

Projects

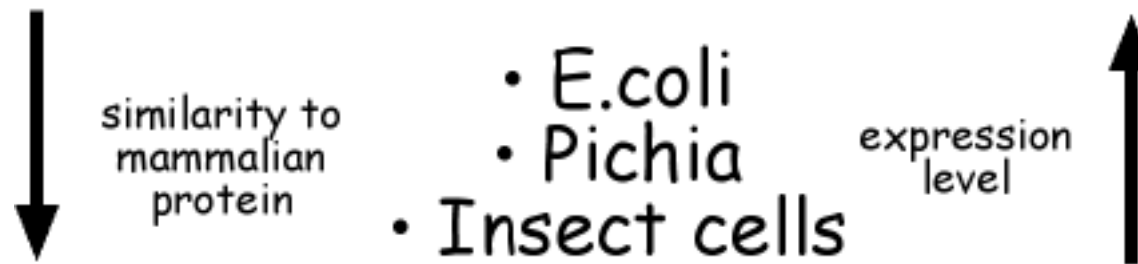
- Recombinant protein production for immunization & X-ray crystal structure (Baculovirus, E.Coli, Pichia)
- Development of a purification procedure for untagged VEGF-C or tag removal
- Further characterization of VEGF-C (confirm? Vladimir's data on affinities and non-Covalent dimerization; heterodimerization with VEGF-D, significance of glycosylation, heparin-affinity, heterogenous signal peptide cleavage/glycosylation/proteolytic processing, comparison with VEGF-D, what is the biologically active form in vivo?)
- Identification of ligand-receptor interaction determinants in VEGF and VEGF-C (hybrid molecule approach)
- (Production of VEGF-B heterodimers in mammalian cells)

Recombinant protein production

Purposes:

1. For immunization of mice (monoclonal antibodies)
2. Source for experiments (CAM assay, rabbit cornea assay, etc.)
3. X-ray Crystal structure

Systems:



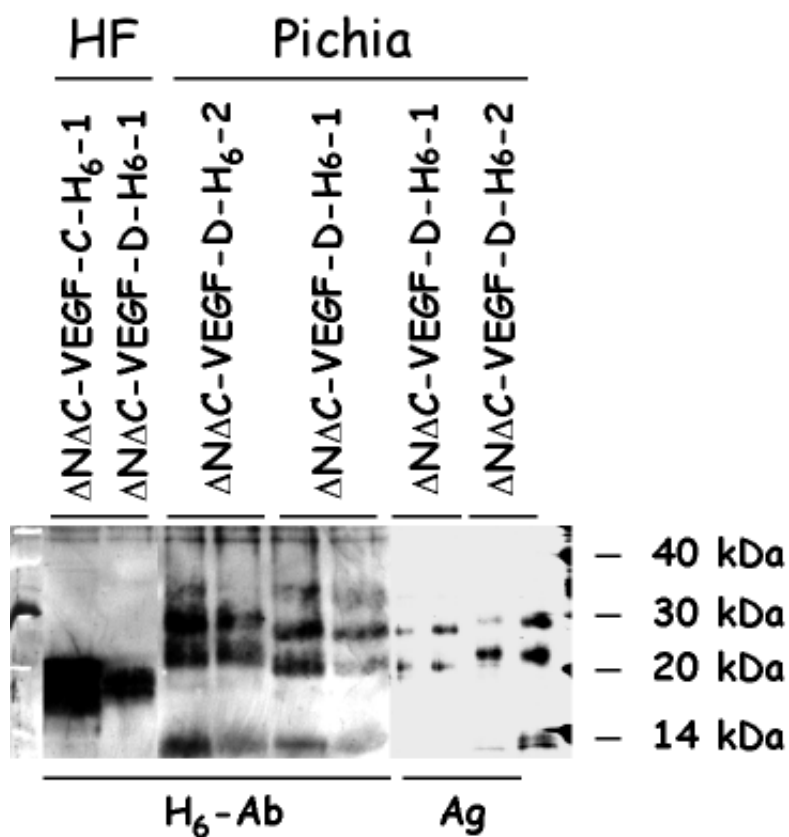
Legend: red = protein successfully expressed, green = these constructs are only intermediates of two-step clonings, but theoretically capable of expression, blue = not yet analyzed, melSP = mellitin signal peptide

Baculovirus construct	Relative amount/remarks
VEGF-A	
pVL1393-hVEGF-A165 (=pVE-165) from Herbert Weich	high (>1 mg/l)
pFB1-hVEGF-A165 from Nina Korsisaari	low (< 0.1 mg/l)
VEGF-B	
167	
pVTBac-melSP-hVEGF-B167	
pFB1-hVEGF-B167 from Nina Korsisaari	low/intracellular
pFB1-melSP-hVEGF-B167	low/intracellular
186	
pFB1-hVEGF-B186-H ₆ (COOH)	medium (0.1-1 mg/l)
pFB1-melSP-hVEGF-B186-H ₆ (COOH)	medium
VEGF-C	
Full length constructs	
pFB1-hVEGF-C-FL	low
pFB1-hVEGF-C-FL-H ₆ (COOH ³)	low
pFB1-hVEGF-C-FL-H ₆ (NH ₂)	low
pFB1-melSP-hVEGF-C-FL-H ₆	
pFB1-hVEGF-C-FL-FLAG(int)	
VEGF homology domain constructs	
pFB1-ΔNΔC-melSP-hVEGF-C-H ₆ (COOH ¹)	high
pFB1-ΔNΔC-melSP-hVEGF-C-H ₆ (COOH ²)	high
pFB1-ΔNΔC-hVEGF-C-H ₆ (COOH ²)	medium
pFB1-ΔNΔC-melSP-hVEGF-C-FLAG(COOH)	

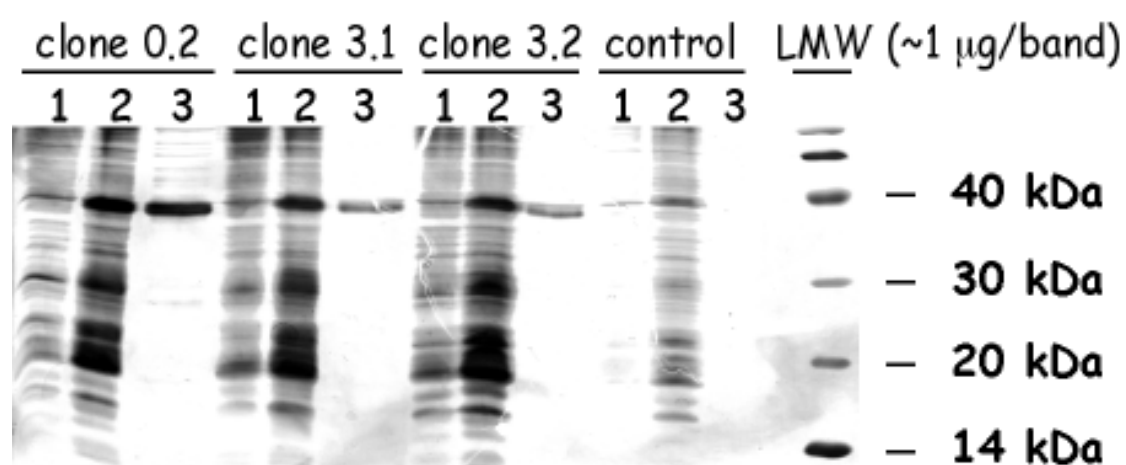
Constructs missing either the NH₂- or COOH-terminal propeptide	
pVTBac-ΔN-melSP-hVEGF-C	
pFB1-ΔC-hVEGF-C-H ₆ (COOH ¹)	medium
pFB1-ΔC-hVEGF-C-H ₆ (COOH ²)	medium
Short splice variants	
pFB1-hVEGF-C-SSV	low
pFB1-ΔN-melSP-hVEGF-C-SSV	not detectable
Mutants	
pFB1-ΔNΔC-hVEGF-C-156C⇒S	medium
pFB1-hVEGF-C-allNXS/T⇒QXS/T	

VEGF-D	
Full length constructs	
pFB1-hVEGF-D-FL	low
pFB1-hVEGF-D-FL-H ₆ (COOH ³)	low
pFB1-melSP-hVEGF-D-FL-H ₆ (COOH ³)	
pFB1-hVEGF-D-FL-MYC(int)	
VEGF homology domain constructs	
pFB1-ΔNΔC-melSP-hVEGF-D(COOH ¹)	medium
pFB1-ΔNΔC-melSP-hVEGF-D(COOH ²)	medium
pFB1-ΔNΔC-melSP-hVEGF-D-MYC(COOH)	
Constructs missing either the NH₂- or COOH-terminal propeptide	
pFB1-ΔC-hVEGF-D(COOH ¹)	
pFB1-ΔC-hVEGF-D(COOH ²)	
pFB1-ΔN-melSP-hVEGF-D(COOH ³)	

Pichia construct	Relative amount/remarks
VEGF-C	
ask Vijay Kumar	
VEGF-D	
<i>Constructs missing the NH₂-terminal propeptide</i>	
pPIC9- α fac- Δ N-hVEGF-D-FL	
pHIL-S1- Δ N-PHO1-hVEGF-D-FL	
pPIC9- α fac- Δ N-hVEGF-D-FL-H6(COOH ³)	
pHIL-S1- Δ N-PHO1-hVEGF-D-FL-H6(COOH ³)	
<i>VEGF homology domain constructs</i>	
pPIC9- Δ N Δ C- α fac-hVEGF-D(COOH ¹)	high/hyperglycosylated?
pPIC9- Δ N Δ C- α fac-hVEGF-D(COOH ²)	high/hyperglycosylated?
pHIL-S1- Δ N Δ C-PHO1-hVEGF-D(COOH ¹)	high/hyperglycosylated?
pHIL-S1- Δ N Δ C-PHO1-hVEGF-D(COOH ²)	high/hyperglycosylated?
pPIC9- Δ N Δ C- α fac-hVEGF-D(COOH ²)	
aII Δ NXS/T $\Rightarrow$ QXS/T	



hVEGF-D-FL-H₆(COOH³) hVEGF-C-FL



1 = lysate of 80 µl culture, no induction

2 = lysate of 80 µl culture, 2 h induction

3 = Ni²⁺NTA miniprep of 1 ml culture, denaturing conditions

<i>E.coli</i> construct	protein	resistance	hosts	expression
-------------------------	---------	------------	-------	------------

full length human VEGF-C (MW_{calc.} ~45.9 kDa including fMet and H₆ tag)

<i>pET15b-hVEGF-C-FL</i>	untagged, starting with MPGPREA	amp	BL21(DE3) AD494(DE3)	very low none
<i>pET15b-hVEGF-C-FL-H₆(COOH³)</i>	his-tagged, starting with MPGPREA	amp	BL21(DE3) AD494(DE3)	very low none
<i>pET28b(+)-hVEGF-C-FL</i>	untagged, starting with MPGPREA	kana	BL21 (DE3)	low
<i>pET28b(+)-hVEGF-C-FL-H₆(COOH³)</i>	his-tagged, starting with MPGPREA	kana	BL21 (DE3)	low
<i>pET15b(-)-hVEGF-C-FL-PPM</i>	untagged, starting with MESGLDL	amp	BL21(DE3)	none
<i>pET28b(+)-hVEGF-C-FL-H₆(COOH³)-PPM</i>	untagged, starting with MESGLDL	kana	BL21(DE3)	none
<i>pET28b(+)-hVEGF-C-FL-H₆(NH₂)-PPM</i>	N-terminally his-tagged starting with MGSSHHHHHHESGLDL	kana		

ΔNΔC human VEGF-C (MW_{calc.} ~38.8 kDa including fMet and H₆ tag)

<i>pET28b(+)-ΔNΔC-hVEGF-C-H₆(COOH¹)</i>	his-tagged VEGF-homology domain starting with MDPTEETIK	kana		
---	--	------	--	--

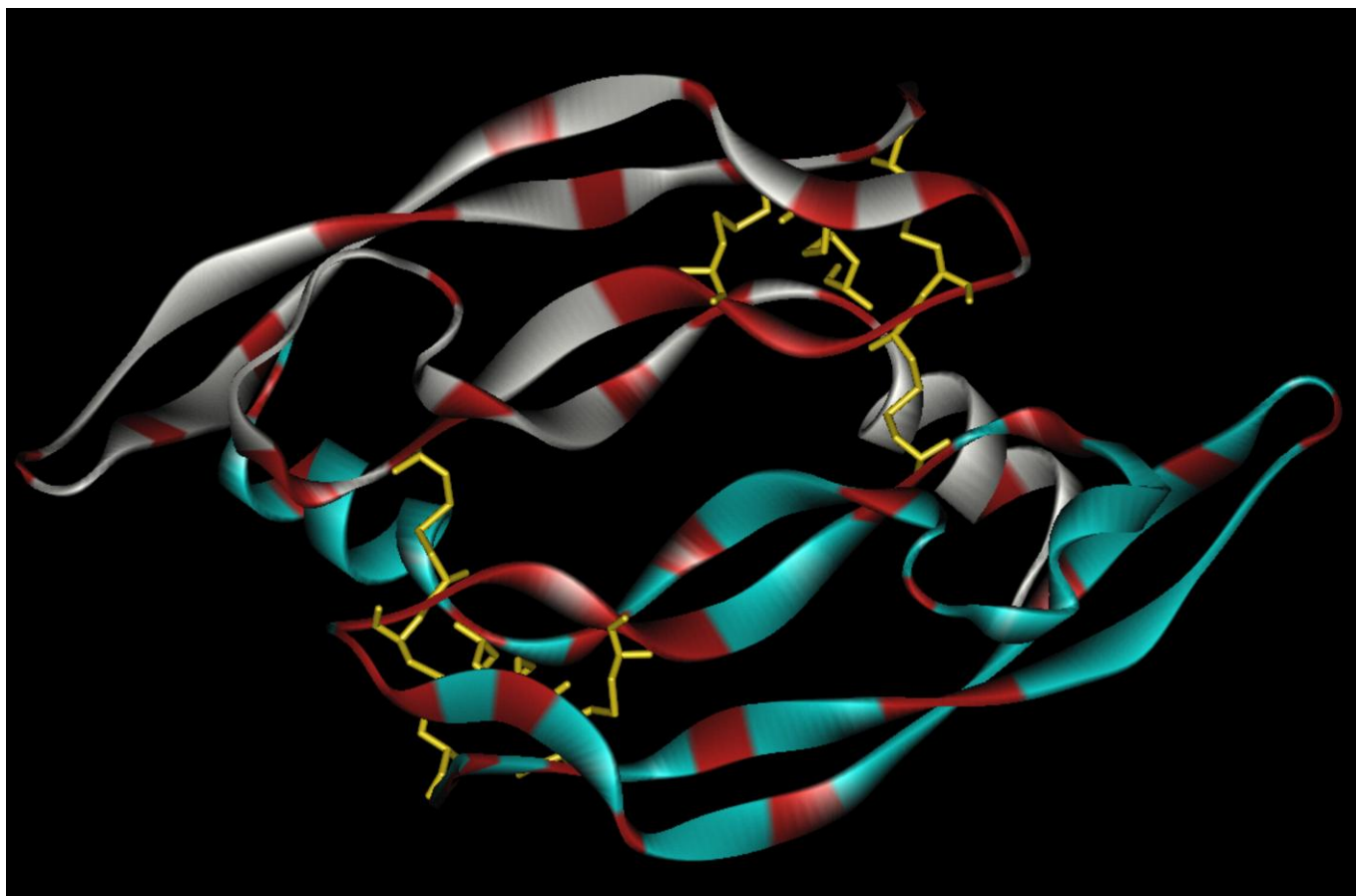
full length human VEGF-D (MW_{calc.} ~38.8 kDa including fMet and H₆ tag)

<i>pET15b-hVEGF-D-FL</i>	untagged, starting with MSSNEHG	amp		
<i>pET15b-hVEGF-D-FL-H₆</i>	his-tagged, starting with MSSNEHG	amp	BL21(DE3) AD494 (DE3)	medium none
<i>pET28b(+)-hVEGF-D-FL</i>	untagged, starting with MSSNEHG	kana		
<i>pET28b(+)-hVEGF-D-FL-H₆</i>	his-tagged, starting with MSSNEHG	kana	BL21 (DE3)	high

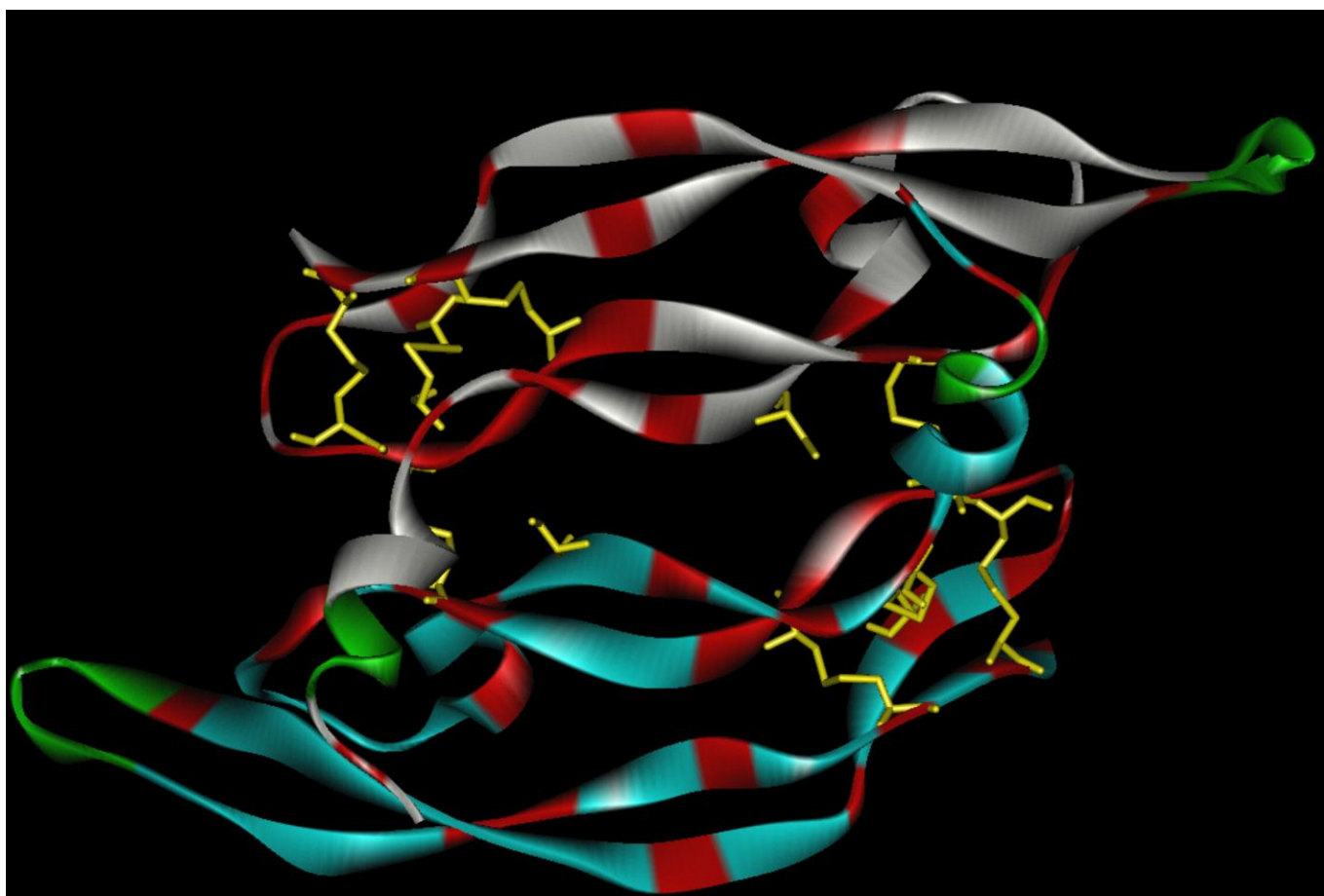
	1				Ab 905	50
hVEGF-C	MHLLGFFSVA	CSLLA..AAL	LPGPREAPAA	AAAF.....	ESGLDLSDAE	
mVEGF-C	MHLLCFLSLA	CSLLA..AAL	IPSPREAPAT	VAAF.....	ESGLGFSEAE	
hVEGF-D	..MYREWVVV	NVFMMLYVQL	VQGSSNEHGP	VK.....RSS	QSTLERS...	
mVEGF-D	..MYGEWGMG	NILMMFHVYL	VQGFRSEHGP	VKDFSFERSS	RSMLERS...	
	51					100
hVEGF-C	PDAGEATAYA	SKDLEEQLRS	VSSVDELMTV	LYPEYWKMYK	CQLRKGGWQH	
mVEGF-C	PDGGEVKAFE	GKDLEEQLRS	VSSVDELMSV	LYPDYWKMYK	CQLRKGGWQ.	
		Ab Achen 1				
hVEGF-D		EQQIRA	ASSLEELLRI	THSEDWKLWR	CRLRLKSF..	
mVEGF-D		EQQIRA	ASSLEELLQI	AHSEDWKLWR	CRLKLKSL..	
	101		Ab 882			150
hVEGF-C	NREQANLNSR	TEE..TIKFA	AAHYNTEILK	SIDNEWKRTQ	CMPREVCIDV	
mVEGF-C	...QPTLNTR	TGD..SVKFA	AAHYNTEILK	SIDNEWKRTQ	CMPREVCIDV	
		Ab 928 & 929				
hVEGF-D	TSMDSR	SASHRSTRFA	ATFYDIETLK	VIDEEWQRTQ	CSPRETCVEV	
mVEGF-D	ASMDSR	SASHRSTRFA	ATFYDTETLK	VIDEEWQRTQ	CSPRETCVEV	
	151					200
hVEGF-C	GKEFGVATNT	FFKPPCVSVY	RCGGCCNSEG	LQCMNTTSTSY	LSKTLFEITV	
mVEGF-C	GKEFGAATNT	FFKPPCVSVY	RCGGCCNSEG	LQCMNTTSTGY	LSKTLFEITV	
hVEGF-D	ASELGKSTNT	FFKPPCVNVF	RCGGCCNEES	LICMNTTSTSY	ISKQLFEISV	
mVEGF-D	ASELGKTTNT	FFKPPCVNVF	RCGGCCNEEG	VMCMNTTSTSY	ISKQLFEISV	
	201		↓C1HIS	↓C2HIS		250
hVEGF-C	PLSQGPKPVT	ISFANHHTSCR	CMSKLDVYRQ	VHSIIRRSPL	.ATLPQCQAA	
mVEGF-C	PLSQGPKPVT	ISFANHHTSCR	CMSKLDVYRQ	VHSIIRRSPL	.ATLPQCQAA	
		Ab Achen 2				
			↓C1HIS	↓C2HIS		
hVEGF-D	PLTSVPPELVP	VKVANHTGCK	CLPTAP..RH	PYSIIRRSIQ	IPEEDRCSHS	
mVEGF-D	PLTSVPPELVP	VKIANHTGCK	CLPTGP..RH	PYSIIRRSIQ	TPEEDECPHS	
	251					300
hVEGF-C	NKTCPTNYMW	NNHICRCLAQ	EDFMFSSDAG	DDSTDGFHDI	CGPNKELDEE	
mVEGF-C	NKTCPTNYVW	NNYMCRCCLAQ	QDFIFYSNVE	DDSTNGFHDV	CGPNKELDED	
hVEGF-D	KKLCPIDMLW	DSNKCKCVLQ	EENPL.AGTE	DHS.....	HLQE.	
mVEGF-D	KKLCPIDMLW	DNTKCKCVLQ	DETPL.PGTE	DHS.....	YLQE.	
	301					350
hVEGF-C	TCQCVCRAGL	RPASCQPHKE	LDRNSCQCVC	KNKLFPSQCG	ANREFDENTC	
mVEGF-C	TCQCVCKGGL	RPSSCQPHKE	LDRDSCQCVC	KNKLFPSQCG	ANREFDENTC	
hVEGF-D				PALCG	PHMMFDEDRC	
mVEGF-D				PTLCG	PHMTFDEDRC	
	351					400
hVEGF-C	QCVCKRTCPR	NQPLNPGKCA	C.ECTESPQK	CLLKGGKFHH	QTCSC.....	
mVEGF-C	QCVCKRTCPR	NQPLNPGKCA	C.ECTENTQK	CFLKGKKFHH	QTCSC.....	
		Ab Achen 3				
hVEGF-D	ECVCKTPCPK	DLIQHPKNCS	CFECKESLET	CCQKHKLFP	DTCSCEDRCP	
mVEGF-D	ECVCKAPCPG	DLIQHPENCS	CFECKESLES	CCQKHKIFHP	DTCSCEDRCP	
	401			↓C3HIS		
hVEGF-C	.YRRPCTNRQ	KACEPGFSYS	EEVCRCVPSY	WKRPMQMS*		
mVEGF-C	.YRRPCANRL	KHCDPGLSFS	EEVCRCVPSY	WKRPHLN*		
				↓C3HIS		
hVEGF-D	FHTRPCASGK	TACAKHCRFP	KEKRAAQGPH	SRKNP*..		
mVEGF-D	FHTRTCASRK	PACGKHWRF	KETR.AQGLY	SQENP*..		

HUMAN-C	1	H	Y	N	T	E	I	L	K	S	I	D	N	E	W	R	K	T	Q	C	M	P	R	E	V	C	I	D	V	G	K	E	F	G	V	A	T	N	T	F	F	K	P	P	C	V	S	V	Y	R	C	G	G	C	C	N	S	
MOUSE-C	1	H	Y	N	T	E	I	L	K	S	I	D	N	E	W	R	K	T	Q	C	M	P	R	E	V	C	I	D	V	G	K	E	F	G	V	A	T	N	T	F	F	K	P	P	C	V	S	V	Y	R	C	G	G	C	C	N	S	
QUAIL-C	1	H	Y	N	A	E	I	L	K	S	I	D	T	E	W	R	K	T	Q	C	M	P	R	E	V	C	V	D	V	G	K	E	F	G	A	T	N	T	F	F	K	P	P	C	V	S	I	Y	R	C	G	G	C	C	N	S		
HUMAN-D	1	F	Y	D	I	E	T	L	K	V	I	D	E	E	W	O	R	T	Q	C	S	P	R	E	T	C	V	E	V	A	S	E	L	G	K	S	T	N	T	F	F	K	P	P	C	V	N	V	F	R	C	G	G	C	C	N	E	
MOUSE-D	1	F	Y	D	T	E	T	L	K	V	I	D	E	E	W	O	R	T	Q	C	S	P	R	E	T	C	V	E	V	A	S	E	L	G	K	S	T	N	T	F	F	K	P	P	C	V	N	V	F	R	C	G	G	C	C	N	E	
HUMAN-A	1	Q	N	H	E	V	V	K	F	M	D	-	V	Y	O	R	S	Y	C	H	P	I	E	T	L	V	D	I	F	O	Q	E	Y	P	D	E	I	E	Y	I	F	K	P	S	C	V	P	L	M	R	C	G	G	C	C	N	D	
PIG-A	1	Q	K	P	H	E	V	V	K	F	M	D	-	V	Y	O	R	S	Y	C	R	P	I	E	T	L	V	D	I	F	O	Q	E	Y	P	D	E	I	E	Y	I	F	K	P	S	C	V	P	L	M	R	C	G	G	C	C	N	D
SHEEP-A	1	Q	K	P	H	E	V	M	K	F	M	D	-	V	Y	O	R	S	Y	C	R	P	I	E	T	L	V	D	I	F	O	Q	E	Y	P	D	E	I	E	Y	I	F	K	P	S	C	V	P	L	M	R	C	G	G	C	C	N	D
BOVINE-A	1	Q	K	P	H	E	V	V	K	F	M	D	-	V	Y	O	R	S	Y	C	R	P	I	E	T	L	V	D	I	F	O	Q	E	Y	P	D	E	I	E	Y	I	F	K	P	S	C	V	P	L	M	R	C	G	G	C	C	N	D
RABBIT-A	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G.PIG-A	1	Q	K	P	R	E	E	V	K	F	M	D	-	V	Y	K	R	S	Y	C	R	P	I	E	M	L	V	D	I	F	O	Q	E	Y	P	D	E	I	E	Y	I	F	K	P	S	C	V	P	L	M	R	C	G	G	C	C	N	D
MOUSE-A	1	Q	K	S	H	E	V	V	K	F	M	D	-	V	Y	O	R	S	Y	C	R	P	I	E	T	L	V	D	I	F	O	Q	E	Y	P	D	E	I	E	Y	I	F	K	P	S	C	V	P	L	M	R	C	A	G	C	C	N	D
RAT-A	1	Q	K	A	H	E	V	V	K	F	M	D	-	V	Y	O	R	S	Y	C	R	P	I	E	T	L	V	D	I	F	O	Q	E	Y	P	D	E	I	E	Y	I	F	K	P	S	C	V	P	L	M	R	C	A	G	C	C	N	D
XENOPUS-A	1	H	K	P	T	E	V	V	K	F	L	K	-	V	Y	E	R	S	M	C	Q	V	R	E	I	L	V	D	I	F	O	Q	E	Y	P	D	E	V	E	Y	I	F	K	P	S	C	V	P	L	M	R	C	A	G	C	C	N	D
QUAIL-A	1	R	K	P	N	E	V	I	K	F	L	E	-	V	Y	E	R	S	F	C	R	T	I	E	T	L	V	D	I	F	O	Q	E	Y	P	D	E	V	E	Y	I	F	R	P	S	C	V	P	L	M	R	C	A	G	C	C	G	D
CHICK-A	1	R	K	P	N	E	V	I	K	F	L	E	-	V	Y	E	R	S	F	C	R	T	I	E	T	L	V	D	I	F	O	Q	E	Y	P	D	E	V	E	Y	I	F	R	P	S	C	V	P	L	M	R	C	A	G	C	C	G	D
ZEBRAF-A	1	K	S	K	N	D	V	I	P	F	M	D	-	V	Y	K	K	S	A	C	K	T	R	E	L	L	V	D	I	I	O	Q	E	Y	P	D	E	I	E	H	T	Y	I	P	S	C	V	V	L	M	R	C	A	G	C	C	N	D
MOUSE-P	1	S	T	E	V	E	V	V	P	F	N	E	-	V	W	G	R	S	Y	C	R	P	M	E	K	L	V	Y	I	A	D	E	H	P	N	E	V	S	H	I	F	S	P	S	C	V	L	L	S	R	C	S	G	C	C	G	D	
RAT-P	1	S	T	E	M	E	V	V	P	F	N	E	-	V	W	G	R	S	Y	C	R	P	M	E	K	L	V	Y	I	A	D	E	H	P	N	E	V	S	H	I	F	S	P	S	C	V	L	L	S	R	C	S	G	C	C	G	D	
HUMAN-P	1	S	S	E	V	E	V	V	P	F	O	E	-	V	W	G	R	S	Y	C	R	A	L	E	R	L	V	D	V	V	S	E	Y	P	S	E	V	E	H	M	F	S	P	S	C	V	S	L	L	R	C	T	G	C	C	G	D	
MOUSE-B	1	S	H	O	K	K	V	V	P	W	I	D	-	V	Y	A	R	A	T	C	O	P	R	E	V	V	V	P	L	S	M	E	L	M	G	N	V	V	K	O	L	V	P	S	C	V	T	V	O	R	C	G	G	C	C	P	D	
RAT-B	1	-	-	K	R	K	V	V	S	W	I	D	-	V	Y	A	R	A	T	C	O	P	R	E	V	V	V	P	L	S	M	E	L	M	G	N	V	V	K	O	L	V	P	S	C	V	T	V	O	R	C	G	G	C	C	P	D	
HUMAN-B	1	G	H	O	R	K	V	V	S	W	I	D	-	V	Y	T	R	A	T	C	O	P	R	E	V	V	V	P	L	T	V	E	L	M	G	T	V	A	K	O	L	V	P	S	C	V	T	V	O	R	C	G	G	C	C	P	D	
POX-N2	1	N	A	D	S	N	T	K	G	W	S	E	-	V	L	K	G	S	E	C	K	P	R	P	I	V	V	P	V	S	E	T	H	P	E	L	T	S	O	R	F	N	P	P	S	C	V	T	L	M	R	C	S	G	C	C	N	D
POX-N7	1	N	D	S	P	P	S	T	N	D	W	M	R	T	L	D	K	S	G	C	K	P	R	D	T	V	V	Y	L	G	E	E	Y	P	E	S	T	N	L	O	Y	N	P	R	C	V	T	V	K	R	C	S	G	C	C	N	G	

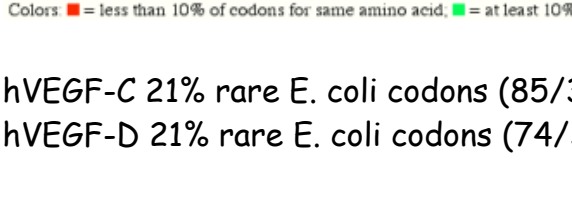
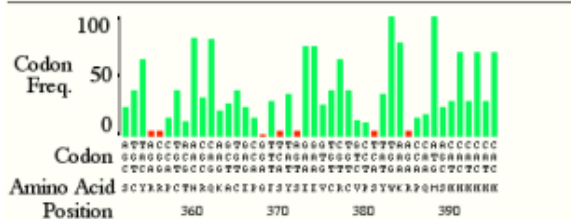
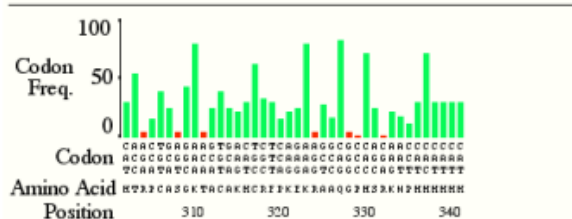
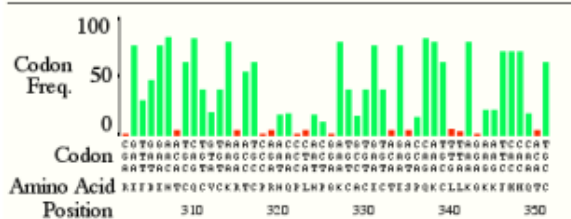
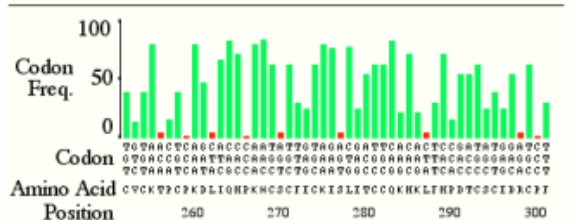
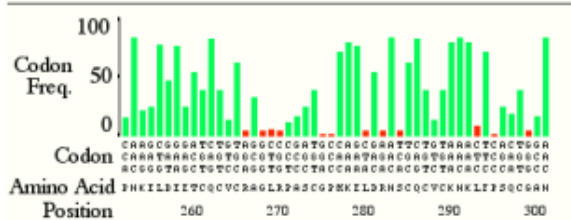
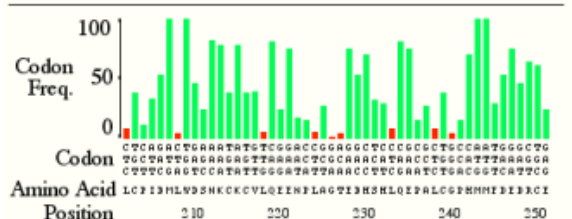
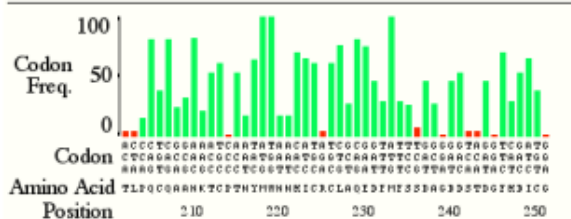
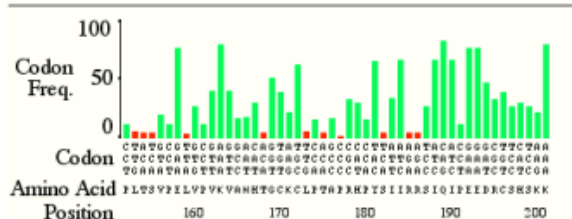
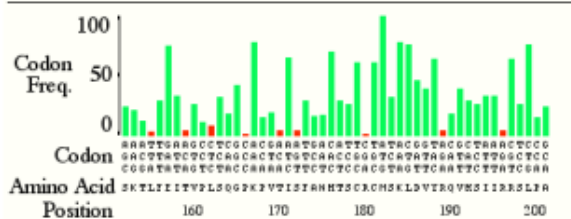
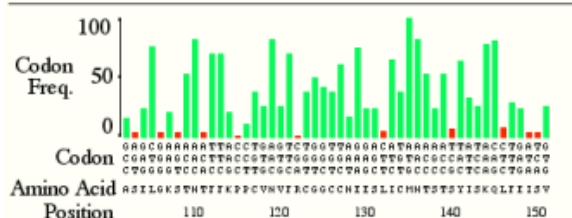
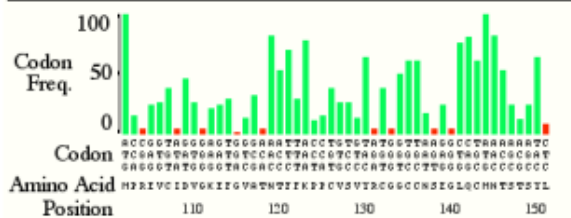
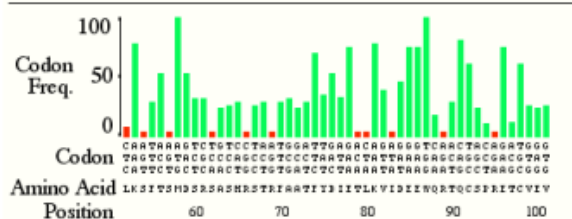
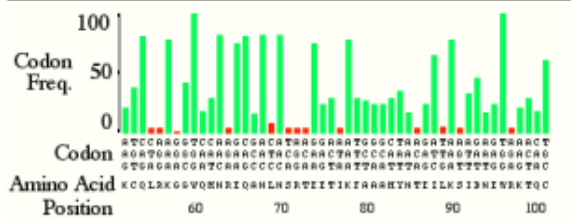
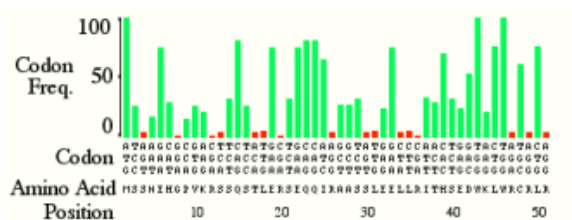
HUMAN-C	57	E	G	L	Q	C	M	N	T	S	T	S	Y	L	S	K	T	L	F	E	I	T	V	P	L	S	Q	-	-	-	-	-	G	P	K	P	V	T	I	S	F	A	N	H	T	S	C	R	C	M	S	K	L	D						
MOUSE-C	57	E	G	L	Q	C	M	N	T	S	T	G	Y	L	S	K	T	L	F	E	I	T	V	P	L	S	Q	-	-	-	-	-	-	G	P	K	P	V	T	I	S	F	A	N	H	T	S	C	R	C	M	S	K	L	D					
QUAIL-C	57	E	G	L	Q	C	M	N	I	S	T	N	Y	I	S	K	T	L	F	E	I	T	V	P	L	S	H	-	-	-	-	-	-	G	P	K	P	V	T	I	S	F	A	N	H	T	S	C	R	C	M	S	K	L	D					
HUMAN-D	57	E	S	L	I	C	M	N	T	S	T	S	Y	I	S	K	Q	L	F	E	I	S	V	P	L	T	S	-	-	-	-	-	-	V	P	E	L	V	P	V	K	V	A	N	H	T	G	C	K	C	L	P	T	A	P					
MOUSE-D	57	E	G	V	M	C	M	N	T	S	T	S	Y	I	S	K	Q	L	F	E	I	S	V	P	L	T	S	-	-	-	-	-	-	V	P	E	L	V	P	V	K	I	A	N	H	T	G	C	K	C	L	P	T	G	P					
HUMAN-A	56	E	G	L	E	C	V	P	T	E	E	S	N	I	T	M	Q	I	M	R	I	K	P	H	Q	G	Q	-	-	-	-	-	-	-	H	I	G	E	M	S	F	L	O	H	N	K	C	E	C	R	P	K	K	D						
PIG-A	56	E	G	L	E	C	V	P	T	E	E	F	N	I	T	M	Q	I	M	R	I	K	P	H	Q	S	Q	-	-	-	-	-	-	-	-	H	I	G	E	M	S	F	L	O	H	N	K	C	E	C	R	P	K	K	D					
SHEEP-A	56	E	S	L	E	C	V	P	T	E	E	F	N	I	T	M	Q	I	M	R	I	K	P	H	Q	S	Q	-	-	-	-	-	-	-	-	H	I	G	E	M	S	F	L	O	H	N	K	C	E	C	R	P	K	K	D					
BOVINE-A	56	E	S	L	E	C	V	P	T	E	E	F	N	I	T	M	Q	I	M	R	I	K	P	H	Q	S	Q	-	-	-	-	-	-	-	-	H	I	G	E	M	S	F	L	O	H	N	K	C	E	C	R	P	K	K	D					
RABBIT-A	32	E	S	L	E	C	V	P	T	E	E	F	N	V	T	M	Q	I	M	R	I	K	P	H	Q	G	Q	-	-	-	-	-	-	-	-	H	I	G	E	M	S	F	L	O	H	N	K	C	E	C	R	P	K	-	-					
G.PIG-A	56	E	S	L	E	C	V	P	T	E	E	F	N	I	T	M	Q	I	M	R	I	K	P	H	Q	G	Q	-	-	-	-	-	-	-	-	H	I	G	E	M	S	F	L	O	H	S	K	C	E	C	R	P	K	K	E					
MOUSE-A	56	E	A	L	E	C	V	P	T	S	E	S	N	I	T	M	Q	I	M	R	I	K	P	H	Q	S	Q	-	-	-	-	-	-	-	-	H	I	E	R	M	S	F	L	O	H	S	R	C	E	C	R	P	K	K	D					
RAT-A	56	E	A	L	E	C	V	P	T	S	E	S	N	V	T	M	Q	I	M	R	I	K	P	H	Q	S	Q	-	-	-	-	-	-	-	-	H	I	G	E	M	S	F	L	O	H	S	R	C	E	C	R	P	K	K	D					
XENOPUS-A	56	E	S	L	E	C	V	P	T	E	C	Y	N	I	T	M	Q	I	M	K	I	K	P	H	I	S	Q	-	-	-	-	-	-	-	-	H	I	M	D	M	S	F	L	O	H	S	Q	C	E	C	R	P	K	K	E					
QUAIL-A	56	E	G	L	E	C	V	P	V	D	V	Y	N	V	T	M	E	I	A	R	I	K	P	H	Q	S	Q	-	-	-	-	-	-	-	-	H	I	A	H	M	S	F	L	O	H	S	K	C	D	C	R	P	K	K	D					
CHICK-A	56	E	G	L	E	C	V	P	V	D	V	Y	N	V	T	M	E	I	A	R	I	K	P	H	Q	S	Q	-	-	-	-	-	-	-	-	H	I	A	H	M	S	F	L	O	H	S	K	C	D	C	R	P	K	K	D					
ZEBRAF-A	56	E	A	L	E	C	V	P	T	E	T	R	N	V	T	M	E	V	L	R	V	K	O	R	V	S	Q	-	-	-	-	-	-	-	-	H	N	F	Q	L	S	F	T	E	H	T	K	C	E	C	R	P	K	A	E					
MOUSE-P	56	E	G	L	H	C	V	P	I	K	T	A	N	I	T	M	Q	I	L	K	I	P	P	N	R	D	P	H	F	Y	-	-	-	-	-	-	-	V	-	E	M	T	F	S	O	D	V	L	C	E	C	R	P	I	L	E				
RAT-P	56	E	G	L	H	C	V	P	A	L	K	T	A	N	I	T	M	Q	I	L	K	I	P	P	N	R	D	P	H	S	Y	-	-	-	-	-	-	-	V	-	E	M	T	F	S	O	D	V	L	C	E	C	R	P	I	L	E			
HUMAN-P	56	E	D	L	H	C	V	P	V	E	T	A	N	V	T	M	Q	L	L	K	I	R	S	G	D	R	P	-	S	Y	-	-	-	-	-	-	-	V	-	E	L	T	F	S	O	H	V	R	C	E	C	R	P	L	R	E				
MOUSE-B	56	D	G	L	E	C	V	P	T	G	O	H	Q	V	R	M	Q	I	L	M	I	Q	Y	-	-	P	S	S	Q	-	-	-	-	-	-	-	-	-	L	G	E	M	S	L	E	E	H	S	Q	C	E	C	R	P	K	K	K	E	D	
RAT-B	54	D	G	L	E	C	V	P	I	G	O	H	Q	V	R	M	Q	I	L	M	I	Q	Y	-	-	P	S	S	Q	-	-	-	-	-	-	-	-	-	-	L	G	E	M	S	L	E	E	H	S	Q	C	E	C	R	P	K	R	K	K	
HUMAN-B	56	D	G	L	E	C	V	P	T	I	G	O	H	Q	V	R	M	Q	I	L	M	I	Q	Y	-	-	P	S	S	Q	-	-	-	-	-	-	-	-	-	-	L	G	E	M	S	L	E	E	H	S	Q	C	E	C	R	P	K	R	K	K
POX-N2	56	E	S	L	E	C	V	P	T	E	E	V	N	V	S	M	E	L	L	M	I	G	A	S	G	-	-	-	S	G	S	N	G	-	-	-	-	-	-	-	M	O	R	L	S	F	V	E	H	K	K	C	D	C	R	P	R	F	T	
POX-N7	57	D	G	O	I	C	T	A	V	E	T	R	N	T	T	V	T	V	S	V	T	G	V	S	S	S	S	G	T	N	S	G	V	S	T	N	L	O	R	I	S	V	T	E	H	T	K	C	D	C	I	G	R	T	T					



VEGF



VEGF-C



Colors: ■ = less than 10% of codons for same amino acid, ■ = at least 10%

Colors: ■ = less than 10% of codons for same amino acid, ■ = at least 10%

hVEGF-C 21% rare E. coli codons (85/394)
hVEGF-D 21% rare E. coli codons (74/340)

NUMBERING FAT BLACK = Overlapping sequence used for annealing short dsDNA fragments
 ACCORDING FAT DARK BLUE = non-CDS of oligos
 TO VEGF RED = cystein
 GREEN = identical amino acid

↓ BAMHI

VEGF-A G| **GAT|CCT|GGG|CAG|AAT|CAT|CAC|GAA|GTG**
 C|CTA|G **GA|CCC|GTC|TTA|GTA|GTG|CTT|CAC**

VEGF-C G| **GAT|CCT|GCA|CAT|TAT|AAT|ACC|GAG|ATC**
 C|CTA|G **GA|CGT|GTA|ATA|TTA|TGG|CTC|TAG**

VEGF-A | D | P | G | Q | N | H | H | E | V
 VEGF-C | D | P | A | H | Y | N | T | E | I

6 7 8 9 10 11 12 13 14

↑↑

6BP-OL 4BP-OL

VEGF-A **GTG|AAA|TTC|ATG|GAT| . |GTC|TAT|CAG|CGC|AGC|TAC|TGC|CAT|CCG|ATC|GAG|ACA|CTG**
CAC|TTT|AAG|TAC|CTA| . |CAG|ATA|GTC|GCG|TCG|ATG|ACG|GTA|GGC|TAG|CAT|TGT|GAC

VEGF-C **CTG|AAA|TCT|ATT|GAT|AAT|GAG|TGG|AGA|AAG|ACT|CAG|TGC|ATG|CCG|AGA|GAG|GTG|TGT**
GAC|TTT|AGA|TAA|CTA|TTA|CTC|ACC|TCT|TTC|TGA|GTC|ACG|TAC|GGC|TCT|CTC|CAC|ACA

VEGF-A V | K | F | M | D | . | V | Y | Q | R | S | Y | C | H | P | I | E | T | L
 VEGF-C L | K | S | I | D | N | E | W | R | K | T | Q | C | M | P | R | E | V | C

15 16 17 18 19 19B 20 21 22 23 24 25 26 27 28 29 30 31 32

SAPI ↓ XHOI KPNI ↓ SAPI

VEGF-A GTG|GAC|ATC|TTC|CAG|G **AATAGAAGAGCTCGAG** **GGTACCGCTCTTCG** **AA|TAC|CCT|GAT**
 CAC|CTG|TAG|AAG|GTC|CTT|A **TCTTCTCGAGCTC** **CCATGGCGAGAAGCTTA** TG|GGA|CTA

VEGF-C ATC|GAC|GTG|GGG|AAG|G **AATAGAAGAGCTCGAG** **GGTACCGCTCTTCG** **AA|TTT|GGA|GTC**
 TAG|CTG|CAC|CCC|TTC|CTT|A **TCTTCTCGAGCTC** **CCATGGCGAGAAGCTTA** AA|CCT|CAG

VEGF-A V | D | I | F | Q E | Y | P | D
 VEGF-C I | D | V | G | K E | F | G | V

33 34 35 36 37 38 39 40 41

11BP-OL ↓ MSCI NCOI

VEGF-A GAG|ATC|GAG|TAC|ATC|TTC|AAG|CCA|TCC|TGC|GTG|CCC|CTG|ATG|AGA|TGT|GG **C|CAT|GG**
 (DS) CTC|TAG|CTC|ATG|TAG|AAG|TTC|GGT|AGG|ACG|CAC|GGG|GAC|TAC|TCT|ACA|CC **G|GTA|CC**

VEGF-C GCG|ACA|AAC|ACC|TTC|TTC|AAG|CCA|CCA|TGT|GTG|TCC|GTG|TAC|AGA|TGT|GG **C|CAT|GG**
 (DS) CGC|TGT|TTG|TGG|AAG|AAG|TTC|GGT|GGT|ACA|CAC|AGG|CAC|ATG|TCT|ACA|CC **G|GTA|CC**

VEGF-A E | I | E | Y | I | F | K | P | S | C | V | P | L | M | R | C | G
 VEGF-C A | T | N | T | F | F | K | P | P | C | V | S | V | Y | R | C | G

42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58

KPNI ↓ SMAI

VEGF-A **GG|TAC|CC** **G|GGT|TGC|TGC|AAT|GAC|GAA|GGG|CTG**
 (DS) **CC|ATG|GG** **C|CCA|ACG|ACG|TTA|CTG|CTT|CCC|GAC**

VEGF-C **GG|TAC|CC** **G|GGT|TGC|TGC|AAT|AGT|GAG|GGG|CTG**
 (DS) **CC|ATG|GG** **C|CCA|ACG|ACG|TTA|TCA|CTC|CCC|GAC**

VEGF-A | G | C | C | N | D | E | G | L
 VEGF-C | G | C | C | N | S | E | G | L

59 60 61 62 63 64 65 66

5BP-OL TSP509I

VEGF-A **GAG|TGC|GTT|CCC|ACC|GAG|GAG|TCC|AAC|ATC|ACC|ATG|CAG|ATT|ATG|AGA|ATT**
 (DS) **CTC|ACG|CAA|GGG|TGG|CTC|CTC|AGG|TTG|TAG|TGG|TAC|GTC|TAA|TAC|TCT|TAA**

VEGF-C **CAG|TGC|ATG|AAC|ACG|TCC|ACG|AGC|TAC|CTC|AGC|AAG|ACG|CTG|TTT|GAA|ATT**
 (DS) **GTC|ACG|TAC|TTG|TGC|AGG|TGC|TGC|ATG|GAG|TCG|TTC|TGC|GAC|AAA|CTT|TAA**

VEGF-A E | C | V | P | T | E | E | S | N | I | T | M | Q | I | M | R
 VEGF-C Q | C | M | N | T | S | T | S | Y | L | S | K | T | L | F | E

67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82

TSP509I 6BP-

VEGF-A **A|ATT|AAA|CCT|CAC| . | . |CAA|GGG|CAG|CAC|ATC|GGA|GAG|ATG|AGC**
 (DS) **T|TAA|TTT|GGA|GTG| . | . |GTT|CCC|GTC|GTG|TAG|CCT|CTC|TAC|TCG**

VEGF-C **A|ATT|ACA|GTG|CCT|CTC|TCT|CAA|GGG|CCC|AAA|CCA|GTG|ACA|ATC|AGC**
 (DS) **T|TAA|TGT|CAC|GGA|GAG|AGA|GTT|CCC|GGG|TTT|GGT|CAC|TGT|TAG|TCG**

VEGF-A I | K | P | H | . | . | Q | G | Q | H | I | G | E | M | S
 VEGF-C I | T | V | P | L | S | Q | G | P | K | P | V | T | I | S

83 84 85 86 86B 86C 87 88 89 90 91 92 93 94 95

OL BBSI NOTI

VEGF-A **TTT|CTC|CAG|CAT|AAC|AAA|TGT|GAA|TGT|AGA|CCA|AAG|AAA|GAT|TGA|GTC|TTC|GCG|GCC|GC**
 (DS) **AAA|GAG|GTC|GTA|TTG|TTT|ACA|CTT|ACA|TCT|GGT|TTC|TTT|CTA|ACT|CAG|AAG|CGC|CGG|CG**

VEGF-C **TTT|GCC|AAT|CAC|ACT|TCC|TGC|CGA|TGC|ATG|TCT|AAG|CTG|GAT|TGA|GTC|TTC|GCG|GCC|GC**
 (DS) **AAA|CGG|TTA|GTG|TGA|AGG|ACG|GCT|ACG|TAC|AGA|TTC|GAC|CTA|ACT|CAG|AAG|CGC|CGG|CG**

VEGF-A F | L | Q | H | N | K | C | E | C | R | P | K | K | D | *
 VEGF-C F | A | N | H | T | S | C | R | C | M | S | K | L | D | *

96 98 98 99 100 101 102 103 104 105 106 107 108 109

Development of a purification procedure for untagged VEGF-C

- Affinity chromatography: VEGF-C Ab or recombinant FLT4-H₆
- "Classical" chromatographic methods (VEGF purification without heparin-column, Siemeister et. al., 1996): cation exchange column & gel filtration; requirement: highly concentrated and sufficiently pure (>90%) starting material (Pichia supernatant, purified E.coli inclusion bodies)
- Tag removal: insert cleavage site (thrombin, factor Xa, rEK) or limited proteolysis (what protease?)

Further Characterization of VEGF-C

4. Confirm? Vladimir's data on affinities for the different forms ("but only the fully processed VEGF-C could activate VEGFR-2")
5. non-covalent dimerization (native PAGE, e.g. blue native PAGE)
6. heterodimerization with VEGF-D (Flag-tagged VEGF-C & Myc-tagged VEGF-D)
7. significance of glycosylation
8. heparin-affinity (full length vs. mature form)
9. heterogeneous signal peptide cleavage/glycosylation/proteolytic processing
10. processing: comparison with VEGF-D
11. what is the biologically active form in vivo????!!!!

Molecular Modelling

- prerequisite for the rational design of site-directed mutations in a protein
- X-ray crystallography and NMR spectroscopy are the only ways to obtain detailed structural information
- In the absence of experimental data, model-building on the basis of the known three dimensional structure of a homologous protein is at present the only reliable method to obtain structural information

At present most model building by homology protocols start from the assumption that, except for the insertions and deletions, the backbone of the model is identical to the backbone of the structure.

9 Steps of Molecular Modelling

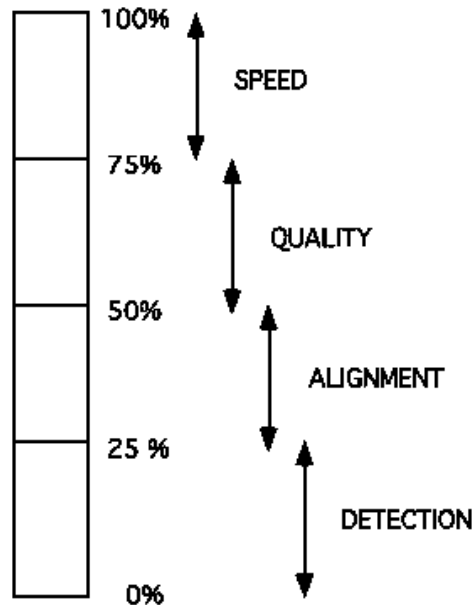
1. Template recognition
2. Alignment
3. Alignment correction
4. Backbone generation
5. Generation of canonical loops
6. Side chain generation plus optimisation
7. Ab initio loop building
8. Overall model optimisation
9. Model verification with optional repeat of previous steps

Template recognition

👍 25-30% sequence homology required

range of sequence
similarity in %
identical residues

key limiting factor
in model building
by homology



Threshold for structural homology

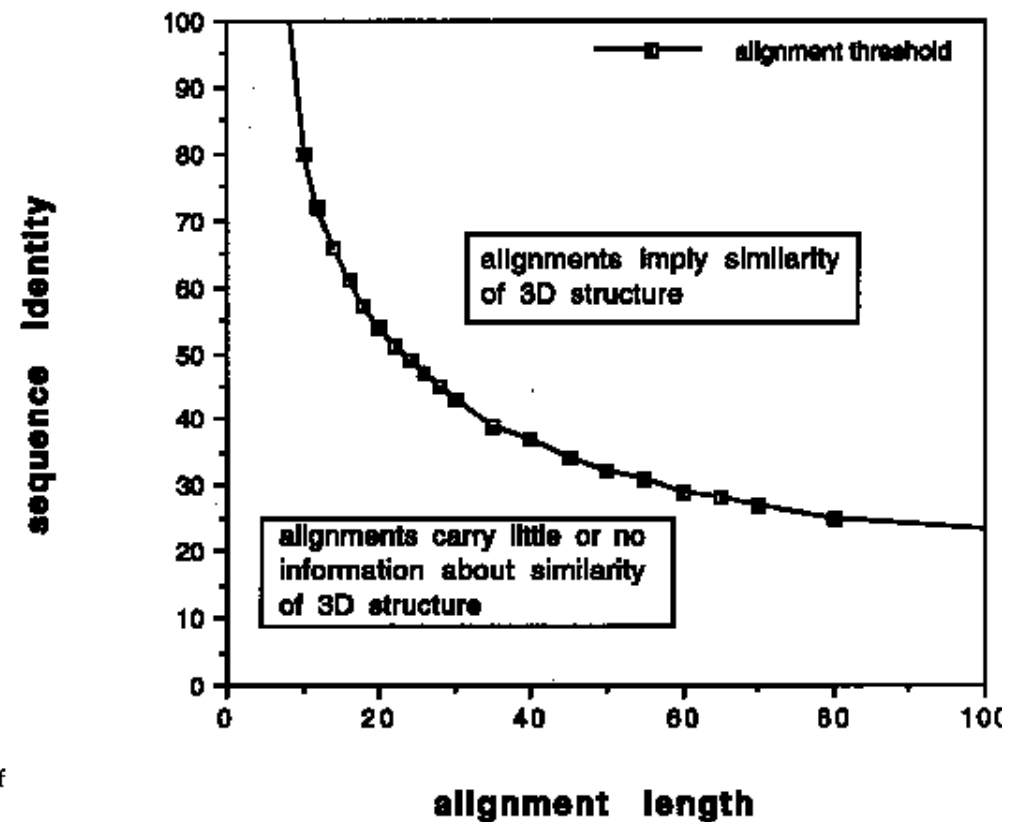


Figure 1. The main limiting steps for model building by homology as function of the percentage sequence identity between the structure and the model.

Alignment

QGQHC FEMSFLQ	QGQH . . CFEMSFLQ	QGQHCFE . . . MSFLQ	QGQ . H . . CFEMSFLQ
GDPHFECVEMTFSQ	GDPHFECVEMTFSQ	. GDPHFECVEMTFSQ	. GDPHFECVEMTFSQ

- Scoring matrices are used to capture evolutionary events in numbers
- Improving the alignment using multiple sequences

sequence 1: LTLTLTLT
sequence 2: YAYAYAYAY

LTLTLTLT.	or	.LTLTLTLT
YAYAYAYAY		YAYAYAYAY

sequence 3: TYTYTYTYT

.LTLTLTLT
TYTYTYTYT
.YAYAYAYAY

- Threading (improving the alignment using the structure)

Example: Ala > Glu substitution (core vs. surface)

BLOSUM62 amino acid substitution matrix (Henikoff, S. and Henikoff, J. G. (1992). Amino acid substitution matrices from protein blocks. Proc. Natl. Acad. Sci. USA 89: 10915-10919.)

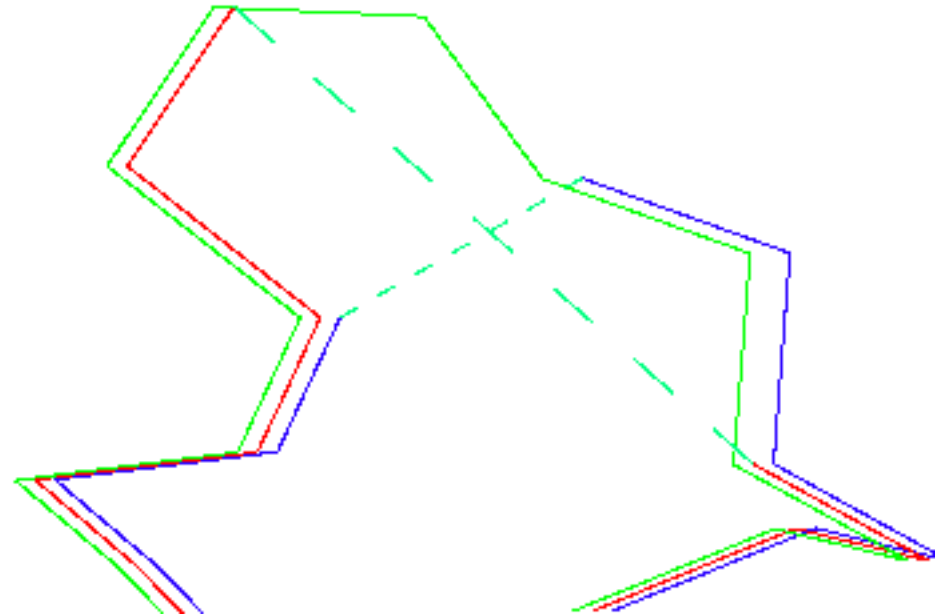
GAP_CREATE 12

GAP_EXTEND 4

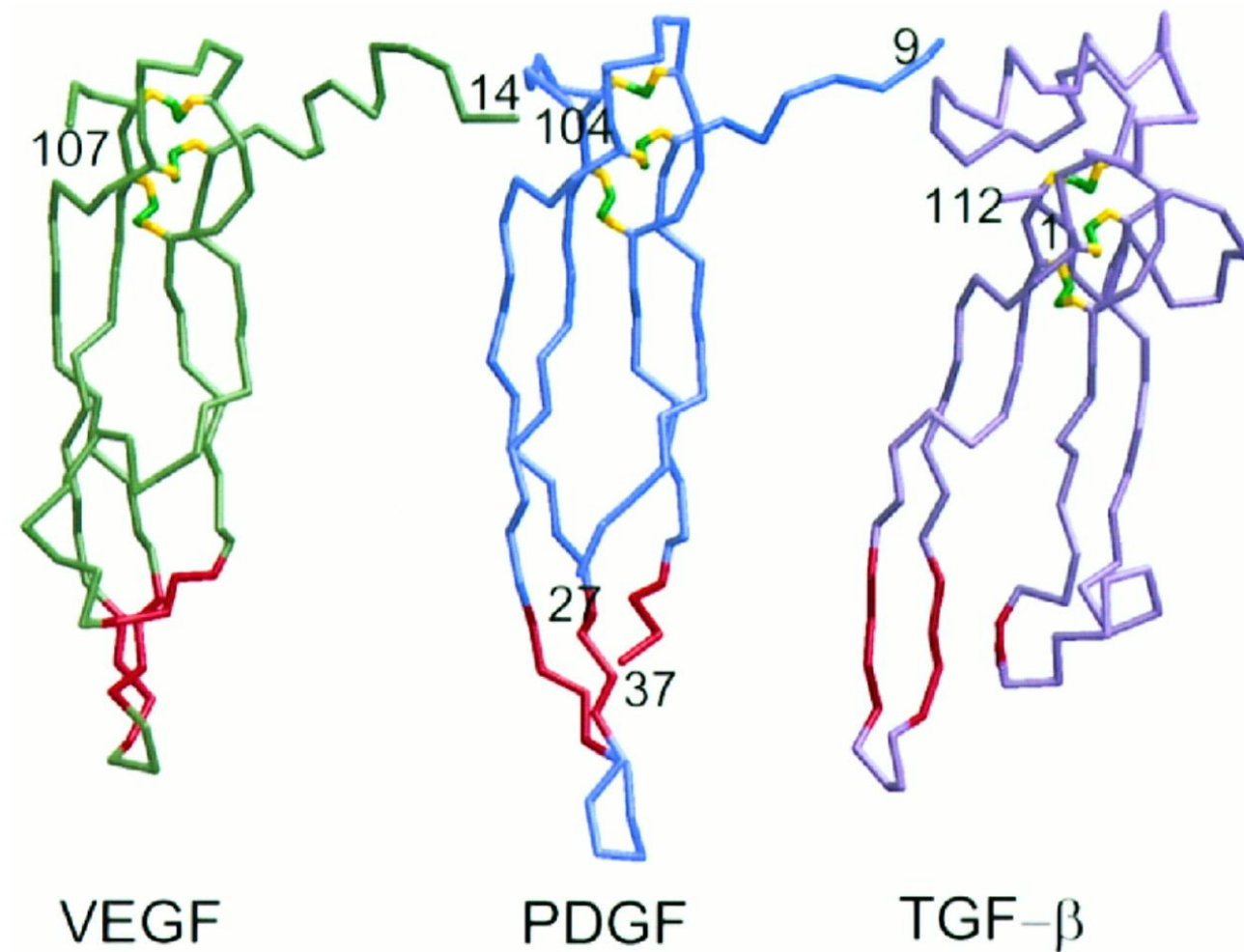
	A	B	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	X	Y	Z
A	4	-2	0	-2	-1	-2	0	-2	-1	-1	-1	-1	-2	-1	-1	-1	1	0	0	-3	-1	-2	-1
B	-2	6	-3	6	2	-3	-1	-1	-3	-1	-4	-3	1	-1	0	-2	0	-1	-3	-4	-1	-3	2
C	0	-3	9	-3	-4	-2	-3	-3	-1	-3	-1	-1	-3	-3	-3	-3	-1	-1	-1	-2	-1	-2	-4
D	-2	6	-3	6	2	-3	-1	-1	-3	-1	-4	-3	1	-1	0	-2	0	-1	-3	-4	-1	-3	2
E	-1	2	-4	2	5	-3	-2	0	-3	1	-3	-2	0	-1	2	0	0	-1	-2	-3	-1	-2	5
F	-2	-3	-2	-3	-3	6	-3	-1	0	-3	0	0	-3	-4	-3	-3	-2	-2	-1	1	-1	3	-3
G	0	-1	-3	-1	-2	-3	6	-2	-4	-2	-4	-3	0	-2	-2	-2	0	-2	-3	-2	-1	-3	-2
H	-2	-1	-3	-1	0	-1	-2	8	-3	-1	-3	-2	1	-2	0	0	-1	-2	-3	-2	-1	2	0
I	-1	-3	-1	-3	-3	0	-4	-3	4	-3	2	1	-3	-3	-3	-3	-2	-1	3	-3	-1	-1	-3
K	-1	-1	-3	-1	1	-3	-2	-1	-3	5	-2	-1	0	-1	1	2	0	-1	-2	-3	-1	-2	1
L	-1	-4	-1	-4	-3	0	-4	-3	2	-2	4	2	-3	-3	-2	-2	-2	-1	1	-2	-1	-1	-3
M	-1	-3	-1	-3	-2	0	-3	-2	1	-1	2	5	-2	-2	0	-1	-1	-1	1	-1	-1	-1	-2
N	-2	1	-3	1	0	-3	0	1	-3	0	-3	-2	6	-2	0	0	1	0	-3	-4	-1	-2	0
P	-1	-1	-3	-1	-1	-4	-2	-2	-3	-1	-3	-2	-2	7	-1	-2	-1	-1	-2	-4	-1	-3	-1
Q	-1	0	-3	0	2	-3	-2	0	-3	1	-2	0	0	-1	5	1	0	-1	-2	-2	-1	-1	2
R	-1	-2	-3	-2	0	-3	-2	0	-3	2	-2	-1	0	-2	1	5	-1	-1	-3	-3	-1	-2	0
S	1	0	-1	0	0	-2	0	-1	-2	0	-2	-1	1	-1	0	-1	4	1	-2	-3	-1	-2	0
T	0	-1	-1	-1	-1	-2	-2	-2	-1	-1	-1	-1	0	-1	-1	-1	1	5	0	-2	-1	-2	-1
V	0	-3	-1	-3	-2	-1	-3	-3	3	-2	1	1	-3	-2	-2	-3	-2	0	4	-3	-1	-1	-2
W	-3	-4	-2	-4	-3	1	-2	-2	-3	-3	-2	-1	-4	-4	-2	-3	-3	-2	-3	11	-1	2	-3
X	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
Y	-2	-3	-2	-3	-2	3	-3	2	-1	-2	-1	-1	-2	-3	-1	-2	-2	-2	-1	2	-1	7	-2
Z	-1	2	-4	2	5	-3	-2	0	-3	1	-3	-2	0	-1	2	0	0	-1	-2	-3	-1	-2	5

Evolutionary homology vs. structural homology

1	2	3	4	5	6	7	8	9	10	11	12	13	14
PHE	ASP	ILE	CYS	ARG	LEU	PRO	GLY	SER	ALA	GLU	ALA	VAL	CYS
PHE	ASN	VAL	CYS	ARG	THR	PRO	---	---	---	GLU	ALA	ILE	CYS
PHE	ASN	VAL	CYS	ARG	---	---	---	THR	PRO	GLU	ALA	ILE	CYS



Backbone generation



Amino acid sequence homology: VEGF-PDGF 19%

- generation of canonical loops (data based, searching libraries)
- side chain generation plus optimisation (rotamers)
- ab initio loop building (energy based, very bad results)
- overall model optimisation (energy minimisation)
- model verification with optional repeat of previous steps