Conformational similarity among amino acid residues: 1. Analysis of protein crystal structure data

A. S. Kolaskar* and V. Ramabrahmam

School of Life Sciences, University of Hyderabad, Hyderabad — 500 001 India

(Received 23 April 1980; revised 9 October 1980)

The probability distribution in the (φ,ψ)-plane obtained for each amino acid residue from crystal structure data of globular proteins is compared. This has shown amino acid residues, Pro and Gly to be conformationally unique. Conformational similarity in the (φ,ψ)-plane of amino acid residues does not necessarily mean that they will have the same chemical or biochemical properties or similar secondary structures. A set of amino acid residues are given which can adopt the conformations of other amino acid residues without much difficulty either in the whole (φ,ψ)-plane or in regions, where the observed conformations are maximum.

Introduction

Prediction of the three dimensional structure of a protein, which depends on its primary structure, is not possible at this stage for several reasons. The major factor is the occurrence of multiple minima in the energy minimization procedure. Statistical methods developed by several groups using the X-ray crystal structure data of globular proteins, are being used to predict secondary structures[1-10] alone, although the accuracy achieved is far from satisfactory. Recently Bourgeios *et al.*[11] have used six different statistical methods to predict the secondary structure of the *lac* repressor and have observed that: 'The large number of disagreements among the results for different methods indicate that only very limited information is provided by each method and the basis on which they operate is not clear.'

A particular amino acid residue, although it has a preferred secondary structure such as α-helix, β-sheet or chain reversals, has got non-negligible probability for other secondary structures as well as in the coil region of globular proteins. Therefore, unless a proper weighting scheme is developed for each residue, predicting whether a particular residue is a part of an expected secondary structure or not will not be successful beyond a certain percentage. Thus it is worthwhile to throw some light on the conformations taken by main chain amino acid residues having side chains which differ not only in their size but also in chemical nature. Earlier efforts in this direction are based on potential energy calculations[4,12] where the reliability is poor because of inaccuracies in the potential functions used[13].

A purely empirical approach seems to be better at this stage since accurate crystal structure data from a large number of globular proteins, having quite different tertiary structures, are available. The (φ,ψ)-values of amino acid residues from protein structure data (φ,ψ being the main chain dihedral angles of the polypeptide chain as defined by the IUPAC-IUB Commission on Biochemical Nomenclature[14]), are used and a very simple algorithm is developed to study the main chain conformational similarity among amino acid residues. The results of our present study indicate that amino acids having entirely different chemical or biochemical properties and also having different preferences for secondary structure, can have similar (φ,ψ)-probability distribution maps. This indicates that the previous classifications of amino acid residues based on the information available from chemical or biochemical studies or even statistical studies made on polypeptides and proteins, do not give sufficient information regarding the conformational preference shown by amino acid residues in three dimensions. In this part, we have discussed the method developed to analyse the protein structure data and the results obtained are given later. However, the applications of these results are presented in the next part of this series.

Experimental

(φ,ψ)-data from the crystal structure of globular proteins mentioned below were collected and analysed. In order to give some idea about the accuracies of the (φ,ψ)-values used in this study the resolutions (in Å) to which these crystal structures are solved or refined are given in brackets. These are: lamprey cynamet haemoglobin (2.0), bovine ferricytochrome b_5(2.0), horse deoxyhaemoglobin dimer (2.8), bonito ferrocytochrome *c* (2.3), tuna ferricytochrome *c* 'outer' (2.0), tuna ferricytochrome *c* 'inner' (2.0), bacterial ferricytochrome c_2 (2.0), bacterial cytochrome c_{550} (2.5), spermwhale metmyoglobin (1.4), bacterial rubredoxin (1.54), bacterial high potential protein (2.0), bacterial ferrodoxin (2.0), subtilisin BPN (2.5), bovine α-chymotrypsin *A* (2.8), bovine chymotrypsinogen *A* (2.5), γ-chymotrypsin *A* (1.9), bovine trypsin (1.9), porcine tosyl elastase (2.5), papain (2.8), bacterial thermolysin (2.3), bovine carboxypeptidase *A* complex (2.0), bovine trypsin–trypsin inhibitor complex (1.9), dog fish lactate dehydrogenase complex (2.8), horse alcohol dehydrogenase complex (2.4), lobster glyceraldehyde-3-P-dehydrogenase

* To whom reprint requests should be addressed.

0141-8130/81/030171-08$02.00

Table 1 Conformational discrepancy index (%) of amino acid residues with respect to that in the far left column

	Ala	Arg	Asn	Asp	Cys	Glu	Gln	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val
N	642	200	336	410	174	352	277	685	171	381	546	547	104	254	313	666	487	136	278	608
Ala	—	37.8	48.2	35.7	41.6	32.2	35.7	60.5	41.2	38.9	35.6	33.1	43.2	38.3	54.8	31.4	38.9	45.6	48.5	40.7
Arg	—	—	45.0	37.1	41.4	41.2	38.2	65.3	38.4	36.5	38.3	36.5	43.2	38.8	61.5	38.9	38.0	47.6	43.3	35.6
Asn	—	—	—	39.7	46.9	45.9	45.4	59.6	47.7	44.3	43.7	44.7	55.7	44.5	65.7	40.1	44.7	46.0	49.3	49.9
Asp	—	—	—	—	44.1	36.6	39.1	62.7	41.4	39.5	39.2	29.7	48.3	38.6	57.5	36.0	38.5	44.8	46.8	44.9
Cys	—	—	—	—	—	48.3	38.7	71.6	40.9	40.1	40.6	41.5	43.2	38.3	66.2	35.3	35.9	40.3	43.5	38.2
Glu	—	—	—	—	—	—	38.3	65.9	42.3	40.7	33.1	36.0	47.6	40.1	59.1	39.7	40.2	48.3	52.7	43.4
Gln	—	—	—	—	—	—	—	67.4	41.1	36.2	31.5	35.1	42.3	35.2	59.2	36.8	33.4	42.1	40.2	35.7
Gly	—	—	—	—	—	—	—	—	70.3	67.9	66.2	61.0	74.0	67.0	76.6	59.1	64.0	71.8	65.7	66.8
His	—	—	—	—	—	—	—	—	—	43.1	39.8	43.5	46.6	43.0	60.9	39.0	39.2	45.5	47.9	45.0
Ile	—	—	—	—	—	—	—	—	—	—	31.2	35.1	41.3	33.2	65.9	36.7	32.6	37.4	37.7	26.0
Leu	—	—	—	—	—	—	—	—	—	—	—	34.3	40.6	33.1	60.5	35.0	32.7	42.6	40.0	33.1
Lys	—	—	—	—	—	—	—	—	—	—	—	—	42.8	34.2	58.1	33.0	35.5	42.6	42.3	38.1
Met	—	—	—	—	—	—	—	—	—	—	—	—	—	40.1	65.0	46.9	42.9	45.1	42.3	38.3
Phe	—	—	—	—	—	—	—	—	—	—	—	—	—	—	61.1	35.6	32.7	40.5	38.5	34.5
Pro	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	58.9	60.7	65.4	66.9	65.7
Ser	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	31.7	43.1	40.8	35.6
Thr	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	39.0	34.6	31.9
Trp	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	45.2	38.9
Tyr	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	33.7
Val	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—

N is the total number of points in (φ,ψ)-plane for each type of amino acid residue

'green' (2.9), bacterial oxidized flavodoxin (1.9), bovine ribonuclease *S* complex (2.0), bacterial nuclease complex (2.0), human Bence–Jones protein (2.0), human immunoglobin *G* 'Fab new' (2.0), jack bean concanavalin *A* (2.4), chicken lysozyme (2.0), chicken triose phosphate isomerase monomer (2.5), carp calcium-binding protein *B* (1.85), human carbonic anhydrase *B* (2.0), human carbonic anhydrase *C* (2.0), human prealbumin dimer (2.5), bacterial semiquinone flavodoxin (1.9).

These data were obtained from AMSOM and were supplied by Richard Fedlmann (National Institutes of Health, USA). Thus (φ,ψ)-values from 38 different proteins having 7567 total amino acid residues were used in this study. The data of each type of amino acid residue were plotted in the (φ,ψ)-plane at a grid interval of 20°. As can be seen from the values of the crystal structure data resolution mentioned, the (φ,ψ)-values obtained are accurate only up to 20°, which prompted us to choose a grid interval of 20° for the plotting. The number of points considered for each amino acid residue are given in the first row of *Table 1*. The number of points in each grid in the (φ,ψ)-plane was normalized. These normalized values are assumed to indicate the probability of occurrence of each conformation. These probability values in percentages are shown in the Appendix. The comparison of normalized (φ,ψ)-map was then carried out. This comparison indicates the main chain conformational similarity among amino acid residues. In order to calculate this quantitatively, the (φ,ψ)-maps of two amino acid residues were compared gridwise and the absolute difference between the probability values of corresponding grids were added up for the whole (φ,ψ)-map to get the discrepancy value between them. Thus, using a simple computer program, the total difference in the probability value for each amino acid residue was calculated by comparing the (φ,ψ)-map of one residue with the (φ,ψ)-map of the remaining residues. The discrepancy values obtained are given as percentages in *Table 1*. The standard deviation associated with each of these discrepancy values was also computed. *Table 1* was first derived using data obtained from the crystal structures of 34 proteins and then extended to include data from 38 proteins. The results of the studies made using these two sets of data have shown that the nature of *Table 1* remained essentially the same even after the addition of (φ,ψ)-data from a few more proteins to the initial set. Thus, results obtained from this study may not get altered even after the addition of (φ,ψ)-data from more proteins and can be assumed to depict the main chain conformational property of amino acid residues.

Results and discussion

A cursory glance at *Table 1* shows that in all cases the least discrepant value of each row is in the range of 25 to 40%, exceptions being Gly and Pro, for which the least discrepant values are more than 50%. This indicates that the probability distributions in the (φ,ψ)-plane for Gly and Pro are distinct from other residues, which means that Gly and Pro take unique main chain conformations in proteins. This is understandable since Gly and Pro have unique side chains. In other words, for all the remaining 18 types of residues there is at least one type of residue whose (φ,ψ)-probability map is similar to the one with which it is compared.

Therefore, if *Table 1* is arranged in order of increasing discrepancy in each row, an idea can be obtained about which amino acid residues are conformationally similar to the residue in the far left column. In order to carry out an objective analysis of this we have adopted the following procedure. The least discrepant value of the row is considered and each residue in the far left column is considered to be conformationally similar to those amino acid residues of the row having discrepancy values between this least value and the value after addition of its standard deviation. Then a confidence limit is set for the discrepancy values of these similar residues, and any other

Table 2 Conformationally similar residues in whole (φ,ψ)-plane

Ala	Ser Glu Lys
Arg	Val Ala Asp Gln His Ile Leu Lys Phe Ser Thr
Asn	Asp Ser
Asp	Lys
Cys	Ser Phe Thr Val
Glu	Ala Leu
Gln	Leu Thr
Gly	—
His	Arg Ala Asp Cys Glu Gln Leu
Ile	Val
Leu	Ile Glu Gln Lys Phe Ser Thr Val
Lys	Asp
Met	Val Ala Arg Cys Gln Ile Leu Lys Phe Thr Tyr
Phe	Leu Gln Ile Lys Ser Thr Val
Pro	—
Ser	Ala Cys Leu Lys Thr
Thr	Ser Gln Ile Leu Phe Tyr Val
Trp	Ile Cys Phe Thr Val
Tyr	Val Thr
Val	Ile

— Indicates that the (φ,ψ)-distribution in the (φ,ψ)-map of the residue in left hand column is distinct from those of the remaining 19 residues

residues whose discrepancy values lie within this limit are also considered conformationally similar to the residue in the far left column. In other words, in each case the probability distribution in the (φ,ψ)-plane of the remaining 19 residues when compared with that of the residue in the far left column of *Table 2*, only residues given in the right hand column are found to have similar probability distribution. Thus, the main chain conformations of amino acid residues given in the far left column of *Table 2*, are similar to the main chain conformations of amino acid residues given in the corresponding rows. It should be mentioned here that the reverse of the above statement, namely the (φ,ψ)-distribution of amino acid residues in the right column in each row is similar to that of (φ,ψ)-distribution of amino acid residue in the far left column, is not generally true. This is illustrated further by considering a few examples from *Table 2*. As can be seen from *Table 2* the probability distribution in (φ,ψ)-plane of Leu when compared with the (φ,ψ)-maps of other residues is similar to that of Ile. However, comparison of the (φ,ψ)-map of Ile with that of other residues indicates that it is similar to that of Val (see *Table 2* and the Appendix). Thus, for residues of the type Leu-Ile the main chain conformational similarity is non-reciprocal. But the reciprocity in the (φ,ψ)-probability maps is exhibited by certain pairs of amino acid residues. They are: Ala–Glu, Ala–Lys, Ala–Ser, Arg–His, Asp–Lys, Cys–Ser, Glu–Leu, Gln–Leu, Gln–Thr, Ile–Val, Leu–Phe, Leu–Thr, Leu–Ser, Phe–Thr, Ser–Thr and Thr–Tyr.

We would like to mention that, as can be seen from the first row of *Table 1*, the total number of data points in the (φ,ψ)-plane considered in the present study are >200 for all amino acid residues except for Arg, Cys, His, Met, and Trp. Hence the results obtained for these residues are subject to higher statistical fluctuations.

The residues similar to Cys are given for academic interest only. In proteins most of the Cys residues are found to form disulphide bridges. Though some residues can take main chain conformations similar to Cys, they can not form disulphide bonds. Met, which is the other sulphur-containing amino acid residue, has (φ,ψ)-probability distribution very different from that of Cys, discrepancy value being (43.2%), indicating that not only the structure of side chain of Met but also its main chain conformations are very different from those of Cys.

Table 2 points out that the two residues which are conformationally similar need not necessarily be similar to each other in chemical or biochemical properties or in their preferences towards secondary structures. This can be illustrated by considering a few examples from *Table 2*. Ala is conformationally similar to Ser; Ala has an aliphatic side chain which is hydrophobic in nature and Ala prefers to exist as an α-helix, while Ser is an hydroxy amino acid which is neutral, polar and prefers to occur in chain reversals. However, when their average main chain conformations are compared, they are very similar. In other words, the replacement of the CH_3 group of the Ala side chain by the CH_2OH group of Ser alters the main chain conformation minimally. A second similar example is that of Asp and Lys, Asp has an acidic side chain and prefers chain reversals while Lys is basic in nature and prefers an α-helical structure. However, the overall (φ,ψ)-distribution of Lys is not very different from that of Asp.

Some of these conformationally similar amino acid residues, though differing in chemical and biochemical nature are found to prefer the same secondary structure. Some examples are: Ala–Glu, Ala–Lys and Gln–Leu (preference for α-helix), Thr–Val, Tyr–Val (preference for β-sheet) and Asn–Asp (both prefer to occur in chain reversals)[15].

Table 2 also indicates that the main chain conformational similarity is not necessarily a direct function of the size and nature of the side chains of amino acid residues. This can be illustrated by Phe and Ile, as can be seen from *Table 2*, Phe is conformationally similar to Ile. The side chains of both these residues are hydrophobic in nature and they both prefer to occur in β-sheets. The only difference is that the side chain of Ile is aliphatic in nature while that of Phe is aromatic.

In a few cases, such as Ile and Val, amino acid residues which are similar chemically, biochemically and with the same preference for secondary structures, are found to be conformationally similar.

From the above analysis one can note that certain amino acid residues, which adopt different secondary structures can take similar main chain conformations. To obtain more insight into this aspect, we have considered two parts of the (φ,ψ)-plane, termed region A and region B. Region A encompasses φ and ψ in the range $-140°$ to $0°$ and $-100°$ to $0°$ respectively. As can be noted, secondary structures α-helix, 3_{10}-helix and chain reversals fall in this region. The second region, namely region B, is defined as one in which φ and ψ vary from $-180°$ to $0°$ and $80°$ to $180°$ respectively. The β-sheet, chain reversals and collagen-type of structures are part of this region. As can be seen from the Appendix the probability of occurrence of observed main chain conformations is very high in these two regions. More than 70% of observed conformations for each residue lie in these two regions, exceptions being Asn and Gly.

Careful analysis of the Appendix indicates that certain residues prefer to occur in region A or region B. Thus if we want to compare the probability distribution values in these heavily populated regions of the (φ,ψ)-map, then these probability values must be renormalized with

Table 3 Conformationally similar residues in region A and region B

Region A	
Ala	Ser Lys
Arg	His Ala Asp Glu Ile Ser Tyr Val
Asn	Ser Asp Gln
Asp	Ser Lys
Cys	Ser
Glu	Gly Leu
Gln	Asp Ala Glu Leu Lys Thr
Gly	Glu
His	Arg Ala Ile
Ile	Val Ala Arg His Lys
Leu	Glu
Lys	Ala Asp Cys Ile Ser Val
Met	Ile Ala Arg Gln Leu Lys Phe
Phe	Lys Ala Asp Gln Gly Ser Tyr
Pro	Lys Ala Arg Asn Asp Gln Phe Ser Thr Val
Ser	Ala Asp
Thr	Gln Asp
Trp	Ile Ala Arg Asp Gln Gly His Lys Phe Ser Thr Tyr Val
Tyr	Arg His Ile Val
Val	Ile Ala Arg Glu Gly Lys

Region B	
Ala	Ser Lys Thr
Arg	Val Asp Lys Thr
Asn	Phe Ile Leu
As	Lys
Cys	Ser Lys Thr Val
Glu	Leu Asn Ala Lys Phe Ser Thr
Gln	Leu Lys Phe Thr
Gly	Ser Ala His Phe Thr Tyr
His	Ser Gln Phe
Ile	Val Thr
Leu	Lys Asn Gln Phe Thr
Lys	Leu Asp Phe Ser Val
Met	Tyr
Phe	Thr Asn Ile Leu Lys Tyr Val
Pro	—
Ser	Thr Ala Cys
Thr	Val Ile Phe Ser
Trp	Ile Asn Thr
Tyr	Met Phe Thr Val
Val	Ile Thr

— Indicates that the (φ,ψ)-distribution in this region of (φ,ψ)-map of the residue in left column is very much distinct from those of the remaining 19 residues

respect to these regions, otherwise an artefact may emerge from the analysis. In other words, we have given proper weighting to observed probability values in the grids forming either region A or region B. The total of these weighted probabilities for both the regions respectively were assumed to be 100. Such weighted probability values are then compared using the method discussed above for region A and region B separately and the results obtained are given in *Table 3*. This table indicates an interesting fact, namely that although the amino acid residue in the far left column prefers a particular secondary structure, those in the right column in the same row need not be the ones which prefer the same secondary structure. This indicates that the (φ,ψ)- distribution of a residue even in this small section of the (φ,ψ)-map is not necessarily a direct function of its preference towards secondary structure.

An examination of *Tables 2* and *3* reveals that for each residue there are some residues similar in region A only, some in region B only but are not similar when the whole (φ,ψ)-planes are compared. Thus Gln is conformationally similar to Leu not only in the whole (φ,ψ)-plane but also in region A and region B, while to Asp in region A only and to Phe in region B only.

One can see from *Table 3* that Gly is conformationally similar to Glu in region A (the discrepancy value being 15.7%). As mentioned earlier, when the whole (φ,ψ)-map for Gly was compared with that of other residues the discrepancy values obtained were very high, particularly in the case of Gly–Glu where the discrepancy value is 65.9% (*Table 1*). But in region A Gly being similar to Glu indicates that a strong α-helix preferrer. Glu, has a (φ,ψ)-distribution similar to that of Gly, which is α-helix indifferent and a strong preferrer of chain reversals. Even though population density for Gly is very much less compared to other residues, the striking similarity in the (φ,ψ)-distribution in this region further strengthens our observation that the α-helix is a combination of continuous chain reversals[16].

Conclusions

Results presented in either *Table 2* or *3* indicate that the main chain conformations of amino acid residues are not direct functions of the nature of their side chains. In fact the analysis of crystal structure data of proteins and of data from the studies made on solutions of polypeptides and proteins clearly indicate that the preferences of amino acid residues towards secondary structures such as α-helix or β-sheet are not directly related to the chemical properties of the residues. Therefore, it should not be surprising that the main chain conformational similarity is also not a function of the chemical property of the amino acid. Our present analysis not only points this out but also indicates that the amino acids which have preferences towards different secondary structures can have similar (φ,ψ)-distribution on average. This is mainly because although a particular amino acid residue might have highest preference for a particular secondary structure, it has considerable probability to exist in other secondary structures as well as in the coil region. In other words, the amino acid residue in the (φ,ψ)-plane will have considerable probability in the allowed regions other than those specified for secondary structures. Thus the observations made earlier that certain residues are conformationally similar though their preferences towards secondary structure differ, clearly indicate that main chain conformational similarity is as much an independent property of a residue as its preference for a particular secondary structure.

We would also like to point out that the replacements carried out using *Table 2* will alter the main chain conformation of polypeptides minimally. Thus in drug design one should not only take into consideration the biochemical and chemical properties of side chains of amino acid residues, but also should give importance to the conformational similarity among the residues. In the next part of the series we have attempted to show how this analysis is also useful in explaining the changes in the primary structures which have occurred during evolution in homologous proteins.

References

1 Chou, P. Y. and Fasman, G. D. *Biochemistry* 1974, **13**, 211
2 Lim, V. I. *J. Mol. Biol.* 1974, **88**, 857
3 Ptitsyn, O. B. and Finkelstein, A. V. *Biophysics* 1970, **15**, 785
4 Burgess, A. W., Ponnuswamy, P. K. and Scheraga, H. A. *Israel J. Chem.* 1974, **12**, 239
5 Wu, T. T., Szu, S. C., Jernigan, R. L., Bilfosky, H. and Kabat, E. A. *Biopolymers* 1978, **17**, 555
6 Wu, T. T. and Kabat, E. A. *Proc. Natl Acad. Sci. USA* 1971, **68**, 1501
7 Kabat, E. A. and Wu, T. T. *Proc. Natl Acad. Sci USA* 1973, **70**, 1473
8 Robson, B. and Suzuki, E. *J. Mol. Biol.* 1976, **107**, 327
9 Bunting, J. R., Athey, T. W. and Cathou, R. E. *Biochim. Biophys. Acta* 1972, **285**, 60
10 Tanaka, S. and Scheraga, H. A. *Macromolecules* 1976, **9**, 168
11 Bourgeios, S., Jernigan, R. L., Szu, S. C., Kabat, E. A. and Wu, T. T. *Biopolymers* 1979, **18**, 2625
12 Ponnuswamy, P. K. and Sasisekharan, V. *Biopolymers* 1971, **10**, 565
13 Ramachandran, G. N. in 'Conformation of Biological Molecules and Polymers' (Ed. E. D. Bergman and B. Pullman), Israel Acad. of Sci. and Humanities, Jerusalem, 1973, pp 1-11
14 IUPAC-IUB Commission on Biochemical Nomenclature, *J. Mol. Biol.* 1970, **52**, 1
15 Levitt, B. *Biochemistry* 1978, **17**, 4277
16 Kolaskar, A. S., Ramabrahmam, V. and Soman, K. V. *Int. J. Peptide Proteins Res.* 1980, **16**, 1

Appendix

The following section contains the main chain conformation probability values (approximated to first decimal) as percentages, obtained from crystal structure data of 38 globular proteins. The regions A and B comprise φ-variations between $-140°$ and $0°$, $-180°$ and $0°$ respectively and ψ-variations between $-100°$ and $0°$ and $80°$ and $180°$ respectively, are also marked to show that maximum probability distribution of (φ,ψ)-map lie in these regions.

Ala

ψ \ φ	-180	-160	-140	-120	-100	-80	-60	-40	-20	0	20	40	60	80	100	120	140	160	180
180	0.3	0.6	0.3	0.3	0.2	0.2	0	0.2	0	0.2	0.2	0	0	0.2	0	0	0.2	0.6	
160	0.2	0	0	0.2	0	0	0	0	0	0	0	0	0.2	0	0	0	0	0	
140	0.2	0	0	0	0	0	0	0	0.2	0	0	0	0	0	0	0	0	0	
120	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
80	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
-20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
-40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
-60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
-80	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
-100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
-120	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
-140	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
-160	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
-180	0.3	0.6	0.3	0.3	0.2	0.2	0	0.2	0	0.2	0.2	0	0	0.2	0	0	0.2	0.6	

Arg

ψ \ φ	-180	-160	-140	-120	-100	-80	-60	-40	-20	0	20	40	60	80	100	120	140	160	180
180	0.5	0.5	2.5	1.0	1.5	0	1.5	0	0	0	0	0	0	0	0.5	0	0	0	
160	0	2.0	2.5	4.0	1.0	0.5	0.5	0.5	0	0	0	0	0.5	0	0	0	0	0	
140	0.5	1.0	1.0	1.5	4.5	2.0	0	0	0	0	0	0	0	0	0	0.5	0	0	
120	0	0.5	1.0	2.5	1.5	1.5	0.5	0.5	0	0	0	0	0	0	0	0	0	0.5	
100	0	0	0.5	0.5	0.5	1.0	0.5	0	0	0	0	0	0	0	0	0	0	0	
80	0	1.0	0.5	0	0	0.5	0	0	0	0.5	0	0	0.5	0	0	0	0	0	
60	0	0.5	0.5	0	0	0	0	0	0	0	0.5	0.5	0	0	0	0	0	0	
40	0	0	0	1.0	1.0	0	0	0	0	0	0	0	0.5	1.0	0	0	0	0	
20	0	0	0.5	0.5	0	0	0.5	0	0	0	0	0.5	0.5	0	0	0	0	0	
0	0	0.5	1.5	1.0	2.0	3.0	1.0	0	0	0	0	0	0	0	0	0	0	0	
-20	0	0	0	0.5	1.5	4.0	2.0	0.5	0	0	0	0	0	0	0	0	0	0	
-40	0	0	0.5	0.5	1.0	7.0	7.0	2.0	0	0	0	0	0	0	0	0	0	0	
-60	0.5	0.5	0	0	0	1.5	3	0.5	0.5	0	1.0	0	0	0	0	0	0	0	
-80	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
-100	0	0	0	0	0	0	0	0.5	0	0	0.5	0	0	0	0	0	0	0	
-120	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0	0	
-140	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
-160	0.5	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0.5	0.5	0	

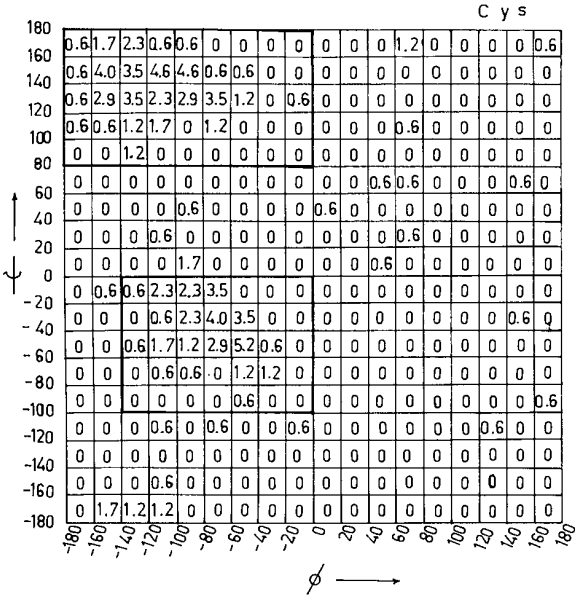
Asn

ψ \ φ	-180	-160	-140	-120	-100	-80	-60	-40	-20	0	20	40	60	80	100	120	140	160	180
180	0.3	1.2	0.3	0.9	0.9	0	0	0	0	0	0	0	0.3	0	0	0	0	0	
160	0	0	1.2	0	0.3	0	0	0	0	0	0	0	0.6	0	0	0	0	0	
140	0.3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
120	0.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
100	0	0	0.3	0	0	0.3	0	0.3	0.3	0	0	0	0	0	0	0	0	0	
80	0	0	0	0	0	0.3	0.3	0.3	0	0	0.3	0.3	0	0.3	0	0	0	0	0.3
60	0	0	0	0	0	1.5	1.2	1.2	0.6	0	0	0	0	0.3	0	0	0	0	
40	0	0	0	0.6	0.6	2.1	2.1	0.9	0	0	0	0	0	0	0	0	0	0	
20	0.6	0	0.3	0	2.7	2.7	2.7	0.3	0	0	0	0	0	0	0	0	0.3	0	
0	0.6	0.3	0.6	0.6	2.7	2.4	0.9	0.3	0	0	0	0	0	0	0	0	0	0	
-20	0	0	0.9	0.6	2.7	2.4	0.9	0.3	0	0	0	0	0	0	0	0	0	0	
-40	0	0	0.3	0.3	0.6	0.6	0.6	0	0	0	0	0	0	0	0	0	0	0	
-60	0	0	0.3	0.3	0.6	0.6	0.6	0	0.3	0	0	0	0	0	0	0	0	0	
-80	0.3	0.6	0.3	0.3	0.2	0.2	0	0.2	0	0.2	0.2	0	0	0.2	0	0	0.2	0.6	
-100	0.3	0	0	0.2	0	0	0	0	0	0	0	0	0.2	0	0	0	0	0	
-120	0	0.2	0	0	0	0	0	0	0.2	0	0	0	0	0	0	0	0	0	
-140	0.2	0	0	0.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
-160	0.3	0.6	0.3	0.3	0.2	0.2	0	0.2	0	0.2	0.2	0	0	0.2	0	0	0.2	0.6	

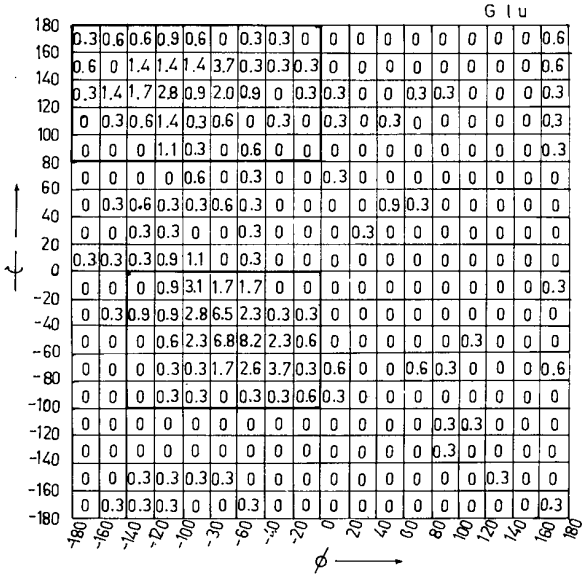
Asp

ψ \ φ	-180	-160	-140	-120	-100	-80	-60	-40	-20	0	20	40	60	80	100	120	140	160	180
180	0.5	0.7	0.5	0.2	0.7	0.2	1.0	0	0	0	0	0	0	0	0	0	0.2	0	
160	0	0.5	0.7	2.0	0.7	1.0	0.7	0.2	0	0	0	0	0	0	0.2	0	0	0	
140	0.2	0.2	1.5	2.4	2.4	0.7	1.0	0	0.2	0	0	0	0	0.5	0	0	0	0	
120	0	0.5	0.5	1.2	1.2	1.0	0.2	0.5	0	0	0	0	0.5	0	0	0	0	0	
100	0	0.7	0.2	0.5	1.2	1.5	0.5	0	0.2	0.2	0.2	0.2	0	0	0	0	0	0	
80	0	0.7	1.0	0.2	1.0	0.5	0	0	0	0	0	1.0	0.2	0	0.2	0	0.5	0	
60	0	0	0.7	0.7	0.5	0	0	0	0	0	0.5	0.5	0.2	0.2	0	0.5	0	0.2	
40	0.5	0	0.7	0.5	0.7	0.2	0.2	0	0	0	0	0	0.5	0	0	0	0.2	0	
20	0.2	1.0	0.5	1.0	1.2	1.0	0.2	0	0	0	0	0.2	0	0.2	0	0	0	0	
0	0	0	0.5	1.7	3.9	2.4	1.0	0	0	0	0	0.2	0	0	0.2	0	0	0	
-20	0.2	0	0	0.2	2.4	3.9	3.2	0.5	0	0	0	0	0	0	0	0	0	0	
-40	0	0.2	0	0	1.7	4.9	5.6	2.0	0	0	0	0.2	0	0	0	0	0	0.2	
-60	0	0	0	0.5	0.7	0.7	3.2	2.0	0	0.2	0.2	0.2	0.2	0	0	0	0	0	
-80	0	0.2	0	0	0	0.2	0	0.7	0.5	0.2	0	0	0	0	0	0.2	0	0	
-100	0	0	0	0	0	0	0.2	0	0	0	0	0	0	0	0	0	0	0	
-120	0	0.2	0	0	0.2	0	0	0	0	0	0.2	0	0	0	0	0	0	0	
-140	0	0.2	0	0.2	0.2	0	0	0	0	0	0	0	0	0	0	0	0	0.2	
-160	0.2	0.7	0.2	1.2	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	
-180																			

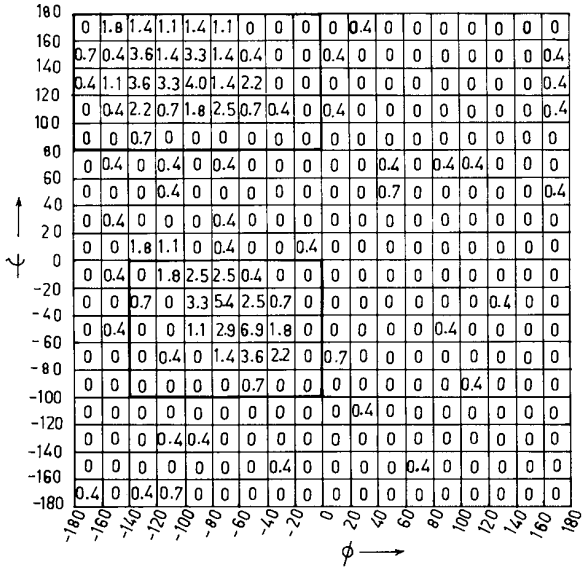
Cys



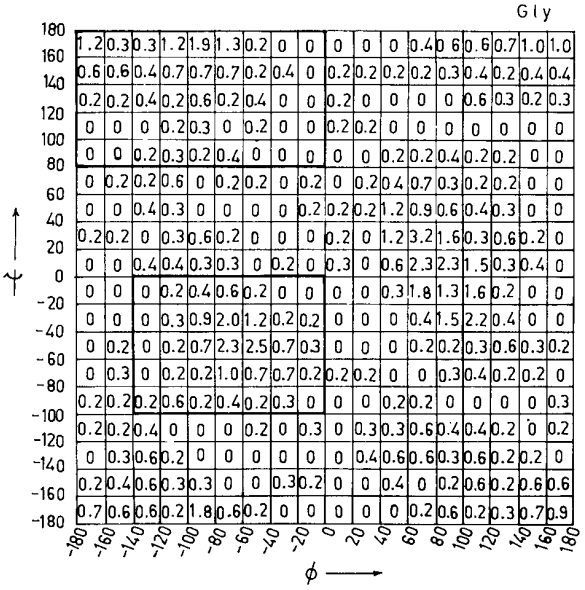
Glu



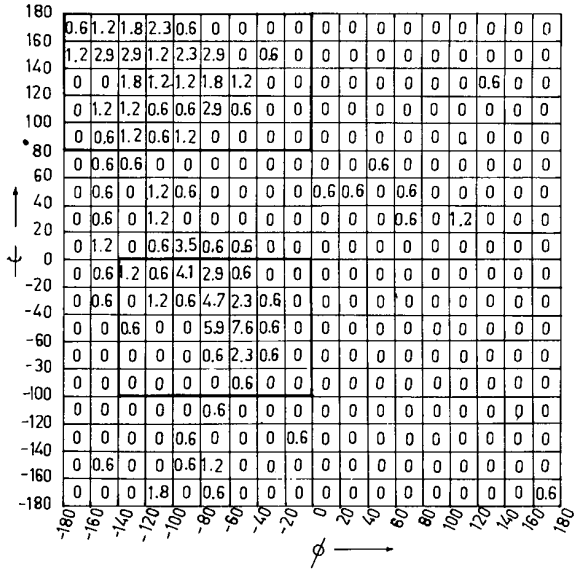
Gln



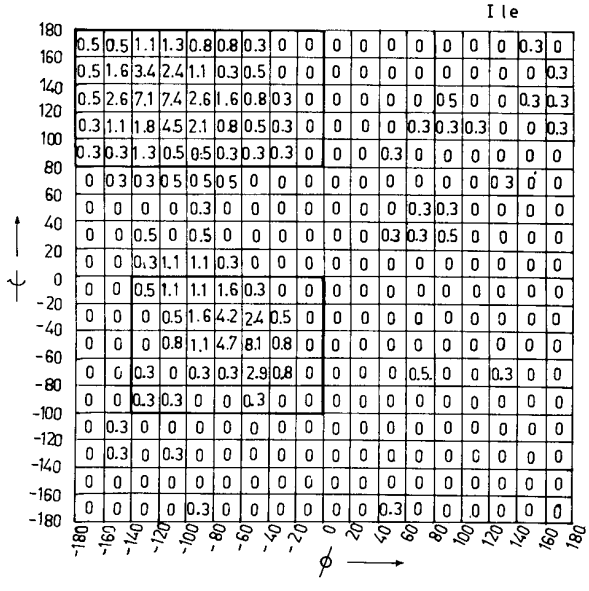
Gly

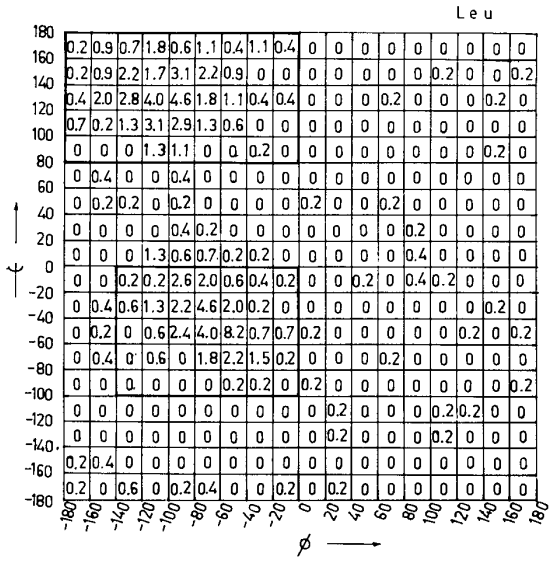


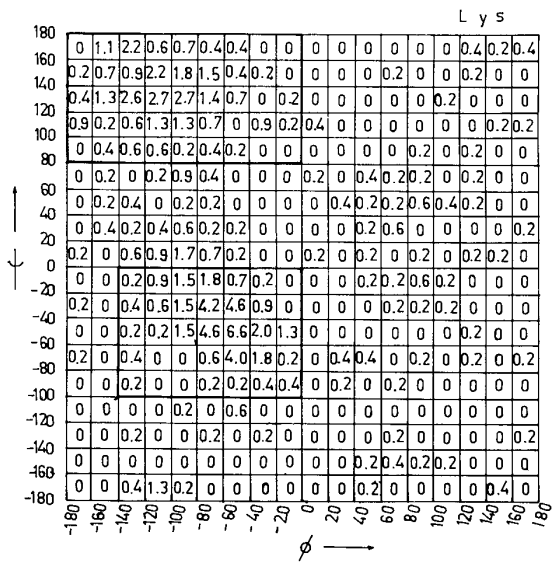
His

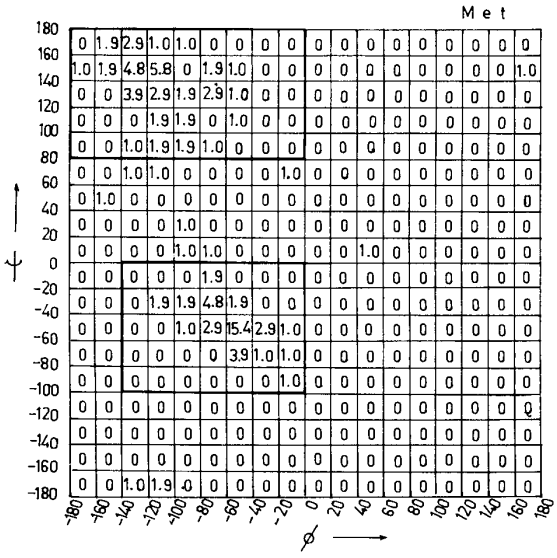


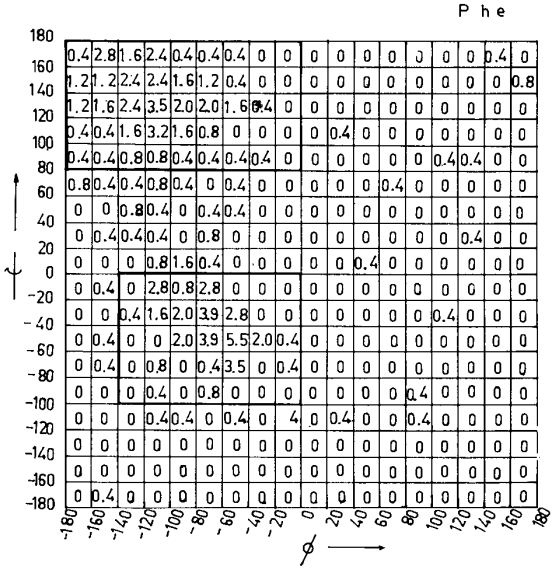
Ile

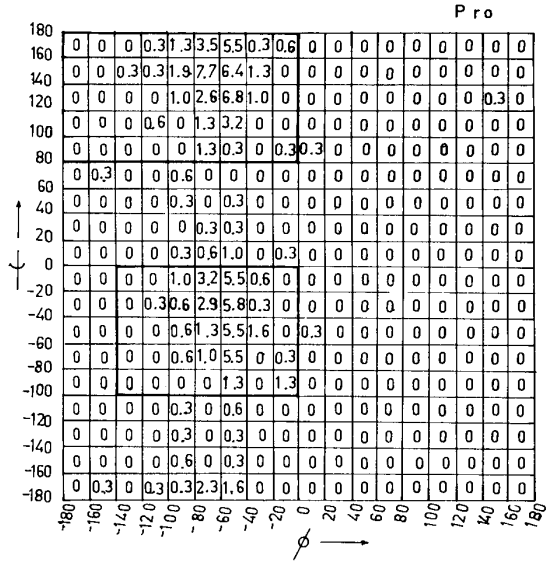


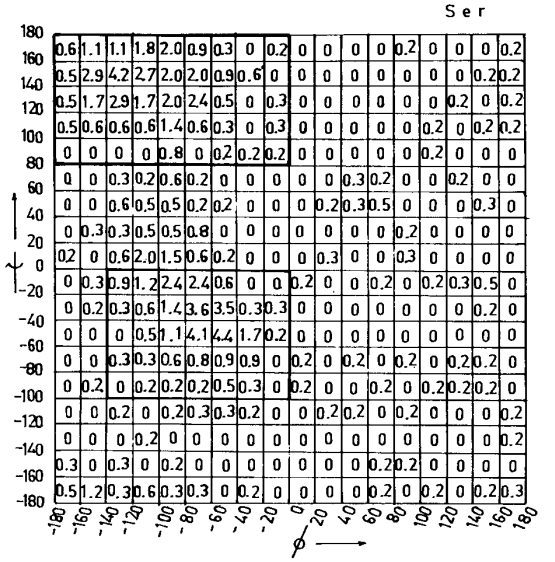

Leu
ψ
φ


Lys
ψ
φ


Met
ψ
φ


Phe
ψ
φ


Pro
ψ
φ


Ser
ψ
φ

Thr

ψ \ φ	-180	-160	-140	-120	-100	-80	-60	-40	-20	0	20	40	60	80	100	120	140	160	180
180	0.4	0.8	2.1	2.3	2.3	1.4	0.4	0	0	0	0	0.2	0	0	0	0	0	0.2	
160	0.4	1.9	2.7	4.1	1.9	1.2	0.4	0.2	0	0	0	0	0	0.6	0	0	0	0	
140	0.2	2.1	3.7	5.5	2.5	1.9	1.2	0	0	0	0	0	0	0.2	0	0	0	0.2	
120	0	0.6	1.6	1.9	1.4	0.6	0.4	0.4	0.2	0	0.2	0.2	0	0	0	0	0	0.2	
100	0.2	0	0	0.4	0.4	0.2	0	0	0	0	0	0	0.2	0.2	0	0	0	0	
80	0	0.2	0	0	0.2	0	0	0	0	0.2	0.2	0	0	0	0	0	0	0	
60	0.2	0.2	0.2	0.4	0.2	0	0	0	0	0	0	0	0	0	0	0	0	0.2	
40	0	0	0.2	0.6	0	0.2	0.2	0	0	0	0	0	0.2	0	0	0	0	0	
20	0	0	0.8	1.2	1.4	0.2	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0.2	0	1.2	3.7	2.3	0.4	0	0	0	0	0	0.2	0	0	0	0	0	
-20	0	0.6	0.2	1.6	2.1	4.7	1.9	1.0	0	0	0	0	0	0.2	0.2	0	0	0	
-40	0	0.4	0.4	0.2	0.4	1.6	4.7	0.6	0	0	0	0.2	0	0	0	0.2	0	0	
-60	0.2	0.2	0.2	0.4	0.6	0.8	3.3	0.4	0	0	0.2	0	0	0.2	0	0	0	0	
-80	0	0	0.2	0	0	0	0.2	0.2	0	0	0	0	0	0	0	0	0	0	
-100	0	0	0	0.4	0	0	0.2	0	0	0	0	0	0	0	0	0	0	0.2	
-120	0	0	0	0	0.2	0	0	0	0	0	0	0	0	0.2	0	0	0	0	
-140	0	0	0.4	0.4	0.4	0	0	0	0	0	0	0	0	0	0	0.2	0	0	
-160	0	0.6	0.2	0.8	0.4	0.6	0	0	0	0	0.2	0	0	0	0	0	0	0.2	
-180																			

Trp

ψ \ φ	-180	-160	-140	-120	-100	-80	-60	-40	-20	0	20	40	60	80	100	120	140	160	180
180	1.5	0	0.7	0.7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
160	0.7	1.5	2.2	2.9	1.5	1.5	1.5	0	0	0	0	0	0	0	0	0	0	0	
140	0.7	2.9	5.2	5.9	1.5	0.7	2.2	0	0	0	0	0	0	0	0	0	0	0	
120	0	1.5	0.7	1.5	0	2.2	0.7	0	0	0	0	0	0	0	0	0	0	0	
100	0	0	2.2	1.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
80	0	0	0.7	0.7	0	0	0	0	0	0.7	0	0	0	0	0	0	0	0	
60	0	0	0	0	0	0	0.7	0	0	0	0	0	0	0	0	0	0	0.7	
40	0.7	1.5	0	2.2	1.5	0.7	0	0	0	0	0	0	0	0	0	0	0	0	
20	0	0	0.7	0	1.5	0.7	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0.7	0.7	0.7	2.2	0	0	0	0	0	0	0	0	0	0	0	0	
-20	0	0	0	1.5	2.9	5.2	4.4	0.7	0	0	0	0	0	0	0	0	0	0	
-40	0	0	0.7	0	1.5	4.4	3.7	0	0	0	0	0	0	0.7	0	0	0	0	
-60	0	0	0	0	0	0	2.9	0.7	0	0	0	0.7	0	0	0	0	0	0	
-80	0	0	0	0	0	0	0	0.7	0	0	0	0	0	0	0	0	0	0	
-100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
-120	0	0	0.7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
-140	0	0.7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
-160	0	2.2	2.2	0.7	0.7	0	0	0	0	0	0	0	0	0	0	0	0	0	
-180																			

Tyr

ψ \ φ	-180	-160	-140	-120	-100	-80	-60	-40	-20	0	20	40	60	80	100	120	140	160	180
180	0	2.9	2.2	0.7	1.8	1.1	0	0	0	0	0	0	0	0	0	0	0	0.4	
160	2.5	1.4	3.6	6.1	2.5	0.7	0.4	0	0	0	0	0	0	0	0.4	0	0.4	0	
140	1.1	0.4	4.7	3.2	1.1	1.8	0.7	0	0	0	0	0	0	0	0	0	0	0	
120	0	0	0.4	3.6	1.8	0.4	1.1	0	0	0	0	0	0	0	0	0	0	0.4	
100	0.4	0	0	1.8	1.1	1.4	0	0	0	0	0	0	0.4	0	0	0.4	0	0	
80	0	0	0	0.4	1.1	0	0	0.4	0	0	0	0	0	0.4	0	0	0	0	
60	0.4	0	0.7	0	0	0	0	0	0	0	0.4	0	0.4	0	0	0	0	0	
40	0	0.4	0	0.7	0.4	0	0	0	0	0	0	0.7	0.4	0	0	0	0	0	
20	0	0.4	1.1	1.4	0.7	0.4	0.4	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	1.4	1.4	1.1	1.1	0	0	0	0	0	0	0.4	0.4	0	0	0	0	
-20	0.4	0	0.4	1.1	0.7	2.2	2.2	0.4	0	0	0	0	0	0	0	0	0	0	
-40	0.4	0.4	0.4	0.4	0.4	3.2	4.3	0.4	0	0	0	0	0	0	0	0	0	0	
-60	0	0.4	0.4	0.4	0	0.7	1.8	0.4	0	0	0	0	0	0	0	0	0	0	
-80	0	0	0	0	0	0	0	0.4	0	0	0	0	0	0	0	0	0	0	
-100	0.4	0	0	0.4	0	0	0	0	0	0	0	0.4	0	0	0	0	0	0	
-120	0	0	0.4	1.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
-140	0	0	0	0	0.4	0.4	0	0	0	0	0	0	0	0	0	0	0	0.4	
-160	0	0	0	1.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
-180																			

Val

ψ \ φ	-180	-160	-140	-120	-100	-80	-60	-40	-20	0	20	40	60	80	100	120	140	160	180
180	0.2	1.2	3.3	1.3	1.0	0	0.2	0	0	0	0	0	0	0	0	0	0	0	
160	0.5	2.5	4.4	4.4	1.8	0.7	0.3	0	0	0	0	0	0.3	0.2	0	0	0	0.3	
140	0.7	2.0	4.3	6.4	3.6	2.0	0.7	0.3	0	0	0	0	0	0.3	0.2	0.2	0	0	
120	0	1.3	2.8	5.4	1.8	0.7	0.3	0	0	0	0	0	0	0	0	0	0	0	
100	0	0.2	0.8	1.6	0.5	0.7	0	0.2	0.2	0	0	0	0.2	0.2	0	0	0	0	
80	0.2	0	0.2	0.3	0.2	0	0	0	0.2	0	0	0	0	0.2	0	0	0.2	0	
60	0	0	0	0.3	0	0	0	0	0	0	0	0.2	0	0	0	0	0	0	
40	0	0	0	0.3	0.3	0.2	0.2	0	0	0	0	0	0	0	0	0	0	0	
20	0	0	0.8	0.2	0	0.3	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0.7	0.3	0.8	1.2	0.6	0.2	0	0	0	0	0	0	0	0	0	0.3	
-20	0	0	0.2	0.3	0.7	4.4	2.5	0.5	0	0	0	0	0	0	0	0	0	0	
-40	0	0	0	0.7	2.1	4.4	5.3	1.3	0.2	0	0	0	0	0.2	0	0	0	0	
-60	0	0	0.2	0.2	0.2	1.2	2.5	0.5	0	0	0	0	0	0	0	0	0.2	0	
-80	0	0	0	0	0.2	0.3	0.3	0.2	0	0.2	0	0	0	0	0.2	0	0	0	
-100	0.2	0	0.2	0	0	0	0	0	0	0	0.2	0	0	0	0	0	0	0	
-120	0	0	0	0.2	0	0.2	0	0	0	0	0	0	0	0.2	0	0	0.2	0	
-140	0	0	0.2	0.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
-160	0.2	0.2	0.7	0.2	0	0	0	0.2	0	0	0	0	0	0.2	0	0	0.2	0	
-180																			