

Supplementary Table S3: Performance of method/parameter combinations on the CSA dataset.
 Within each group of predictors, the one with the highest AUC was awarded a point. In the case of ties, the point would be evenly split between the methods which performed equally well.

	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V	% win	
Sequence	1	.	.	$\frac{1}{2}$	7.5	
Sequence + bin	1	1	1	.	1	1	.	1	1	1	1	1	.	1	1	$\frac{1}{2}$	1	.	.	.	82.5	
FEATURE	1	.	1	10
FEATURE 6-12-18	1	1	1	1	1	1	1	1	1	1	1	1	.	1	.	1	1	1	1	1	90	
FEATURE 18
No reduction	1	.	.	$\frac{1}{2}$	1	1	.	17.5	
BLOSUM 15	1	$\frac{1}{2}$	1	1	1	.	.	1	27.5	
BLOSUM 10	.	1	$\frac{1}{2}$	1	.	.	.	12.5	
BLOSUM 8	.	.	1	1	1	15	
BLOSUM 6	1	$\frac{1}{2}$	7.5	
BLOSUM 5	
BLOSUM 4	$\frac{1}{2}$	2.5	
BLOSUM 3	1	1	10	
BLOSUM 2	$\frac{1}{2}$	1	7.5	
Unlabeled	
No reduction, norm	.	$\frac{1}{2}$.	$\frac{1}{2}$	1	1	.	1	1	1	1	1	1	.	.	1	.	1	1	.	60	
BLOSUM 15, norm	.	.	1	$\frac{1}{2}$.	.	1	1	17.5	
BLOSUM 10, norm	.	$\frac{1}{2}$	1	.	.	.	7.5	
BLOSUM 8, norm	1	1	10	
BLOSUM 6, norm	
BLOSUM 5, norm	
BLOSUM 4, norm	1	5	
BLOSUM 3, norm	
BLOSUM 2, norm	
Unlabeled, norm	
Selected Sequence	1	1	1	15	
Selected FEATURE	1	.	.	1	.	.	1	1	1	.	.	1	30	
Selected kernel	1	5	
Selected norm kernel	.	1	1	.	1	1	.	1	1	.	1	1	1	1	.	50		
Best Sequence	1	1	10	
Best FEATURE	1	.	.	1	1	1	.	.	.	20	
Best kernel	1	5	
Best norm kernel	.	1	1	.	1	1	1	1	.	.	1	1	.	1	.	.	1	1	1	65		