

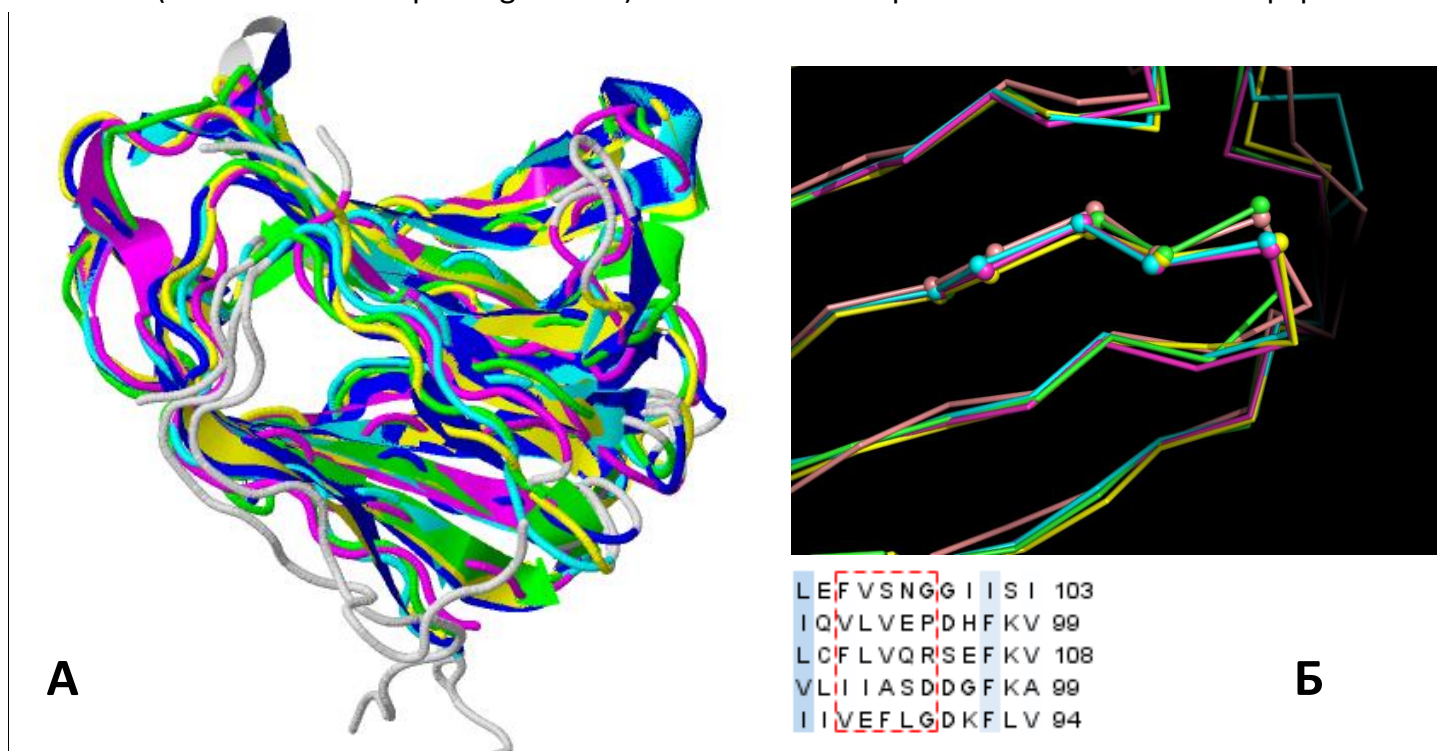
# Демонстрация умений

## Задание d2: совмещение структур

### Совмещение структур своего белка и четырёх гомологов

Для поиска гомологов галектина LEC-6 (PDB-код: 3VV1) был использован сервис PDBeFold; поиск производился структуре цепи B.

Впоследствии в качестве гомологов были выбраны следующие структуры: 2NMO (цепь A), 2D6K (цепь A), 1BKZ( цепь B), 2YMZ (цепь D); все они являются галектинами разных организмов. Затем была выбрана функция “Submit for Multiple Alignment” → “view superposed”, откуда скачаны выровненные 3D-структуры (Рис.1А). Также со странички описания (“Submit for Multiple Alignment”) было скачано выравнивание в текстовом формате.



**Рис.1. Суперпозиция структур глутаматдегидрогеназы из различных организмов. Разные структуры выделены разными цветами. (А) – вся цепь; (Б) – атомы СА невыравненных по последовательности остатков**

В качестве примера разных остатков, выровненных в структуре можно привести конец одного из тяжёлых  $\beta$ -листа (Рис.1Б), в которой не имеются консервативные позиции (см. выравнивание в углу), однако структуры накладываются довольно неплохо. Да и в целом, имеется множество позиций, несовпадающих по типу остатков, но совпадающих по положению в структуре (Рис.2А).

**А**

```

PDB:3vv1:B/1-140 1 - g i s f r n e f f n p q T P V N I P - V Q g - F S N G A R L R L V L L P T S A D S R F H I N L R T P - - - - D D 50
PDB:2nmo:A/1-138 1 - - - - - - - - - - p l i V P Y N L P - L P g g V V P R M L I T I L G T V K P N A N R I A L D F Q R G - - - - N D 42
PDB:2d6k:A/1-147 1 a l f s a q s p y i n p i l P F T G P - I Q g g L Q E G L Q V T L Q G T T K S F A Q R F V V N F Q N S g - - - - N D 53
PDB:1bkz:B/1-135 1 - - - - - - - - - - s n V P H K S S - L P e g I R P G T V L R I R G L V P P N A S R F H V N L L C G e e q g S D 45
PDB:2ymz:D/1-130 1 - - - - - - - - - - a R M F E M F n L D - - W K S G G T M K I K G H I S E D A E S F A I N L G C K s - - - - S D 40

PDB:3vv1:B/1-140 51 I V L H F N A R F D E - - - G A V V N N S T S G G G W Q S E D R H A - N P F Q Q N K I Y T L E F V S N G G I I S I 103
PDB:2nmo:A/1-138 43 V A F H F N P R F N E n n r R V I V C N T K L D N N W G R E E R Q S v F P F E S G K P F K I Q V L V E P D H F K V 99
PDB:2d6k:A/1-147 54 I A F H F N P R F E E g - - G Y V V C N T K Q N G Q W G P E E R K M q M P F Q K G M P F E L C F L V Q R S E F K V 108
PDB:1bkz:B/1-135 46 A A L H F N P R L D T - - - S E V V F N S K E Q G S W G R E E R G P g V P F Q R G Q P F E V L I I A S D D G F K A 99
PDB:2ymz:D/1-130 41 L A L H F N P R F N E - - - S V I V C N S L C S D N W Q Q E Q R D K h F N F Y K G S T V K I I V E F L G D K F L V 94

PDB:3vv1:B/1-140 104 F V N G A H F A D F V E R T P - S H g V H L I E I E G G V H V H S A H V S h - - 140
PDB:2nmo:A/1-138 100 A V N D A H L L Q Y N H R V K k L N e I S K L G I S G D I D L T S A S Y T m i - 138
PDB:2d6k:A/1-147 109 M V N K K F F V Q Y Q H R V P - Y H I V D T I A V S G C L K L S F I T F Q t q n 147
PDB:1bkz:B/1-135 100 V V G D A Q Y H H F R H R L P - L A r V R L V E V G G D V Q L D S V R I F - - - 135
PDB:2ymz:D/1-130 95 K L P D G H E V E F P N R H G y D K - I S Y L N I L G G F K V T S F K V E - - - 130
  
```

**Б**

```

PDB:3vv1:B/1-140 1 - G I S F R N E F F N P Q T P V N I P V Q - G F S N G A R L R L V L L P T S A D S R F H I N L - R - - - T P D D I 51
PDB:2nmo:A/1-138 1 - - - - - - - - - - P L I V P Y N L P L P G G V V P R M L I T I L G T V K P N A N R I A L D F - Q - - - R G N D V 43
PDB:2d6k:A/1-147 1 A L F S A Q S P Y I N P I I P F T G P I Q G G L Q E G L Q V T L Q G T T K S F A Q R F V V N F Q N - - - S G N D I 54
PDB:1bkz:B/1-135 1 - - - - - - - - - - S N V P H K S S L P E G I R P G T V L R I R G L V P P N A S R F H V N L L C G E E Q G S D A 46
PDB:2ymz:D/1-130 1 - - - - - - - - - - A R M F E M F N L - - - - D W K S G G T M K I K G H I S E D A E S F A I N L G C - - - K S S D L 41

PDB:3vv1:B/1-140 52 V L H F N A R F D E G - - - A V V N N S T S G G G W Q S E D R H A N - P F Q Q N K I Y T L E F V S N G G I I S I F 104
PDB:2nmo:A/1-138 44 A F H F N P R F N E N N R R V I V C N T K L D N N W G R E E R Q S V F P F E S G K P F K I Q V L V E P D H F K V A 100
PDB:2d6k:A/1-147 55 A F H F N P R F E E G G - - Y V V C N T K Q N G Q W G P E E R K M Q M P F Q K G M P F E L C F L V Q R S E F K V M 109
PDB:1bkz:B/1-135 47 A L H F N P R L D T S - - - E V V F N S K E Q G S W G R E E R G P G V P F Q R G Q P F E V L I I A S D D G F K A V 100
PDB:2ymz:D/1-130 42 A L H F N P R F N E S - - - V I V C N S L C S D N W Q Q E Q R D K H F N F Y K G S T V K I I V E F L G D K F L V K 95

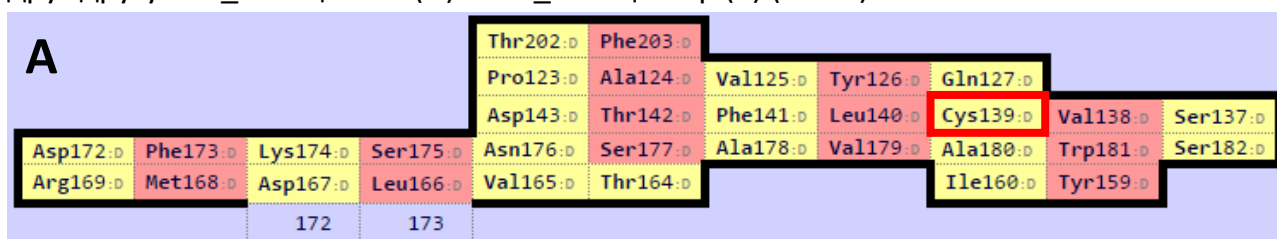
PDB:3vv1:B/1-140 105 V N G A H F A D F V E R T - P S H G V H L I E I E G G V H V H S A H V S H - - 140
PDB:2nmo:A/1-138 101 V N D A H L L Q Y N H R V K K L N E I S K L G I S G D I D L T S A S Y T M I - 138
PDB:2d6k:A/1-147 110 V N K K F F V Q Y Q H R V - P Y H L V D T I A V S G C L K L S F I T F Q T Q N 147
PDB:1bkz:B/1-135 101 V G D A Q Y H H F R H R L - P L A R V R L V E V G G D V Q L D S V R I F - - - 135
PDB:2ymz:D/1-130 96 L P D G H E V E F P N R H - G Y D K I S Y L N I L G G F K V T S F K V E - - - 130
  
```

**Рис.2. Множественные выравнивания последовательностей глутаматдегидрогеназы из различных организмов. (А) Выравнивание, скачанное с сервера PDBeFold; (Б) выравнивание, полученное в результате работы алгоритма Muscle. Раскраска остатков – Clustalx, консервативность – 50%.**

Что касается сравнения выравниваний, полученных PDBeFold и Muscle, многие позиции являются выровненными в обоих случаях (Рис.2А,Б). Тем не менее, имеются множество позиций, которые выровнены по последовательности, но не выровнены по структуре (реже наоборот), поэтому рассматриваемые множественные выравнивания всё-таки устроены по-разному.

**Совмещение по заданному выравниванию**

С помощью SCOP была найдена структура 1KGC, содержащая обе цепи Т-клеточного рецептора. В результате получим карты β-листов, из которых две должны соответствовать друг другу: sht\_2 из цепи α (D) и sht\_6 из цепи β (E) (Рис.3).



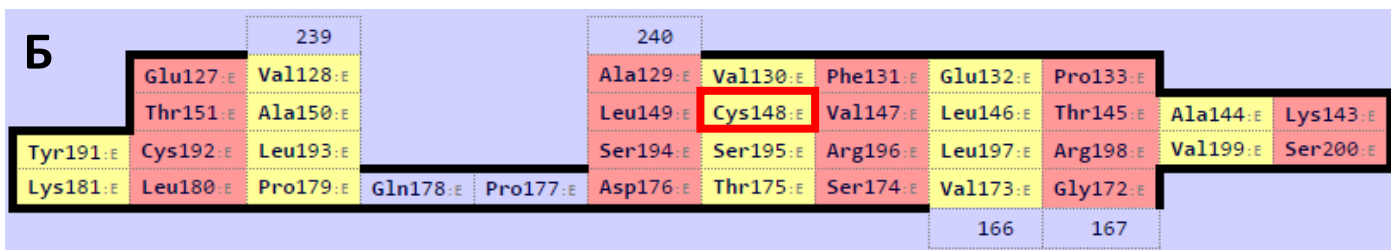


Рис.3. Карты  $\beta$ -листов (sht\_2 (А) и sht\_6 (Б)), полученные в результате работы сервиса SheeP. Выделен консервативный остаток Cys, образующий дисульфидную связь.

В дальнейшем вырежем из PDB-файла координаты атомов, соответствующие доменам SCOP: (118-206) для цепи  $\alpha$  (D) и (119-247) для цепи  $\beta$  (E). При использовании команды «pair\_fit /alpha//D/127+138-140+180+160/CA, /beta//E/130+147-149+195+175/CA» получим довольно неплохое выравнивание структур (Рис.4). Полученное выравнивание позволяет сделать вывод о сходстве топологий двух доменов, поскольку белковая цепь в обоих случаях имеет одинаковый ход (однако в случае цепи  $\alpha$  образуется длинный нерегулярный участок между тяжами, а в случае цепи  $\beta$  – тяж другого  $\beta$ -листа).

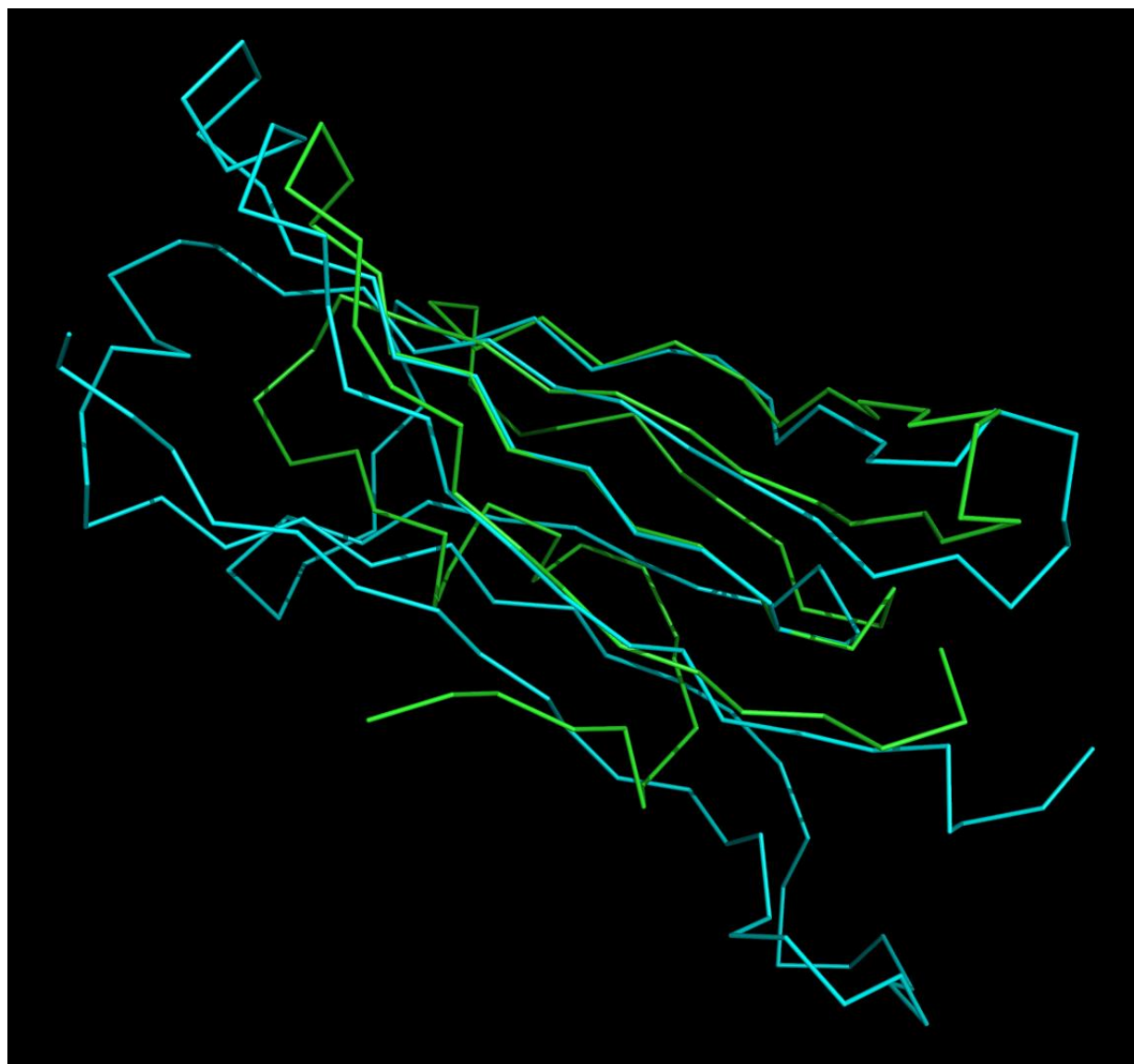


Рис.4. Выравнивание структур константных доменов  $\alpha$  (зеленые тяжи) и  $\beta$  (циановые тяжи) цепей Т-клеточного рецептора.