Anirudh Rao
BE21B004

BT3040 – Bioinformatics

Practical 7

1

The Python code to compute the amino acid composition of the given sequences is shown below:

```
import pandas as pd

seq1 = "RATPTRWPVGCFNRPWTKWSYDEALDGIKAAGYAWTGLLTASKPSLHHATATPEYLAALKQKSRHAA"
seq2 = "AAAVMMGLAAIGAAIGIGILGGKFLEGAARQPDLIPLLRTQFFIVMGLVDAIPMIAVGLGLYVMFAVA"
seq3 =
"AADVSAAVGATGQSGMTYRLGLSWDWDKSWWQTSTGRLTGYWDAGYTYWEGGDEGAGKHSLSFAPVFVYEFAGDSIKPFIEAGIGV
AAFSGTRVGDQNLGSSLNFEDRIGAGLKFANGQSVGVRAIHYSNAGLKQPNDGIESYSLFYKIPI"


def composition(sequence):
    residues = ["G", "A", "V", "L", "I", "P", "F", "Y", "W", "S", "T", "C", "M", "N",
"Q", "D", "E", "K", "R", "H"]
    composition = {residue: round(100 * sequence.count(residue) / len(sequence), 2) for
residue in residues}

    return composition


comp1 = composition(seq1)
comp2 = composition(seq2)
comp3 = composition(seq3)

df = pd.DataFrame([comp1, comp2, comp3]).transpose()
df.columns = ["Seq. 1", "Seq. 2", "Seq. 3"]
print(df)
```

The output of this is:

```
Run - Anirudh-Python
Run:  workboard
C:\Users\anira\PycharmProjects\Anir
     Seq. 1  Seq. 2  Seq. 3
G      5.97   14.71   15.23
A     17.91   19.12   10.60
V      1.49    8.82    5.30
L      8.96   13.24    5.96
I      1.49   11.76    5.30
P      7.46    4.41    2.65
F      1.49    5.88    5.30
Y      4.48    1.47    5.30
W      5.97    0.00    3.97
S      5.97    0.00    9.93
T     10.45    1.47    4.64
C      1.49    0.00    0.00
M      0.00    7.35    0.66
N      1.49    0.00    3.31
Q      1.49    2.94    3.31
D      2.99    2.94    5.96
E      2.99    1.47    3.97
K      7.46    1.47    3.97
R      5.97    2.94    3.31
H      4.48    0.00    1.32

Process finished with exit code 0
```

From this, we can see that:

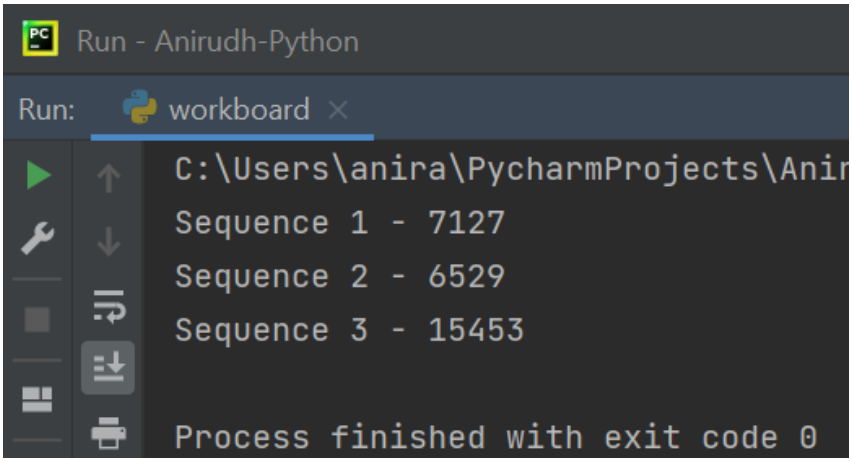
- All three sequences are rich in hydrophobic residues
- Sequence 1 is rich in threonine, making it slightly hydrophilic
- Sequence 1 is richer in basic amino acids than the other two sequences
- Sequence 3 is richer in acidic amino acids than the other two sequences
- Sequence 2 is poorer in aromatic amino acids compared to the other two sequences

2

The Python code to compute the molecular weights of the given sequences is shown below:

```
seq1 = "RATPTRWPVGCFNRPWTKWSYDEALDGIKAAGYAWTGLLTASKPSLHHATATPEYLAALKQKSRHAA"
seq2 = "AAAVMMGLAAIGAAIGIGILGGKFLEGAARQPDLIPLLRTQFFIVMGLVDAIPMIAVGLGLYVMFAVA"
seq3 =
"AADVSAAVGATGQSGMTYRLGLSWDWDKSWWQTSTGRLTGYWDAGYTYWEGGDEGAGKHSLSFAPVFVYEFAGDSIKPFIEAGIGV
AAFSGTRVGDQNLGSSLNFEDRIGAGLKFANGQSVGVRAIHYSNAGLKQPNDGIESYSLFYKIPI"

def mol_weight(sequence):
    weights = {"A": 85,"C": 115,"D": 130,"E": 145,"F": 160,"G": 70,"W": 200,"H":
150,"I": 125,"K": 145,"L": 125,"M": 143,"N": 130,"Y": 175,"P": 110,"Q": 140,"R":
170,"S": 100,"T": 115,"V": 110}
    weight = 0
    for residue in sequence:
        weight += weights[residue]
    water = 18*(len(sequence) - 1)
    return weight - water


count = 1
for sequence in [seq1,seq2,seq3]:
    print(f"Sequence {count} - {mol_weight(sequence)}")
    count += 1
```

The output of this is:

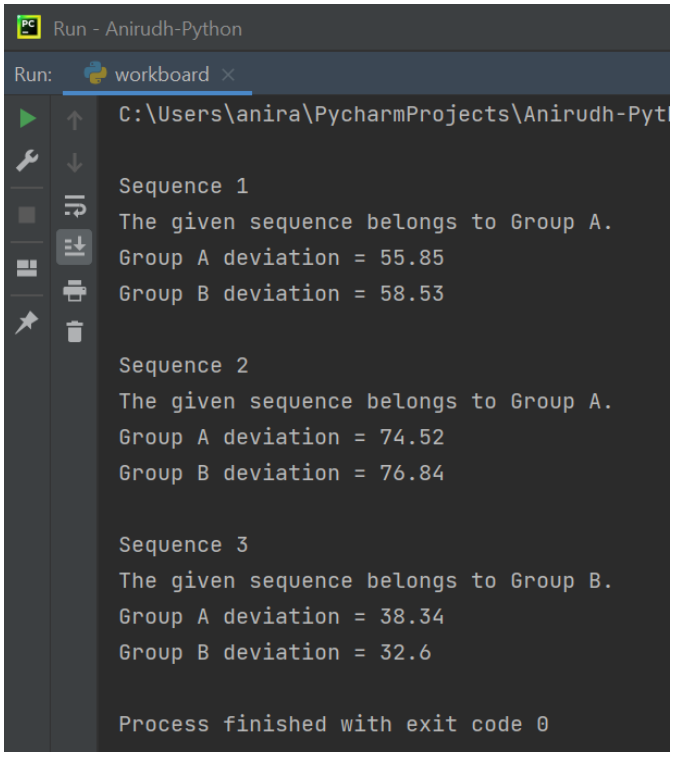


3

The Python code to classify the given sequences under Group A or Group B is shown below:

```
seq1 = "RATPTRWPVGCFNRPWTKWSYDEALDGIKAAGYAWTGLLTASKPSLHHATATPEYLAALKQKSRHAA"
seq2 = "AAAVMMGLAAIGAAIGIGILGGKFLEGAARQPDLIPLLRTQFFIVMGLVDAIPMIAVGLGLYVMFAVA"
seq3 =
"AADVSAAVGATGQSGMTYRLGLSWDWDKSWWQTSTGRLTGYWDAGYTYWEGGDEGAGKHSLSFAPVFVYEFAGDSIKPFIEAGIGV
AAFSGTRVGDQNLGSSLNFEDRIGAGLKFANGQSVGVRAIHYSNAGLKQPNDGIESYSLFYKIPI"

standards = {
    "A": (8.47, 8.95),
    "D": (5.97, 5.91),
    "C": (1.39, 0.47),
    "E": (6.32, 4.78),
    "T": (5.79, 6.54),
    "F": (3.91, 3.68),
    "G": (7.82, 8.54),
    "H": (2.26, 1.25),
    "I": (5.71, 4.77),
    "V": (7.02, 6.76),
    "K": (5.76, 4.93),
    "L": (8.48, 8.78),
    "M": (2.21, 1.56),
    "N": (4.54, 5.74),
    "W": (1.44, 1.24),
    "P": (4.63, 3.74),
    "Q": (3.82, 4.75),
    "R": (4.93, 5.24),
    "S": (5.94, 8.05),
    "Y": (3.58, 4.13),
}


def composition(sequence):
    residues = ["G", "A", "V", "L", "I", "P", "F", "Y", "W", "S", "T", "C", "M", "N",
"Q", "D", "E", "K", "R", "H"]
    composition = {residue: round(100 * sequence.count(residue) / len(sequence), 2) for
residue in residues}

    return composition


def classify(sequence):
    group_a_deviation = 0
    group_b_deviation = 0

    comp = composition(sequence)

    for residue in standards.keys():
        group_a_deviation += abs(comp[residue] - standards[residue][0])
        group_b_deviation += abs(comp[residue] - standards[residue][1])

    if group_a_deviation <= group_b_deviation:
        print("The given sequence belongs to Group A.")
    else:
        print("The given sequence belongs to Group B.")
    print(f"Group A deviation = {round(group_a_deviation, 2)}")
    print(f"Group B deviation = {round(group_b_deviation, 2)}")


count = 1
for sequence in [seq1, seq2, seq3]:
    print("")
    print(f"Sequence {count}")
    classify(sequence)
    count += 1
```

The output of this is:



4

The Python code to compute the three different residue pair preferences is given below:

```
import pandas as pd


def composition(sequence):
    residues = ["G", "A", "V", "L", "I", "P", "F", "Y", "W", "S", "T", "C", "M", "N",
"Q", "D", "E", "K", "R", "H"]
    composition = {residue: round(100 * sequence.count(residue) / len(sequence), 2) for
residue in residues}

    return composition


def dipeptide_composition_1(sequence):
    residues = ["G", "A", "V", "L", "I", "P", "F", "Y", "W", "S", "T", "C", "M", "N",
"Q", "D", "E", "K", "R", "H"]
    dipeptide_counts = {residue1: {residue2: 0 for residue2 in residues} for residue1
in residues}

    for i in range(len(sequence) - 1):
        dipeptide = sequence[i:i + 2]
        dipeptide_counts[dipeptide[0]][dipeptide[1]] += 1

    comp = composition(sequence)
    dipeptide_composition = {residue1: {residue2: round(
        100 * dipeptide_counts[residue1][residue2] / ((len(sequence) / 100) *
(comp[residue1] + comp[residue2])), 2) if comp[residue1] + comp[residue2] != 0 else 0
for residue2 in residues} for residue1 in residues}

    return pd.DataFrame(dipeptide_composition).transpose()


def dipeptide_composition_2(sequence):
    residues = ["G", "A", "V", "L", "I", "P", "F", "Y", "W", "S", "T", "C", "M", "N",
"Q", "D", "E", "K", "R", "H"]
```

```
    dipeptide_counts = {residue1: {residue2: 0 for residue2 in residues} for residue1
in residues}

    for i in range(len(sequence) - 1):
        dipeptide = sequence[i:i + 2]
        dipeptide_counts[dipeptide[0]][dipeptide[1]] += 1

    dipeptide_composition = {
        residue1: {residue2: round(100 * dipeptide_counts[residue1][residue2] /
(len(sequence) - 1), 2) for residue2 in
                   residues} for residue1 in residues}

    return pd.DataFrame(dipeptide_composition).transpose()


def dipeptide_composition_3(sequence):
    residues = ["G", "A", "V", "L", "I", "P", "F", "Y", "W", "S", "T", "C", "M", "N",
"Q", "D", "E", "K", "R", "H"]
    dipeptide_counts = {residue1: {residue2: 0 for residue2 in residues} for residue1
in residues}

    for i in range(len(sequence) - 1):
        dipeptide = sequence[i:i + 2]
        dipeptide_counts[dipeptide[0]][dipeptide[1]] += 1

    comp = composition(sequence)
    dipeptide_composition = {residue1: {residue2: round(100 *
dipeptide_counts[residue1][residue2] / (
                (len(sequence) ** 2 / 100 ** 2) * (comp[residue1] * comp[residue2])),
2) if comp[residue1] * comp[
        residue2] != 0 else 0 for residue2 in residues} for residue1 in residues}

    return pd.DataFrame(dipeptide_composition).transpose()
```

<u>Sequence 1 – Score 1</u>

	G	A	V	L	I	P	F	Y	W	S	T	C	M	N	Q	D	E	K	R	H
G	0.00	0.00	0.00	10.00	20.01	0.00	0.00	14.28	0.00	0.00	0.00	20.01	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
A	6.25	12.50	0.00	11.11	0.00	0.00	0.00	0.00	6.25	6.25	15.79	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
V	20.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
L	0.00	5.55	0.00	8.33	0.00	0.00	0.00	0.00	0.00	0.00	7.69	0.00	0.0	0.00	0.00	12.49	0.00	9.09	0.00	11.11
I	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	16.68	0.00	0.00
P	0.00	0.00	16.68	0.00	0.00	0.00	0.00	0.00	11.11	11.11	8.33	0.00	0.0	0.00	0.00	0.00	14.28	0.00	0.00	0.00
F	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	50.09	0.00	0.00	0.00	0.00	0.00	0.00
Y	0.00	6.67	0.00	11.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	19.98	0.00	0.00	0.00	0.00
W	0.00	0.00	0.00	0.00	0.00	11.11	0.00	0.00	0.00	12.50	18.18	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S	0.00	0.00	0.00	10.00	0.00	0.00	0.00	14.28	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	11.11	12.50	0.00
T	9.09	10.53	0.00	0.00	0.00	16.67	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	8.33	9.09	0.00
C	0.00	0.00	0.00	0.00	0.00	0.00	50.09	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
M	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
N	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	20.01	0.00
Q	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	16.68	0.00	0.00
D	16.66	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	24.96	0.00	0.00	0.00
E	0.00	7.14	0.00	0.00	0.00	0.00	0.00	19.98	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
K	0.00	5.88	0.00	0.00	0.00	10.00	0.00	0.00	11.11	11.11	0.00	0.00	0.0	0.00	16.68	0.00	0.00	0.00	0.00	0.00
R	0.00	6.25	0.00	0.00	0.00	11.11	0.00	0.00	12.50	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	14.28
H	0.00	13.33	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	16.66

Here, the residue in the row index indicates position i and the residue in the column index indicates position $i + 1$.

Sequence 1 – Score 2

	G	A	V	L	I	P	F	Y	W	S	T	C	M	N	Q	D	E	K	R	H
G	0.00	0.00	0.00	1.52	1.52	0.00	0.00	1.52	0.00	0.00	0.00	1.52	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
A	1.52	4.55	0.00	3.03	0.00	0.00	0.00	0.00	1.52	1.52	4.55	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
V	1.52	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
L	0.00	1.52	0.00	1.52	0.00	0.00	0.00	0.00	0.00	0.00	1.52	0.00	0.0	0.00	0.00	1.52	0.00	1.52	0.00	1.52
I	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	1.52	0.00	0.00
P	0.00	0.00	1.52	0.00	0.00	0.00	0.00	0.00	1.52	1.52	1.52	0.00	0.0	0.00	0.00	0.00	1.52	0.00	0.00	0.00
F	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	1.52	0.00	0.00	0.00	0.00	0.00	0.00
Y	0.00	1.52	0.00	1.52	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	1.52	0.00	0.00	0.00	0.00
W	0.00	0.00	0.00	0.00	0.00	1.52	0.00	0.00	0.00	1.52	3.03	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S	0.00	0.00	0.00	1.52	0.00	0.00	0.00	1.52	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	1.52	1.52	0.00
T	1.52	3.03	0.00	0.00	0.00	3.03	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	1.52	1.52	0.00
C	0.00	0.00	0.00	0.00	0.00	0.00	1.52	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
M	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
N	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	1.52	0.00
Q	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	1.52	0.00	0.00
D	1.52	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	1.52	0.00	0.00	0.00
E	0.00	1.52	0.00	0.00	0.00	0.00	0.00	1.52	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
K	0.00	1.52	0.00	0.00	0.00	1.52	0.00	0.00	1.52	1.52	0.00	0.00	0.0	0.00	1.52	0.00	0.00	0.00	0.00	0.00
R	0.00	1.52	0.00	0.00	0.00	1.52	0.00	0.00	1.52	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	1.52
H	0.00	3.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	1.52

Sequence 1 – Score 3

	G	A	V	L	I	P	F	Y	W	S	T	C	M	N	Q	D	E	K	R	H
G	0.00	0.00	0.00	4.16	25.04	0.00	0.00	8.33	0.00	0.00	0.00	25.04	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
A	2.08	2.08	0.00	2.78	0.00	0.00	0.00	0.00	2.08	2.08	3.57	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
V	25.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
L	0.00	1.39	0.00	2.77	0.00	0.00	0.00	0.00	0.00	0.00	2.38	0.00	0.0	0.00	0.00	8.32	0.00	3.33	0.00	5.55
I	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	20.04	0.00	0.00
P	0.00	0.00	20.04	0.00	0.00	0.00	0.00	0.00	5.00	5.00	2.86	0.00	0.0	0.00	0.00	0.00	9.99	0.00	0.00	0.00
F	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	100.34	0.00	0.00	0.00	0.00	0.00	0.00
Y	0.00	2.78	0.00	5.55	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	16.63	0.00	0.00	0.00	0.00
W	0.00	0.00	0.00	0.00	0.00	5.00	0.00	0.00	0.00	6.25	7.14	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S	0.00	0.00	0.00	4.16	0.00	0.00	0.00	8.33	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	5.00	6.25	0.00
T	3.57	2.38	0.00	0.00	0.00	5.72	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	2.86	3.57	0.00
C	0.00	0.00	0.00	0.00	0.00	0.00	100.34	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
M	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
N	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	25.04	0.00
Q	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	20.04	0.00	0.00
D	12.48	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	24.92	0.00	0.00	0.00
E	0.00	4.16	0.00	0.00	0.00	0.00	0.00	16.63	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
K	0.00	1.67	0.00	0.00	0.00	4.00	0.00	0.00	5.00	5.00	0.00	0.00	0.0	0.00	20.04	0.00	0.00	0.00	0.00	0.00
R	0.00	2.08	0.00	0.00	0.00	5.00	0.00	0.00	6.25	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	8.33
H	0.00	5.55	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	11.10

Sequence 2 – Score 1

	G	A	V	L	I	P	F	Y	W	S	T	C	M	N	Q	D	E	K	R	H
G	5.00	8.69	0.00	21.05	11.11	0.00	0.00	0.0	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.0	9.09	0.00	0.0
A	0.00	19.23	15.79	0.00	14.29	0.00	0.00	0.0	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.0	0.00	6.67	0.0
V	6.25	5.26	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.0	0.00	0.0	27.28	0.0	0.00	12.51	0.0	0.00	0.00	0.0
L	10.52	4.54	6.67	5.55	5.88	0.00	0.00	10.0	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	10.0	0.00	9.09	0.0
I	16.67	4.76	7.15	5.88	0.00	18.19	0.00	0.0	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.0	0.00	0.00	0.0
P	0.00	0.00	0.00	8.33	0.00	0.00	0.00	0.0	0.0	0.0	0.00	0.0	12.51	0.0	0.00	20.01	0.0	0.00	0.00	0.0
F	0.00	5.88	0.00	7.69	8.34	0.00	12.51	0.0	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.0	0.00	0.00	0.0
Y	0.00	0.00	14.29	0.00	0.00	0.00	0.00	0.0	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.0	0.00	0.00	0.0
W	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.0	0.00	0.00	0.0
S	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.0	0.00	0.00	0.0
T	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.0	0.00	0.0	0.00	0.0	33.35	0.00	0.0	0.00	0.00	0.0
C	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.0	0.00	0.00	0.0
M	13.33	0.00	0.00	0.00	7.70	0.00	11.12	0.0	0.0	0.0	0.00	0.0	10.00	0.0	0.00	0.00	0.0	0.00	0.00	0.0
N	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.0	0.00	0.00	0.0
Q	0.00	0.00	0.00	0.00	0.00	20.01	16.67	0.0	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.0	0.00	0.00	0.0
D	0.00	6.67	0.00	9.09	0.00	0.00	0.00	0.0	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.0	0.00	0.00	0.0
E	9.09	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.0	0.00	0.00	0.0
K	0.00	0.00	0.00	0.00	0.00	0.00	20.01	0.0	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.0	0.00	0.00	0.0
R	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.0	33.35	0.0	0.00	0.0	25.01	0.00	0.0	0.00	0.00	0.0
H	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.0	0.00	0.00	0.0

Sequence 2 – Score 2

	G	A	V	L	I	P	F	Y	W	S	T	C	M	N	Q	D	E	K	R	H
G	1.49	2.99	0.00	5.97	2.99	0.00	0.00	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	1.49	0.00	0.0
A	0.00	7.46	4.48	0.00	4.48	0.00	0.00	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	0.00	1.49	0.0
V	1.49	1.49	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.00	0.0	4.48	0.0	0.00	1.49	0.00	0.00	0.00	0.0
L	2.99	1.49	1.49	1.49	1.49	0.00	0.00	1.49	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	1.49	0.00	1.49	0.0
I	4.48	1.49	1.49	1.49	0.00	2.99	0.00	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.0
P	0.00	0.00	0.00	1.49	0.00	0.00	0.00	0.00	0.0	0.0	0.00	0.0	1.49	0.0	0.00	1.49	0.00	0.00	0.00	0.0
F	0.00	1.49	0.00	1.49	1.49	0.00	1.49	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.0
Y	0.00	0.00	1.49	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.0
W	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.0
S	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.0
T	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.00	0.0	0.00	0.0	1.49	0.00	0.00	0.00	0.00	0.0
C	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.0
M	2.99	0.00	0.00	0.00	1.49	0.00	1.49	0.00	0.0	0.0	0.00	0.0	1.49	0.0	0.00	0.00	0.00	0.00	0.00	0.0
N	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.0
Q	0.00	0.00	0.00	0.00	0.00	1.49	1.49	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.0
D	0.00	1.49	0.00	1.49	0.00	0.00	0.00	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.0
E	1.49	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.0
K	0.00	0.00	0.00	0.00	0.00	0.00	1.49	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.0
R	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	1.49	0.0	0.00	0.0	1.49	0.00	0.00	0.00	0.00	0.0
H	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.0

Sequence 2 – Score 3

	G	A	V	L	I	P	F	Y	W	S	T	C	M	N	Q	D	E	K	R	H
G	1.00	1.54	0.00	4.44	2.50	0.00	0.00	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	10.0	0.00	0.0
A	0.00	2.96	3.85	0.00	2.89	0.00	0.00	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	0.0	3.85	0.0
V	1.67	1.28	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.00	0.0	10.01	0.0	0.00	8.34	0.00	0.0	0.00	0.0
L	2.22	0.85	1.85	1.23	1.39	0.00	0.00	11.11	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	11.11	0.0	5.56	0.0
I	3.75	0.96	2.09	1.39	0.00	8.34	0.00	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	0.0	0.00	0.0
P	0.00	0.00	0.00	3.70	0.00	0.00	0.00	0.00	0.0	0.0	0.00	0.0	6.67	0.0	0.00	16.68	0.00	0.0	0.00	0.0
F	0.00	1.92	0.00	2.78	3.13	0.00	6.26	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	0.0	0.00	0.0
Y	0.00	0.00	16.68	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	0.0	0.00	0.0
W	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	0.0	0.00	0.0
S	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	0.0	0.00	0.0
T	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.00	0.0	0.00	0.0	50.04	0.00	0.00	0.0	0.00	0.0
C	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	0.0	0.00	0.0
M	4.00	0.00	0.00	0.00	2.50	0.00	5.00	0.00	0.0	0.0	0.00	0.0	4.00	0.0	0.00	0.00	0.00	0.0	0.00	0.0
N	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	0.0	0.00	0.0
Q	0.00	0.00	0.00	0.00	0.00	16.68	12.51	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	0.0	0.00	0.0
D	0.00	3.85	0.00	5.56	0.00	0.00	0.00	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	0.0	0.00	0.0
E	10.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	0.0	0.00	0.0
K	0.00	0.00	0.00	0.00	0.00	0.00	25.02	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	0.0	0.00	0.0
R	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	50.04	0.0	0.00	0.0	25.02	0.00	0.00	0.0	0.00	0.0
H	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.00	0.0	0.00	0.0	0.00	0.00	0.00	0.0	0.00	0.0

Sequence 3 – Score 1

	G	A	V	L	I	P	F	Y	W	S	T	C	M	N	Q	D	E	K	R	H
G	2.17	7.69	6.45	9.38	6.45	0.00	0.00	6.45	0.00	2.63	3.33	0.0	4.17	0.00	7.14	9.38	0.00	3.45	3.57	0.00
A	15.38	9.37	4.17	0.00	4.17	5.00	4.17	0.00	0.00	0.00	4.35	0.0	0.00	4.76	0.00	4.00	0.00	0.00	0.00	0.00
V	9.68	4.17	0.00	0.00	0.00	0.00	6.25	6.25	0.00	4.35	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	7.69	0.00
L	6.25	0.00	0.00	0.00	0.00	0.00	5.88	0.00	0.00	8.34	6.25	0.0	0.00	7.14	0.00	0.00	0.00	13.34	0.00	0.00
I	6.45	0.00	0.00	0.00	0.00	8.33	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	14.29	7.14	0.00	10.00
P	0.00	0.00	8.33	0.00	8.33	0.00	8.33	0.00	0.00	0.00	0.00	0.0	0.00	11.11	0.00	0.00	0.00	0.00	0.00	0.00
F	0.00	12.50	6.25	0.00	6.25	0.00	0.00	6.25	0.00	4.35	0.00	0.0	0.00	0.00	0.00	0.00	7.14	0.00	0.00	0.00
Y	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	14.29	8.70	6.66	0.0	0.00	0.00	0.00	0.00	7.14	7.14	7.69	0.00
W	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8.34	0.00	0.00	0.0	0.00	0.00	9.10	20.01	8.34	0.00	0.00	0.00
S	5.26	3.23	4.35	12.50	4.35	0.00	4.35	4.35	9.53	3.33	4.55	0.0	0.00	5.00	0.00	0.00	0.00	0.00	0.00	0.00
T	10.00	0.00	0.00	0.00	0.00	0.00	0.00	13.32	0.00	4.55	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	8.33	0.00
C	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
M	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	12.50	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
N	3.57	4.76	0.00	7.14	0.00	0.00	7.69	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	7.14	0.00	0.00	0.00	0.00
Q	0.00	0.00	0.00	0.00	0.00	11.11	0.00	0.00	0.00	10.00	8.33	0.0	0.00	10.00	0.00	0.00	0.00	0.00	0.00	0.00
D	3.13	4.00	5.88	0.00	0.00	0.00	0.00	0.00	6.67	4.17	0.00	0.0	0.00	0.00	7.14	0.00	6.67	6.67	7.14	0.00
E	6.90	4.55	0.00	0.00	0.00	0.00	7.14	0.00	0.00	4.76	0.00	0.0	0.00	0.00	0.00	6.67	0.00	0.00	0.00	0.00
K	0.00	0.00	0.00	0.00	7.14	10.00	7.14	0.00	0.00	4.76	0.00	0.0	0.00	0.00	9.10	0.00	0.00	0.00	0.00	12.52
R	0.00	4.76	7.69	14.29	7.69	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
H	0.00	0.00	0.00	0.00	0.00	0.00	0.00	10.00	0.00	5.89	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Sequence 3 – Score 2

	G	A	V	L	I	P	F	Y	W	S	T	C	M	N	Q	D	E	K	R	H
G	0.67	2.00	1.33	2.00	1.33	0.00	0.00	1.33	0.00	0.67	0.67	0.0	0.67	0.00	1.33	2.00	0.00	0.67	0.67	0.00
A	4.00	2.00	0.67	0.00	0.67	0.67	0.67	0.00	0.00	0.00	0.67	0.0	0.00	0.67	0.00	0.67	0.00	0.00	0.00	0.00
V	2.00	0.67	0.00	0.00	0.00	0.00	0.67	0.67	0.00	0.67	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.67	0.00
L	1.33	0.00	0.00	0.00	0.00	0.00	0.67	0.00	0.00	1.33	0.67	0.0	0.00	0.67	0.00	0.00	0.00	1.33	0.00	0.00
I	1.33	0.00	0.00	0.00	0.00	0.67	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	1.33	0.67	0.00	0.67
P	0.00	0.00	0.67	0.00	0.67	0.00	0.67	0.00	0.00	0.00	0.00	0.0	0.00	0.67	0.00	0.00	0.00	0.00	0.00	0.00
F	0.00	2.00	0.67	0.00	0.67	0.00	0.00	0.67	0.00	0.67	0.00	0.0	0.00	0.00	0.00	0.00	0.67	0.00	0.00	0.00
Y	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.33	1.33	0.67	0.0	0.00	0.00	0.00	0.00	0.67	0.67	0.67	0.00
W	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.67	0.00	0.00	0.0	0.00	0.00	0.67	2.00	0.67	0.00	0.00	0.00
S	1.33	0.67	0.67	2.00	0.67	0.00	0.67	0.67	1.33	0.67	0.67	0.0	0.00	0.67	0.00	0.00	0.00	0.00	0.00	0.00
T	2.00	0.00	0.00	0.00	0.00	0.00	0.00	1.33	0.00	0.67	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.67	0.00
C	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
M	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.67	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
N	0.67	0.67	0.00	0.67	0.00	0.00	0.67	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.67	0.00	0.00	0.00	0.00
Q	0.00	0.00	0.00	0.00	0.00	0.67	0.00	0.00	0.00	1.33	0.67	0.0	0.00	0.67	0.00	0.00	0.00	0.00	0.00	0.00
D	0.67	0.67	0.67	0.00	0.00	0.00	0.00	0.00	0.67	0.67	0.00	0.0	0.00	0.00	0.67	0.00	0.67	0.67	0.67	0.00
E	1.33	0.67	0.00	0.00	0.00	0.00	0.67	0.00	0.00	0.67	0.00	0.0	0.00	0.00	0.00	0.67	0.00	0.00	0.00	0.00
K	0.00	0.00	0.00	0.00	0.67	0.67	0.67	0.00	0.00	0.67	0.00	0.0	0.00	0.00	0.67	0.00	0.00	0.00	0.00	0.67
R	0.00	0.67	0.67	1.33	0.67	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
H	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.67	0.00	0.67	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Sequence 3 – Score 3

	G	A	V	L	I	P	F	Y	W	S	T	C	M	N	Q	D	E	K	R	H
G	0.19	0.82	1.09	1.45	1.09	0.00	0.00	1.09	0.00	0.29	0.62	0.0	4.36	0.00	1.74	1.45	0.00	0.73	0.87	0.00
A	1.63	1.17	0.78	0.00	0.78	1.56	0.78	0.00	0.00	0.00	0.89	0.0	0.00	1.25	0.00	0.69	0.00	0.00	0.00	0.00
V	1.63	0.78	0.00	0.00	0.00	0.00	1.56	1.56	0.00	0.83	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	2.50	0.00
L	0.97	0.00	0.00	0.00	0.00	0.00	1.39	0.00	0.00	1.48	1.59	0.0	0.00	2.22	0.00	0.00	0.00	3.71	0.00	0.00
I	1.09	0.00	0.00	0.00	0.00	3.12	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	4.17	2.08	0.00	6.27
P	0.00	0.00	3.12	0.00	3.12	0.00	3.12	0.00	0.00	0.00	0.00	0.0	0.00	5.00	0.00	0.00	0.00	0.00	0.00	0.00
F	0.00	2.34	1.56	0.00	1.56	0.00	0.00	1.56	0.00	0.83	0.00	0.0	0.00	0.00	0.00	0.00	2.08	0.00	0.00	0.00
Y	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.17	1.67	1.78	0.0	0.00	0.00	0.00	0.00	2.08	2.08	2.50	0.00
W	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.78	0.00	0.00	0.0	0.00	0.00	3.34	5.56	2.78	0.00	0.00	0.00
S	0.58	0.42	0.83	2.22	0.83	0.00	0.83	0.83	2.23	0.44	0.95	0.0	0.00	1.33	0.00	0.00	0.00	0.00	0.00	0.00
T	1.86	0.00	0.00	0.00	0.00	0.00	0.00	3.57	0.00	0.95	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	2.86	0.00
C	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
M	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	14.32	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
N	0.87	1.25	0.00	2.22	0.00	0.00	2.50	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	2.22	0.00	0.00	0.00	0.00
Q	0.00	0.00	0.00	0.00	0.00	5.00	0.00	0.00	0.00	2.67	2.86	0.0	0.00	4.00	0.00	0.00	0.00	0.00	0.00	0.00
D	0.48	0.69	1.39	0.00	0.00	0.00	0.00	0.00	1.85	0.74	0.00	0.0	0.00	0.00	2.22	0.00	1.85	1.85	2.22	0.00
E	1.45	1.04	0.00	0.00	0.00	0.00	2.08	0.00	0.00	1.11	0.00	0.0	0.00	0.00	0.00	1.85	0.00	0.00	0.00	0.00
K	0.00	0.00	0.00	0.00	2.08	4.17	2.08	0.00	0.00	1.11	0.00	0.0	0.00	0.00	3.34	0.00	0.00	0.00	0.00	8.37
R	0.00	1.25	2.50	4.45	2.50	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
H	0.00	0.00	0.00	0.00	0.00	0.00	0.00	6.27	0.00	3.35	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

The Python code to list the top 10 preferred residues from each of the three pair preferences is given below:

```
from heapq import nlargest


def find_top_ten(dip_comp):
    values = {residue1 + residue2: dip_comp.loc[residue1][residue2] for residue1 in
dip_comp.index for residue2 in
              dip_comp.columns}
    top_ten = nlargest(10, values, key=values.get)

    return top_ten
```

Using this, we get the top 10 preferred residues from each of the three pair preferences as:

	Sequence 1	Sequence 2	Sequence 3
Score 1	'FN', 'CF', 'DE', 'GI', 'GC', 'VG', 'NR', 'YD', 'EY', 'WT'	'TQ', 'RT', 'VM', 'RQ', 'GL', 'PD', 'QP', 'KF', 'AA', 'IP'	'WD', 'AG', 'IE', 'YW', 'RL', 'LK', 'TY', 'KH', 'FA', 'SL'
Score 2	'AA', 'AT', 'AL', 'WT', 'TA', 'TP', 'HA', 'GL', 'GI', 'GY'	'AA', 'GL', 'AV', 'AI', 'VM', 'IG', 'GA', 'GI', 'LG', 'IP'	'AG', 'GA', 'GL', 'GD', 'AA', 'VG', 'FA', 'WD', 'SL', 'TG'
Score 3	'FN', 'CF', 'GI', 'GC', 'VG', 'NR', 'DE', 'IK', 'PV', 'QK'	'TQ', 'RT', 'KF', 'RQ', 'PD', 'YV', 'QP', 'QF', 'LY', 'LE'	'MT', 'KH', 'IH', 'HY', 'WD', 'PN', 'QP', 'RL', 'GM', 'IE'

5

The Python code to compute the average hydrophobicity, helical contact area, and total non-bonded energy of the given sequences is shown below:

```
import pandas as pd


def hydrophobicity(sequence):
    data = {'A': 13.85, 'D': 11.61, 'C': 15.37, 'E': 11.38, 'F': 13.93, 'G': 13.34,
'H': 13.82, 'I': 15.28, 'K': 11.58,
            'L': 14.13, 'M': 13.86, 'N': 13.02, 'P': 12.35, 'Q': 12.61, 'R': 13.10,
'S': 13.39, 'T': 12.70, 'V': 14.56,
            'W': 15.48, 'Y': 13.88}

    total = 0

    for residue in sequence:
        total += data[residue]

    return round(total / len(sequence), 2)


def helical_contact_area(sequence):
    data = {'A': 20.0, 'D': 26.0, 'C': 25.0, 'E': 33.0, 'F': 46.0, 'G': 13.0, 'H':
37.0, 'I': 39.0, 'K': 46.0,
            'L': 35.0, 'M': 43.0, 'N': 28.0, 'P': 22.0, 'Q': 36.0, 'R': 55.0, 'S':
20.0, 'T': 28.0, 'V': 33.0,
            'W': 61.0, 'Y': 46.0}

    total = 0
```

```python
    for residue in sequence:
        total += data[residue]

    return round(total, 2)


def total_nonbonded_energy(sequence):
    data = {'A': 1.9, 'D': 1.52, 'C': 2.04, 'E': 1.54, 'F': 1.86, 'G': 1.9, 'H': 1.76,
'I': 1.95, 'K': 1.37, 'L': 1.97,
            'M': 1.96, 'N': 1.56, 'P': 1.7, 'Q': 1.52, 'R': 1.48, 'S': 1.75, 'T': 1.77,
'V': 1.98, 'W': 1.87, 'Y': 1.69}

    total = 0

    for residue in sequence:
        total += data[residue]

    return round(total, 2)


seq1 = "RATPTRWPVGCFNRPWTKWSYDEALDGIKAAGYAWTGLLTASKPSLHHATATPEYLAALKQKSRHAA"
seq2 = "AAAVMMGLAAIGAAIGIGILGGKFLEGAARQPDLIPLLRTQFFIVMGLVDAIPMIAVGLGLYVMFAVA"
seq3 =
"AADVSAAVGATGQSGMTYRLGLSWDWDKSWWQTSTGRLTGYWDAGYTYWEGGDEGAGKHSLSFAPVFVYEFAGDSIKPFIEAGIGV
AAFSGTRVGDQNLGSSLNFEDRIGAGLKFANGQSVGVRAIHYSNAGLKQPNDGIESYSLFYKIPI"

properties =
pd.DataFrame({seq:[hydrophobicity(seq),helical_contact_area(seq),total_nonbonded_energy
(seq)] for seq in [seq1,seq2,seq3]})
properties.columns = ["Seq. 1", "Seq. 2", "Seq. 3"]
properties = properties.transpose()
properties.columns = ["Average Hydrophobicity", "Helical Contact Area", "Total Non-
Bonded Energy"]
print("")
print(properties)
```

The output of this is:

```
Run - Anirudh-Python
Run:  workboard
C:\Users\anira\PycharmProjects\Anirudh-Python\venv\Scripts\python.exe C:/Users/anir

        Average Hydrophobicity  Helical Contact Area  Total Non-Bonded Energy
Seq. 1                   13.35                2156.0                   117.74
Seq. 2                   13.77                2067.0                   126.66
Seq. 3                   13.42                4616.0                   267.75

Process finished with exit code 0
```

- The sequences have similar average hydrophobicity values.
- Sequence 3 has the greatest helical contact area.
- Sequence 3 has the greatest total non-bonded energy.