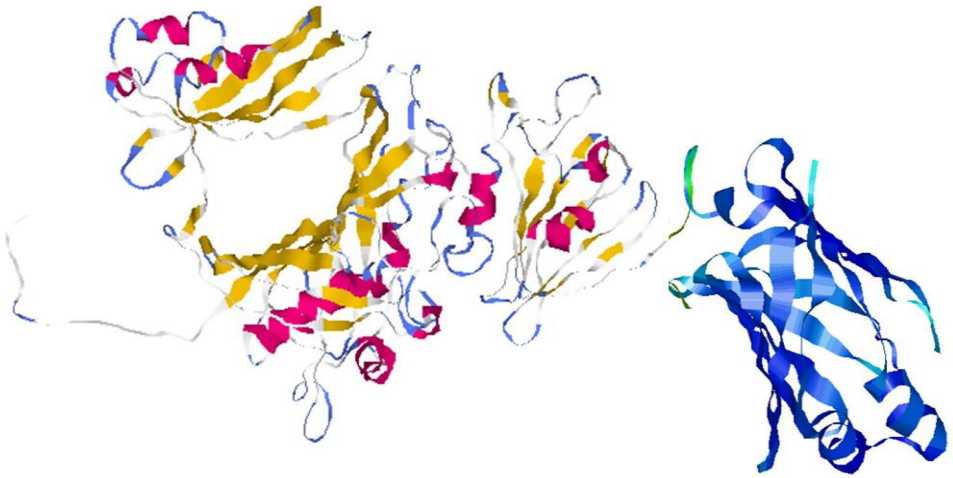
Solution Properties of the new Fusion Protein HUG



CASO STUDIO 2:
Ricerche svolte dal 2010 ad oggi

HELP (*Human Elastin-like Polypeptide*), a synthetic polypeptide based on the **VAPGVG** hexapeptidic motif that is found in the human elastin repetition domains, shows very interesting chemical-physical properties in solution and in particular it has the peculiar thermal behavior defined as the **reverse phase transition**:

- at temperatures below the transition temperature (T_t), the biopolymer is soluble in aqueous solutions where the free chains of HELP exist in a disordered and completely hydrated state.

- at temperatures above T_t, these chains show a more orderly structure (ß-spiral) stabilized by intramolecular hydrophobic interactions that favor their association and the formation of an amorphous solid phase.

HELP

HELP

MW 44885.7 536 aa, Theoretical pI: 11.68

MRGSHHHHHHGSAAAAAAAAKAAAKAAQFGLVPGVGVAPGVGVAPGVGVAPGVGLAPGVGVAPGV
GVAPGVGVAPGIAPAAAAAAKAAAKAAQFGLVPGVGVAPGVGVAPGVGVAPGVGLAPGVGVAPGV
GVAPGVGVAPGIAPAAAAAAKAAAKAAQFGLVPGVGVAPGVGVAPGVGVAPGVGLAPGVGVAPGV
GVAPGVGVAPGIAPAAAAAAKAAAKAAQFGLVPGVGVAPGVGVAPGVGVAPGVGLAPGVGVAPGV
GVAPGVGVAPGIAPAAAAAAKAAAKAAQFGLVPGVGVAPGVGVAPGVGVAPGVGLAPGVGVAPGV
GVAPGVGVAPGIAPAAAAAAKAAAKAAQFGLVPGVGVAPGVGVAPGVGVAPGVGLAPGVGVAPGV
GVAPGVGVAPGIAPAAAAAAKAAAKAAQFGLVPGVGVAPGVGVAPGVGVAPGVGLAPGVGVAPGV
GVAPGVGVAPGIAPAAAAAAKAAAKAAQFGLVPGVGVAPGVGVAPGVGVAPGVGLAPGVGVAPGV
GVAPGVGVAPGIAP**GV**

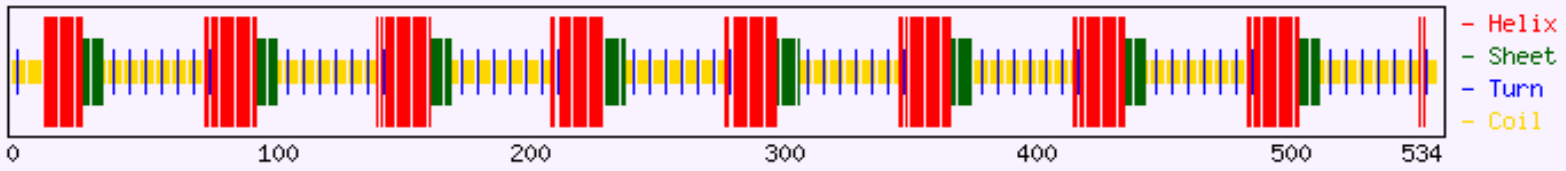
Alpha helix	(Hh)	:	137	is	25.56%
3_{10} helix	(Gg)	:	0	is	0.00%
Pi helix	(Ii)	:	0	is	0.00%
Beta bridge	(Bb)	:	0	is	0.00%
Extended strand	(Ee)	:	0	is	0.00%
Beta turn	(Tt)	:	0	is	0.00%
Bend region	(Ss)	:	0	is	0.00%
Random coil	(Cc)	:	399	is	74.44%
Ambiguous states	(?)	:	0	is	0.00%
Other states		:	0	is	0.00%

Total number of negatively charged residues (Asp + Glu): 0
Total number of positively charged residues (Arg + Lys): 17

HELP

Target Sequence:

```
        10         20         30         40         50         60         70
MRGSHHHHHH GSAAAAAAAA KAAAKAAQFG LVPGVGVAPG VGVAPGVGVA PGVGLAPGVG VAPGVGVAPG

        80         90        100        110        120        130        140
VGVAPGIAPA AAAAAKAAAK AAQFGLVPGV GVAPGVGVAP GVGVAPGVGL APGVGVAPGV GVAPGVGVAP

       150        160        170        180        190        200        210
GIAPAAAAAA KAAAKAAQFG LVPGVGVAPG VGVAPGVGVA PGVGLAPGVG VAPGVGVAPG VGVAPGIAPA

       220        230        240        250        260        270        280
AAAAAKAAAK AAQFGLVPGV GVAPGVGVAP GVGVAPGVGL APGVGVAPGV GVAPGVGVAP GIAPAAAAAA

       290        300        310        320        330        340        350
KAAAKAAQFG LVPGVGVAPG VGVAPGVGVA PGVGLAPGVG VAPGVGVAPG VGVAPGIAPA AAAAAKAAAK

       360        370        380        390        400        410        420
AAQFGLVPGV GVAPGVGVAP GVGVAPGVGL APGVGVAPGV GVAPGVGVAP GIAPAAAAAA KAAAKAAQFG

       430        440        450        460        470        480        490
LVPGVGVAPG VGVAPGVGVA PGVGLAPGVG VAPGVGVAPG VGVAPGIAPA AAAAAKAAAK AAQFGLVPGV

       500        510        520        530
GVAPGVGVAP GVGVAPGVGL APGVGVAPGV GVAPGVGVAP GIAP
```



Secondary Structure:

In **2013**, a new protein, called **UnaG**, was identified and isolated from the muscle of the Japanese eel (*Anguilla japonica*) (N:B: è la prima proteina fluorescente derivata dai vertebrati).

This free-fatty acid binding protein binds unconjugated bilirubin (Br) in a highly specific and selective way, emitting strong fluorescence (Kumagai et al., 2013).

The UnaG gene was been cloned to the C-terminal of the gene encoding the HELP polypeptide obtaining a new functionalized synthetic polypeptide called **HUG**, an acronym that indicates **HELP-UnaG** sequence.

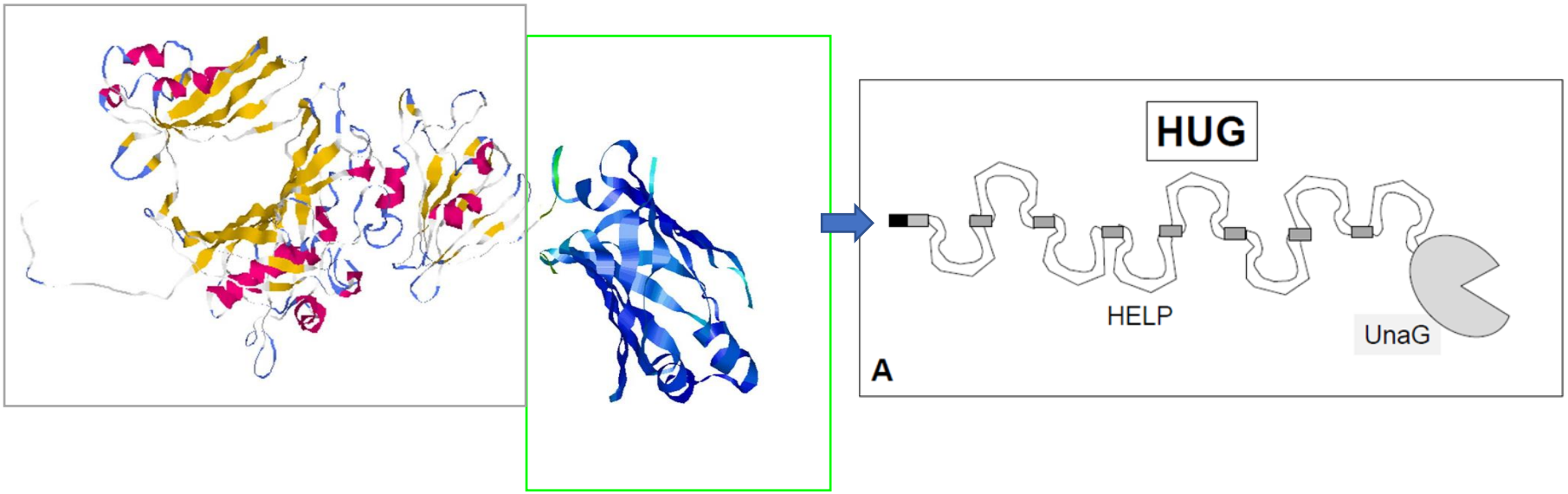
HUG (HELP-UnaG) was been produced from a synthetic gene of the HELP polypeptide fused with the **139** amino acids coding sequence of the UnaG bilirubin-binding protein

Its molecular property appears highly interesting because unconjugated bilirubin (indirect bilirubin) is hardly measured with traditional colorimetric assay.

HUG

A bi-functional, synthetic protein

HUG is composed of a bilirubin-binding domain (UnaG) fused with a scaffold (HELP)



HUG is the acronym of **H**ELP-**U**na**G**

HUG

HELP-UnaG (clon. /10/2014) Number of amino acids: 675, Molecular weight: 60406.47, Theoretical pI: 9.88

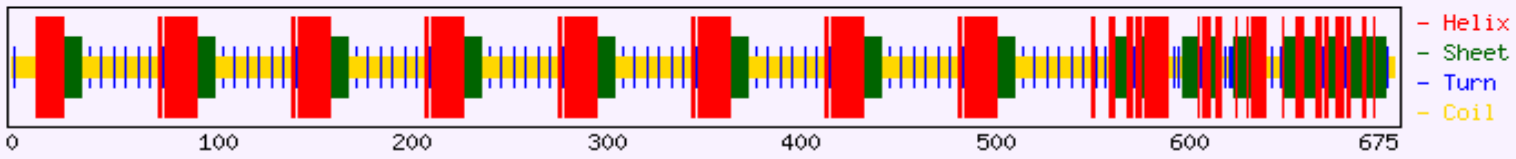
MRGSHHHHHHGSAAAAAAAAKAAAKAAQFGLVPGVGVAPGVGVAPGVGVAPGVGLAPGVGVAPGVGVAP
GVGVAPGIAPAAAAAAKAAAKAAQFGLVPGVGVAPGVGVAPGVGVAPGVGLAPGVGVAPGVGVAPGVGV
APGIAPAAAAAAKAAAKAAQFGLVPGVGVAPGVGVAPGVGVAPGVGLAPGVGVAPGVGVAPGVGVAPGI
APAAAAAAKAAAKAAQFGLVPGVGVAPGVGVAPGVGVAPGVGLAPGVGVAPGVGVAPGVGVAPGIAPAA
AAAAKAAAKAAQFGLVPGVGVAPGVGVAPGVGVAPGVGLAPGVGVAPGVGVAPGVGVAPGIAPAAAAAA
KAAAKAAQFGLVPGVGVAPGVGVAPGVGVAPGVGLAPGVGVAPGVGVAPGVGVAPGIAPAAAAAAKAAA
KAAQFGLVPGVGVAPGVGVAPGVGVAPGVGLAPGVGVAPGVGVAPGVGVAPGIAPAAAAAAKAAAKAAQ
FGLVPGVGVAPGVGVAPGVGVAPGVGLAPGVGVAPGVGVAPGVGVAPGIAP
GGMVEKFVGTWKIADSHNFGEYLKAIGAPKELSDGGDATTPTLYISQKDGDKMTVKIENGPPTFLDTQV
KFKLGEEFDEFPSDRRKGVKSVVNLVGEKLVYVQKWDGKETTYVREIKDGKLVVTLTMGDVVAVRSYRR
ATE

Alpha helix	(Hh) :	198 is	22.84%
3_{10} helix	(Gg) :	0 is	0.00%
Pi helix	(Ii) :	0 is	0.00%
Beta bridge	(Bb) :	0 is	0.00%
Extended strand	(Ee) :	45 is	5.19%
Beta turn	(Tt) :	0 is	0.00%
Bend region	(Ss) :	0 is	0.00%
Random coil	(Cc) :	624 is	71.97%
Ambiguous states (?)	:	0 is	0.00%
Other states	:	0 is	0.00%

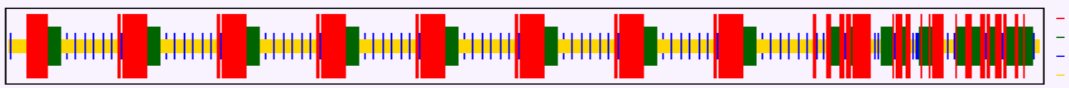
Total number of negatively charged residues (Asp + Glu): 22
Total number of positively charged residues (Arg + Lys): 39

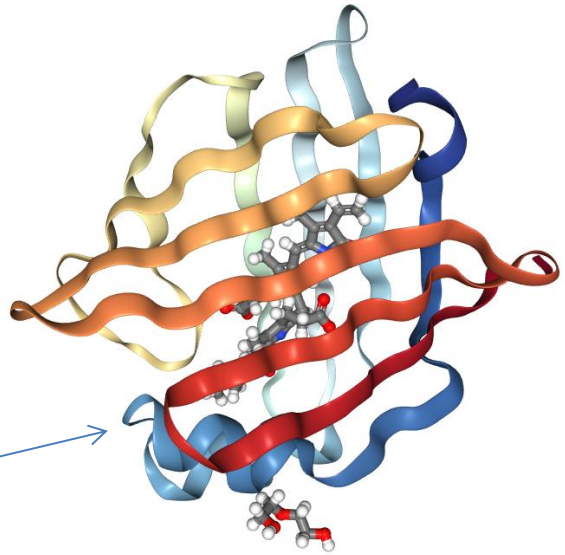
HUG

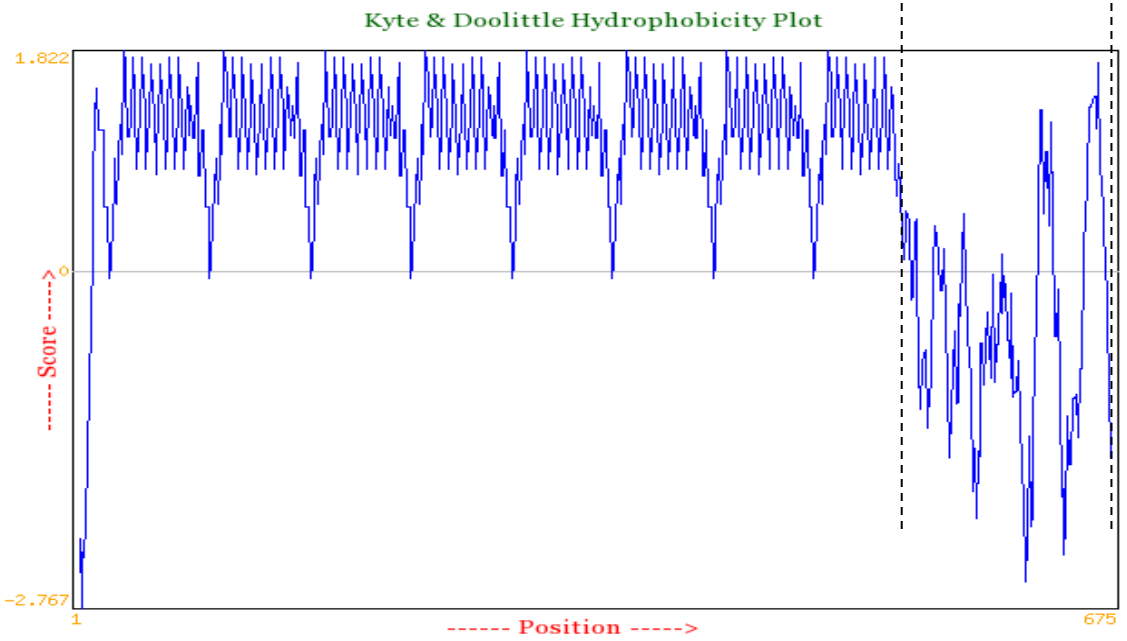
```
        10         20         30         40         50         60         70
MRGSHHHHHH GSAAAAAAAA KAAAKAAQFG LVPGVGVAPG VGVAPGVGVA PGVGLAPGVG VAPGVGVAPG

        80         90        100        110        120        130        140
VGVAPGIAPA AAAAAKAAAK AAQFGLVPGV GVAPGVGVAP GVGVAPGVGL APGVGVAPGV GVAPGVGVAP

       150        160        170        180        190        200        210
GIAPAAAAAA KAAAKAAQFG LVPGVGVAPG VGVAPGVGVA PGVGLAPGVG VAPGVGVAPG VGVAPGIAPA

       220        230        240        250        260        270        280
AAAAAKAAAK AAQFGLVPGV GVAPGVGVAP GVGVAPGVGL APGVGVAPGV GVAPGVGVAP GIAPAAAAAA

       290        300        310        320        330        340        350
KAAAKAAQFG LVPGVGVAPG VGVAPGVGVA PGVGLAPGVG VAPGVGVAPG VGVAPGIAPA AAAAAKAAAK

       360        370        380        390        400        410        420
AAQFGLVPGV GVAPGVGVAP GVGVAPGVGL APGVGVAPGV GVAPGVGVAP GIAPAAAAAA KAAAKAAQFG

       430        440        450        460        470        480        490
LVPGVGVAPG VGVAPGVGVA PGVGLAPGVG VAPGVGVAPG VGVAPGIAPA AAAAAKAAAK AAQFGLVPGV

       500        510        520        530        540        550        560
GVAPGVGVAP GVGVAPGVGL APGVGVAPGV GVAPGVGVAP GIAPGGMVEK FVGTWKIADS HNFGEYLKAI

       570        580        590        600        610        620        630
GAPKELSDGG DATTPTLYIS QKDGDKMTVK IENGPPTFLD TQVKFKLGEE FDEFPSDRRK GVKSVVNLVG

       640        650        660        670
EKLVYVQKWD GKETTYVREI KDGKLVVTLT MGDVVAVRSY RRATE
```



```
        10         20         30         40         50         60         70
MRGSHHHHHH GSAAAAAAAA KAAAKAAQFG LVPGVGVAPG VGVAPGVGVA PGVGLAPGVG VAPGVGVAPG

        80         90        100        110        120        130        140
VGVAPGIAPA AAAAAKAAAK AAQFGLVPGV GVAPGVGVAP GVGVAPGVGL APGVGVAPGV GVAPGVGVAP

       150        160        170        180        190        200        210
GIAPAAAAAA KAAAKAAQFG LVPGVGVAPG VGVAPGVGVA PGVGLAPGVG VAPGVGVAPG VGVAPGIAPA

       220        230        240        250        260        270        280
AAAAAKAAAK AAQFGLVPGV GVAPGVGVAP GVGVAPGVGL APGVGVAPGV GVAPGVGVAP GIAPAAAAAA

       290        300        310        320        330        340        350
KAAAKAAQFG LVPGVGVAPG VGVAPGVGVA PGVGLAPGVG VAPGVGVAPG VGVAPGIAPA AAAAAKAAAK

       360        370        380        390        400        410        420
AAQFGLVPGV GVAPGVGVAP GVGVAPGVGL APGVGVAPGV GVAPGVGVAP GIAPAAAAAA KAAAKAAQFG

       430        440        450        460        470        480        490
LVPGVGVAPG VGVAPGVGVA PGVGLAPGVG VAPGVGVAPG VGVAPGIAPA AAAAAKAAAK AAQFGLVPGV

       500        510        520        530        540        550        560
GVAPGVGVAP GVGVAPGVGL APGVGVAPGV GVAPGVGVAP GIAPGGMVEK FVGTWKIADS HNFGEYLKAI

       570        580        590        600        610        620        630
GAPKELSDGG DATTPTLYIS QKDGDKMTVK IENGPPTFLD TQVKFKLGEE FDEFPSDRRK GVKSVVNLVG

       640        650        660        670
EKLVYVQKWD GKETTYVREI KDGKLVVTLT MGDVVAVRSY RRATE
```





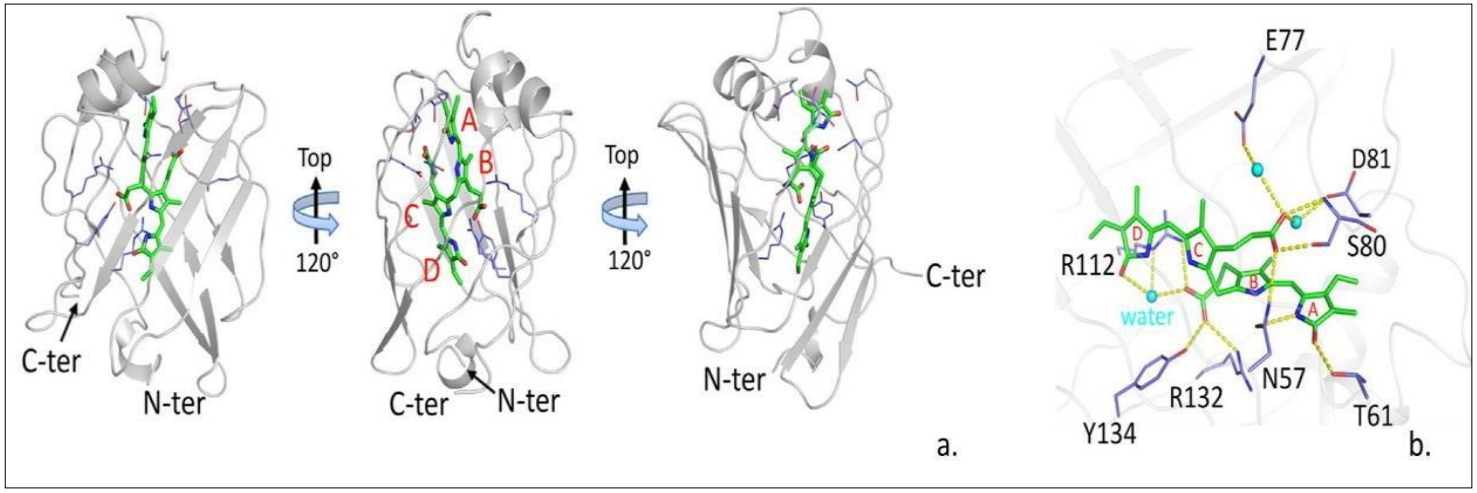


The **biosensor HUG** for the quantitative analysis of unconjugated bilirubin was developed in our laboratory (Dr. Antonella Bandiera and co-workers):

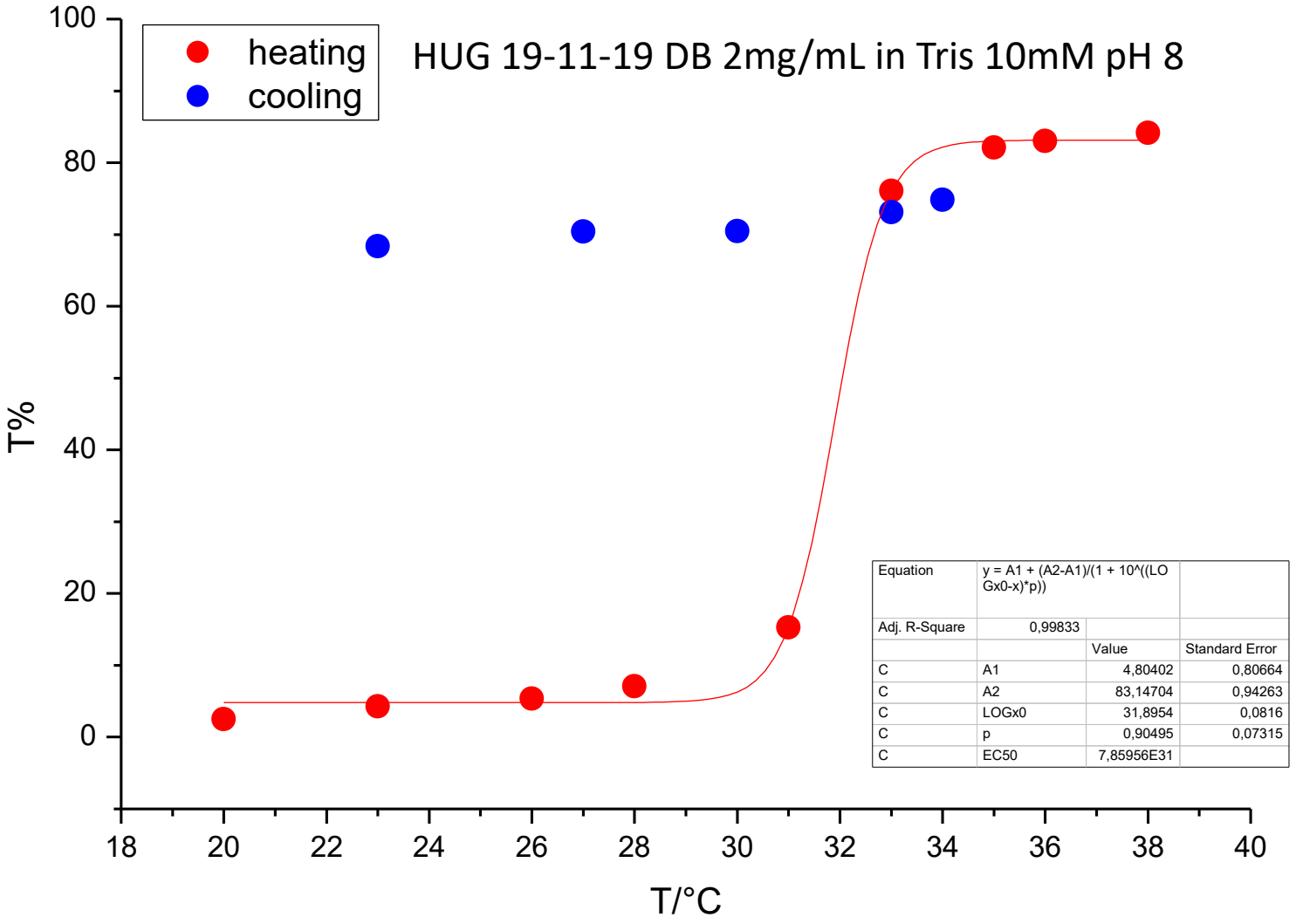
This new fusion protein preserves:

1) the reverse phase transition property of HELP

2) the highly specific capacity of UnaG to bind bilirubin and to become fluorescent.



The reverse thermal transition of HUG



Equation	y = A1 + (A2-A1)/(1 + 10^((LOGx0-x)*p))		
Adj. R-Square	0,99833		
		Value	Standard Error
C	A1	4,80402	0,80664
C	A2	83,14704	0,94263
C	LOGx0	31,8954	0,0816
C	p	0,90495	0,07315
C	EC50	7,85956E31	

CHARACTERIZATION OF HUG

- **Molecular mechanics and dynamic simulation**
- **Circular Dichroism (CD)**
- **Differential Scanning Calorimetry (DSC)**
- **Potentiometric titration of HUG and HELP**

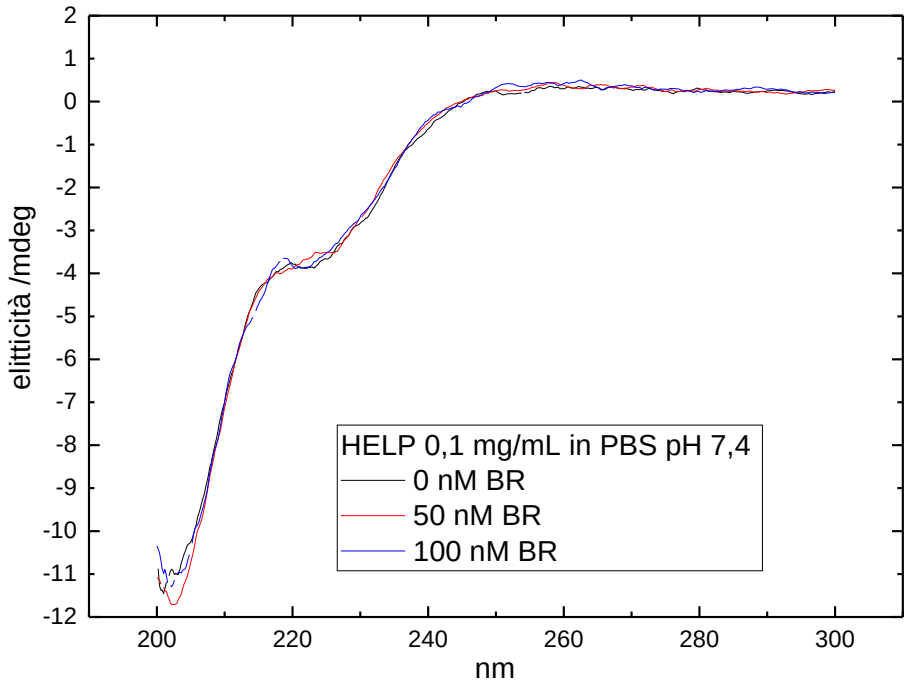
The physical and chemical characterization of the new fluorescent biosensor HUG. This step is very important for optimization of liquid assay of free bilirubin in biological samples

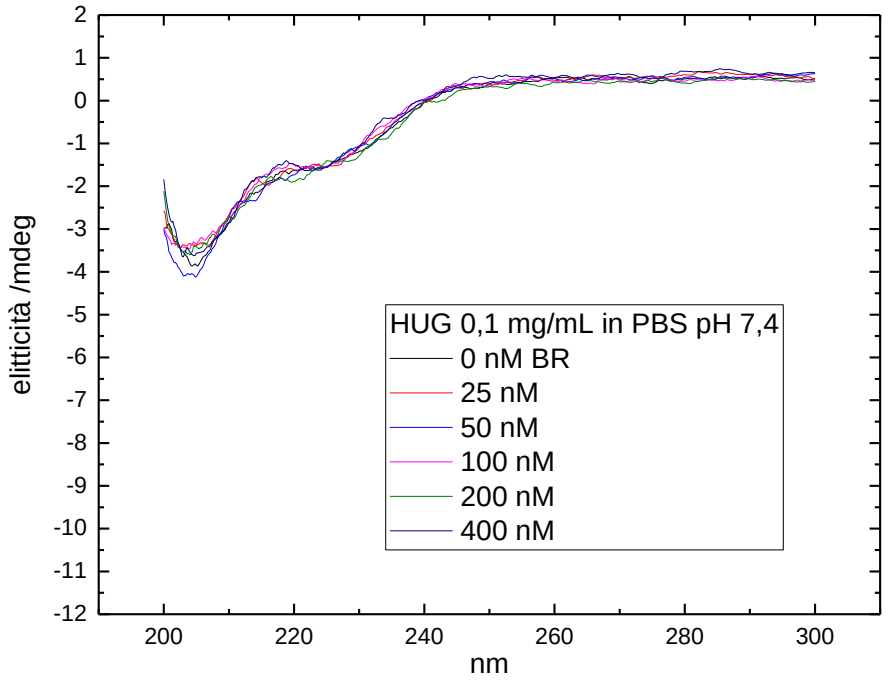
Analyzing quantitatively the ligand binding to macromolecule it is possible to focus on the application and importance of HUG biopolymer for bilirubin quantitative assessment in biological systems

Fluorescence spectroscopy is one of the most convenient methods for the evaluation of binding processes. Since the intensity and wavelength are very sensitive to the change of the environment due to the ligand binding, variation in fluorescence intensity as a function of ligand concentration provide information about the strength of the protein-ligand interaction.

CIRCULAR DICHROISM

CD is an excellent tool for rapid determination of the secondary structure and folding properties of proteins that have been obtained using recombinant techniques. The most widely used applications of protein CD are to determine whether an expressed protein is folded. In addition, it can be used to study protein-ligand interactions.

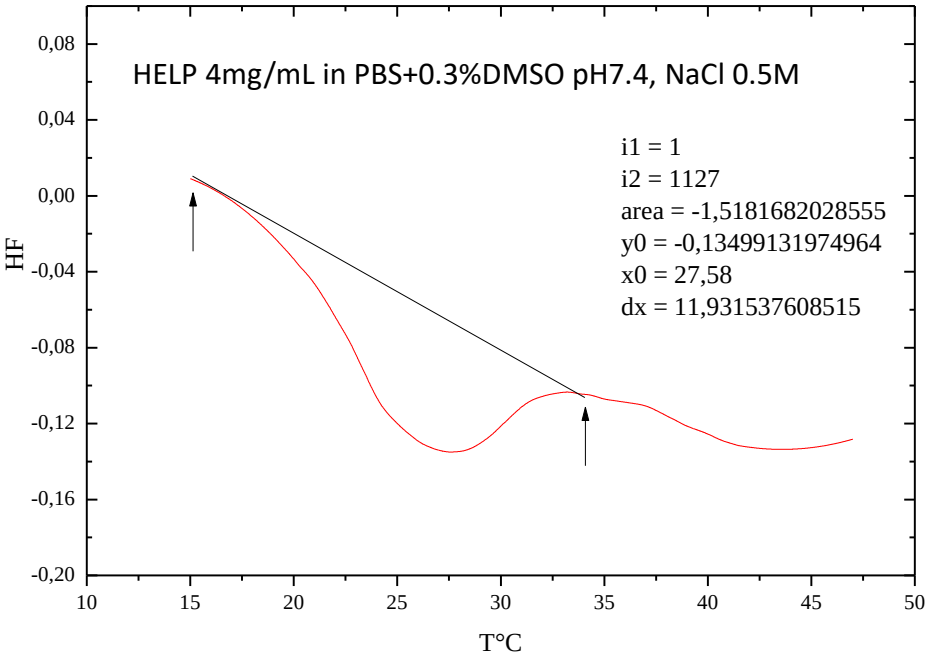


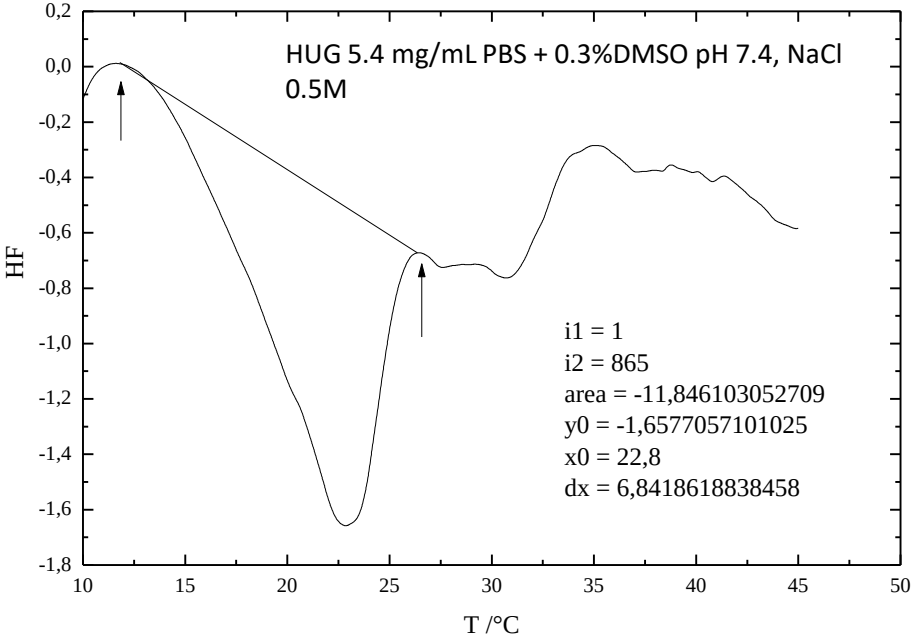


DSC SETARAM (preliminary results)

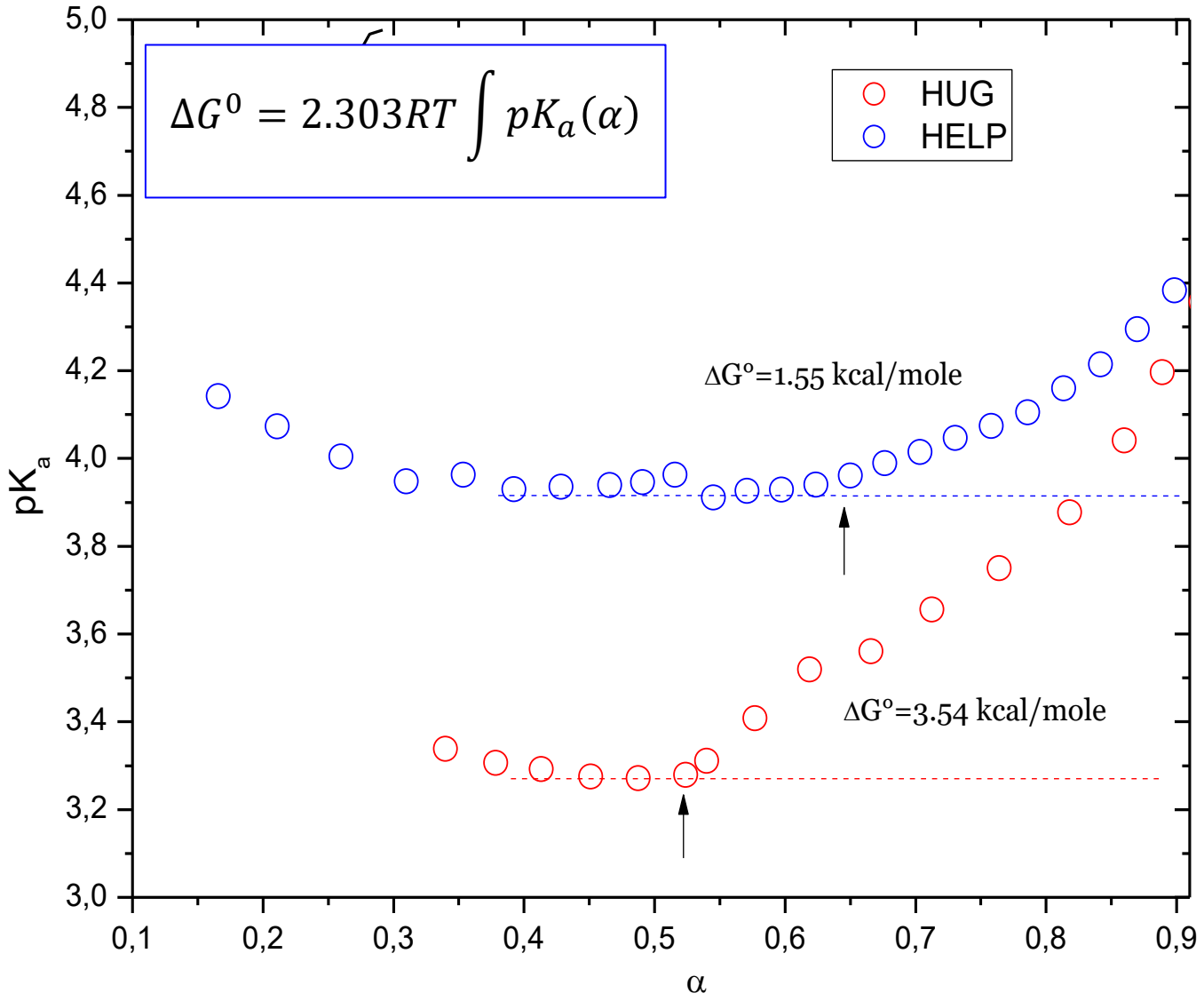
The DSC approach employs a reference cell and a sample cell heated at a controlled rate. It can detect some of transitions such as melts, glass transitions, phase changes.

Calorimetric methods demonstrate an advantage over other methods in their ability to provide a relatively full picture of thermodynamic parameters during binding reactions, including the equilibrium binding constants, the enthalpy of binding reactions (ΔH)and the entropy change (ΔS).



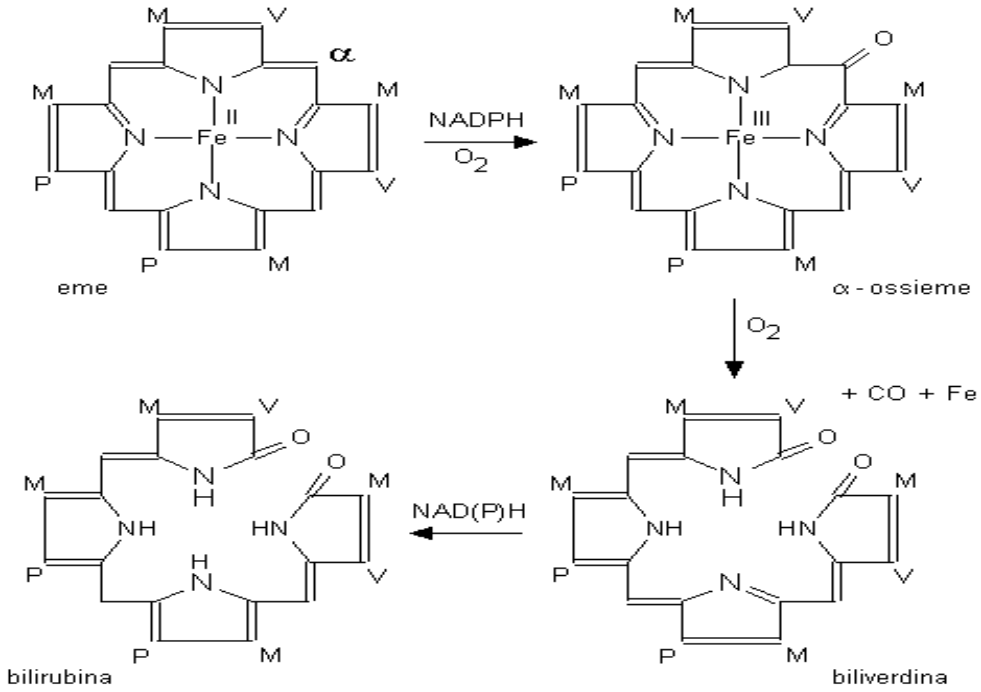


Potentiometric titration of HUG and HELP



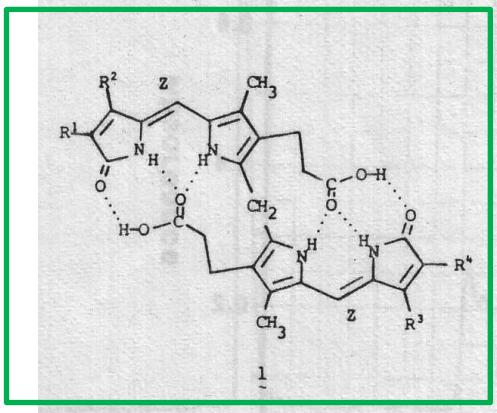
La bilirubina (dal latino bilis = bile e ruber = rosso) è un pigmento di colore rosso-arancione che deriva per l'80% dal catabolismo dell'eme dell'emoglobina (Hb) e per il restante 20% da altre emoproteine, come le mioglobine e i citocromi.

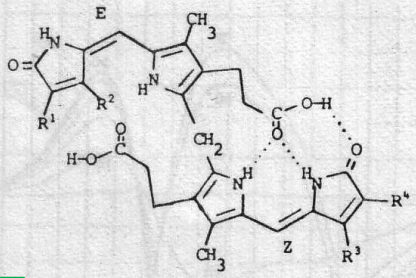
La bilirubina si ottiene dal processo di degradazione della ferroporfirina, o gruppo eme, dell'emoglobina, rilasciata durante la distruzione dei globuli rossi, o eritrociti, nella milza.



il metabolismo della bilirubina non coniugata avviene a livello epatico, dopo essere stata trasportata nel sangue da una proteina di trasporto, l'albumina, in virtù della sua insolubilità in acqua (liposolubilità);

BILIRUBIN

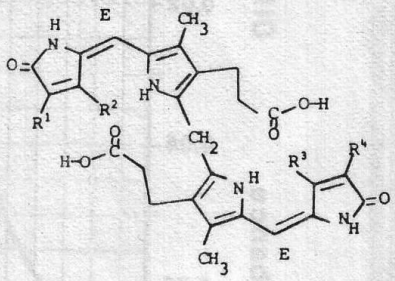




IXα: $R^1=R^3=CH_3$, $R^2=R^4=CH=CH_2$

IIIα: $R^1=R^4=CH=CH_2$, $R^2=R^3=CH_3$

XIIIα: $R^1=R^4=CH_3$, $R^2=R^3=CH=CH_2$



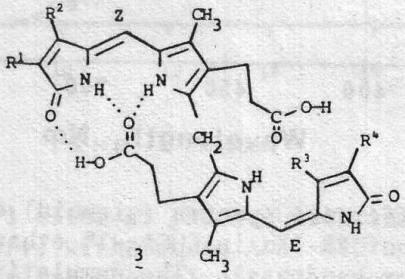
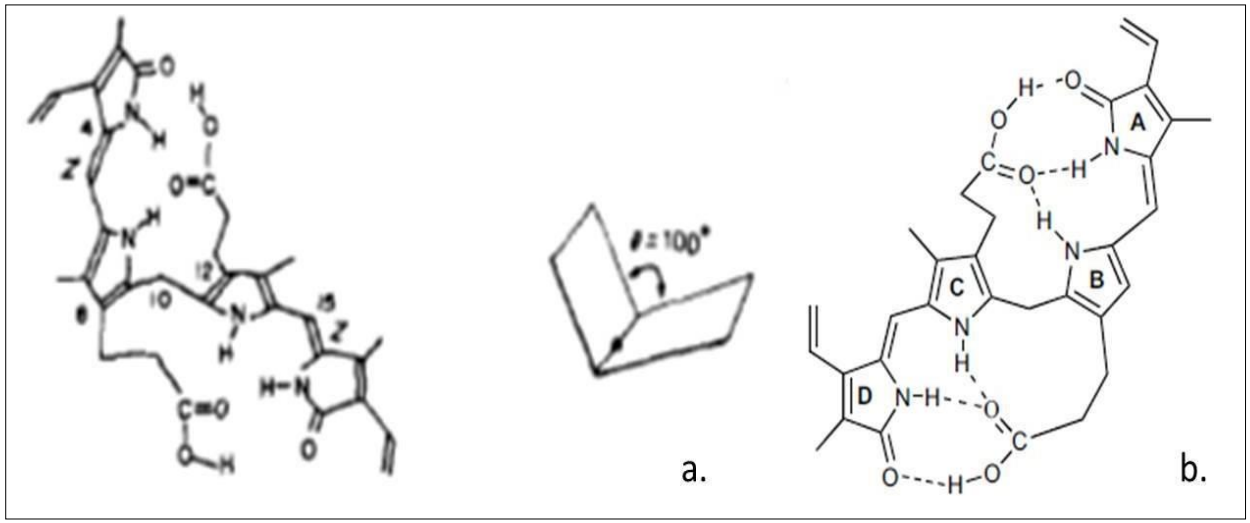
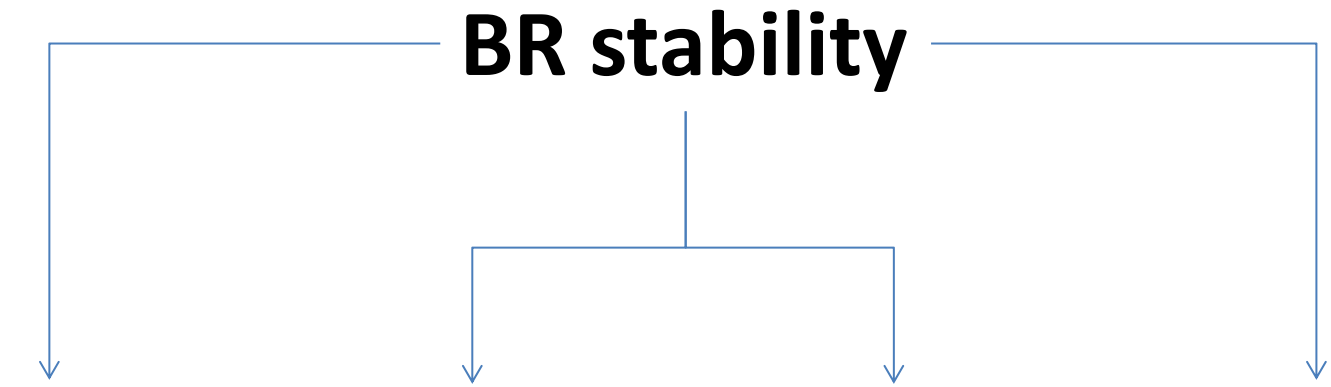


Figure 1 Configurational isomers of BR showing geometric isomerism about the 5,15 *meso* double bonds: 1 = Z-Z, 2 = E-Z, 3 = Z-E and 4 = E-E. The Z-Z configuration (1) is the stable ground state structure of BR (1). The E-Z (2), Z-E (3) and E-E (4) structures are the photochemically accessible isomers which, in the case of BR-IXα, we call collectively PBR. The E-Z (2) and Z-E (3) isomers are identical for (symmetrical) BR-IIIα and BR-XIIIα but not for BR-IXα. In the corresponding BR dimethyl esters, the propionic acid groups become methyl propionate groups.


θ = 100°
a.
A
B
C
D
b.

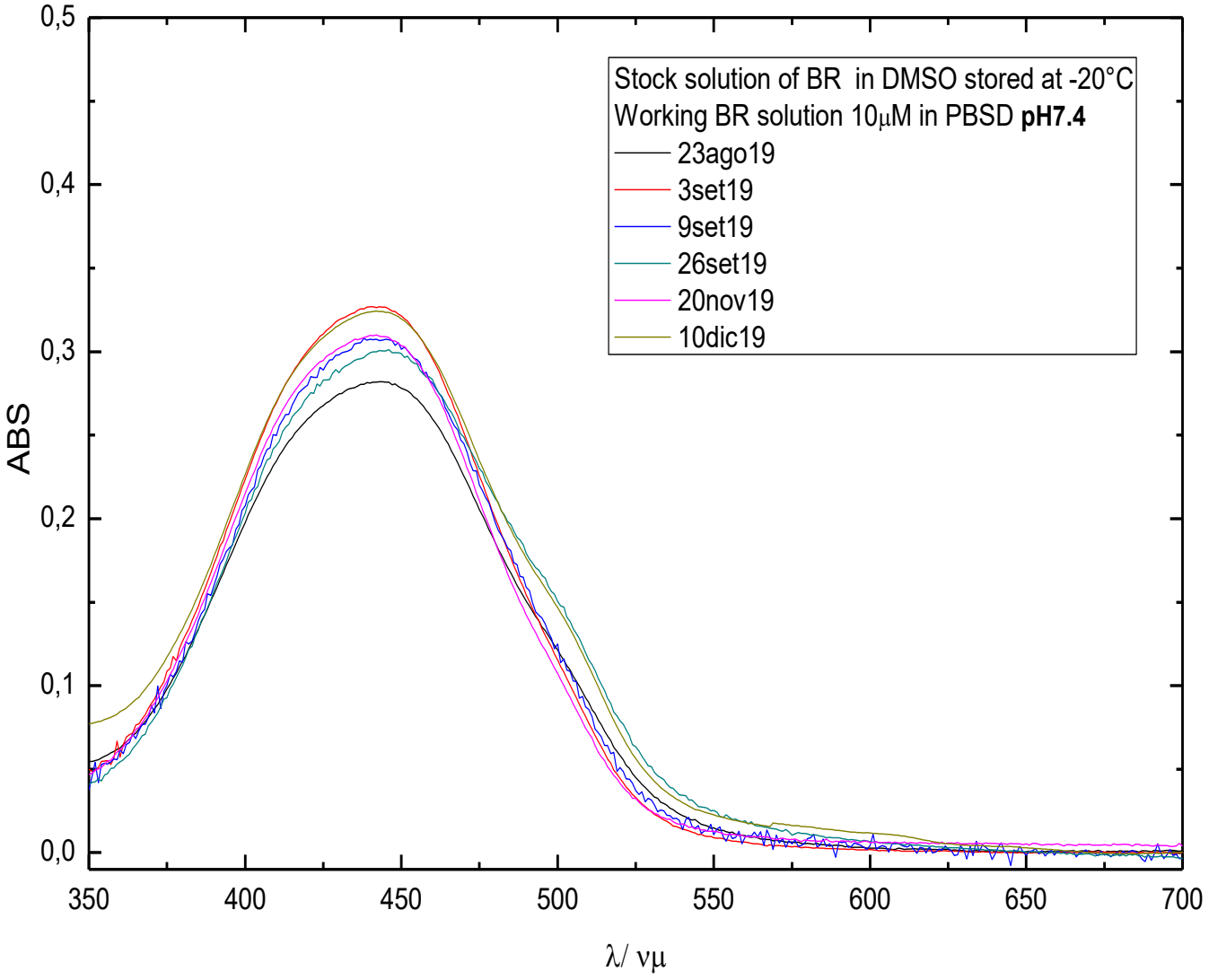
we assessed the bilirubin stability at given experimental conditions:



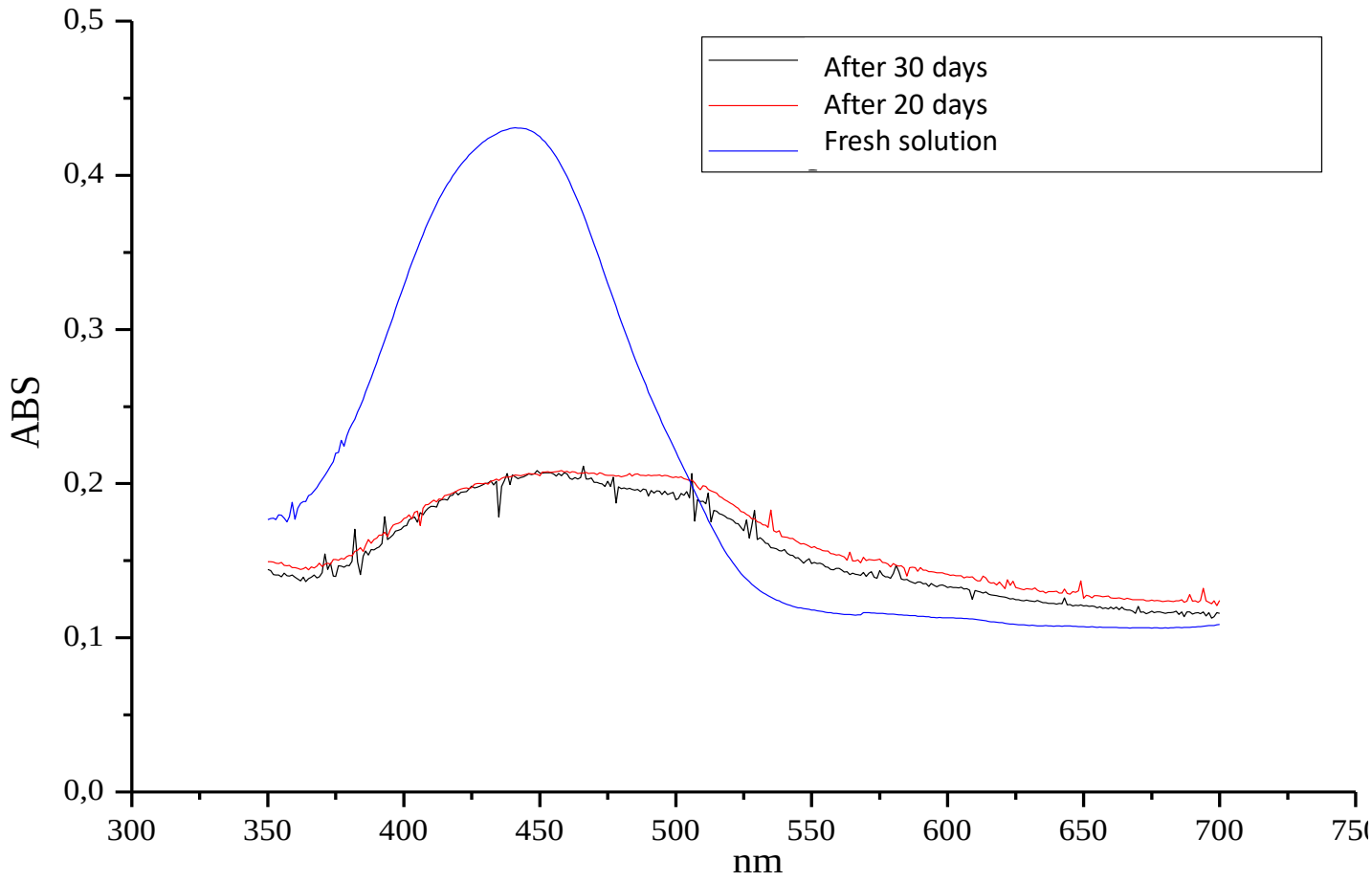
Stock solution of BR is prepared in 100% DMSO and it is stored up to -20°C. Stability assay was done on working bilirubin solution (**BR 10µM in PBS + 0.3%DMSO**). This mother solution is used to prepare the standard solutions.

NB "To prevent the denaturation of UnaG, the final concentrations of DMSO were at most 0.2% (vol/vol) in all experiments." Shitashima 2017

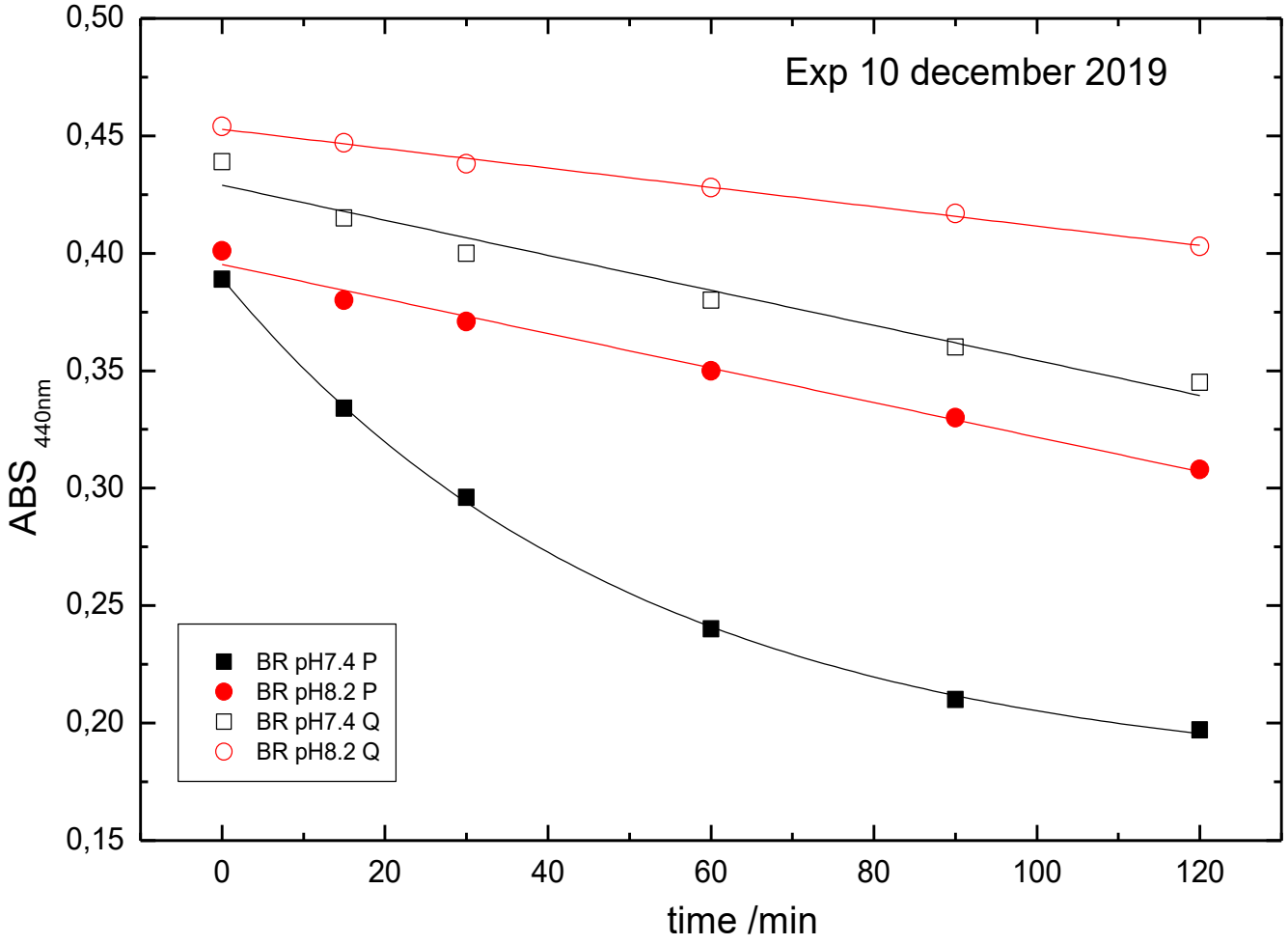
BR is stable when the solution is stored at -20°C



Working solution (PBSD pH 7.4) is not stable if it is stored at -20°C



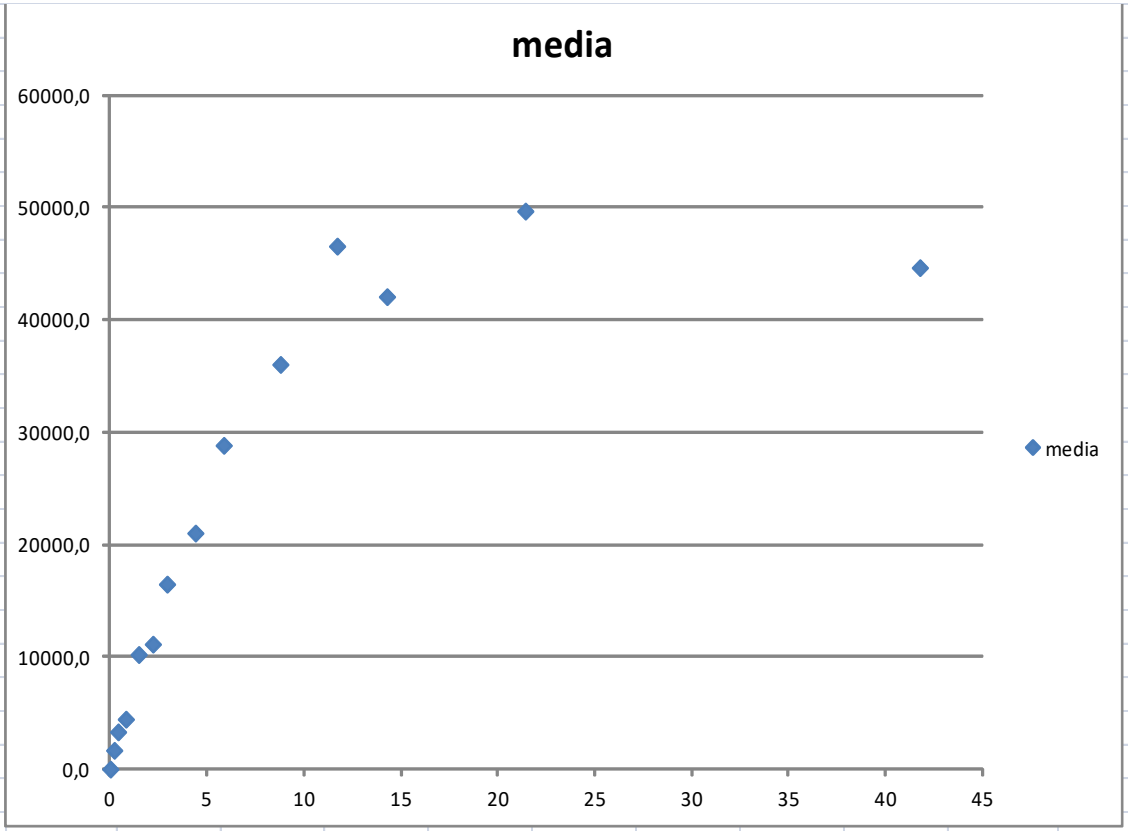
BR stability is influenced by surface effects (plastic or quartz cuvettes)



HUG-BR interaction

The estimate of unconjugated bilirubin concentration has been performed on 96-wells plate in liquid or solid phase (PBSD pH 7.4)

ng BR	media	stdev
0	34,6	7,0
0,2	1708,75	0,0
0,4	3350,5	0,0
0,8	4480,5	0,0
1,5	10235,2	2585,1
2,2	11158,5	6690,1
2,9	16497,6	3515,5
4,4	21042,4	6179,8
5,9	28855,5	3333,5
8,8	36055,4	4107,6
11,7	46582,4	5827,7
14,274	42084	0,0
21,411	49702	0,0
41,769	44667,75	0,0



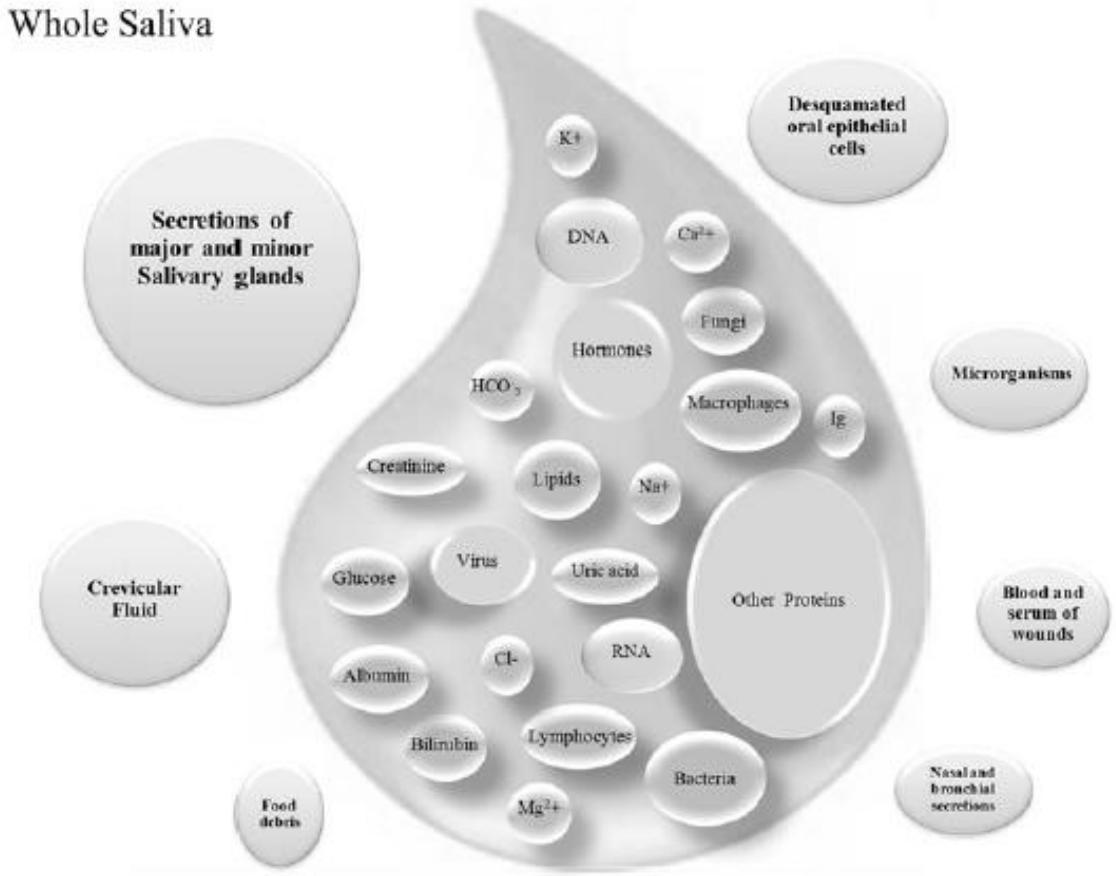


FIG. 1. Schematic representation of the components of whole saliva. The size of each individual droplet is an approximate representation of its concentration in whole saliva.

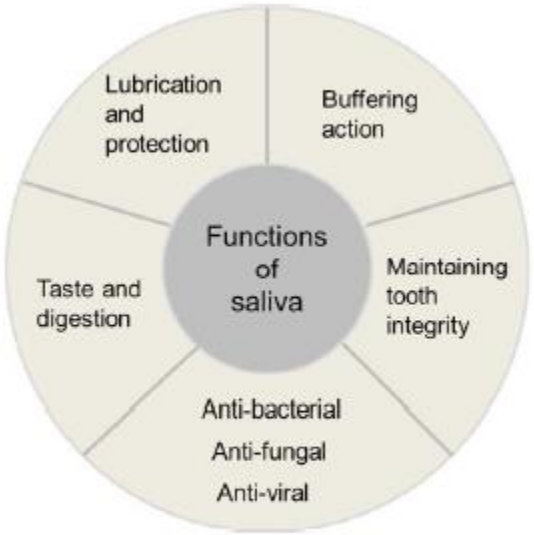


Figure 1: Main functions of saliva

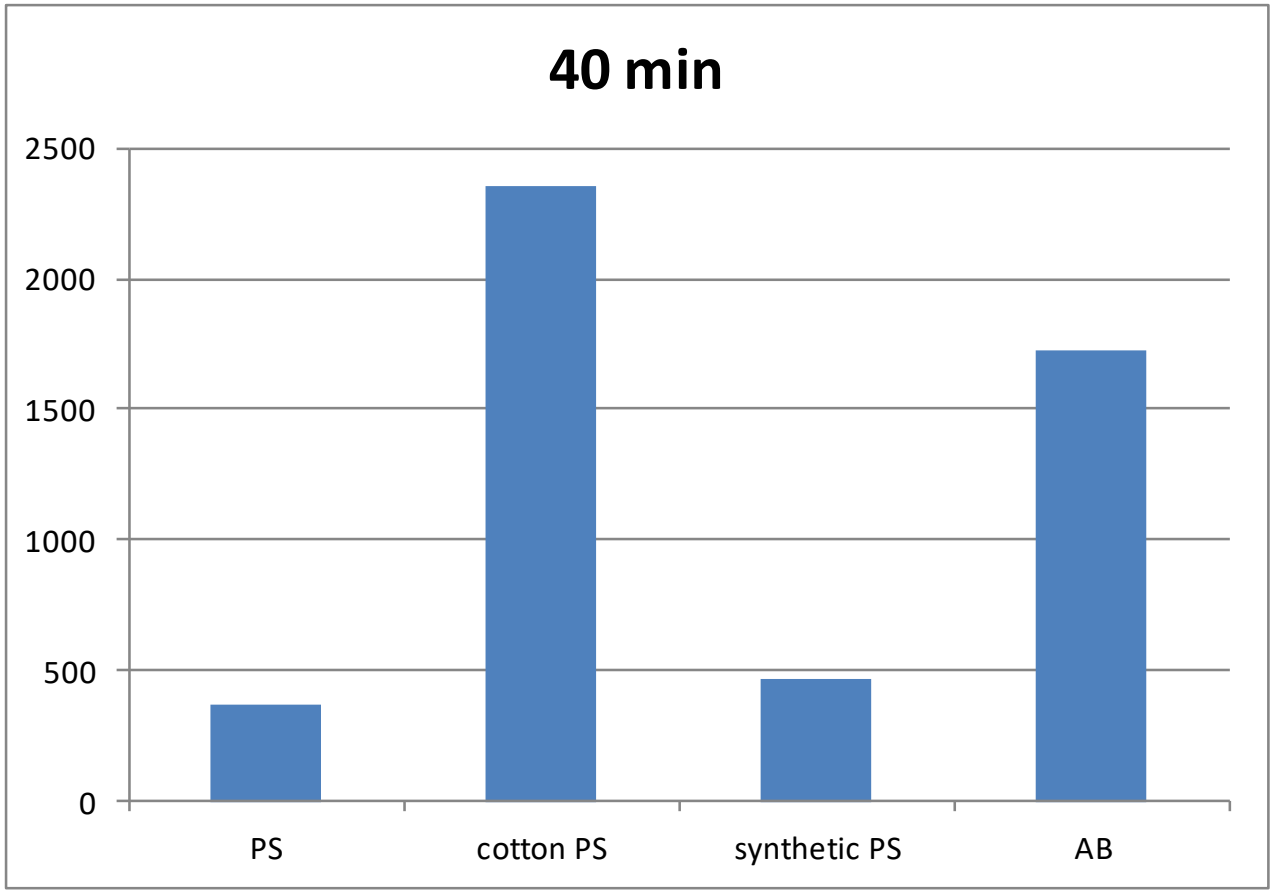
Table 1 **Components of authentic human saliva and a comparison of the normal range of the concentrations between saliva and other biological fluids**

Real saliva compositions		Normal range		Ref
		Saliva	Other biological fluids	
1. Inorganic compounds	Na^+	20-80 mmol/L	Plasma 145 mmol/L	[11]
	K^+	20 mmol/L	4 mmol/L	
	Ca^{2+}	1-4 mmol/L	2.2 mmol/L	
	Cl^-	30-100 mmol/L	120 mmol/L	
	HCO_3^-	15-80 mmol/L	25 mmol/L	
	Phosphate	4 mmol/L	1.2 mmol/L	
	Mg^{2+}	0.2 mmol/L	1.2 mmol/L	
	SCN^-	2 mmol/L	< 0.2 mmol/L	
	NH_3	3 mmol/L	0.05 mmol/L	
2. Organic compounds (non-protein and lipids)	Uric acid	3.38 ± 0.21 mg/dL	Serum 6.31±0.24 mg/dL	[21-23]
		217.2±110.3 mol/L		
		0.1–7.5 mg/dL		
	Bilirubin	0.5-5.0 µmol/L	Serum 0.2-1.2 mg/dL	[24]
	Creatinine	0.12 ± 0.06 mg/dL	Serum 0.89 ± 0.17 mg/dL	[25]
		0.05–0.2 mg/dL	Serum 0.6–1.5 mg/dL	
	Glucose	91.3±10.1 mg/dL	Plasma 80-120 mg/dL	[22, 26]
		4-13 mg/dL		
	Cholesterol	0.02-5.46 µmol/L	Serum <5 mmol/L	[27]
	Lactate	0.3-1.8 mM	Serum 0.5-1.0 mM	[28, 29]
		0.1 to 2.5 mmol/L		
3. Protein/Polypeptide compounds	a-Amylase	19-308 U/mL*	Serum 0.05-0.125 U/mL*	[23, 30]
		93±62 U/L *		
		2.64±1.8 mg/mL		
	Albumin	0.2±0.1 mg/mL	Serum 3.5-5.5 g/dL	[31]
	Secretory-IgA	80–717 mg/dL	Serum 70-400 mg/dL	[23, 32]
		124.3-333.5 µg/mL		
	Mucins group	MUC5B: 2.4±1.7 U/mL	Serum 9.9 ± 0.8 ng/ml	[31, 33]
		1.19±0.17 mg/mL		
	Lysozyme	3–50 µg/mL	Serum 7.4 ± 1.8 mg/mL	[23, 32, 34]
		59.7 to 1062.3 µg/ml	Serum 4-9 µg/mL	
	Total proteins	7.1–223.2 mg/dL	Serum 6-8 g/dL	[23, 31]
		0.9±0.2 mg/mL		
4. Hormones	Cortisol	3.5-27.0 mg/dL	Serum 2-25 mg/dL	[35]
	Testosterone	32-55 pg/mL	Serum 320-600 ng/dL	[36]
	Progesterone	Luteal phase 436 ±34 pmol/L	Serum Male: < 1 ng/mL	[37]
		Follicular phase 22.1±2.7 pmol/L	Serum Female: 0.1-20 ng/mL	
	Estrogen(Estradiol)	Luteal phase 20.6 ±0.4 pmol/L	Serum Male: 15-60 pg/mL	[37]
			Serum Female: 15-370 pg/mL	

*U/mL: enzymatic activity per unit (mL) of saliva

Preliminary tests on saliva

Test with 10 uL of HUG and 150 uL of saliva



ATTENZIONE: non e' stata considerata la fluorescenza spontanea della saliva!!!

Binding of bilirubin to HUG protein

The Br-HUG interaction was studied by mean of the fluorescence titration technique assessing the intensity enhancement during ligand addition

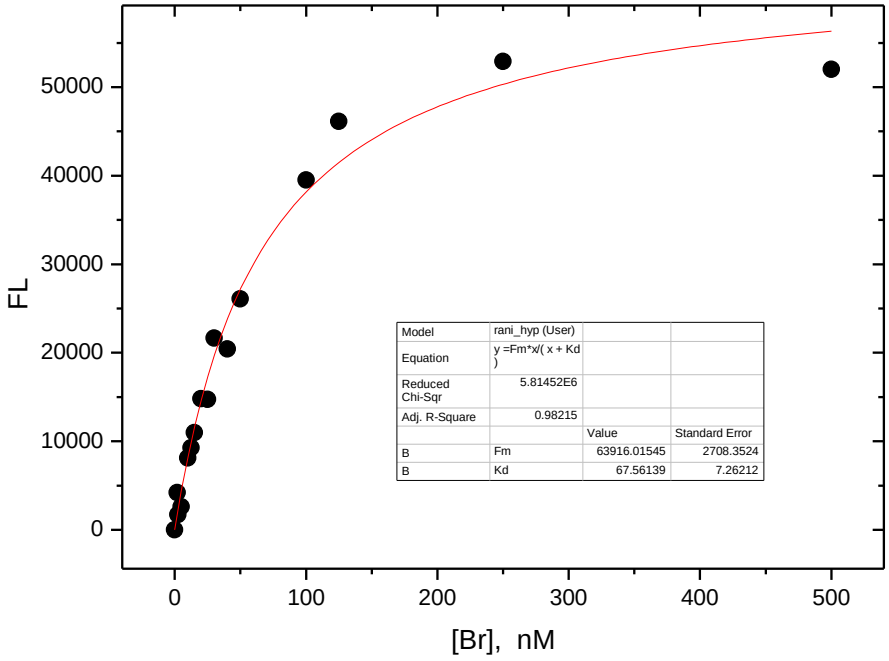
$$Br + Alb \overset{K^A_a}{\leftrightharpoons} (Br - Alb)$$

+

$$HUG$$

$\downarrow\uparrow$ $K_a = 1/K_D$

$$(Br - HUG)$$ **fluorescence**



For each fluorescence value (F), the fractional enhancement (Y) was computed by the equation:

$$Y = \frac{F}{F_0}$$

where Y is the fractional saturation related to the extent of binding and F_o the fluorescence intensity at the Br/HUG ratio greater than 1 (the asymptotic value).

$$K_D = \frac{[HUG][Br]}{[Br - HUG]}$$

During the binding titration the Br concentration is increased so that saturation Y is expressed in terms of the Br-HUG complex concentration, $[Br\text{-}HUG]$ as:

$$Y = \frac{[Br - HUG]}{[HUG]_T} \qquad (1)$$

where $[HUG]_T$ is the total HUG concentration used for the measurement. Then:

$$[Br - HUG]_{\square} = \frac{[HUG][Br]}{K_D}$$

$$[HUG]_T = [HUG] + [Br - HUG] = P_T$$

$$Y = \frac{[HUG][Br]/K_D}{[HUG] + [HUG][Br]/K_D}$$

$$Y = \frac{[Br]}{K_D + [Br]}$$

Since in the binding measurement instead of free concentrations the total protein and ligand concentrations are known, an expression of Y as a function of total quantities is derived as follows:

$$[Br] = [Br]_T - [Br - HUG] = L_T - [Br - HUG]$$

$$[HUG] = P_T - [Br - HUG]$$

then:

$$K_D = \frac{(L_T - [Br - HUG])(P_T - [B - rHUG])}{[Br - HUG]}$$

$$K_D[Br - HUG] = P_T L_T - (P_T + L_T)[Br - HUG] + [Br - HUG]^2$$

$$[Br - HUG]^2 - (P_T + L_T + K_D)[Br - HUG] + P_T L_T = 0$$

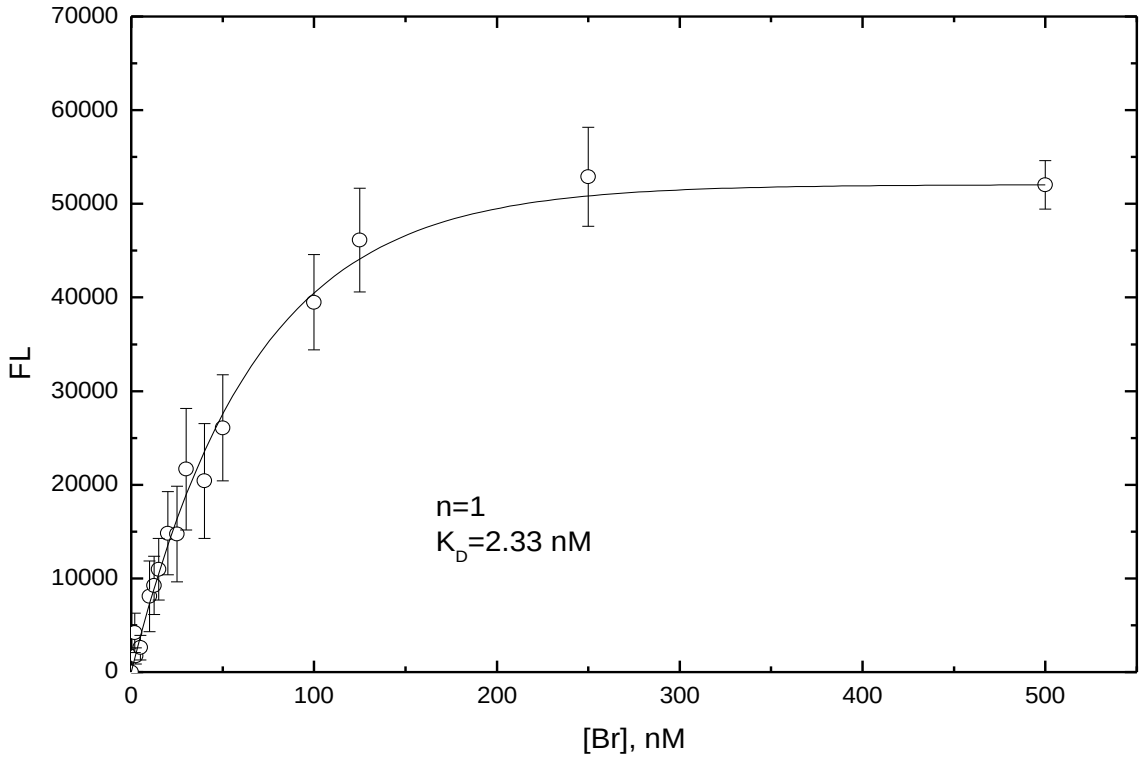
The root of the equation is:

$$[Br - HUG] = \frac{(P_T + L_T + K_D) - \sqrt{\left\{\left(-(P_T + L_T + K_D)\right)^2 - 4P_T L_T\right\}}}{2}$$

By substitution in the equation (1):

$$Y = \frac{(P_T + L_T + K_D) - \sqrt{\left\{\left(-(P_T + L_T + K_D)\right)^2 - 4P_T L_T\right\}}}{2P_T} \qquad (2)$$

The plot of $F=F_oY$ versus $L_T = [Br]$ is shown in Figure where average values of all the results obtained by several experiments are reported.



By nonlinear least-squares fitting of the hyperbolic curve the best evaluation of equilibrium dissociation constant (K_D=2.33 nM) and of the maximal fluorescence value of the ligand-bound protein (F_o=56300) was obtained

$$Br + Alb \overset{K^{A}_{a}}{\leftrightharpoons} (Br - Alb)$$

+

$$HUG$$

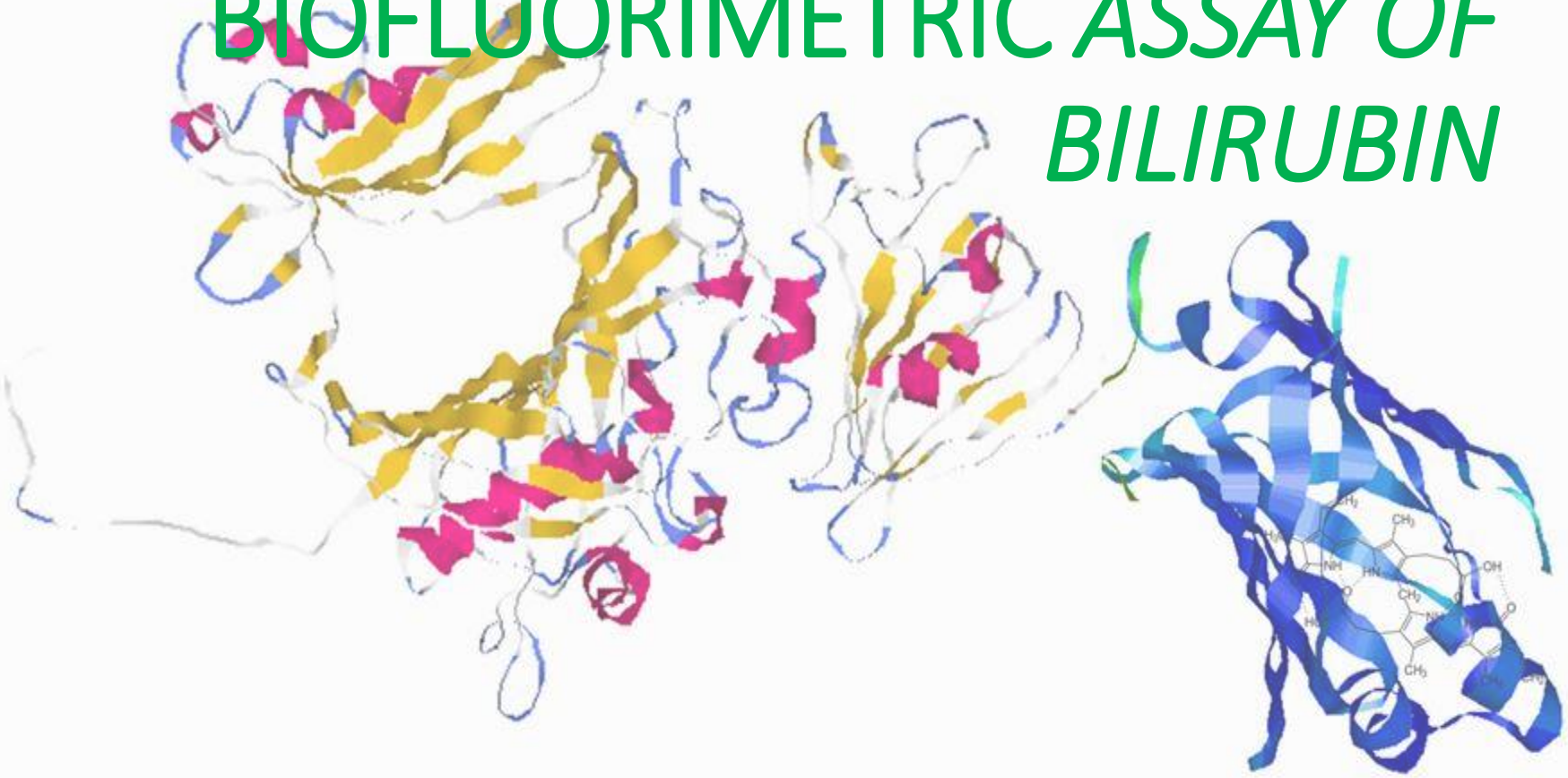
$\downarrow\uparrow$ $K_a = 1/K_D$

$$(Br - HUG) \quad \textbf{fluorescence}$$

$K_a = 1/K_D$ = **4.3 x 10^8 M^{-1}** per *Br-HUG*

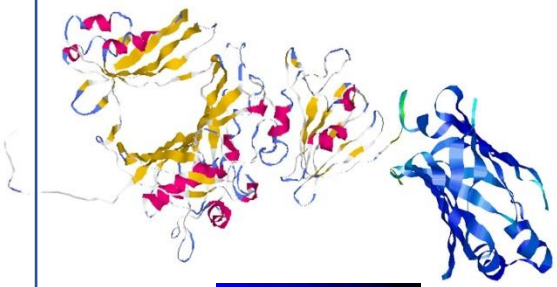
$K^{A}_{a} = 1/K_D$ = **ca 10^6 M^{-1}** per *Br-Alb*

BIOFLUORIMETRIC *ASSAY OF BILIRUBIN*



BIOFLUORIMETRIC *ASSAY OF SERUM BILIRUBIN*

- Direct analysis of indirect bilirubin by HUG



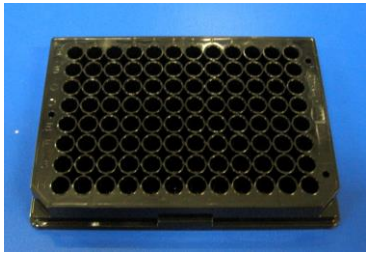
1

HUG is a synthetic protein with a bilirubin-binding domain (UnaG) fused with Human Elastin-like Polypeptide

HUG solution

Add sample

2. HUG binds bilirubin in a sample (< 30 µL) and emits fluorescence



3. Samples are placed in multiwell plates

4. Fluorescence is measured by reader

The emission fluorescence of the complex was detected at λ = 528 nm following excitation at λ = 485 nm

Future goals:

1. Molecular Dynamics simulation of the single protein (HUG and HELP),

2. Molecular Dynamics of binding process (HUG-BR),

3. A thorough study of the reverse thermal transition of HUG in comparison to HELP (which is a critical property for a future upscale of biopolymer production),

4. Bilirubin assay on fish serum samples.