1,000	5,000	20,000	50,000	100,000
	dRMSD (Å)	12.8	12.3	11.5	10.7	9.4
ProteinNet 11	Iteration	1,000	5,000	10,000	100,000	150,000
	dRMSD (Å)	13.7	13.5	13.2	12.1	11.4
ProteinNet 12	Iteration	1,000	5,000	20,000	50,000	100,000
	dRMSD (Å)	13.5	12.6	12.2	11.4	10.6

Table S3. Validation set milestones for training RGNs, related to Table 1.

RGN validation performance was monitored during training, and if the shown accuracy milestones were not achieved by the given iteration number, training was terminated and a new model started.