Supplementary Table S3. Benchmark results for the recall and precision of CS prediction in Sec/SPII in four tolerance windows. The PEFT-SP (LoRA) represents PEFT-SP using LoRA with the ESM2-3B backbone. The bold value represents the highest recall or precision among the predictors in a particular tolerance window. The n.d. represents the model was not trained on the data. The * denotes the performance reported in the SignalP 6.0.

	Archaea				**Gram-negative bacteria**				**Gram-positive bacteria**			
Method	**± 0**	**± 1**	**± 2**	**± 3**	**± 0**	**± 1**	**± 2**	**± 3**	**± 0**	**± 1**	**± 2**	**± 3**
CS recall												
PEFT-SP (LoRA)	0.764	0.764	0.764	0.764	**0.858**	**0.895**	**0.900**	**0.908**	0.813	0.832	0.852	0.860
SignalP 6.0 retrained	**0.806**	**0.806**	**0.806**	**0.806**	0.855	0.886	0.892	0.898	0.766	0.779	0.794	0.812
SignalP 6.0 *	0.778	0.778	0.778	0.778	0.852	0.852	0.856	0.864	**0.875**	**0.883**	**0.883**	**0.883**
SignalP 5.0 *	0.778	0.778	0.778	0.778	0.895	0.895	0.895	0.907	0.900	0.900	0.900	0.900
LipoP *	0.778	0.778	0.778	0.778	0.837	0.837	0.837	0.837	0.700	0.700	0.700	0.700
PRED - LIPO *	0.556	0.556	0.556	0.556	0.646	0.646	0.646	0.646	0.767	0.767	0.767	0.767
SPEPlip *	n.d.	n.d.	n.d.	n.d.	0.887	0.887	0.891	0.891	0.850	0.850	0.850	0.850
CS precision												
PEFT-SP (LoRA)	0.744	0.744	0.744	0.744	0.707	0.736	0.743	0.750	0.715	0.734	0.755	0.763
SignalP 6.0 retrained	**0.883**	**0.883**	**0.883**	**0.883**	0.708	0.735	0.743	0.747	0.671	0.684	0.700	0.710
SignalP 6.0 *	0.583	0.583	0.583	0.583	0.913	0.913	0.917	0.925	0.929	0.938	0.938	0.938
SignalP 5.0 *	0.583	0.583	0.583	0.583	0.895	0.895	0.895	0.907	0.931	0.931	0.931	0.931
LipoP *	0.636	0.636	0.636	0.636	0.951	0.951	0.951	0.951	0.955	0.955	0.955	0.955
PRED - LIPO *	0.714	0.714	0.714	0.714	**0.954**	**0.954**	0.954	0.954	0.939	0.939	0.939	0.939
SPEPlip *	n.d.	n.d.	n.d.	n.d.	**0.954**	**0.954**	**0.958**	**0.958**	**0.962**	**0.962**	**0.962**	**0.962**