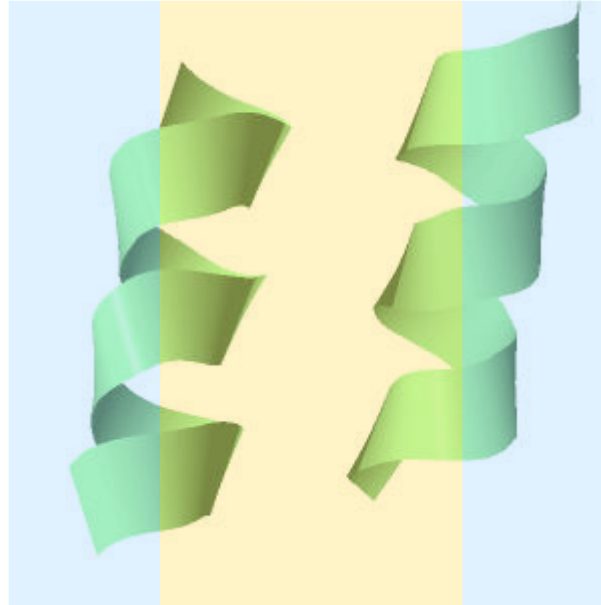




Politechnika
Wroclawska



Modelowanie interakcji helis transmembranowych



Witold Dyrka, Jean-Christophe Nebel, Małgorzata Kotulska

Instytut Inżynierii Biomedycznej i Pomiarowej, Politechnika Wroclawska
Faculty of Computing, Information Systems and Mathematics, Kingston University, Londyn

Cel projektu

Przewidywanie konformacji przestrzennej

interakcji helisa-helisa

w białkach transmembranowych

na podstawie sekwencji aminokwasowej

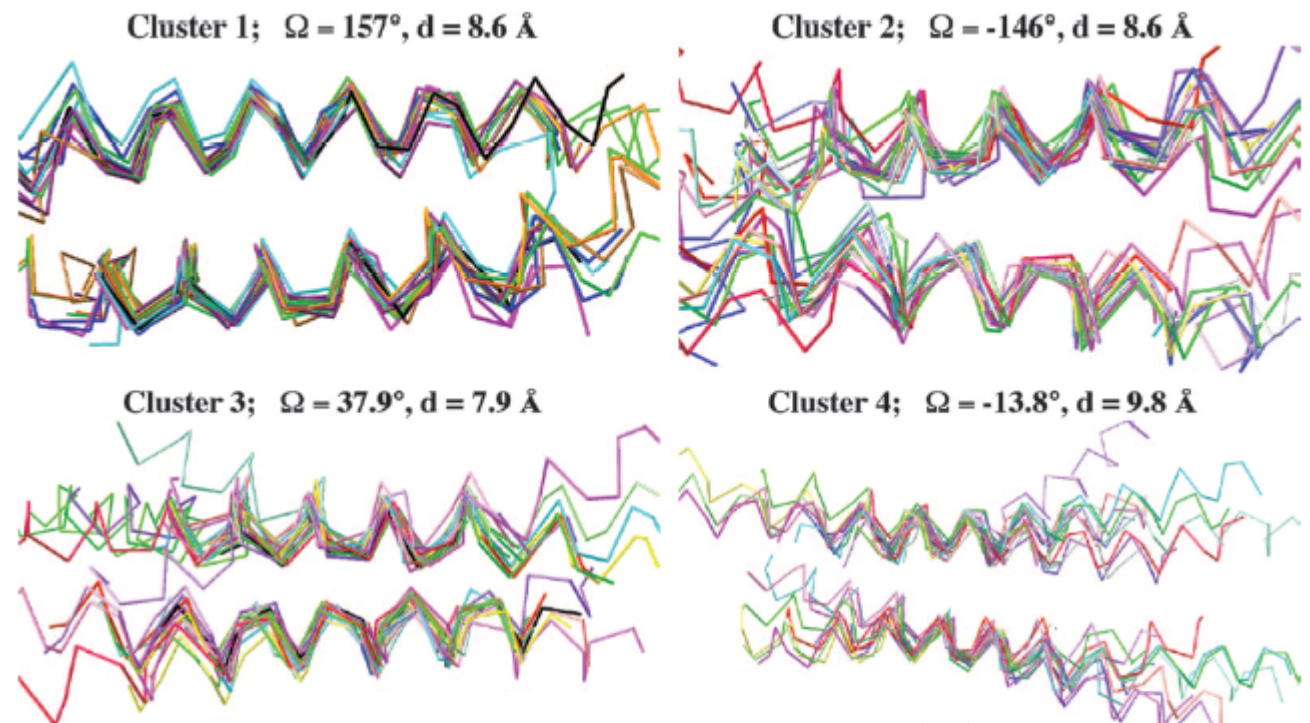
w oparciu o model gramatyczny

Interakcje helisa-helisa w białkach transmembranowych

Typy konformacji

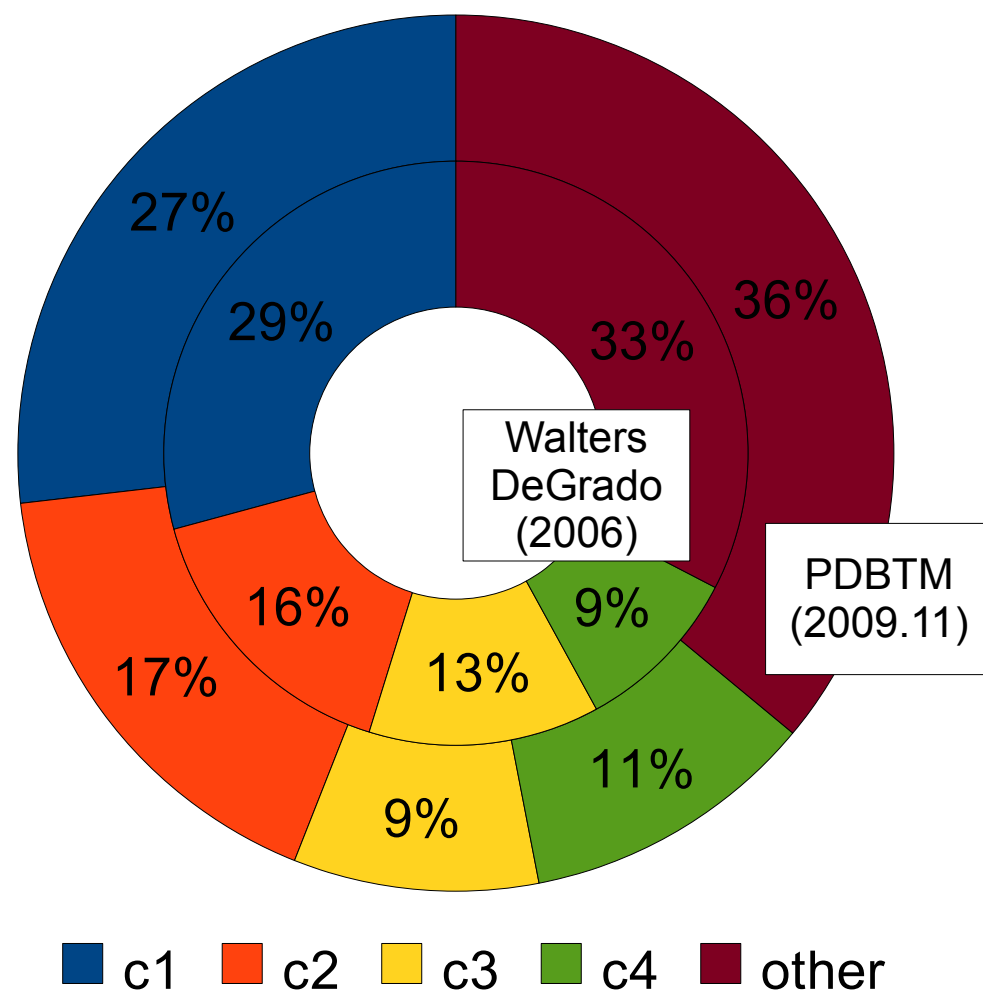
- 2/3 interfejsów helisa-helisa można pogrupować w 4 zbiory o charakterystycznej konformacji przestrzennej (Walters&DeGrado'06)

podział
na podstawie
RMSD
pomiędzy
fragmentami
helis w kontakcie



Zbiory danych uczących i testowych

- **uczenie:** po 20 fragmentów par helis z PDBTM najlepiej dopasowanych do wzorca typu
- **walidacja:** oryginalny zbiór Waltersa i DeGrado z 2006 r. niehomologiczny (<40%) ze zbiorem uczącym
 - c1 – 92, c2 – 49, c3 – 37, c4 – 27 par
- **test:** wszystkie pary helis z nieredundantnego PDBTM w kontakcie typu c1-c4



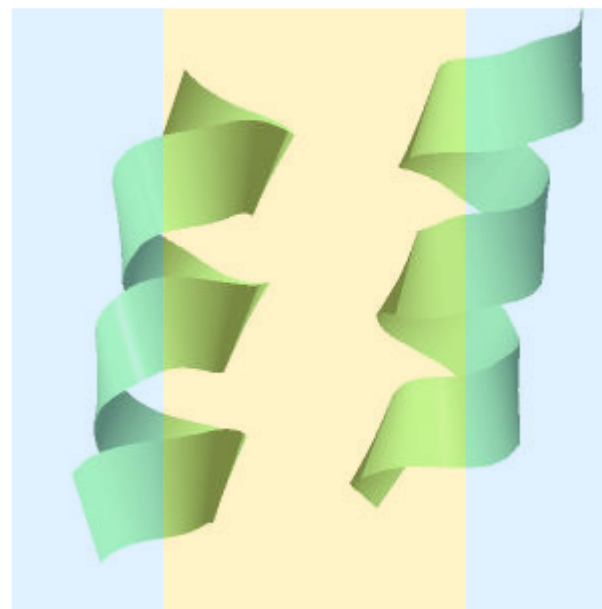
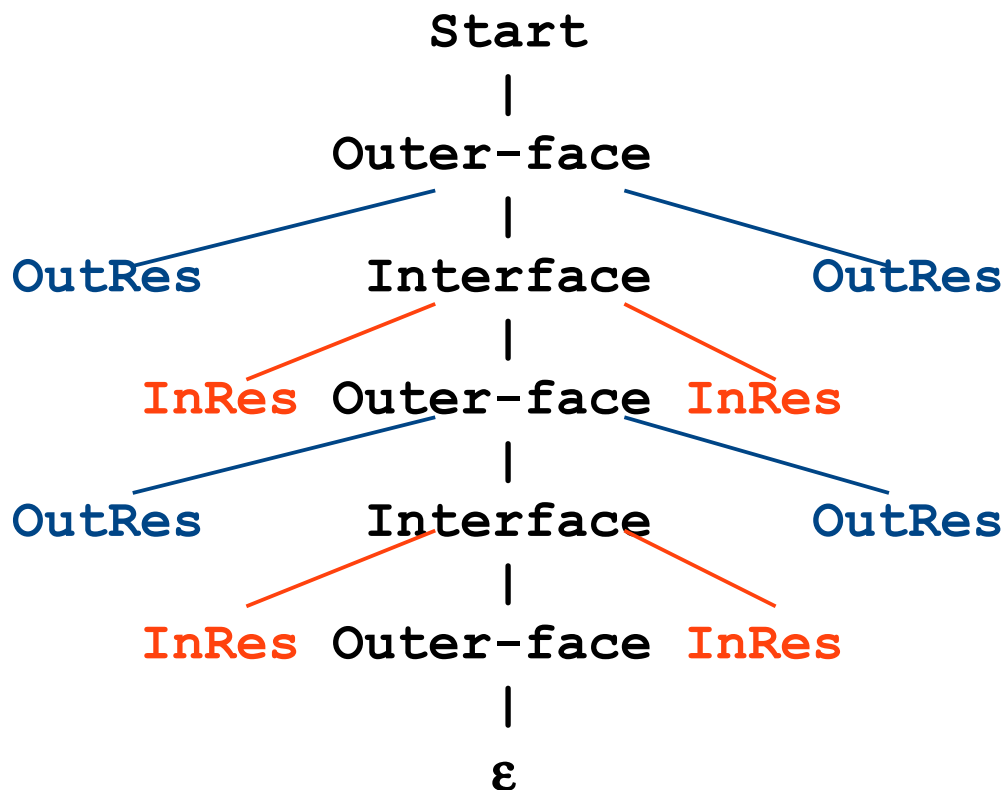
Interakcje helisa-helisa – model gramatyczny (1)

Start → Outer-face

Outer-face → **OutsideResidues1** Interface **OutsideResidues2** | ε

Interface → **InsideResidues1** Outer-face **InsideResidues2** | ε

na podstawie: Waldispuehl J, Steyaert J-M. TCS 335:67-92 (2005)



Interakcje helisa-helisa – model gramatyczny (2)

Start → Outer-face

Outer-face → **OutsideResidues1** Interface **OutsideResidues2** | ε

Interface → **InsideResidues1** Outer-face **InsideResidues2** | ε

OutsideResidues1, OutsideResidues2,

InsideResidues1, InsideResidues2 - reprezentują po 1 lub 2 aminokwasy

- Każdemu aminokwasowi przypisany poziom cechy fizykochemicznej, np. **hydrofobowość**, **rozmiar** itp.
- Reguła gramatyczna dla każdej kombinacji
 - ilości aminokwasów (1 lub 2)
 - poziomu cechy (**Low**, **Medium**, **High**)

np.

Interface → **Low** **Medium** Outer-face **Medium** **Low**

Interface → **Medium** Outer-face **Low** **High**

Interakcje helisa-helisa – model gramatyczny (3)

1. Interface → **Low Medium** Outer-face **Medium Low**, Pr: 0.10
2. Interface → **Medium** Outer-face **Low High**, Pr: 0.20

-
- Każdej regule przypisane jest prawdopodobieństwo jej użycia, np. wychodząc od symbolu `Interface`:
 - regułę 1 użyjemy z prawdopodobieństwem 0.10,
 - regułę 2 użyjemy z prawdopodobieństwem 0.20.
 - **Uczenie** - dobór optymalnego zestawu prawdopodobieństw dla danego zbioru uczącego
 - algorytm genetyczny

Zadanie

Klasyfikacja –

rozdzielenie do którego z 4 klastrów należy
zadany kontakt helisa-helisa

na podstawie zwracanej przez parser
liczby opisującej dopasowanie
analizowanej sekwencji do gramatyki

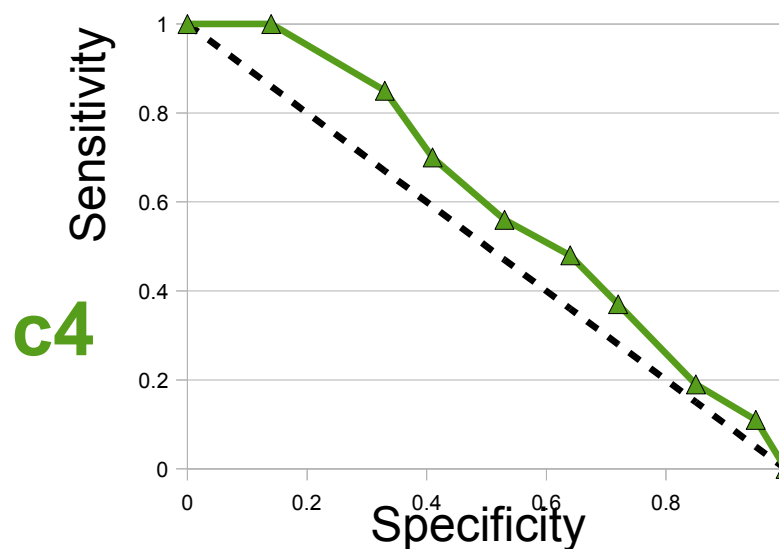
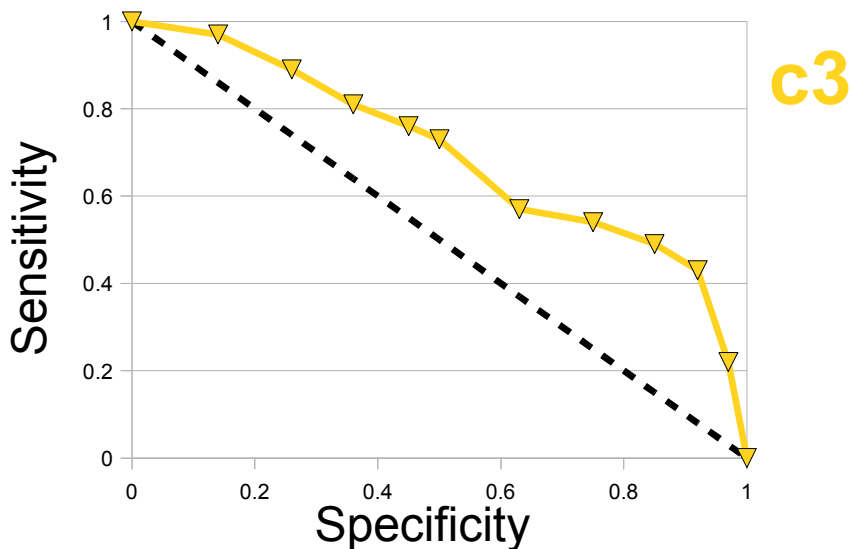
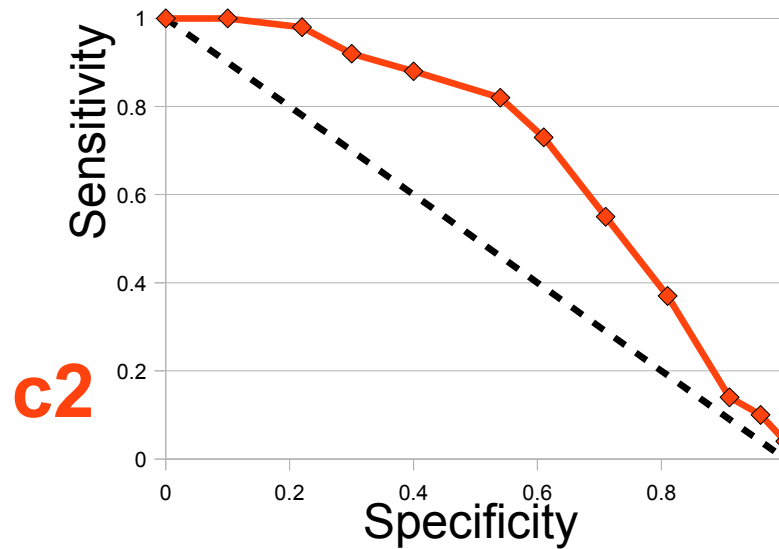
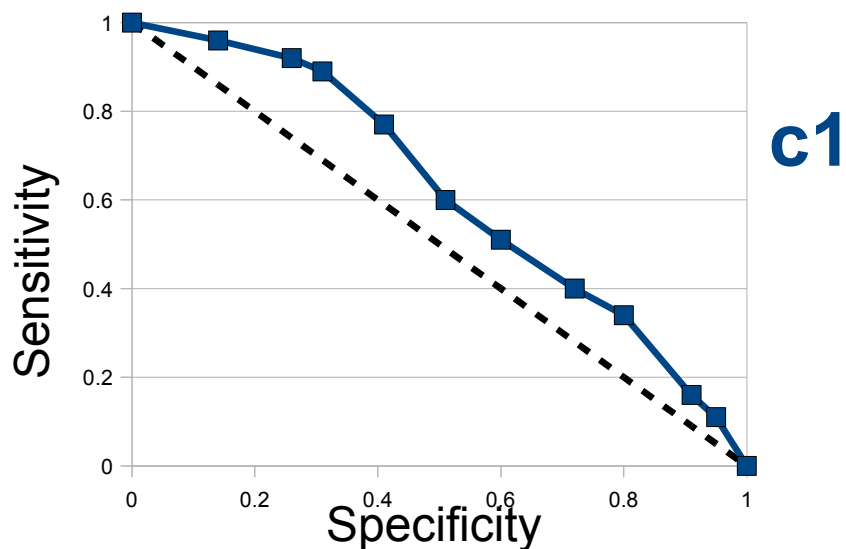
Walidacja: wyniki klasyfikacji fragmentów kontaktowych par helis(1)

Trained for	Using property	Tested against	AUC ROC	Sensitivity	Specificity	Accuracy
c1	accessibility	c2	0.60	0.78	0.43	0.66
		c3	0.60	0.47	0.70	0.53
		c4	0.54	0.77	0.41	0.69
		c2+c3+c4	0.59	0.76	0.42	0.57
c2	van der Waals volume	c1	0.68	0.73	0.62	0.66
		c3	0.54	0.76	0.46	0.63
		c4	0.71	0.55	0.78	0.63
		c1+c3+c4	0.65	0.73	0.57	0.61
c3	van der Waals volume	c1	0.67	0.51	0.80	0.72
		c2	0.52	0.43	0.78	0.63
		c4	0.69	0.49	0.85	0.64
		c1+c2+c4	0.63	0.49	0.79	0.74
c4	accessibility	c1	0.43	0.44	0.54	0.52
		c2	0.54	0.81	0.39	0.54
		c3	0.60	0.48	0.73	0.62
		c1+c2+c3	0.50	0.44	0.60	0.58

Walidacja: wyniki klasyfikacji fragmentów kontaktowych par helis(2)

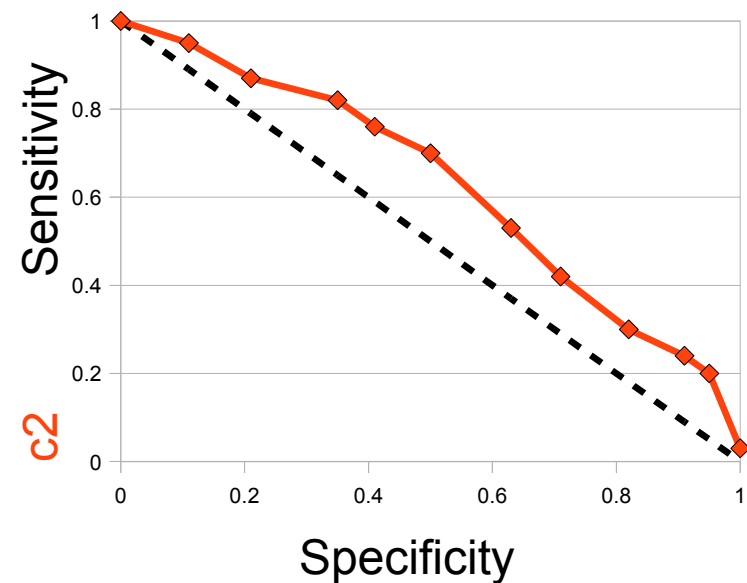
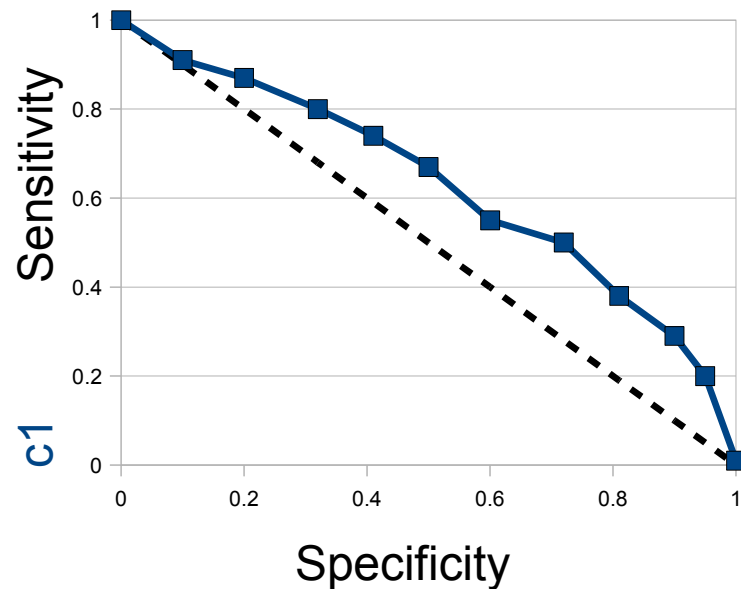
Trained for	Using property	Tested against	AUC ROC	Sensitivity	Specificity	Accuracy
c1	accessibility	c2	0.60	0.78	0.43	0.66
		c3	0.60	0.47	0.70	0.53
		c4	0.54	0.77	0.41	0.69
		c2+c3	0.60	0.76	0.42	0.60
c2	van der Waals volume	c1	0.68	0.73	0.62	0.66
		c3	0.54	0.76	0.46	0.63
		c4	0.71	0.55	0.78	0.63
		c1+c4	0.69	0.73	0.61	0.64
c3	van der Waals volume	c1	0.67	0.51	0.80	0.72
		c2	0.52	0.43	0.78	0.63
		c4	0.69	0.49	0.85	0.64
		c1+c4	0.67	0.49	0.85	0.76
c4	accessibility	c1	0.43	0.44	0.54	0.52
		c2	0.54	0.81	0.39	0.54
		c3	0.60	0.48	0.73	0.62
		c2+c3	0.57	0.48	0.64	0.60

Walidacja: krzywe ROC klasyfikatorów - fragmenty kontaktowe par helis



Test: wyniki anotacji par helis pozycja kontaktu nieznana

Test	AUC ROC	Sensitivity	Specificity	Accuracy
c1 vs c2+c3 accessibility	0.62	0.50	0.72	0.61
c2 vs c1+c4 vdW volume	0.62	0.67	0.52	0.57



Podsumowanie

- Pierwsza metoda przewidująca konformację interakcji helisa-helisa reprezentująca zależności pomiędzy dwoma elementami pary
- Skuteczność klasyfikacji została sprawdzona na niezależnym zbiorze danych
 - pole pod krzywą klasyfikatora (AUC): 0.57 - 0.69
- Zastosowania:
 - punkt startowy modelowania struktury lokalnej
 - ograniczenie przestrzeni poszukiwań metod globalnej predykcji ab initio

Dziękuję za uwagę

Praca wspierana przez granty:
MNiSW N N519 401537,
UE(POKL)/UMarsz. Dolnego Śląska GRANT II/37/2009,
Politechniki Wrocławskiej,
British Council Young Scientists Programme WAR/342/108.